Nucleic Acids Research, 2026, **54**, gkag750
https://doi.org/10.1093/nar/gkag750
Synthetic Biology and Bioengineering

OXFORD

In vitro properties of large serine integrase hybrids derived from ϕC31 and TG1 integrases

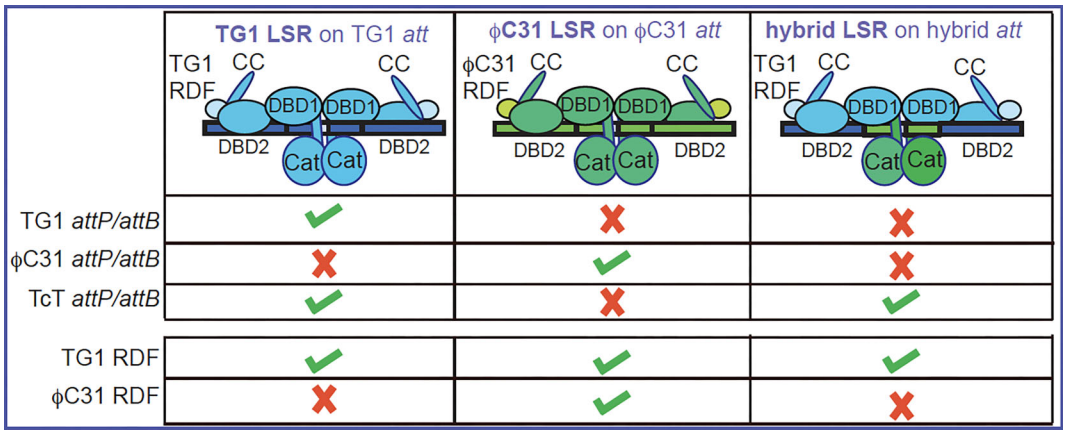
Makeba Lawson-Williams[1], Alexandria Holland[1], Adebayo J Bello[1], Phoebe A Rice[2], Femi J Olorunniji[1,*]

[1]School of Pharmacy and Biomolecular Sciences, Faculty of Science, Liverpool John Moores University, James Parsons Building, Byrom Street, Liverpool L3 3AF, United Kingdom
[2]Department of Biochemistry & Molecular Biology, The University of Chicago, Chicago IL 60637, United States
*To whom correspondence should be addressed. Email: F.J.Olorunniji@ljmu.ac.uk

Abstract

Phage-derived serine integrases are site-specific recombinases from the large serine recombinase (LSR) family that catalyse precise DNA rearrangements. Their high specificity and unidirectional activity make them attractive tools for genome engineering and synthetic biology. However, their strict sequence requirements limit application to predefined target sites. Here, we report *in vitro* activities of hybrid integrases derived from ϕC31 and TG1 integrases that efficiently recombine hybrid *attP* × *attB* substrates while preserving interaction with the cognate recombination directionality factor (RDF) via the TG1-derived coiled-coil domain. These hybrid recombinases retain site discrimination and catalyse attL × attR recombination exclusively in the presence of the appropriate RDF, indicating preserved directionality control. Analysis of the activities of hybrid integrases across a panel of hybrid substrates highlight the importance of DNA-binding domains-*att* site recognition motifs in governing specificity. This study presents a modular framework for engineering programmable serine integrases with tailored specificity. Combined with structural prediction tools and expanded LSR discovery and characterization, this strategy holds promise for rational design of recombinases targeting custom genomic sites.

Graphical abstract



Introduction

Large serine recombinases (LSRs), also known as serine integrases, catalyse site-specific recombination through a coordinated, multistep mechanism. Following binding to recombination attachment (att) sites, integrase dimers form a tetrameric synaptic complex. Within this synapse, the att sites are aligned and cleaved by active-site serine residues via nucleophilic attack on the scissile phosphates. Subsequent strand exchange is achieved through subunit rotation and re-ligation of DNA ends to yield precise recombination products [1].

Site-specific recombination reactions catalysed by LSRs have the unique properties of unidirectionality with the excision reaction requiring the presence of a second protein, the recombination directionality factor (RDF) [2]. This unidirectional nature allows most LSR-catalysed reactions to go to near completion both *in vivo* and *in vitro* [3]. In addition, LSRs like other members of the serine recombinase

Received: December 15, 2025. **Revised:** July 5, 2026. **Accepted:** July 7, 2026

family catalyse complete breaking and ligation of all four DNA strands in the recombining duplex DNA partners. The domain architecture of LSRs and their interactions with their recombination sites are illustrated in Fig. 1. Attachment site recognition and binding are mediated primarily by the two DNA-binding domains (DBD1 and DBD2; also called the recombinase domain and the zinc-binding domain, respectively) while the catalytic domain, the E helix and the coiled-coil (CC) inserted within DBD2 mediate interactions necessary for synapsis (pairing two recombining sites) and catalysis [1, 4].

Recent cryo-EM structures of SPβ integrase [5] and ϕC31 integrase [6] reveal LSR binding and synapsis to their *att* sites in both pre- and post-cleavage recombination states, largely agreeing with the model proposed by Rutherford *et al.* [7] and based on a half-site structure of LI integrase–*attP* complex. Several biochemical studies showed that RDFs bind to the CC and/or DBD2 of their cognate LSRs [8–11]. In addition, the cryo-EM structures of SPβ integrase fused with its cognate RDF in integrative and excisive complexes [5] support the AlphaFold multimer-based model that RDFs bind to their LSR at the junction of the DBD2 and the base of the CC motif (Fig. 1B; [12, 13]).

Compared to other serine recombinase subfamilies (e.g. resolvases and invertases), LSRs exhibit exceptional sequence specificity and inherently unidirectional activity. They must discriminate among four *att* sites, *attP*, *attB*, *attL*, and *attR*, and selectively bring together compatible pairs of *att* sites. This is particularly challenging for discriminating between *attL* and *attR*, which each contain half-sites derived from both *attP* and *attB*. Although synapsis and recognition mechanisms have been partly characterized [14], the full rules governing site selectivity remain unresolved (Fig. 1D). Minor deviations from *att* consensus sequences can stall recombination or result in generation of incomplete products, underscoring the stringency of sequence requirements [14]. A deeper understanding of these constraints is essential for expanding the utility of LSRs in genome engineering [15, 16].

LSRs confer several advantages for genome editing: Like other members of the serine recombinase family, they catalyse complete breaking and ligation of all four DNA strands in the recombining duplex DNA partners. This is a unique advantage over the CRISPR–Cas system since there is no dependence on host's systems for repairing double strand breaks. They also function in diverse cell types and require short target sites (~40–55 bp).

Recent work has explored the engineering of LSRs to enhance their efficiency *in vivo* for genome editing applications through codon optimization and the incorporation of nuclear localzation signals and has demonstrated improved efficiency of transgene integration into human pluripotent stem cells [17]. Bxb1 integrase has been successfully adapted for targeted integration in both mouse and human stem cells, with significant improvements in recombination efficiency observed at safe harbour loci [18]. Additionally, engineered variants of Bxb1 integrase have demonstrated improved specificity and reduced off-target effects [19]. The development of bridge recombinases and their applications in mammalian systems have also highlighted the capability of using LSRs as tools for genome editing [20, 21]. Better understanding of such systems will enhance our ability to design more bespoke genome editing tools [22].

Despite the high specificity of LSRs, recent studies have highlighted the potential for off-target recombination at cryptic genomic sites. Advanced sequencing approaches, such as HIDE-Seq and Cryptic-Seq, have been used to comprehensively map and validate off-target events [23]. In addition to this, IntQuery, a deep learning model, has been developed to predict the genome-wide off-target recombination potential of LSRs [24]. These findings highlight the importance of continued effort to optimize the specificity of LSRs for use in genome engineering. Fauser *et al.* [25] developed a Modular Integrase (MINT) platform that allows for the reprogramming of Bxb1 to target specific genomic loci, such as the AAVS1 and TRAC sites, where wild-type Bxb1 shows no activity. Using directed evolution and structural modelling to identify critical protein–DNA interactions, the authors designed Bxb1 integrase variants with improved specificity and activity achieving targeted integration efficiencies of up to 35% in human cells.

A key challenge to the use of LSRs in genome editing is that their strict sequence specificity limits their direct use at endogenous genomic loci. Although hybrid systems combining LSRs with CRISPR–Cas technologies such as prime editing [26], STRAIGHT-IN [27], and SYMBIOSIS [28] have been developed to enable targeted integration, they still rely on engineered 'landing pads' pre-inserted into the genome. A more flexible solution would be the development of programmable LSRs capable of recognizing arbitrary sequences without auxiliary targeting systems. Despite sharing a conserved catalytic mechanism, LSRs are functionally orthogonal: each integrase recognizes unique *att*P/attB sites. This specificity has been critical for synthetic biology applications that exploit orthogonal recombination circuits [3]. However, current structural and biochemical data are insufficient for designing fully customized integrases with novel sequence recognition properties.

To address this, we explored a rational design approach to engineer hybrid integrases that recombine synthetic att sites. Chimeric integrases were generated by domain-swapping between related LSRs, ϕC31 and TG1, and their activities were tested on corresponding hybrid *att* substrates. This modular strategy seeks to uncouple and reassign site recognition properties by reassembling functional domains. Hybrid recombinases have previously been used to alter specificity or introduce regulatory control into serine recombinases. For example, Zif268–resolvase fusions retarget recombination to zinc finger binding motifs [29–31], and ϕC31-intein constructs enabled post-translational control of integrase reconstitution [32]. Farruggio and Calos [33] reported that domain-swapped LSRs derived from ϕC31, ϕBT1, and TG1 integrases could catalyse *in vivo* recombination at native and hybrid *att* sites, maintaining partial specificity in *Escherichia coli* and HeLa cells.

In this study, as a starting point for designing hybrid integrases and characterizing their activities both *in vivo* and *in vitro*, we chose a ϕC31–TG1 hybrid integrase (CT8,8) described by Farruggio and Calos [33]. In this construct, the catalytic domain and the E-helix of ϕC31 integrase were swapped into the same position in TG1 integrase to give a hybrid that recognizes hybrid sites made from the *attP* and *attB* sites of the two parent integrases. We reengineered this construct to examine its activity *in vitro* and designed complementary hybrid *att* sites containing hybrid sequence features from both parent integrases.

