Gated rotation mechanism of site-specific recombination by ϕC31 integrase

Femi J. Olorunniji[a], Dorothy E. Buck[b], Sean D. Colloms[a], Andrew R. McEwan[c], Margaret C. M. Smith[c,1], W. Marshall Stark[a,2], and Susan J. Rosser[a,2]

[a]Institute of Molecular, Cell, and Systems Biology, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom; [b]Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom; and [c]Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, Scotland, United Kingdom

Edited by Arthur Landy, Brown University, Providence, RI, and approved October 12, 2012 (received for review June 27, 2012)

Integrases, such as that of the *Streptomyces* temperate bacteriophage ϕC31, promote site-specific recombination between DNA sequences in the bacteriophage and bacterial genomes to integrate or excise the phage DNA. ϕC31 integrase belongs to the serine recombinase family, a large group of structurally related enzymes with diverse biological functions. It has been proposed that serine integrases use a "subunit rotation" mechanism to exchange DNA strands after double-strand DNA cleavage at the two recombining *att* sites, and that many rounds of subunit rotation can occur before the strands are religated. We have analyzed the mechanism of ϕC31 integrase-mediated recombination in a topologically constrained experimental system using hybrid "*phes*" recombination sites, each of which comprises a ϕC31 *att* site positioned adjacent to a regulatory sequence recognized by Tn*3* resolvase. The topologies of reaction products from circular substrates containing two *phes* sites support a right-handed subunit rotation mechanism for catalysis of both integrative and excisive recombination. Strand exchange usually terminates after a single round of 180° rotation. However, multiple processive "360° rotation" rounds of strand exchange can be observed, if the recombining sites have nonidentical base pairs at their centers. We propose that a regulatory "gating" mechanism normally blocks multiple rounds of strand exchange and triggers product release after a single round.

The large serine recombinase ϕC31 integrase (605 amino acids) promotes integration of the bacteriophage ϕC31 genome into the *Streptomyces* host chromosome by recombination between a phage site *attP* and a bacterial site *attB*, resulting in an integrated prophage flanked by two recombinant sites, *attL* and *attR* (1, 2). The prophage can remain dormant for many generations in this lysogenic state until it is excised by integrase-mediated recombination between *attL* and *attR*, reforming *attP* and *attB* sites. The *att* sites (~50 bp) each comprise sequences recognized by integrase flanking a central 2-bp overlap sequence at which crossover occurs (3–5) (Fig. 1*C* and Fig. S1). Two integrase subunits bind to each *att* site, forming on-site dimers. Two *att* sites that are to recombine are then brought together by dimer–dimer interactions. ϕC31 integrase and related serine integrases catalyze efficient recombination between *attP* and *attB*, but not between the prophage sites *attL* and *attR*, or any other pairs of sites. However, in the presence of a phage-encoded recombination directionality factor, integrase specificity is altered; *attL* × *attR* recombination is efficient and *attP* × *attB* recombination is inhibited (4, 6, 7). There is considerable interest in this group of recombinases as tools for applications in biotechnology and genetic manipulation (8).

Previous studies on the mechanism of recombination by small serine recombinases have led to a "subunit rotation" model for strand exchange (9–11). In this model, cleavage of all four DNA strands in a synaptic complex of the two recombining sites creates double-strand breaks with a recombinase subunit covalently attached to each DNA 5′ end. A 180° rotation of one pair of half-sites relative to the other pair then positions the half-sites in a recombinant configuration, and the strands are religated (Fig. 1*A*). The sequence and structural similarities between the N-terminal domains of the small serine recombinases and the much larger serine integrases (2, 12) suggest that the serine integrases might also use a subunit rotation mechanism of strand exchange.

An important implication of the subunit rotation model is that iteration of the mechanism should be feasible; the configurations of the cleaved synaptic complex before and after one "round" of 180° rotation are predicted to be similar, permitting further rounds of rotation. Experimental analysis of iteration product topologies provided crucial evidence in support of subunit rotation by small serine recombinases (9, 13–17). Recently, data consistent with extensive repeated rounds of subunit rotation by the serine integrase Bxb1 were obtained by single-molecule and gel electrophoretic analysis (18). It was concluded that the synaptic intermediate containing two cleaved *att* sites and Bxb1 integrase can remain rotationally open for many minutes, allowing many rounds of rotation to take place. However, the biological functions of site-specific recombinases typically require a single round of reaction, yielding recombinants. Repeated rounds of reaction and/or persistently cleaved intermediates could be disadvantageous, leading to genetic instability. Previous in vitro analysis of ϕC31 integrase activity is consistent with single-round reactions and efficient ligation of recombinant sites (19). We therefore set out to design a system that would allow us to detect single or repeated rounds of ϕC31 integrase-mediated recombination by topological analysis of the reaction products, and thus determine whether the mechanism is gated (that is, preferentially terminates after a single round of strand exchange) or ungated, involving a rotationally open intermediate.

Results

Hybrid *phes* Sites for Topologically Selective ϕC31 Integrase-Mediated Recombination. Determination of the precise topological changes accompanying strand exchange by serine integrases is challenging because, in contrast to small serine recombinases such as Tn*3* resolvase, synapsis of sites by these enzymes is not topologically selective. Random collision of two *att* sites within a supercoiled substrate molecule leads to many different synapse topologies, and thus a complex mixture of product topologies (Fig. 1*B*) (19, 20). To restrict the topology of the ϕC31 system, we made

Author contributions: F.J.O., D.E.B., S.D.C., W.M.S., and S.J.R. designed research; F.J.O. performed research; A.R.M. and M.C.M.S. contributed new reagents/analytic tools; F.J.O., D.E.B., S.D.C., W.M.S., and S.J.R. analyzed data; and F.J.O., S.D.C., W.M.S., and S.J.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]Present address: Department of Biology, University of York, York YO10 5DD, United Kingdom.

