Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

Making serine integrases work for us

W Marshall Stark

DNA site-specific recombinases are enzymes (often associated with mobile DNA elements) that catalyse breaking and rejoining of DNA strands at specific points, thereby bringing about precise genetic rearrangements. Serine integrases are a group of recombinases derived from bacteriophages. Their unusual properties, including directionality of recombination and simple site requirements, are leading to their development as efficient, versatile tools for applications in experimental biology, biotechnology, synthetic biology and gene therapy. This article summarizes our current knowledge of serine integrase structure and mechanism, then outlines key factors that affect the performance of these phage recombination systems. Recently published studies, that have expanded the repertoire of available systems and reveal system-specific characteristics, will help us to choose the best integrases for envisaged applications.

Address
Institute of Molecular, Cell and Systems Biology, University of Glasgow, Bower Building, Glasgow G12 8QQ, Scotland, United Kingdom

Corresponding author: Stark, W Marshall (Marshall.Stark@glasgow.ac.uk)

Current Opinion in Microbiology 2017, **38**:130–136

This review comes from a themed issue on **Mobile genetic elements and HGT in prokaryotes**

Edited by **Andrew Lang**, **J Thomas Beatty** and **Phoebe Rice**

http://dx.doi.org/10.1016/j.mib.2017.04.006

Introduction

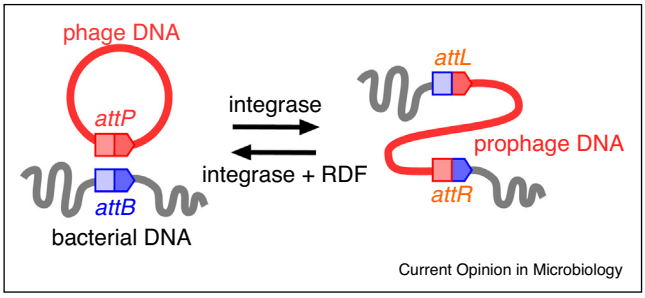
Since the early days of molecular biology, the world of prokaryote mobile genetic elements and the cellular defences against them has proved to be a fertile source of tools for manipulation of DNA molecules [1,2]. Examples include restriction enzymes, DNA polymerases and ligases, transposases, and CRISPR-Cas systems. Another large class of useful systems that is prevalent in mobile elements carry out precise, programmed DNA rearrangements by the process known as site-specific recombination. Each site-specific recombination system requires a recombinase enzyme that recognizes specific sites in the DNA and catalyses strand cleavage and rejoining. One class of recombinases called the serine integrases is currently attracting much interest from those seeking practical methods for precise manipulation of DNA molecules or sequences.

Many bacteriophages with double-stranded DNA genomes have the capacity to integrate their genomic DNA into that of a host cell, thereby entering a lysogenic state where the prophage DNA is passively replicated along with the host DNA over many generations. At some point the prophage DNA is excised and the phage enters the lytic phase, whereupon it replicates its DNA genome repeatedly and packages it into phage particles prior to cell lysis. Integration and excision are mediated by a phage-encoded integrase enzyme. Integrases belong to one of the two large families of site-specific recombinases, called the serine recombinases and the tyrosine recombinases according to the nucleophilic active site amino acid residue that attacks specific DNA phosphodiesters to cleave strands. Integration and excision of the archetypal lysogenic phage, λ, are promoted by a tyrosine integrase. Serine integrases were discovered relatively recently (in the 1990s), one of the first examples to be studied being that of the *Streptomyces* phage ϕC31 [3,4•,5].

Phage integration is brought about by integrase-promoted recombination between an *attP* site in the circularized phage DNA and a bacterial genomic site *attB*. Excision is essentially the reverse of integration; integrase promotes recombination between the two sites flanking the prophage DNA (called *attL* and *attR*) [5] (Figure 1). Integrase thus promotes both integration and excision, but phage biology demands that only one 'direction' is preferred at any particular time. Phages have therefore evolved mechanisms to achieve directionality; that is, to drive recombination towards integration when entering lysogeny, and towards excision when entering the lytic state. Another phage-encoded protein is required for specific stimulation of *attL* × *attR* recombination; often called Xis ('excisionase') for λ integrase and its tyrosine integrase relatives, or RDF (Recombination Directionality Factor) for serine integrases [5,6].

Although the tyrosine and serine integrases both carry out essentially the same function, their modes of action are very different. The *attB* sites for tyrosine integrases are short (~25–40 bp) and are bound by an integrase dimer, but the *attP* sites are typically long, with multiple binding sites for integrase subunits as well as other 'accessory' proteins which may be required to stimulate recombination (in the case of the 240-bp phage λ *attP*, these include the bacterial host-encoded architectural proteins IHF and

Figure 1



Bacteriophage integration and excision. See text for details.