The choice of ϕC31 and TG1 reflects their high sequence similarity (~50% identity, Fig. 2) and their well-characterized functional crosstalk showing evidence of partial

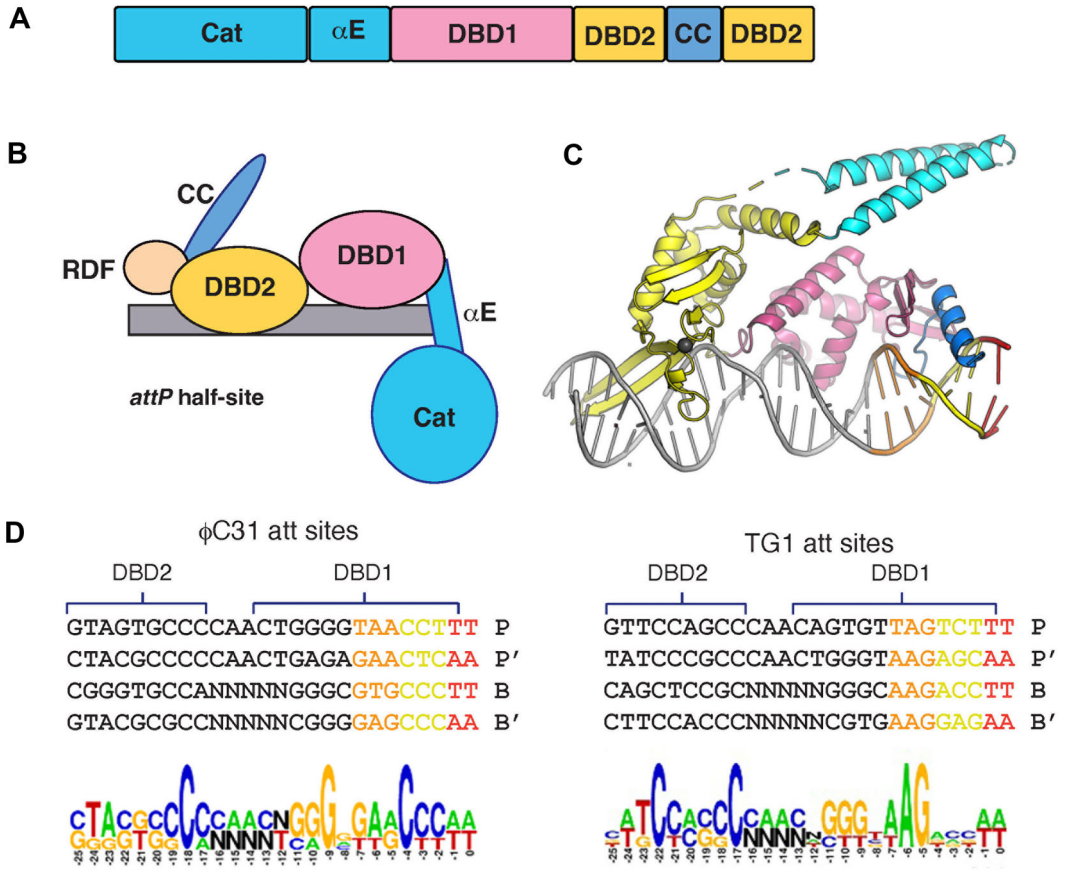


Figure 1. Domain architecture of LSRs and interaction with recombination sites *attP* and *attB*. (**A**) Domain architecture of LSRs. The catalytic domain, cat (blue) and the αE helix (blue), mediate catalysis and provide the subunit rotation interface. The two DNA-binding domains (DBD1 and DBD2) mediate DNA binding and sequence specificity. DBD1, also known as 'recombinase domain' or RD is shown in pink. DBD2, also known as the zinc-binding domain or ZD is shown in yellow. The CC domain (teal) is embedded within DBD2 and mediate directionality control. (**B**) Schematic illustration of LSR domain interactions with the recombination site (*att*P) and the RDF based on structure of LI integrase–*att*P complex [13, 34] and AlphaFold multimer-predicted interaction of the RDF with DBD2-proximal region of the CC domain [7]. (**C**) Structure of LI integrase–*att*P complex [1]. (**D**) Annotation of binding motifs on *att*P and *att*B recognized by DBD1 and DBD2 of ϕC31 and TG1 integrases [1]. The appropriate motifs bound by DBD1 and DBD2 are indicated and the central dinucleotides in *att*P and *att*B are shown in red font. The 5 Ns are inserted in *attB* to align the sequence motifs for DBD2 binding with the same region in *attP* [47]. The three base pairs closest to the central dinucleotide are highlighted in yellow, and the next three are highlighted in orange, to match the corresponding bases in the DNA structure shown in panel (C). The sequence logos for these alignments are shown below to highlight patterns of base conservation at each position across *attP* and *attB* [48].

orthogonality: TG1 is activated exclusively by its cognate RDF (gp25), while ϕC31 can be activated by either ϕC31 RDF (gp3) or gp25 [34]. In addition, the two integrases are specific to their cognate *attP* and *attB* sites.

By engineering both the LSR and substrate, we aimed to dissect the modular determinants of *att* site recognition and directionality. These findings offer a generalisable strategy for programmable recombination and provide structural insights into domain-specific contributions to LSR–*att* site interactions and help lay a foundation for modular genome engineering platforms.

Materials and methods

AlphaFold protein structure prediction

Three-dimensional structures were predicted using Alphafold3 [35]. Wild-type amino acid sequences were retrieved from the NCBI database (Accession numbers: TG1 integrase, UZP69254.1; ϕC31, BAM36534.1) and hybrid constructs were generated by *in silico* splicing of parental sequences at defined junctions. Each sequence was submitted individually using default parameters. Predicted structures were visualized and edited using PyMOL (https://pymol.org). Structure-based alignments were performed using PyMOL's super command, which applies sequence-independent structural fitting with iterative refinement. Where necessary, domain components were aligned independently and manually adjusted for optimal overlay.

In vitro plasmid substrates for intramolecular recombination

Recombination substrates were prepared by cloning synthetic double-stranded *att* site oligonucleotides into the pC31-delPBX plasmid backbone [36]. Each *att* site was flanked by unique restriction enzyme sites to facilitate directional insertion via ligation. Substrate plasmids contain a ColE1 origin of replication, an ampicillin resistance cassette, and two *att* sites in direct ('head-to-tail') orientation, allowing recombination to resolve the parent plasmid into two smaller product plasmids (Fig. 5). Supercoiled plasmids were purified from

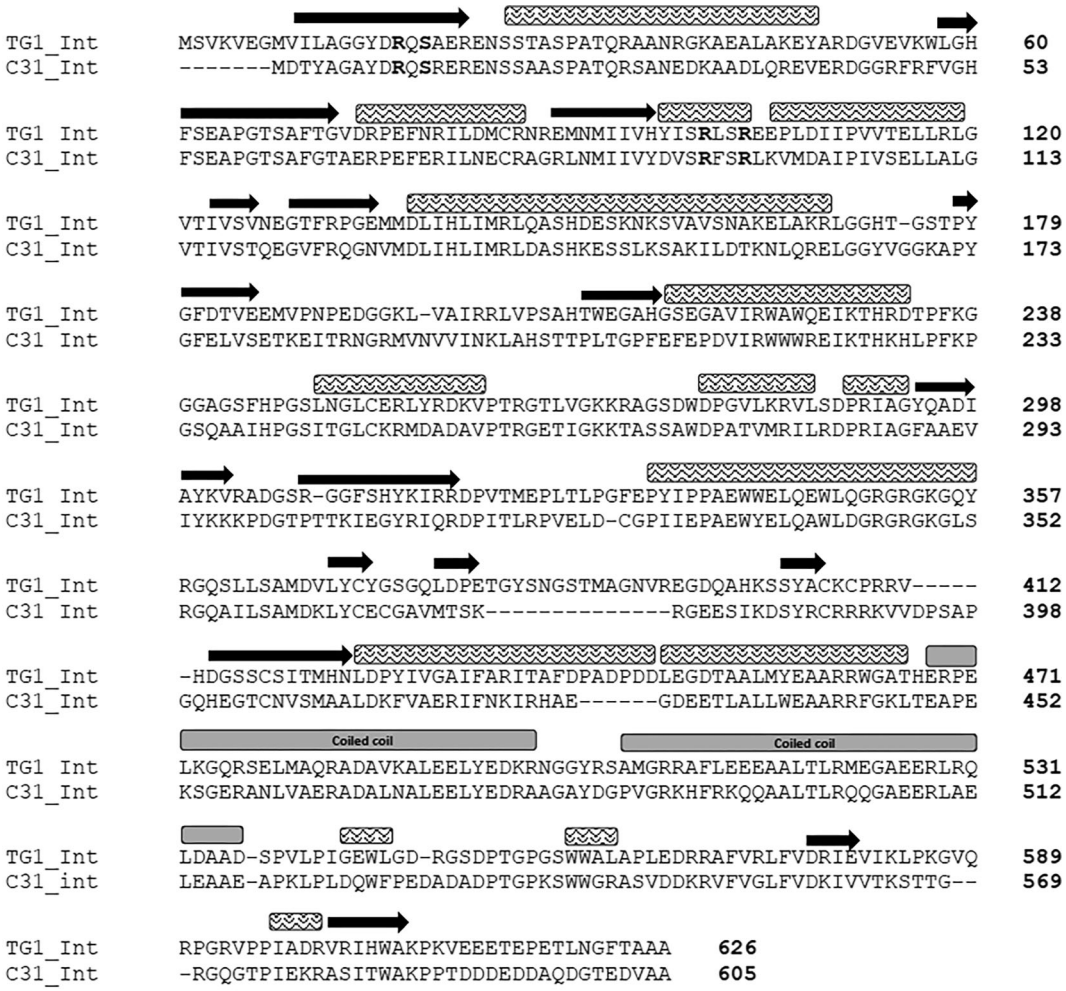


Figure 2. Sequence alignments of TG1 and ϕC31 integrases. Secondary structures are annotated as shown. Key amino acid residues in the active site including the nucleophilic serine are highlighted in bold. Multiple sequence alignment was generated using Clustal Omega [7] and adjusted based on predicted structures. Arrows, beta strands; rectangles, helices. The CC rectangles are shaded grey and labelled as shown.