[2]To whom correspondence may be addressed. E-mail: Marshall.Stark@glasgow.ac.uk or Susan.Rosser@glasgow.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1210964109/-/DCSupplemental.

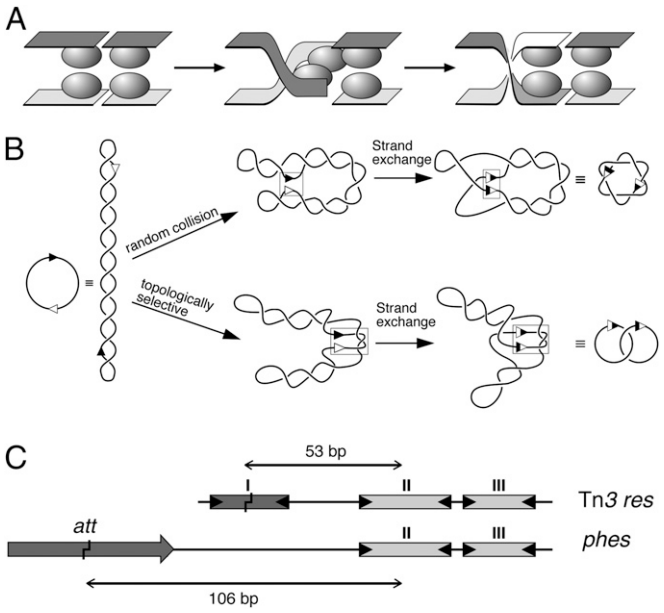

Fig. 1. Mechanism and topology of site-specific recombination by ϕC31 integrase and other serine recombinases. (*A*) Proposed subunit rotation mechanism for strand exchange by serine recombinases. The cleaved double-stranded DNA is represented by ribbons and the recombinase subunits by balls. The cartoon shows right-handed 180° rotation, starting from a nonrecombinant configuration (*Left*) and proceeding to recombinant (*Right*). (*B*) Pathways for synapsis and recombination in a supercoiled plasmid substrate. The upper pathway illustrates recombination following synapsis of two recombination crossover sites (represented by arrowheads) by random collision (only one of many possible synapse topologies is shown), leading to topologically complex products (in this example, a six-noded catenane). The lower pathway illustrates topologically selective synapsis, which in the case of the Tn*3 res* accessory sites traps three negative nodes (boxed). A specific mechanism of strand exchange at the crossover sites leads to a single product topology (here, right-handed rotation as shown in *A* gives a two-noded catenane). In the Tn*3* system, the arrowheads correspond to *res* binding site I; in the hybrid *phes* system described here, they correspond to *att* sites for ϕC31 integrase. (*C*) Structures of Tn*3 res* and *phes* recombination sites. The boxes represent binding sites for resolvase dimers or an integrase dimer (*att* site). The center-center distance between *res* sites I and II is indicated, as is the corresponding *att*-site II distance in the optimal *phes* sites. Full sequences of *phes* sites and further details are in Fig. S1.

substrates containing a pair of hybrid "*phes*" recombination sites. A *phes* site comprises a ϕC31 *att* site and the accessory sites II and III of *res* (the recombination site for Tn*3* resolvase), separated by a spacer that can be varied in length or sequence (Fig. 1*C* and Fig. S1). When resolvase is added to a supercoiled plasmid substrate containing two compatible *phes* sites (*phesP* × *phesB* or *phesL* × *phesR*) in direct repeat (Fig. 2*A*), the *res* accessory sites are predicted to synapse and intertwine in a topologically selective manner, trapping three negative interdomainal nodes between the *att* sites when they are subsequently synapsed on addition of ϕC31 integrase (Fig. 1*B*) (21). In preliminary experiments, we determined a suitable length for the spacer DNA between the *att* and *res* components of *phes*. If this spacer is too short, or if the helical phase of the *res* and *att* components is inappropriate, resolvase-mediated synapsis of the *res* accessory sites might block integrase-mediated intramolecular synapsis of the *att* sites (21). We found that recombination was efficient, predominantly intramolecular, and topologically specific (as described in detail below) when the length of DNA between the centers of *att* and *res* site II was 106 bp (10 helical turns) (Fig. S2). This spacer length was therefore used in all of the *phes* sites for our subsequent experiments.

Topology of *phes* × *phes* Recombination. When a plasmid substrate with two $phes^{TT}$ sites ($phesP^{TT} \times phesB^{TT}$) (superscripts denote the overlap sequences) was treated with ϕC31 integrase in the absence of Tn*3* resolvase, we observed a complex mixture of product topological species (Fig. 2*B*, lane 3). Treatment of the products with DNase I removes supercoiling, allowing electrophoretic separation of the nicked product species according to the number of topological nodes (13–17) (Fig. 2*B*, lane 7). The main products were thus seen to be recombinant DNA circles, and catenanes comprising the two recombinant circles linked with even numbers of nodes (2, 4, 6…), as predicted if synapsis of sites is by random collision, followed by a single round of strand exchange by a rotation mechanism (Fig. 1*B*, and Figs. S3–S8) (20). Digestion with a restriction enzyme showed that most of the products were recombinants (Fig. 2*B*, lane 11).

When the $phesP^{TT} \times phesB^{TT}$ substrate was preincubated with Tn*3* resolvase for 10 min before addition of integrase, recombination efficiency (as visualized in the restriction digest) was just as high as it was without resolvase, but the product topology was strikingly different (Fig. 2*B*, lanes 4, 8, and 12). The predominant product species was a two-noded catenane, and formation of topologically complex products was much reduced. The two-noded catenane is the predicted product if topologically selective synapsis by resolvase (trapping three negative nodes) is followed by a single round of ϕC31 integrase-mediated strand exchange by a right-handed rotation mechanism (Fig. 1*B*). The two-noded catenane accumulated during a time course experiment (Fig. 2*C*; for comparison, Fig. S4 shows an analogous time course in the absence of resolvase). The complex catenane products (labeled "cp" in Fig. 2*B*) formed in the presence of resolvase may indicate a minor fraction of the substrate that has undergone repeated rounds of strand exchange (see *Discussion*). Products with double-strand breaks at one or both *att* sites were not detected in these experiments.

To investigate the apparent distinctions between our results with ϕC31 integrase and results of analysis of the Bxb1 integrase system (18), we performed an analogous series of experiments using Bxb1 integrase and a plasmid substrate identical to that used in Fig. 2*B*, except that the ϕC31 *attP* and *attB* sites were substituted by those of Bxb1. Our results with the two systems were very similar; in both cases the major product was the two-noded catenane when resolvase was added before the integrase, consistent with a single round of strand exchange (Fig. S6).

The ϕC31 integrase, like other serine recombinases, makes double-strand breaks at the center of each *att* site, leaving half-sites with 2-nt extensions at their 3′ ends. Ligation of the ends to form recombinants takes place only if these 3′ extensions can base pair with each other. Natural *att* sites have an asymmetric "overlap sequence" at the central 2 bp, and as a result, the two half-sites formed upon cleavage by integrase are different (Fig. 2*D*). Only complementary half-sites can be religated. Synapsis of an *attP* and an *attB* site can occur in either "parallel" or "antiparallel" orientations (19, 22) (Fig. 2*D*). Exchange of half-sites by rotation following parallel synapsis will align complementary ends that can form recombinants, but exchange following antiparallel synapsis will align noncomplementary half-sites that cannot ligate (Fig. 2*D*). However, if the overlap sequence has twofold symmetry (for example, TA/TA), parallel and antiparallel synaptic alignments can both yield base-paired recombinants. Therefore, as expected, a very complex mixture of knots and catenanes was obtained when a $phesP^{TA} \times phesB^{TA}$ substrate was treated with integrase in the absence of resolvase (Fig. 2*B*, lane 14) consisting of resolution and inversion products as well as nonrecombinant molecules (Fig. 2*A*; Fig. 2*B*, lane 17). In contrast, when resolvase was added before integrase, inversion was suppressed, and the predominant product was a two-noded catenane (Fig. 2*B*, lanes 15 and 18), just as for the $phesP^{TT} \times phesB^{TT}$ reaction (Fig. 2*B*, lanes 8 and 12). We conclude that synapsis of the *phes* accessory sites by resolvase specifies