E. coli DS941 using standard alkaline lysis protocols and validated by agarose gel electrophoresis (1.0%). DNA concentration was quantified by measuring absorbance at 260 nm.

Expression constructs for protein production

Coding sequences for wild-type and hybrid integrases were codon-optimized and inserted into the T7-based pET-28a(+) expression vector (Novagen) using restriction enzyme cloning between NdeI and XhoI sites. All constructs included N-terminal His-tags present in the cloning vector for the purpose of purification via affinity chromatography. Recombinant proteins were overexpressed in *E. coli* BL21(DE3)pLysS strain and purified using nickel–NTA chromatography as previously described [36, 37].

In vitro recombination reactions and product analysis

Purified integrases were diluted in storage buffer (25 mM Tris–HCl, pH 7.5, 1 mM dithiothreitol, 1 M NaCl, 50% glycerol). Substrate plasmids were diluted in Qiagen EB buffer to working concentrations. Reactions (35 µl total) contained 0.5 µM of protein (integrase ± RDF), 25 µg/ml plasmid substrate, 100 µg/ml bovine serum albumin, 50 mM Tris–HCl (pH 7.5), 5 mM spermidine, and 0.1 mM ethylenediaminetetraacetic acid (pH 8.0). Incubation was performed at 30°C for 2 or 8 h and terminated by heating at 80°C for 10 min to denature the integrase. Post-reaction, samples were digested with 40 units of NruI (New England Biolabs) in 26 µl of buffer B103 (90 mM Tris–HCl, pH 7.5, 20 mM $MgCl_2$) at 37°C for 2 h. Digestion was quenched by adding 20 µl of loading buffer (25 mM Tris–HCl, pH 8.2, 25% Ficoll, 6% sodium dodecyl sulfate, 1 mg/ml proteinase K, 0.23 mg/ml bromophenol blue). Reaction products were analysed by electrophoresis on 1.25% agarose gels.

The intensities of DNA bands on the agarose gel corresponding to substrate and products were quantitated with the volume analysis tool of Image Lab software (BioRad) using automatic band detection. These were manually corrected to exclude artifact bands. The molar proportions of recombinant products were determined by factoring in the molecular weight of the DNA in each band [34, 37]. The mean extent of recombination and standard deviation (%) from quantitation of triplicate experiments are given below each lane.

Results

Structural comparison of ϕC31 and TG1 integrases by AlphaFold3 prediction

To explore the structural determinants of sequence specificity, AlphaFold3 [35] was used to model ϕC31 and TG1 integrases docked onto DNA (Fig. 3A). Both predictions recapitulated domain architectures typical of serine recombinases: the catalytic domain and E-helix resembled those of resolvases [38–41], while DBD1 and DBD2 have folds similar those seen in the A118 integrase structure [7] and the recent cryo-EM structures of SPβ integrase [5] and ϕC31 integrase [6]. In all models, the C-terminal portion of helix E and the loop that follows it bind in the minor groove at the centre of the *att* site (similar to γδ resolvase [7]).

Design of ϕC31/TG1 hybrid integrases

To dissect the contribution of specific domain–DNA contacts, two hybrid integrases were generated by domain swapping between the catalytic domain of ϕC31 and the DNA binding regions of TG1. Junctions were placed at motifs conserved in sequence and predicted structure to minimize disruption. The hybrids differ in whether the C-terminus of helix E and the loop that follows it derive from TG1 (hCT1) or from ϕC31 (hCT2). The latter is based Farruggio and Calos [33]. The hCT1 junction was made in the conserved 'ES' motif within helix E and the hCT2 junction just before the conserved 'PYGF' motif, which is common to both LSRs and marks the beginning of DBD1.

Design of hybrid att sites

Hybrid *att* sites were constructed by substituting ϕC31-specific bases into the TG1 *att* site in 3-bp increments (Fig. 4), following a strategy similar to that adopted by Farruggio and Calos [33]. The central dinucleotide (TT) was left unmodified, as it is conserved across both LSRs. Base numbering was centred around this dinucleotide (position 0). These hybrid *att* sites were designed to match the architecture of the hybrid integrases, with TG1-derived DBD1 and DBD2 motifs recognizing TG1-like flanking regions, and the ϕC31-derived catalytic domain engaging ϕC31-like core motifs. This rational design allowed targeted probing of domain-specific DNA interactions and facilitated analysis of junction-specific recognition events (Fig. 3).

In vitro activities of hybrid integrases on wild-type and hybrid sites

The catalytic activities of hCT1 and hCT2 were evaluated alongside the parent LSRs using standard recombination assays [34] (Fig. 5). As expected, ϕC31 and TG1 integrases recombined their respective wild-type *attP* × *attB* substrates with no cross-reactivity. Neither hybrid integrase was active on wild-type ϕC31 or TG1 sites (Fig. 5), confirming functional orthogonality and the specificity of the parental LSRs. When tested on hybrid *att* sites (Fig. 6A), both hCT1 and hCT2 recombined the substrates with 3 base substitutions around the central dinucleotide of *attB*, *att*P3 × *att*B3 and *attP6* × *attB3*. However, they failed to recombine *attP3* × *attB6* and *attP6* × *attB6*, substrates containing 6-bp substitutions in attB (Fig. 6B). TG1 integrase, by contrast, successfully recombined all four hybrid substrates, while ϕC31 integrase showed no activity on any hybrid site. These results indicate that the hybrid LSRs retain sequence specificity and recognize hybrid *att* sites in a manner consistent with their domain composition—that is, DBD1 and DBD2 of TG1, included in the hybrids, are primarily responsible for DNA sequence recognition, but helix E and the loop that follows it, which derive from ϕC31 sites in the hybrids, is primarily responsible for recognizing the innermost 3 bp of each half-site. Importantly, no recombination was observed at wild-type TG1 and ϕC31 sites, supporting the fidelity of the hybrid recombinases—that is, WT ϕC31 integrase was inactive on all combinations of hybrid sites, which were expected to be incompatible with ϕC31 integrase's DBD1 and DBD2.

A distinct difference between the hybrids was the accumulation of cleaved, unligated reaction intermediates in hCT2 reactions (Fig. 6). These linear products suggest that DNA cleavage occurred, but religation failed, possibly due to instability in the post-cleavage synaptic complex or reduced domain compatibility during strand rotation [42]. This contrasts with hCT1, which produced fewer incomplete products, and with TG1 integrase, which efficiently completed recombination on all substrates.

Excisive recombination of hybrid attL × attR sites requires TG1 RDF

Recombination between *attL* and *attR* sites by LSRs requires activation by an RDF, which reconfigures the integrase to favour excision over integration. To determine whether ϕC31/TG1 hybrid LSRs retain directionality control, we tested *attL* × *attR* activity of hCT1 and hCT2 *in vitro*. Because RDF interacts with the CC motif embedded in DBD2, and this domain in both hybrids is TG1-derived, we predicted that TG1 RDF (gp25), but not ϕC31 RDF (gp3), would activate the hybrids (Fig. 7A). This is consistent with prior observations that TG1 integrase is not significantly activated by ϕC31 RDF [34]. To test this, we engineered hybrid *attR* and *attL* sites using *attP3* and *attB3*, substrates previously shown to be efficiently recombined by both hybrids. The hybrid *att*R substrate (*P3/B3attR*) was made using the left half site of *attP3* and the right half site of *attB3*. Conversely, the hybrid *attL* substrate (*P3/B3attL*) was made using the left half site of *attB3* and the right half site of *attP3*. The sequences of the two hybrid *attR* and *attL* sites are shown in Fig. 7B.

As shown in Fig. 8, both hCT1 and hCT2 catalysed recombination of *attR3* × *attL3* only in the presence of TG1 RDF, giving activities of 18% and 50%, respectively. Neither hybrid showed activity in reactions containing ϕC31 RDF, confirming the specificity of LSR–RDF interactions and the retention of regulated directionality in both constructs.

Site-specific DNA recognition is determined mainly by the two DNA-binding domains

Finally, we sought to establish the priority of DBD1 and DBD2 over the catalytic domain in att site recognition. Results presented in Figs 5 and 6 show that TG1 integrase displayed higher activity at hybrid sites than the two hybrid LSRs in all reactions studied. In contrast, ϕC31 integrase showed no activity towards the hybrid sites. While this could be a unique property of TG1 integrase resulting from its intrinsic high activity [34, 37], an alternative explanation is that the DBD1 and DBD2 motifs of the hybrid sites are