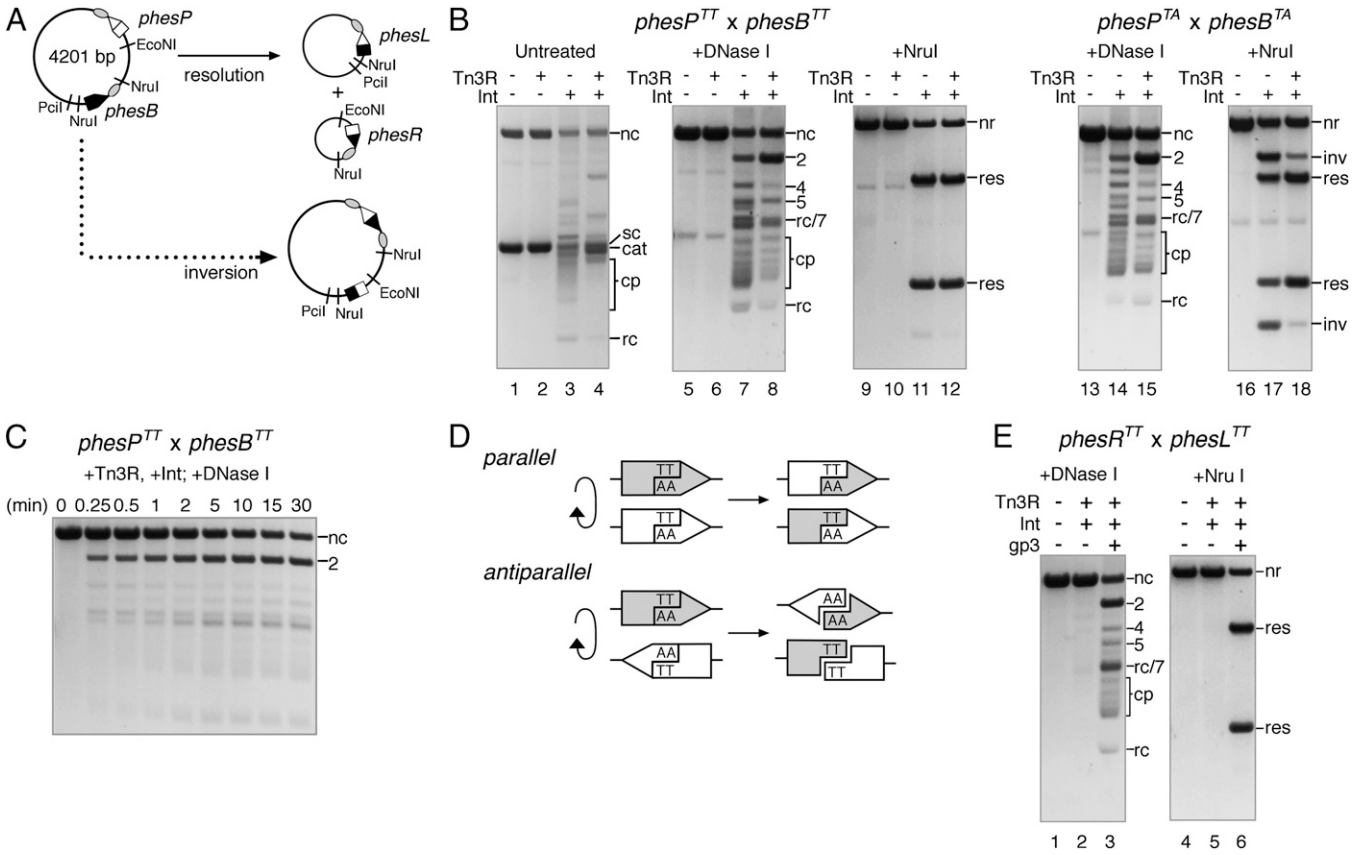

Fig. 2. Recombination between *phes* sites. (*A*) Plasmid substrates for *phes* recombination. The *phes* sites are in direct repeat. The *att* moieties are represented as pointed boxes, and the *res* accessory sites as ovals. Approximate positions of restriction sites used in product analysis are shown. The top line of the diagram shows the expected resolution (deletion) products from a substrate containing *att* sites with asymmetric overlap sequences. The lower part of the diagram (dashed arrow) shows the inversion recombination product which can be formed when the overlap sequences of the recombining *att* sites are palindromic. See main text for details. (*B*) *phes* recombination products and product topologies. (Lanes 1–12) A $phesP^{TT} \times phesB^{TT}$ substrate was treated with resolvase or integrase as indicated above each lane, and the products were separated by electrophoresis either without further treatment (*Left*), or after nicking with DNase I to remove supercoiling (*Center*), or after digestion with NruI (*Right*). NruI digestion of the substrate gives fragments of 3,707 and 494 bp (the small fragment has been cropped from the gel images). Recombination gives two product circles of 2,688 and 1,513 bp, which may be catenated. Each circle is cut once by NruI. (Lanes 13–18) A $phesP^{TA} \times phesB^{TA}$ (palindromic overlap sequence) substrate was treated the same as the samples in lanes 1–12; products after DNase I or NruI treatment are shown. Bands on the gels are annotated as follows: sc, supercoiled circle substrate; nc, nicked circle substrate; cat, supercoiled two-noded catenane; cp, topologically complex recombination products (≥ nine nodes); rc, circular recombinant product; 2, 4, 5, etc., nicked products with the indicated number of topological nodes; nr, nonrecombinant restriction fragment; res, resolution product restriction fragment; inv, inversion product restriction fragment. Further details on the assignment of product topologies are given in Figs. S3 and S8. (*C*) Time course of $phesP^{TT} \times phesB^{TT}$ recombination by integrase following synapsis by resolvase. Reaction samples were stopped at the times (minutes) indicated above the lanes. The gel shows products after nicking with DNase I. (*D*) Parallel and antiparallel alignments of *att* sites. In the cartoon, the sites both have an asymmetric (TT/AA) overlap sequence. When the sites are aligned in parallel, cleavage and rotation juxtaposes half-sites with complementary 3′ extensions which ligate to give recombinants. Rotation following antiparallel alignment juxtaposes half-sites with noncomplementary 3′ extensions that do not ligate. Sites with symmetric overlap sequences (e.g., TA/TA; not shown for clarity) will be cleaved to give half-sites with identical, complementary extensions and thus produce recombinants from either parallel or antiparallel alignments (Fig. 2*B*, lanes 13–18). (*E*) Products of *phesL* × *phesR* recombination. Proteins were added as indicated above the lanes. Products were treated with DNase I (lanes 1–3) or NruI (lanes 4–6) before gel electrophoresis. The annotation of the bands is as in Fig. 2*B*.

parallel alignment of the *att* sites (as shown in Fig. 1*B*), such that rotational strand exchange yields the two-noded catenane product.

We also analyzed the topology of *phesL* × *phesR* recombination by ϕC31 integrase in the presence of the ϕC31 recombination directionality factor gp3 (see introductory paragraphs). When a $phesL^{TT} \times phesR^{TT}$ substrate was treated with Tn*3* resolvase followed by ϕC31 integrase and gp3, the predominant product was a two-noded catenane (Fig. 2*E*), just as for the $phesP^{TT} \times phesB^{TT}$ reaction. The topological mechanism of *attL* × *attR* recombination in this system is therefore the same as for *attP* × *attB* recombination (right-handed rotation) (Fig. 1 *A* and *B*).

Topology of Multiple Rounds of Strand Exchange in Mismatched Substrates. Multiple rounds of strand exchange by serine recombinases can happen when a single round would yield recombinants with mismatched base pairs. This situation occurs when the overlap sequences of the two sites are different, or when sites with the same, asymmetric overlap sequence are aligned in an antiparallel way (see above). Recombinant products are usually not observed; instead, a second round of strand exchange may restore the nonrecombinant configuration, allowing religation but altering the DNA topology (14, 15, 18, 19, 22, 23) (Fig. 3*A*).

As expected, no recombinant products were observed when a mismatched $phesP^{TT} \times phesB^{TC}$ plasmid was treated with resolvase, then integrase (Fig. 3*B*, *Right*). Some knotted products were formed, but the efficiency was low and substantial amounts of cleavage products accumulated (Fig. 3*B*, lanes 2 and 8). We speculated that the DNA molecules were "stalled" in a cleaved *phesL* × *phesR* configuration, and that the addition of gp3 would license further rounds of strand exchange, as its natural function is