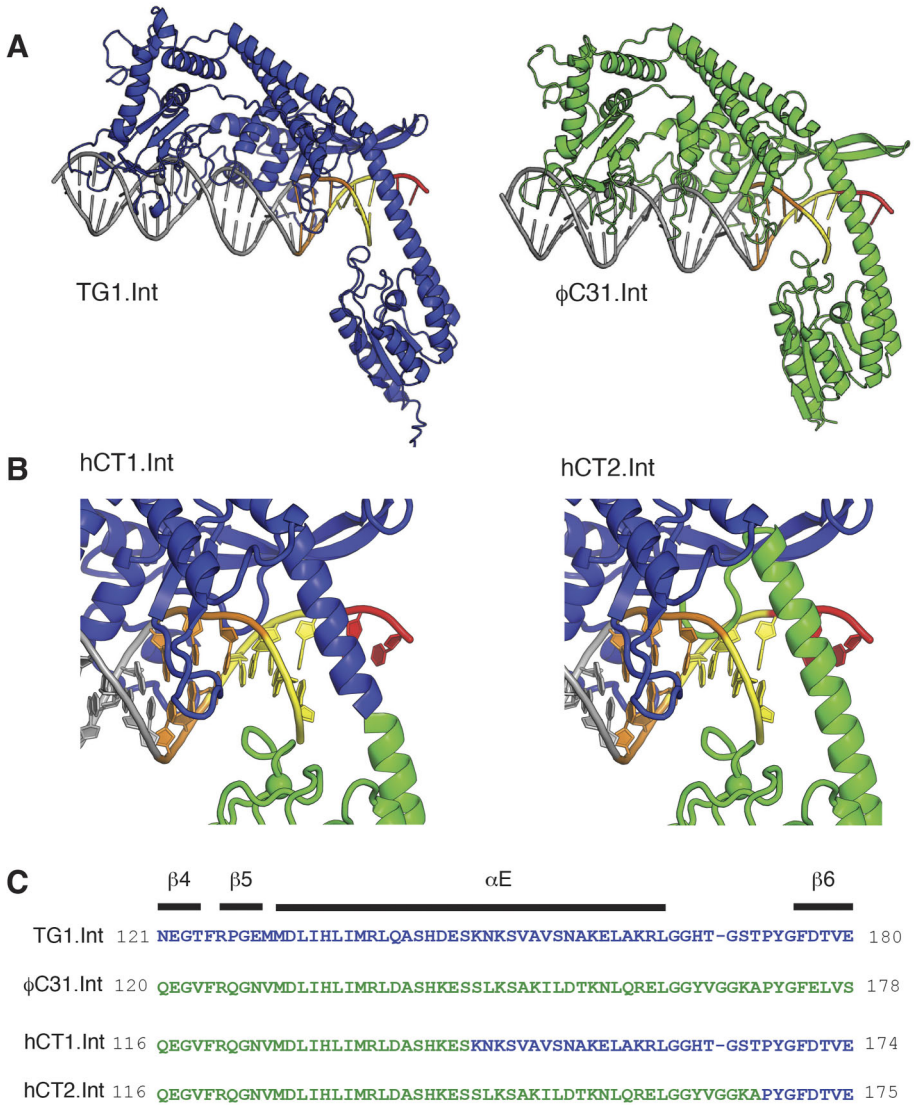


Figure 3. Design of ϕC31/TG1 hybrid integrases. (**A**) AlphaFold3-predicted 3D structures of wild-type TG1 and ϕC31 integrases, shown in blue and green, respectively. The structures show the LSRs bound to *attP* based on the structure of LI-*attP* integrase [47] (**B**) The hybrid integrases (hCT1 and hCT2) are coloured by parental origin. The N-terminal end of both hCT1 and hCT2 integrases contain ϕC31 integrase N-terminal catalytic domain. TG1 integrase sequence is introduced at position K144 for hCT1, and P168 for hCT2. In both instances, the junction in the hybrids where TG1 integrase sequence is introduced is within helix E. The subsequent protein sequence consists solely of TG1 integrase. (**C**) Protein sequence alignment of the junction regions in hCT1 and hCT2 integrases and the corresponding regions within the parent sequences.

recognized by TG1 integrase. To clarify this, we designed hybrid *att* sites with TG1 *att* site sequence in the central part, and the flanking sequences recognized by DBD1 and DBD2 of ϕC31 integrase (Fig. 9) and tested the recombination activities of TG1 and ϕC31 integrases on these substrates. As shown in Fig. 9C, ϕC31 integrase recognized and catalyse a limited recombination (19%) of CtC (*att*P3 × *att*B3), but not CtC (*attP6* × *attB6*). In contrast, TG1 integrase had no activity on either hybrid sequences. This allows unambiguous conclusion about the higher priority of the DNA-binding domains (DBD1 and DBD2) over the catalytic domain and helix E in recognizing *att* sites.

Table 1 shows a summary of the extent of recombination activities of the wild-type and hybrid integrases on all the *att* sites tested.

Table 1. Summary of the extent of recombination activities of wild-type and hybrid integrases on *att* sites. The values shown are obtained from Figs 5, 6, 8, and 9

	Integrases			
***att* sites**	**WT ϕC31**	**WT TG1**	**hCT1**	**hCT2**
ϕC31 *attP* × *attB*	76	0	0	0
TG1 *attP* × *attB*	0	89	0	0
TcT *attP3* × *attB3*	0	88	22	41
TcT *attP6* × *attB3*	0	82	18	32
TcT *attP3* × *attB6*	0	82	8	9
TcT *attP6* × *attB6*	0	80	4	0
CtC *attP3* × *attB3*	19	0	ND	ND
TcT *attR3* × *attL3*	ND	ND	18	50

All values are expressed as percentage of starting substrate converted to product.

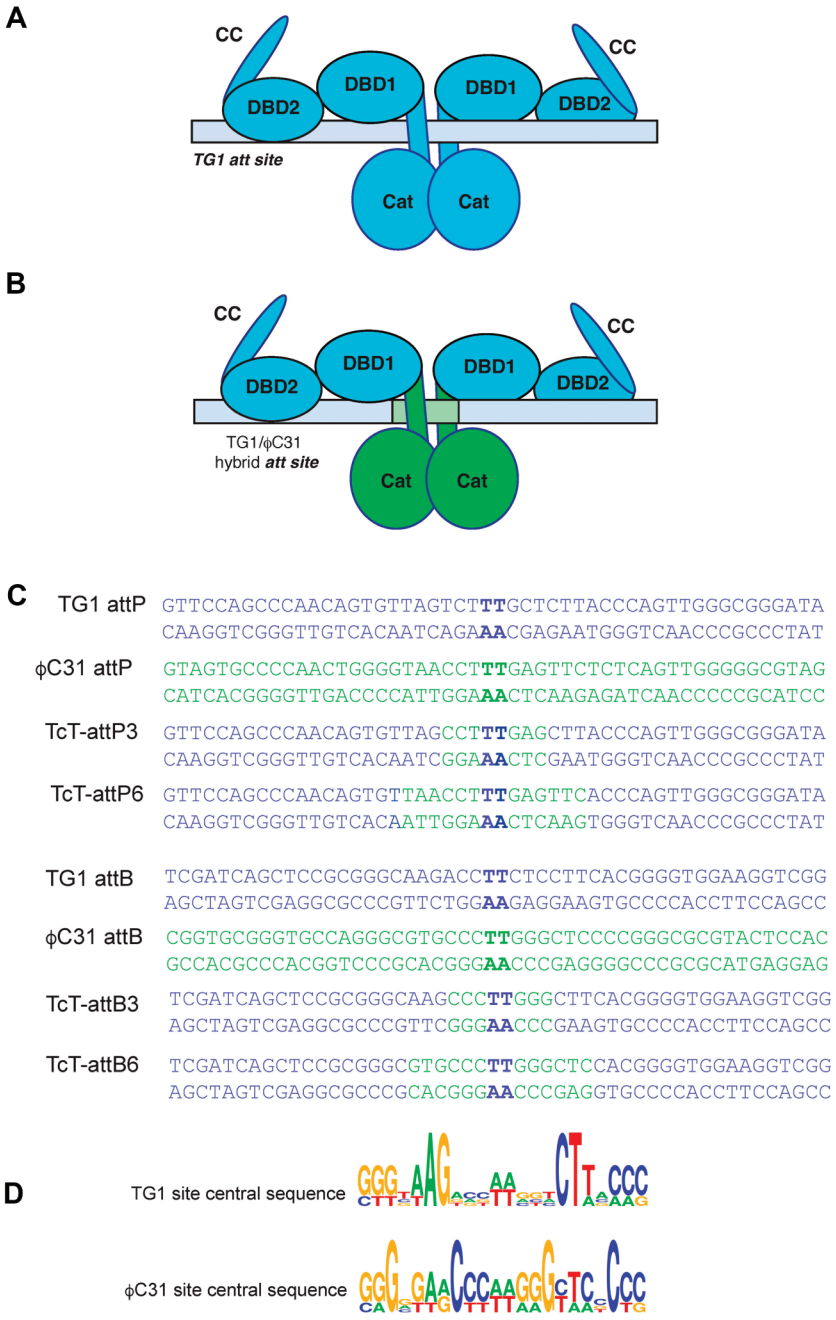


Figure 4. Architecture of wild-type and hybrid integrase and *att* sites. (**A**) Schematic illustration of integrase bound to *attP* site using TG1 integrase as example (**B**) Schematic illustration of hybrid integrase (hCT) bound to hybrid TcT-*attP* site. (**C**) Sequences of recombination *att* sites. ϕC31 integrase and TG1 integrase sequences are shown in green and blue, respectively. Hybrid attachment sites contain ϕC31 *att* site sequence at the central dinucleotides extending outwards in both directions. The hybrid site nomenclature describes the number of substitutions made to each site; *attP3* being TG1 *attP* sequence with 3 nucleotides either side of the central dinucleotides substituted for ϕC31 *attP* sequence. (**D**) Sequence logos for the central sequences of the parent and hybrid sites are shown to highlight patterns of base conservation across *attP* and *attB* [7].

Discussion

This study demonstrates that engineered hybrid ϕC31/TG1 integrases retain sequence-specific recombination activity *in vitro* and can discriminate between compatible *att* site pairs. The hybrids catalysed site-specific recombination only when substrate sequences met their specificity constraints (Table 1), and their activity on *attL* × *attR* sites required the TG1-specific RDF an indication that any interactions made between the regulatory CC domain from TG1 and the catalytic domain and/or helix E of ϕC31 remained functionally intact. These findings provide direct *in vitro* validation that LSRs can be rationally engineered to recognize novel DNA sequences through modular domain exchange and structure-guided design. These findings align with recent advances in LSR engineering, where modifications to catalytic and DNA-binding domains have been shown to enhance specificity and reduce off-target effects [19].