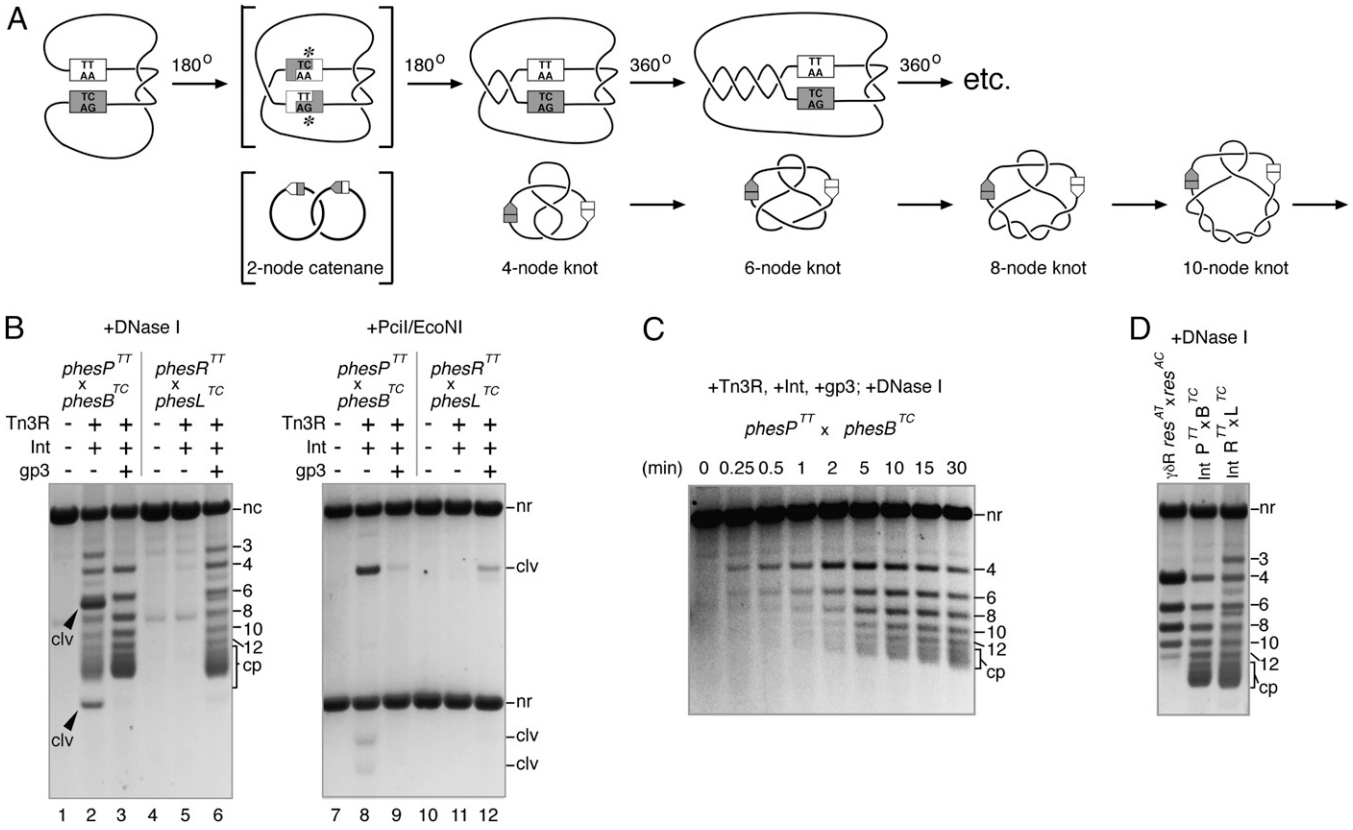

Fig. 3. Reactions of substrates with mismatched *phes* sites. (*A*) Cartoon showing predicted products from a mismatched *phes* × *phes* substrate. The first-round recombinant product would have mismatched base pairs (marked with asterisks), and is not observed; instead, a further 180° rotation yields a non-recombinant, four-noded knot. Further rounds of 360° rotation (without synapse dissociation) yield more complex knotted products, as illustrated. An expanded version of this scheme is shown in Fig. S9. (*B*) Reactions of $phesP^{TT} \times phesB^{TC}$ and $phesL^{TC} \times phesR^{TT}$ mismatched substrates. Proteins were added as indicated above the lanes. Products were nicked with DNase I (lanes 1–6) or digested with PciI and EcoNI (lanes 7–12). PciI + EcoNI cut the substrate to give fragments of 3,089 and 1,112 bp. If integrase cleaves the DNA at one or both *att* sites, additional fragments (clv) of 2,247, 842, 671 and 441 bp are produced (the smallest fragment is not shown on this gel, for clarity). The hypothetical recombinant product bands (2,688 and 1,513 bp) are undetected. The nicked products are thus nonrecombinant knots, the numbers indicating the number of topological nodes. (*C*) Time course of $phesP^{TT} \times phesB^{TC}$ recombination (with resolvase, integrase and gp3). Reaction samples were stopped at the times (minutes) indicated above the lanes. (*D*) Comparison of the electrophoretic mobilities of knots produced by $phesP^{TT} \times phesB^{TC}$ and $phesL^{TC} \times phesR^{TT}$ reactions (following addition of resolvase, gp3, and integrase) with knots made by the action of γδ resolvase on an identical-sized mismatched $res^{AT} \times res^{AC}$ substrate. An annotated version of this gel along with further lanes providing size markers is shown in Fig. S8.

to promote *attL* × *attR* reactivity. The presence of gp3 did indeed result in higher levels of knotting and greatly reduced amounts of cleavage products (Fig. 3*B*, lanes 3 and 9). The "ladder" of knotted products seen in Fig. 3*B*, lane 3, which starts at the four-node (4, 6, 8…), is consistent with processive double rounds of right-handed rotational strand exchange within the *phes* synapse (Fig. 3*A*). In time-course experiments with the $phesP^{TT} \times phesB^{TC}$ substrate (Fig. 3*C* and Fig. S7), the four-noded knot product was most abundant at early times, but its amount decreased as it was replaced by more complex products (6, 8, 10 nodes, and so forth) at later time points, in accordance with the predicted pathway for the knotting reaction shown in Fig. 3*A*.

A similar ladder of nonrecombinant knotted products (4, 6, 8… nodes) was seen when a $phesL^{TC} \times phesR^{TT}$ substrate was treated with resolvase, integrase, and gp3 (Fig. 3*B*, lane 6). However, additional weaker interleaved bands were evident corresponding to topologies of three, five, and seven nodes. The three- and five-node bands can also be seen in the products of the $phesP^{TT} \times phesB^{TC}$ reaction in the absence of gp3 (Fig. 3*B*, lane 2). We speculate that these minor products are formed when cleaved intermediates accumulate (Fig. 3*B*, lanes 8 and 12); dissociation of one of the half-sites from the synaptic complex and partial untangling of the DNA before ligation might then lead to aberrant product topologies.

If ϕC31 integrase uses a subunit rotation mechanism, the knots made by repeated double rounds of strand exchange of a mismatched *phes* substrate should be topologically identical to knots made by resolvase on a mismatched *res* substrate (15). To facilitate a direct comparison, we made a $res^{AT} \times res^{AC}$ substrate that was exactly the same size as our *phes* substrates. The knots made from this substrate by γδ resolvase [for which a processive rotation mechanism has been previously reported (15)] and the knots made from the $phesP^{TT} \times phesB^{TC}$ or $phesL^{TC} \times phesR^{TT}$ substrates by ϕC31 integrase and gp3 in the presence of Tn*3* resolvase had identical mobilities (Fig. 3*D* and Fig. S8).

Discussion

Rotation Mechanism for ϕC31 Integrase. Synapsis of the *res* accessory sites by Tn*3* resolvase is topologically selective and traps three negative supercoil nodes (20, 21, 24). If ϕC31 integrase has a right-handed rotational mechanism like that of resolvase, the first round recombination product from a circular *phes* × *phes* substrate is predicted to be a two-noded catenane (Fig. 1*B*), and this is exactly what we observe experimentally for sites with either asymmetric or symmetric overlap sequences (*Results*). The equivalence of the resolvase and ϕC31 integrase mechanisms was confirmed by comparison of the topologies of processive double-

rounds of strand exchange in mismatched *phes* and *res* substrates (Fig. 3*D* and Fig. S9). Our data therefore support a right-handed subunit rotation mechanism (Fig. 1*A*) for ϕC31 integrase by analogy with resolvase (Fig. S9). We also show that the *phesL* × *phesR* reaction (catalyzed by integrase plus gp3) proceeds by a topologically equivalent pathway (Figs. 2*E*, 3*D*, and Fig. S8). We hypothesize that subunit rotation by integrase can have either a right-handed or left-handed sense, as is the case for other serine recombinases, but that right-handed rotation is favored in our experiments by loss of negative supercoiling from the substrate plasmid DNA (9).

Gated Rotation Mechanism. Two scenarios might account for the specificity of ϕC31 integrase for unidirectional recombination (*attP* × *attB* to *attL* × *attR*, or vice versa):

i) Cleavage of all four DNA strands in the two recombining *att* sites is followed by unconstrained rotation of the two halves of the intermediate relative to each other (Fig. 1*A*). However, ligation of the half-sites to form recombinants is favored, and further reaction is then prohibited. In this scenario, a wide range of complex catenane *phes* product topologies is predicted.

ii) Cleavage of the substrate *att* sites is followed by a single 180° rotation, and ligation of the half-sites to form recombinants takes place before any further rotation can occur. In this scenario, *phes* × *phes* recombination in the presence of resolvase should yield a specific two-noded catenane product (Fig. 1*B*).

A recent single-molecule and topological analysis of the reactions catalyzed by the serine integrase of mycobacteriophage Bxb1 provided evidence of unconstrained rotation consistent with the first scenario (18). It was concluded that the cleaved intermediates remained "rotationally open" for periods of up to several minutes, allowing equilibrium between negative supercoiling and topological entanglement to be reached. However, a two-noded catenane is the predominant product in ϕC31 integrase-catalyzed reactions of matched *phesP* × *phesB* circular substrates (Fig. 2) and analogous reactions catalyzed by Bxb1 integrase (Fig. S6), consistent with the second scenario. In addition, we do not see evidence of persistently cleaved intermediates in these reactions. We conclude that the integrase subunit rotation mechanism is normally gated, stopping after a single (180°) round.

The reasons for the inconsistency between our results and those of Bai et al. (18) are unclear. One possible explanation is that the different conditions and substrates used by Bai et al. for single-molecule analysis have somehow inhibited ligation of cleaved sites and thereby suppressed the gating mechanism. We do observe minor *phesP* × *phesB* products with more complex (>two node) topologies (Fig. 2*B*, lane 8), which we assign as catenanes made by 180° rotation followed by one or more further 360° rotations before ligation in the recombinant (*attL* × *attR*) configuration. The gating mechanism after 180° rotation has thus been bypassed in formation of these minor products, although they have fewer topological nodes than would be expected if there were a fully rotationally open intermediate.

Products of Mismatched Substrates. Knotting of mismatched substrates, indicating double (360°) rounds of strand exchange, is stimulated by the presence of gp3; in the absence of gp3, products of double-strand cleavage at both sites accumulate (Fig. 3*B* and Fig. S6). These observations are also consistent with a gated mechanism: strand exchange normally stops after one round to give recombinant products with *attL* and *attR* sites, but ligation in this configuration is blocked when the sites are mismatched, so the cleaved intermediate accumulates. The presence of gp3 licenses another round of strand exchange starting from the (cleaved) *attL* × *attR* configuration, giving nonrecombinant cleaved *attP* and *attB* sites that can base pair correctly and religate. The slowness of reactions of mismatched $phesP^{TT} \times phesB^{TC}$ substrates relative to the corresponding matched substrates (compare Fig. 2*C* with Fig. S7) is also consistent with stalling of rotation in a mismatched recombinant configuration after one round of rotation.