In this work, we used *in vitro* characterization to delve deeper into the properties of hybrid LSRs. Our findings show that hybrid integrases retain features associated with wild-

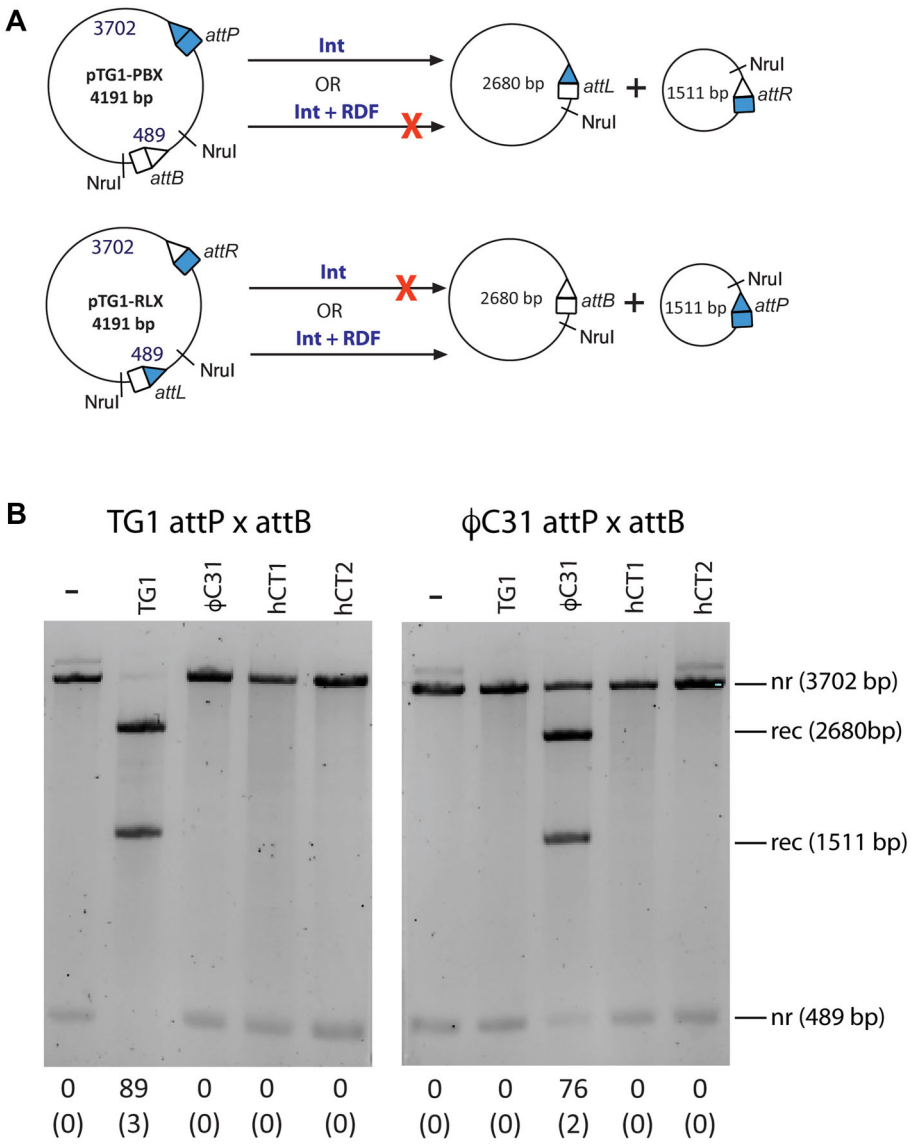


Figure 5. *In vitro* recombination of TG1 and ϕC31 wild-type *att* sites. (**A**) Schematic illustration of the recombination assay. The substrate plasmid shown, pTG1PBX (4191 bp), is a deletion substrate for *attP* × *attB* recombination by TG1 integrase. Other substrates have the same overall architecture. The site orientation promotes the deletion of smaller fragment (1781 bp). Recombination by wild-type integrases, TG1 and ϕC31, or hybrid integrases hCT1 and hCT2, results in two recombination products (rec); the larger 2680 bp *attR* containing plasmid and the deleted 1511 bp *attL* containing plasmid. Non-recombined DNA (nr) are the 3702 bp and 489 bp fragments. (**B**) Recombination activities of integrases on TG1 and ϕC31 wild-type *att* sites. Reactions were incubated for 2 h in the reaction buffer described in the 'Materials and methods' section. Reaction products were digested with the restriction endonuclease NruI prior to visualization on 1.2% agarose gel electrophoresis. Left panel, TG1 *attP* × *attB*; right panel ϕC31 *attP* × *attB*. The mean extent of recombination and standard deviation (%) from quantitation of triplicate experiments are given below each lane.

type parental integrases, and that they are subject to the same rules governing directionality and sequence specificity [34]. We showed that specificity of hybrid integrases is largely determined by the DNA-binding domains of the parental integrases, providing a framework for future development of bespoke integrases that can recognize and recombine novel sites. We also demonstrated that the interaction of hybrid integrases with the RDF can be modularly engineered since RDF specificity is determined by the parental integrase that donates the DBD2/CC interface.

Although prior work by Faruggio and Calos [33] demonstrated *in vivo* activity of similar constructs, this study is the first to define their catalytic potential and RDF dependence in a purified, reconstituted system. In that work, the authors reported the activities of the hybrid LSRs against native and hybrid *att* sites *in vivo*. Some of the hybrid LSRs were catalytically active and retained a level of site-specificity being able to discriminate between *att* sites. Specifically, hybrid LSRs derived from ϕBT1 and ϕC31 integrases, and those from ϕC31 and TG1 integrases catalysed recombination *in vivo* in *E. coli* and in *HeLa* cells [33].

The success of hybrid LSR construction was facilitated by the structural compatibility between ϕC31 and TG1 integrases: both are derived from the same phage family, exhibit high sequence conservation (Fig. 2), and share similar domain organization (Fig. 3). Their *att* sites also display substantial

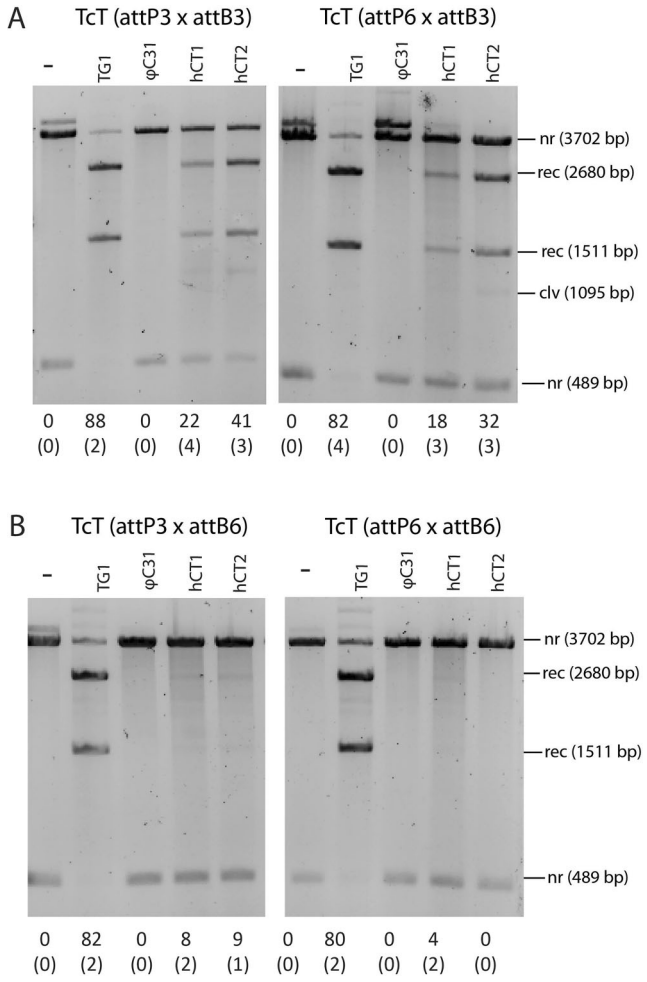


Figure 6. *In vitro* recombination of hybrid TcT *attP* × *attB* sites. (**A**) Left panel, TcT (*attP3* × *attB3*); right panel, TcT (*attP6* × *attB3*). (**B**) Left panel, TcT (*attP3* × *attB6*); right panel, TcT (*attP6* × *attB6*). Reaction conditions, product analysis, and quantitation are as described in Fig. 5. Recombined (rec) and non-recombined (nr) fragments are described in Fig. 5. The 1095 bp cleaved unligated fragment (clv) is visible in panel (B).

homology, minimizing structural perturbations following domain swapping and enabling productive interactions with partially hybrid substrates. This compatibility reduces the likelihood of steric clashes and supports efficient binding and recombination with minimal sequence redesign.

In some reactions, unrecombined cleavage intermediates accumulated, particularly in hCT2-mediated recombination (Fig. 6). These products, arising from successful DNA cleavage but failed strand ligation, appeared as discrete bands following NruI digestion. Since wild-type TG1 integrase efficiently recombined the same substrates without incomplete products, it is likely that cleavage was intact and the religation defect stemmed from instability during or after synapsis. Structural mismatches at the ϕC31/TG1 junction, specifically within the E-helix, may impair rotation or alignment required for ligation. Furthermore, recent structures [5, 6] have shown unexpected interactions between catalytic domains and CC tips, which may stabilize the synaptic complex and would be mismatched in our hybrids. Similar phenotypes have been observed in deregulated resolvase mutants and in chimaeric recombinases [29, 30], where activating mutations or domain mismatches at the E-helix (αE) alter recombination directionality and/or sequence specificity. Future efforts may benefit from refining the splice junction to better preserve ϕC31 domain architecture.

Hybrid LSRs successfully recombined *attP3* and *attP6* but failed to recombine *attB6*, where substitutions may have disrupted key interactions required for synapsis. The additional 3bp segments that were changed in the *attP6* and *attB6* sites are contacted not only by the loop following helix E, but also by other portions of DBD1 [5–7]. The observed pattern implies that changes in *att*B are more disruptive to activity, perhaps because even the wild-type *attB* site, which is found in the bacterial genome, may be suboptimal. Alternatively, the differences between activity on *attP6* vs. *attB6* sites may reflect the fact that only *attB6* changed the identity of the central of these three base pairs, which is conserved in all TG1 sites.