The more complex (≥6 nodes) knotted products from $phesP^{TT} \times phesB^{TC}$ substrates (Fig. 3 *B* and *C*) are as expected from "processive" multiple 360° rotation events; that is, the second and subsequent reactions follow the first without dissociation of the synapsed Tn*3 res* accessory sites (15). We speculate that a relatively strong interaction between the ligated integrase-bound recombinant *att* sites stabilizes the product synaptic intermediate, favoring multiple rounds of 360° rotation without synapse dissociation. In a $phesP^{TT} \times phesB^{TC}$ time course (with gp3), there was no accumulation of the cleaved DNA intermediates that would be expected if a long-lived rotating intermediate were being formed (Fig. 3*C*). Furthermore, the product of one 360° rotation event (four-noded knot) appears at early time points, whereas more complex products predominate at later time points, suggesting that the sites are ligated after each 360° rotation, and the products can then react again.

Mechanistic Implications. In their natural context, small serine recombinases, such as the transposon resolvases and DNA invertases, stop after one round of strand exchange (1, 25, 26), even though the catalytic module (comprising two crossover sites synapsed by a recombinase tetramer) is fully competent for indefinite rounds of rotation (27). This selectivity is imposed by regulatory components involving accessory DNA sites and proteins (see introductory paragraphs). A parsimonious explanation is that cleavage and ligation of the crossover sites are rapid compared with rotation, and the equilibrium strongly favors ligation. Upon ligation of the recombinant sites, a conformational change involving the synapsed accessory sites might prohibit further reaction. Although the natural ϕC31 integrase system does not involve any accessory sites or (for *attP* × *attB* recombination) proteins, our results are largely consistent with a similar interpretation: rapid cleavage-ligation equilibrium in the rotating intermediate, strongly biased toward ligation. As soon as ligated recombinant *att* sites are formed, the selectivity of integrase against recombination between these sites (e.g., *attL* and *attR*, in the absence of gp3) may be manifested, and further rounds of reaction blocked. However, our results with mismatched *phesP* × *phesB* substrates showing accumulation of cleaved sites in the absence of gp3 (Fig. 3*B*) suggest that there is also a barrier to multiple rounds of rotation before ligation of the sites. We hypothesize that the "rotation-ready" cleaved intermediate might have two types of protein–protein interaction between the halves that will rotate relative to each other. The tetramer of integrase catalytic domains is predicted to have a flat interface between the rotating dimers, similar to that seen in resolvase crystal structures (11, 12). A second interface involving the large C-terminal parts of the integrase subunits might operate the gate mechanism, inhibiting more than one round of rotation of the half-sites and their attached catalytic domains. As yet, there is no informative high-resolution structural information on the complete integrase-DNA synaptic complex, but biochemical studies have revealed a synaptic interaction involving the C-terminal part of integrase, which is consistent with this hypothesis (28).

Materials and Methods

Plasmids and DNA. Plasmids containing a single *phes* site were constructed by replacing the *loxP* site in a *lesB96* plasmid (21) with synthetic double-stranded oligonucleotides containing *att* sequences. Sequences of the *phes* sites (with the overlap sequence TT) are given in Fig. S1. Plasmids containing two *phes* sites were constructed by the method described in ref. 21. Full sequences of the plasmids are available on request. Supercoiled plasmid DNA was purified from transformed *Escherichia coli* DH5 cells, using a Qiagen HiSpeed Midi kit according to the manufacturer's instructions. DNA concentrations were estimated by measuring absorbance at 260 nm.

ϕC31 Integrase and gp3. ϕC31 integrase overexpression was induced by adding isopropylthio-β-galactoside to exponentially growing cultures of BL21(DE3)pLysS transformed with pARM010, an overexpression vector of polyhistidine-tagged ϕC31 integrase, and purified using a His-Trap FF column (GE Healthcare) as previously described (29). The recombination directionality factor (gp3) was purified as previously described (6). Concentration and purity of the proteins were estimated by SDS-PAGE and measurement of absorbance at 280 nm. Methods for Bxb1 integrase reactions are in the legend to Fig. S6.

In Vitro Recombination Reactions and Product Analysis. Integrase, Tn*3* resolvase, and gp3 were each diluted at 0 °C in a buffer containing 25 mM Tris·HCl (pH 7.5), 1 mM DTT, 1 M NaCl, and 50% (vol/vol) glycerol. Dilutions were stored at −20 °C. In a typical *phes* reaction, diluted Tn*3* resolvase (6 μL, ∼4 μM) was added to a solution of substrate plasmid DNA (60 μL, 20 μg/mL) in a buffer containing 50 mM Tris·HCl (pH 7.5), 0.1 mM EDTA, 5 mM spermidine, and 0.1 mg/mL BSA, and the sample was incubated at 30 °C for 10 min. Recombination was then initiated by the addition of integrase (6 μL, ∼4 μM). For reactions in the presence of gp3, equal volumes of integrase (∼8 μM) and gp3 (∼8 μM) were mixed thoroughly and kept on ice for 15 min, then 6 μL of this mixture was added to the reactions. Reactions were for 30 min at 30 °C unless stated otherwise, and were terminated by heating at 80 °C for 10 min. Aliquots from the reaction samples (20 μL) were treated with restriction enzymes or DNase I as described (15, 27). Loading buffer [25 mM Tris·HCl (pH 8.2), 20% (wt/vol) Ficoll, 0.5% sodium dodecyl sulphate, 5 mg/ml protease K, 0.25 mg/ml bromophenol blue; 5 μl] was added to each sample. The products were then separated by agarose gel electrophoresis and visualized as previously described (15). Samples treated with DNase I were run on 0.7% agarose gels; all other samples were run on 1.2% (wt/vol) gels. Digital images of ethidium-stained gels were recorded using a Bio-Rad GelDoc apparatus, and are shown in reverse contrast. The topologies of nicked product bands were assigned by comparison of their mobilities with those of product species of known topology, made by Tn*3* resolvase or γδ resolvase from *res* × *res* substrate plasmids of identical size to the *phes* substrates. Further details on assignment of product topology are in the legend to Fig. S8.