As demonstrated here, functional hybrid integrases are more likely to be formed when the parental integrases are from the same clade. Hybrids made from ϕC31 family including ϕC31, TG1, and BT1 integrases showed varying degrees of recombination activities in an *in vivo* assay [33]. In a different study, it was shown that wild-type Bxz2 and Peaches integrases, which also share a clade, can recombine the same *attB* site, despite sharing just 59% amino acid sequence identity and ∼50% sequence similarity in their *attP* sites [43]. In these two examples, most of the similarities between the two *attP*s are found in the central 6 bp and the DBD2 motif suggesting that these two regions are key determinants of specificity and may indicate which integrase pairs that could be used as starting materials for building functional hybrids. Based on these two published examples, we expect that hybrids made from integrases that belong to the same clade (e.g. ϕC31 family or Bxb1 family) are more likely to show recombination activities especially if there are conserved sequences across DBD2 motifs in their *attP* sites.

Our findings provide some insights into design principles for generating functional hybrids with other integrases. Accurate mapping of sequences contributing to the DNA-binding domains and the junction of the catalytic domain and the E helix is essential in determining the most appropriate splice points. The limited available structural data shows that helix E in particular recognizes DNA sequences through indirect readout, with rules that are difficult to decipher. While structure prediction tools may make this easier, experimental testing of the activities of hybrids on several hybrid sites will be required to refine the process.

Importantly, both hybrid LSRs retained directionality control, catalysing *att*L × *att*R recombination only in the presence of TG1 RDF, consistent with RDF binding via DBD2 and the CC domain. As shown previously, TG1 integrase discriminates between its own RDF and ϕC31 RDF showing limited *attR* × *attL* activities in the presence of the latter [34]. In contrast, ϕC31 integrase is readily activated by non-cognate RDFs from TG1 and ϕBT1. Based on predicted protein structures and protein–protein interaction models, it was speculated that this stricter specificity shown by TG1 integrase derives from a unique DPDD residues found at the base of the CC motif in DBD2, which forms an insertion that interacts with its RDF [34]. TG1 RDF is more different from ϕC31 and BT1 RDFs (61%–63% identity) in contrast to the other two which share 85% sequence similarity to each other. Mutations or protein domain swapping can result in loss of recombination directionality among the serine recombinase family [29, 30,

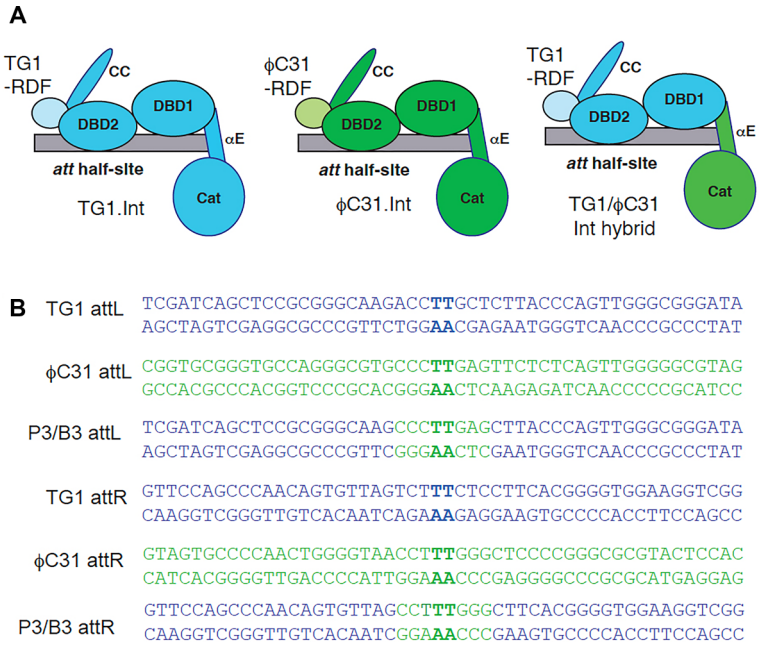


Figure 7. LSR–RDF interactions in wild-type and hybrid integrases. (**A**) Schematic illustration of LSR domain interactions with the recombination site and the RDF based on structure of LI integrase–*att*P complex [13, 34] and AlphaFold multimer-predicted interaction of the RDF with DBD2-proximal region of the CC domain [13, 34]. The hybrid integrases used in this study (hCT1 and hCT2) are predicted to recognize TG1 RDF as illustrated in the right panel. (**B**) *attR* and *attL* sites. The nomenclature of the hybrid *attR* and *attL* sites follows the pattern in Fig. 4 with a description of the number of substitutions made to each site. For example, *P3/B3attL* and *P3/B3attR* being the *attL* and *attR* sequences that results from recombination between *attP3* and *attB3*.

44, 45]. This deregulation usually results from changes on protein interfaces that mediate protein-protein contacts for controlling the directionality of reactions and underscores the importance of the LSR–RDF interface in preserving regulated excision activity.

TG1 integrase recombined all the four hybrid *att* sites shown in Fig. 5 due to similarity between the hybrid sites and TG1 natural *att* sites. Eight of the central 14 base pairs between TG1 and ϕC31 *attP* sites are identical; hence the hybrid sites TcT-*attP*3 and TcT-*attP*6 share sequence similarity to wild-type TG1 *attP*. This similarity explains why TG1 integrase recombines the hybrid sites and supports the conclusion that specificity lies mostly in the DBD motifs of the *att* sites. In contrast, TG1 integrase was unable to recombine CtC *att* sites (Fig. 9), while ϕC31 showed 19% activity on CtC-*attP3*, and no activity on CtC-*attP6*. As shown in Fig. 5, ϕC31 integrase did not recombine any of the TcT hybrid sites. These observations show that while wild-type integrases may recombine hybrid or bespoke sites that bear sequence similarities to their native att sites, carefully engineered hybrid integrases are required to recombine DNA sequences that differ from canonical att sites of natural integrases.

A notable and unexpected outcome was the ability of wild-type TG1 integrase to recombine all hybrid *att* sites tested. TG1 integrase catalyses limited recombination of its attB with ϕC31 attP but is unable to recombine ϕC31 attB with its own attP (F.J. Olorunniji, unpublished). Hence efficient recombination of hybrid *att*B sites by TG1 integrase (Fig. 6), despite inability to recombine wild-type ϕC31 *att*B, suggests that the *att*P-binding event may be the primary determinant of productive synapsis. High-affinity binding to *att*P may compensate for suboptimal *att*B recognition, enabling recombination despite sequence divergence. This could be a determining factor in choosing LSRs for applications requiring integration into non-cognate pseudo-attB sites. The broad activity of TG1 integrase across diverse hybrid *att* sites may reflect how the catalytic domain and helix E interact with DNA and suggests that DBD1 and DBD2 are primarily responsible for properly positioning the LSR on its target site. In contrast, ϕC31 integrase showed no activity on the hybrid sequences, an expected finding given the TG1-like composition of the DBD1 and DBD2-binding portions of the *att* sites, but consistent with prior observations that ϕC31 integrase fails to recombine TG1 *attP* or *attB* sites, even when paired with ϕC31 *attB* or *attP*, respectively (F.J. Olorunniji, unpublished). In both *attP* and *attB*, 5 out of the 6 bases flanking the central dinucleotides are identical between ϕC31 and TG1 sites. Hence, the hybrid site CtC (*attP3* × *attB3*) may not be significantly functionally different from ϕC31 *attP* and *attB* sites. Yet, ϕC31 integrase showed a limited recombination activity (19%) on CtC (*attP3* × *attB3*). In contrast, TG1 integrase achieved 88% recombination of hybrid TcT (*attP3* × *attB3*) (Fig. 6). This may reflect an inherent higher affinity (binding and/or synapsis) for cognate *attP*/*attB* sites by TG1 integrase than ϕC31 integrase.

These findings underscore the functional discrimination retained by the hybrids and suggest that substrate recognition is dominantly governed by *att*–DBD interactions and domain-specific alignments. It is also noteworthy that ϕC31 integrase showed limited activity on CtC hybrid substrates while TG1 integrase efficiently recombined all TcT hybrid substrates tested. This difference might reflect the inherent higher catalytic efficiency of TG1 integrase in comparison to ϕC31 integrase [34, 37]. Another explanation for the low activity of ϕC31 integrase on CtC (*attP3* × *attB3*) could be related to

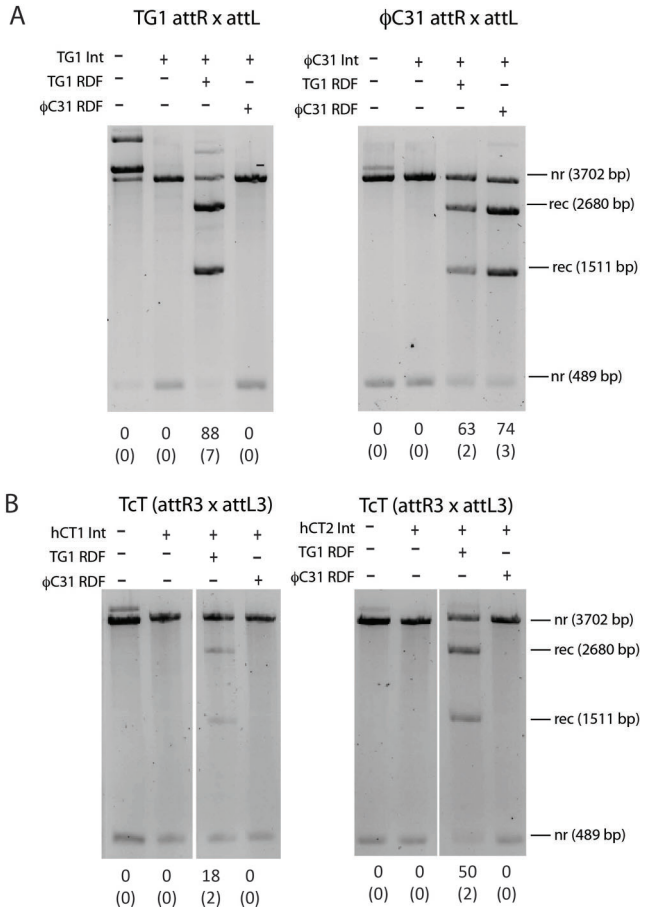


Figure 8. *In vitro* recombination of hybrid TcT *attR* × *attL* sites by hybrid integrases. The substrate architecture, reaction conditions, product analysis, and quantitation are as described in Fig. 5. (**A**) Left panel, TG1 *attR* × *attL*; right panel, ϕC31 *attR* × *attL*. (**B**) Left panel, hCT1 reaction on TcT (*attR3* × *attL3*); right panel, hCT2 reaction on TcT (*attR3* × *attL3*). Reaction conditions are as described in Fig. 5.

the effect of replacing bases 1–3 with TG1 sequences in both *attP* and *attB*. The changes resulted in fewer identical bases at DBD1 positions across P, P', B, and B' half sites compared to wild-type *attP* and *attB* (Fig. 9). These changes in the sequence around the central dinucleotide disrupts the aligned symmetry of bases across P, P', B, and B' half sites [1]. Such changes could result in reduced binding affinity and hence less optimal recombination activity. In contrast, the changes in TG1 *attP* and *attB* made in TcT (*attP*3 × attB3) did not alter the overall aligned symmetry in the DBD1 motif of the *att* sites (Fig. 4), making it catalytically accessible to TG1 integrase.

The inability of hCT1 and hCT2 to recombine TG1 *attP* × *attB* may result from the contributions of Helix E and the catalytic domain to interactions that determine the recognition of bases around the central core of *attP* and *attB* sites. Recent cryoEM structures of SPβ integrase [5] and ϕC31 integrase [6] bound to *att* sites in synaptic tetramer complexes reveal direct contact between the E Helices and the core central 8 bases in both integrases. We speculate that inability of the ϕC31 E helix in hCT1 and hCT2 to recognize the central core of TG1 *attP* × *attB* is the primary reason for the lack of activity as shown in Fig. 5, despite having the cognate DBD1 and DBD2 for those sites. This might also account for the observed differences in activities between hCT1 and hCT2 (Figs 6 and 8). Generally, hCT2 with more ϕC31 integrase component in the E helix showed higher activity than hCT1, presumably allowing better recognition of TcT sequences with ϕC31 sequences around the central core of *attP* and *attB*. These patterns suggest that future design of hybrid integrases need to include considerations of the contribution of Helix E to site recognition and specificity. It may also be that, despite engineering the hCT1 splice site within a conserved sequence, it disrupted a stabilizing salt bridge between the D on one side of the splice and a K 3 residues later on the other side of the splice site. This may also explain the generally low activity of the hCT1 hybrid and suggests strategies for improving it.

Our findings show that the DNA-binding domains (DBD1 and DBD2), and to some degree helix E and the loop that follows it, but not the catalytic domain, of LSRs are primarily important for site-specific DNA recognition and recombination (Fig. 9). It follows that to make bespoke recombinases, it may be possible to select a highly active LSR whose domains recognize the target DNA sequences, and with some targeted mutagenesis optimize the activity and specificity of the LSR. Selection of LSR pairs with highly similar sequences in the centres of their *att* sites may also facilitate design of highly active hybrids. This approach could be an effective strategy for targeting and recombining user-defined *att* sites in DNA using naturally existing LSRs.

Recently, McClain *et al.* [46] developed integrase on-demand (IOD), a novel method for systematically identifying native LSRs and their corresponding *attB* sites within a target organism's genome, an approach that can potentially eliminate the need for synthetic landing pads. The study used computational predictions and experimental validation to demonstrate the functionality of predicted integrase-*attB* pairs in various bacterial strains, including *Pseudomonas putida* and *Synechococcus elongatus*. Our synthetic hybrid LSR approach could complement IOD by designing bespoke LSRs that can integrate payloads into pseudo-*attB* sites in target genomes.

Conclusion

This study defines a modular strategy for engineering LSRs capable of catalysing site-specific recombination on synthetic hybrid substrates. Beyond demonstrating the feasibility of *in vitro* hybrid integrase activity, the results highlight the crucial role of *attB*–RD domain interactions in determining recombination specificity. In particular, the findings raise important questions about the structural constraints that govern *attB* recognition and the sequence elements required beyond the proximal bases flanking the central dinucleotide. It is likely that contacts between the catalytic domain and DBD1 may contribute to sequence specificity [5, 6], necessitating the need for careful selection of LSRs for use in domain swapping strategies. Addressing these questions through further structural and mutational analysis will refine the rules of *att* site targeting and accelerate the development of programmable recombination tools for synthetic biology and genome engineering.

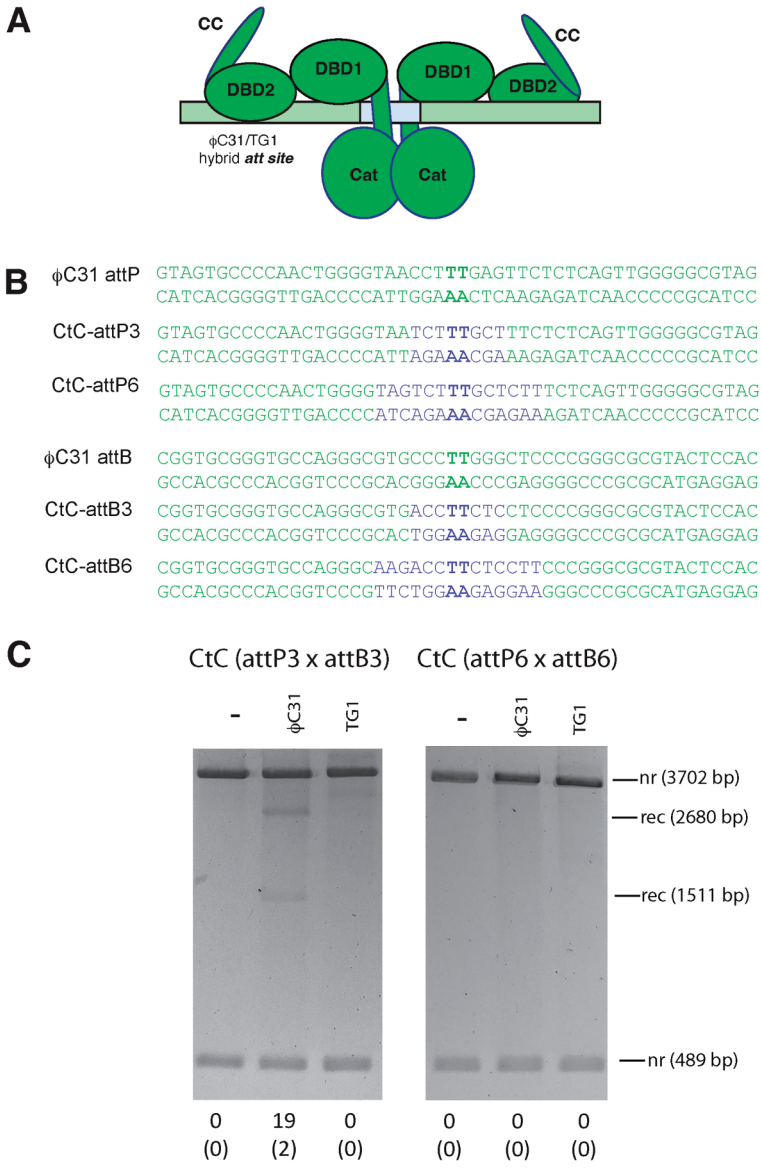


Figure 9. Recombination activities of ϕC31 integrase on hybrid CtC *att* site. (**A**) Schematic illustration of ϕC31 integrase bound to hybrid CtC-*attP* site. (**B**) Sequences of hybrid CtC recombination *att* sites. ϕC31 integrase and TG1 integrase sequences are shown in green and blue, respectively. Hybrid attachment sites containTG1 *att* site sequence at the central dinucleotides extending outwards in both directions replacing corresponding ϕC31 *att* sequences. The hybrid site nomenclature describes the number of substitutions made to each site; CtC *attP3* being ϕC31 *attP* sequence with 3 nucleotides either side of the central dinucleotides substituted for TG1 *attP* sequence. (**C**) Recombination activities of ϕC31 integrase on hybrid CtC sites. Reaction conditions, product analysis, and quantitation are as described in Fig. 5, except that reactions were carried out for 8 h.

Acknowledgements

We thank Abdulrazak Alsaleh for conducting some of the *in vitro* reactions that supported this study.

Author contributions: Makeba Lawson-Williams (Investigation [equal], Methodology [equal], Writing – original draft [supporting]), Alexandria Holland (Investigation [supporting], Writing – original draft [supporting]), Adebayo J. Bello (Investigation [supporting]), Phoebe A Rice (Funding acquisition [equal], Methodology [equal], Writing – review & editing [equal]), and Femi John Olorunniji (Conceptualization [equal], Funding acquisition [lead], Methodology [lead], Writing – original draft [lead], Writing – review & editing [lead])

Conflict of interest

None declared.

Funding

This work was supported by the UK Research and Innovation (UKRI) and National Science Foundation (NSF) (collaborative grant NSF/BIO 2223480 and UKRI/BBSRC BB/X012085/1 to P.A.R. and F.J.O.), and Wellcome Trust Accelerator Award 316534/Z/24/Z to F.J.O. Funding for open access charge: UKRI/BBSRC. Funding to pay the Open Access publication charges for this article was provided by UKRI/BBSRC.

Data availability

All the data underlying this article are contained within the article.