ACKNOWLEDGMENTS. We thank Martin Boocock for helpful discussions and comments on the manuscript, and Arlene McPherson for technical support. This research was funded by Engineering and Physical Sciences Research Council (EPSRC) Grant EP/H031367/1; EPSRC Grant EP/H019154/1 (to S.D.C.); research in the M.C.M.S. laboratory is supported by the Biotechnology and Biological Sciences Research Council Grants BB/H001212/7 and BB/H005447/1; and D.E.B's research is supported by EPSRC Grant EP/G039585/1.

1. Grindley NDF, Whiteson KL, Rice PA (2006) Mechanisms of site-specific recombination. *Annu Rev Biochem* 75:567–605.
2. Smith MCM, Thorpe HM (2002) Diversity in the serine recombinases. *Mol Microbiol* 44(2):299–307.
3. Smith MCM, Brown WR, McEwan AR, Rowley PA (2010) Site-specific recombination by φC31 integrase and other large serine recombinases. *Biochem Soc Trans* 38(2):388–394.
4. Thorpe HM, Wilson SE, Smith MCM (2000) Control of directionality in the site-specific recombination system of the *Streptomyces* phage φC31. *Mol Microbiol* 38(2):232–241.
5. Groth AC, Olivares EC, Thyagarajan B, Calos MP (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci USA* 97(11): 5995–6000.
6. Khaleel T, Younger E, McEwan AR, Varghese AS, Smith MCM (2011) A phage protein that binds φC31 integrase to switch its directionality. *Mol Microbiol* 80(6):1450–1463.
7. Stark WM (2011) Cutting out the φC31 prophage. *Mol Microbiol* 80(6):1417–1419.
8. Brown WRA, Lee NCO, Xu Z, Smith MCM (2011) Serine recombinases as tools for genome engineering. *Methods* 53(4):372–379.
9. Stark WM, Sherratt DJ, Boocock MR (1989) Site-specific recombination by Tn*3* resolvase: Topological changes in the forward and reverse reactions. *Cell* 58(4):779–790.
10. Dhar G, Sanders ER, Johnson RC (2004) Architecture of the hin synaptic complex during recombination: The recombinase subunits translocate with the DNA strands. *Cell* 119(1):33–45.
11. Li W, et al. (2005) Structure of a synaptic γδ resolvase tetramer covalently linked to two cleaved DNAs. *Science* 309(5738):1210–1215.
12. Yuan P, Gupta K, Van Duyne GD (2008) Tetrameric structure of a serine integrase catalytic domain. *Structure* 16(8):1275–1286.
13. Wasserman SA, Dungan JM, Cozzarelli NR (1985) Discovery of a predicted DNA knot substantiates a model for site-specific recombination. *Science* 229(4709):171–174.
14. Kanaar R, et al. (1990) Processive recombination by the phage Mu Gin system: Implications for the mechanisms of DNA strand exchange, DNA site alignment, and enhancer action. *Cell* 62(2):353–366.
15. Stark WM, Grindley NDF, Hatfull GF, Boocock MR (1991) Resolvase-catalysed reactions between *res* sites differing in the central dinucleotide of subsite I. *EMBO J* 10(11): 3541–3548.
16. Stark WM, Boocock MR (1994) The linkage change of a knotting reaction catalysed by Tn*3* resolvase. *J Mol Biol* 239(1):25–36.
17. Rowland SJ, Stark WM, Boocock MR (2002) Sin recombinase from *Staphylococcus aureus*: Synaptic complex architecture and transposon targeting. *Mol Microbiol* 44(3): 607–619.
18. Bai H, et al. (2011) Single-molecule analysis reveals the molecular bearing mechanism of DNA strand exchange by a serine recombinase. *Proc Natl Acad Sci USA* 108(18): 7419–7424.
19. Smith MCA, Till R, Smith MCM (2004) Switching the polarity of a bacteriophage integration system. *Mol Microbiol* 51(6):1719–1728.
20. Stark WM, Boocock MR (1995) Topological selectivity in site-specific recombination. *Mobile Genetic Elements*, ed Sherratt DJ (Oxford Univ Press, Oxford), pp 101–129.
21. Kilbride EA, Burke ME, Boocock MR, Stark WM (2006) Determinants of product topology in a hybrid Cre-Tn*3* resolvase site-specific recombination system. *J Mol Biol* 355(2):185–195.
22. Ghosh P, Kim AI, Hatfull GF (2003) The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of *attP* and *attB*. *Mol Cell* 12(5):1101–1111.
23. Heichman KA, Moskowitz IPG, Johnson RC (1991) Configuration of DNA strands and mechanism of strand exchange in the Hin invertasome as revealed by analysis of recombinant knots. *Genes Dev* 5(9):1622–1634.
24. Grainge I, Buck D, Jayaram M (2000) Geometry of site alignment during int family recombination: Antiparallel synapsis by the Flp recombinase. *J Mol Biol* 298(5):749–764.
25. Stark WM, Boocock MR, Sherratt DJ (1992) Catalysis by site-specific recombinases. *Trends Genet* 8(12):432–439.
26. Dhar G, Heiss JK, Johnson RC (2009) Mechanical constraints on Hin subunit rotation imposed by the Fis/enhancer system and DNA supercoiling during site-specific recombination. *Mol Cell* 34(6):746–759.
27. Olorunniji FJ, He J, Wenwieser SVCT, Boocock MR, Stark WM (2008) Synapsis and catalysis by activated Tn*3* resolvase mutants. *Nucleic Acids Res* 36(22):7181–7191.
28. McEwan AR, Rowley PA, Smith MCM (2009) DNA binding and synapsis by the large C-terminal domain of φC31 integrase. *Nucleic Acids Res* 37(14):4764–4773.
29. Rowley PA, Smith MC, Younger E, Smith MCM (2008) A motif in the C-terminal domain of φC31 integrase controls the directionality of recombination. *Nucleic Acids Res* 36(12):3879–3891.