References

1. Van Duyne GD, Rutherford K. Large serine recombinase domain structure and attachment site binding. *Crit Rev Biochem Mol Biol* 2013;48:476–91.
2. Lewis JA, Hatfull GF. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. *Nucleic Acids Res* 2001;29:2205–16. https://doi.org/10.1093/nar/29.11.2205
3. Stark WM. Making serine integrases work for us. *Curr Opin Microbiol* 2017;38:130–6. https://doi.org/10.1016/j.mib.2017.04.006
4. Smith MCM, Thorpe HM. Diversity in the serine recombinases. *Mol Microbiol* 2002;44:299–307. https://doi.org/10.1046/j.1365-2958.2002.02891.x
5. Shin H, Pigli Y, Peña Reyes T *et al.* Structural basis of directionality control in large serine integrases. bioRxiv, https://doi.org/10.1101/2025.01.03.631226, 5 February 2026, preprint; not peer reviewed.
6. Sun YE, Aspinall L, Joseph AP *et al.* Structural basis of DNA recombination catalysis and regulation by ϕC31 integrase. bioRxiv, https://doi.org/10.1101/2025.05.02.651858, 2 May 2025, preprint; not peer reviewed.
7. Rutherford K, Yuan P, Perry K *et al.* Attachment site recognition and regulation of directionality by the serine integrases. *Nucleic Acids Res* 2013;41:8341–56. https://doi.org/10.1093/nar/gkt580
8. Rowley PA, Smith MCA, Younger E *et al.* A motif in the C-terminal domain of phiC31 integrase controls the directionality of recombination. *Nucleic Acids Res* 2008;36:3879–91. https://doi.org/10.1093/nar/gkn269
9. Mandali S, Gupta K, Dawson AR *et al.* Control of recombination directionality by the listeria phage A118 protein Gp44 and the coiled-coil motif of its serine integrase. *J Bacteriol* 2017;199:e00019–17. https://doi.org/10.1128/JB.00019-17
10. Fogg PCM, Younger E, Fernando BD *et al.* Recombination directionality factor gp3 binds ϕC31 integrase via the zinc domain, potentially affecting the trajectory of the coiled-coil motif. *Nucleic Acids Res* 2018;46:1308–20. https://doi.org/10.1093/nar/gkx1233
11. Abe K, Takahashi T, Sato T. Extreme C-terminal element of SprA serine integrase is a potential component of the 'molecular toggle switch' which controls the recombination and its directionality. *Mol Microbiol* 2021;115:1110–21. https://doi.org/10.1111/mmi.14654
12. Alsaleh A, Holland A, Shin H *et al.* Large serine integrases utilise scavenged phage proteins as directionality cofactors. *Nucleic Acids Res* 2025;53:gkaf050. https://doi.org/10.1093/nar/gkaf050
13. Shin H, Holland A, Alsaleh A *et al.* Identification of cognate recombination directionality factors for large serine recombinases by virtual pulldown. *Nucleic Acids Res* 2025;53:gkaf691. https://doi.org/10.1093/nar/gkaf691
14. Ghosh P, Bibb LA, Hatfull GF. Two-step site selection for serine-integrase-mediated excision: dNA-directed integrase conformation and central dinucleotide proofreading. *Proc Natl Acad Sci USA* 2008;105:3238–43. https://doi.org/10.1073/pnas.0711649105
15. Olorunniji FJ, Rosser SJ, Stark WM. Site-specific recombinases: molecular machines for the genetic revolution. *Biochem J* 2016;473:673–84. https://doi.org/10.1042/BJ20151112
16. Ba F, Zhang Y, Wang L *et al.* Applications of serine integrases in synthetic biology over the past decade. *SynBio* 2023;1:172–89. https://doi.org/10.3390/synbio1020012
17. Autio MI, Perrin A. Highly enhanced transgene integration in human pluripotent stem cells by serine integrase BxbI. bioRxiv, https://doi.org/10.1101/2024.10.18.619063, 19 October 2024, preprint; not peer reviewed.
18. Rath P, Kramer P, Biggs D *et al.* Optimizing approaches for targeted integration of transgenic cassettes by integrase-mediated cassette exchange in mouse and human stem cells. *Stem Cells* 2025;43:sxae092. https://doi.org/10.1093/stmcls/sxae092
19. Rose J, Kim H, Von Stetina J *et al.* Engineered Bxb1 variants improve integrase activity and fidelity. bioRxiv, https://doi.org/10.1101/2024.10.21.619419, 21 October 2024, preprint; not peer reviewed.
20. Durrant MG, Fanton A, Tycko J *et al.* Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat Biotechnol* 2023;41:488–99. https://doi.org/10.1038/s41587-022-01494-w
21. Fanton A, Bartie LJ, Martins JQ *et al.* Site-specific DNA insertion into the human genome with engineered recombinases. *Nat Biotechnol* 2025. https://doi.org/10.1038/s41587-025-02895-3
22. Hiraizumi M, Perry NT, Durrant MG *et al.* Structural mechanism of bridge RNA-guided recombination. *Nature* 2024;630:994–1002. https://doi.org/10.1038/s41586-024-07570-2
23. Hazelbaker DZ, Mehta JB, McGinnis C *et al.* Large serine integrase off-target discovery and validation for therapeutic genome editing. bioRxiv, https://doi.org/10.1101/2024.08.23.609471, 24 August 2024, preprint; not peer reviewed.
24. Bakalar MH, Biondi T, Liang X *et al.* Large serine integrase off-target discovery with deep learning for genome wide prediction. bioRxiv, https://doi.org/10.1101/2024.10.10.617699, 13 October 2024, preprint; not peer reviewed.
25. Fauser F, Arangundy-Franklin S, Davis JE *et al.* Retargeted serine integrases for one-step, precise integration of large DNA sequences in human cells. *Nat Biotechnol* 2026. https://doi.org/10.1038/s41587-026-03186-1
26. Anzalone AV, Gao XD, Podracky CJ *et al.* Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat Biotechnol* 2021;40:731–40. https://doi.org/10.1038/s41587-021-01133-w
27. Blanch-Asensio A, Grandela C, Brandão KO *et al.* STRAIGHT-IN enables high-throughput targeting of large DNA payloads in human pluripotent stem cells. *Cell Reports Methods* 2022;2:100300. https://doi.org/10.1016/j.crmeth.2022.100300
28. Ba F, Liu Y, Liu W-Q *et al.* SYMBIOSIS: synthetic manipulable biobricks via orthogonal serine integrase systems. *Nucleic Acids Res* 2022;50:2973–85. https://doi.org/10.1093/nar/gkac124
29. Akopian A, He J, Boocock MR *et al.* Chimeric recombinases with designed DNA sequence recognition. *Proc Natl Acad Sci USA* 2003;100:8688–91. https://doi.org/10.1073/pnas.1533177100
30. Prorocic MM, Wenlong D, Olorunniji FJ *et al.* Zinc-finger recombinase activities *in vitro*. *Nucleic Acids Res* 2011;39:9316–28. https://doi.org/10.1093/nar/gkr652
31. Proudfoot C, McPherson AL, Kolb AF *et al.* Zinc finger recombinases with adaptable DNA sequence specificity. *PLoS One* 2011;6:e19537. https://doi.org/10.1371/journal.pone.0019537
32. Olorunniji FJ, Lawson-Williams M, McPherson AL *et al.* Control of ϕC31 integrase-mediated site-specific recombination by protein trans-splicing. *Nucleic Acids Res* 2019;47:11452–60. https://doi.org/10.1093/nar/gkz936
33. Farruggio AP, Calos MP. Serine integrase chimeras with activity in *E. coli* and HeLa cells. *Biol Open* 2014;3:895–903. https://doi.org/10.1242/bio.20148748
34. MacDonald AI, Baksh A, Holland A *et al.* Variable orthogonality of serine integrase interactions within the ϕC31 family. *Sci Rep* 2024;14:26280. https://doi.org/10.1038/s41598-024-77570-9
35. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
36. Olorunniji FJ, McPherson AL, Rosser SJ *et al.* Control of serine integrase recombination directionality by fusion with the

directionality factor. *Nucleic Acids Res* 2017;45:8635–45. https://doi.org/10.1093/nar/gkx567
37. Abioye J, Lawson-Williams M, Lecanda A *et al.* High fidelity one-pot DNA assembly using orthogonal serine integrases. *Biotechnol J* 2023;18:2200411. https://doi.org/10.1002/biot.202200411
38. Yang W, Steitz TA. Crystal structure of the site-specific recombinase γδ resolvase complexed with a 34 by cleavage site. *Cell* 1995;82:193–207. https://doi.org/10.1016/0092-8674(95)90307-0
39. Li W, Kamtekar S, Xiong Y *et al.* Structure of a synaptic γδ resolvase tetramer covalently linked to two cleaved DNAs. *Science* 2005;309:1210–5. https://doi.org/10.1126/science.1112064
40. Mouw KW, Rowland S-J, Gajjar MM *et al.* Architecture of a serine recombinase–DNA regulatory complex. *Mol Cell* 2008;30:145–55. https://doi.org/10.1016/j.molcel.2008.02.023
41. Trejo CS, Rock RS, Stark WM *et al.* Snapshots of a molecular swivel in action. *Nucleic Acids Res* 2018;46:5286–96. https://doi.org/10.1093/nar/gkx1309
42. Olorunniji FJ, Buck DE, Colloms SD *et al.* Gated rotation mechanism of site-specific recombination by ϕC31 integrase. *Proc Natl Acad Sci USA* 2012;109:19661–6. https://doi.org/10.1073/pnas.1210964109
43. Singh S, Rockenbach K, Dedrick RM *et al.* Cross-talk between diverse serine integrases. *J Mol Biol* 2014;426:318–31. https://doi.org/10.1016/j.jmb.2013.10.013
44. Arnold PH. Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. *EMBO J* 1999;18:1407–14. https://doi.org/10.1093/emboj/18.5.1407
45. Olorunniji FJ, He J, Wenwieser SVCT *et al.* Synapsis and catalysis by activated Tn3 resolvase mutants. *Nucleic Acids Res* 2008;36:7181–91. https://doi.org/10.1093/nar/gkn885
46. McClain HM, Lowrey LC, Quinto LB *et al.* Integrase-On-Demand: bioprospecting integrases for targeted genomic insertion of genetic cargo. *Nucleic Acids Res* 2026;54:gkag106. https://doi.org/10.1093/nar/gkag106
47. Crooks GE, Hon G, Chandonia J-M *et al.* WebLogo: a sequence logo generator: Fig. 1. *Genome Res* 2004;14:1188–90. https://doi.org/10.1101/gr.849004
48. Sievers F, Wilm A, Dineen D *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011;7:539. https://doi.org/10.1038/msb.2011.75

Received: December 15, 2025. **Revised:** July 5, 2026. **Accepted:** July 7, 2026