Fis; [7]). The *attL* and *attR* sites, made by splicing *attB* with *attP*, are thus also long and complex. In contrast, the *attP* and *attB* sites for serine integrases are both short (*attP* ~50 bp; *attB* ~40 bp); each site is bound by a dimer of integrase, and no host-encoded proteins are required. The mechanisms of DNA strand exchange also differ; tyrosine integrases make single-strand breaks in the DNA and exchange a single strand of each site to form an intermediate Holliday junction-like DNA structure, which is resolved to recombinants by cleavage and exchange of the other pair of strands [7], whereas serine integrases make double-strand breaks in the DNA and exchange strands by a rotational mechanism [3,5] (Figure 2).

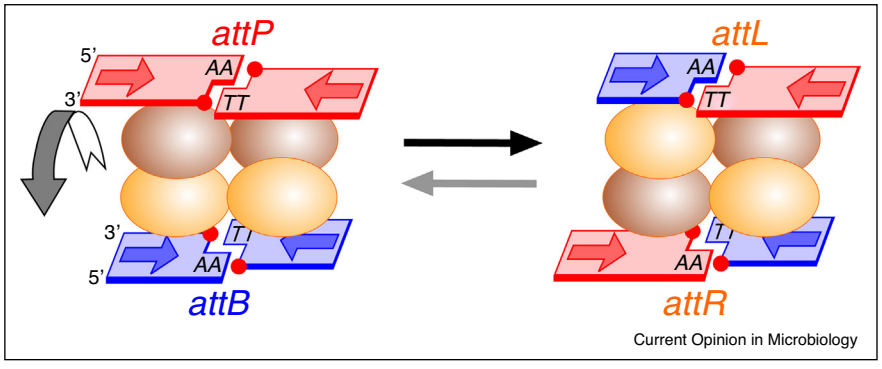
Both families of integrases have been widely adopted for applications in biotechnology and synthetic biology that demand precise, efficient DNA rearrangements [8,9]. Notably, the phage λ tyrosine integrase (Int) is the basis of the very popular Gateway cloning system [10]. However, the serine integrase systems, with apparently simpler DNA components and without requirement for host factors, have recently moved into the limelight, with numerous actual or mooted applications.

Serine integrase mechanism

Following recognition of the *att* sites by integrase, dimers bound at an *attP* and an *attB* interact to form a synaptic complex (Figure 2). Integrase then promotes double-strand cleavage at the centres of both sites. After strand cleavage, each DNA 'half-site' is covalently attached to an integrase subunit, via a phosphodiester linking the active site serine residue to the recessed 5′ DNA end. The cleaved half-sites along with the attached integrase subunits then swap positions by a rotational movement, and the DNA ends are rejoined in this recombinant configuration by reversal of the cleavage mechanism, giving *attL* and *attR* sites [4•,11•,12]. The *attP* and *attB* sites (and thus the integrase-*att* complexes) have imperfect two-fold symmetry, which suggests that they could recombine in two different ways, with the sites aligned in 'parallel' or 'antiparallel'. However, the top and bottom strand cleavages of the sites are at each end of an asymmetric central 2-bp sequence, creating 2-nt overhangs on the cleaved DNA ends which must correctly base-pair in the recombinant sites. Recombination thus demands parallel alignment (as in Figure 2) [3,5,13•].

Integrase binds to all four types of *att* sites, but in the absence of RDF it characteristically promotes recombination only between *attP* and *attB*. Excisive (*attL* × *attR*) recombination additionally requires the presence of the RDF, but the basic catalytic mechanism is thought to be otherwise essentially the same as for *attP* × *attB* recombination. The asymmetric central 2 bp of the sites enforces parallel recombination as described above

Figure 2



Mechanism of strand exchange by serine integrases, The *att* sites are shown with double-strand breaks, after strand cleavage (left) and before ligation (right). The ends are exchanged by a rotational movement of DNA with attached integrase subunits (orange ovals). The recessed 5′ ends, attached to the integrase via a phosphodiester to the active site serine residue, are shown as red dots. The blue and red arrowheads indicate similar sequence motifs giving the sites imperfect two-fold symmetry. The asymmetric central dinucleotide sequence shown (AA/TT) is that of the ϕC31 *att* sites. Note that, if the sites were aligned 'antiparallel' (*e.g.*, by reversing the orientation of *attB*), the product sites would contain mismatched basepairs; such products are not formed (see text for details). The 'reverse' reaction (grey arrow) additionally requires RDF, which interacts with integrase subunits.

[13•,14]. RDF reprogrammes integrase to favour *attL* × *attR* recombination by binding directly to the integrase protein [5,15,16] (not the DNA, as is the case for Xis in tyrosine integrase systems; [7]).

Unlike many other site-specific recombinases, serine integrases do not display a strong specificity for the structure of the substrate DNA molecules; they will recombine sites in the same DNA molecule (in either head-to-head or head-to-tail orientation) or in different molecules, and the molecules themselves can be supercoiled, or linear, or even double-stranded oligonucleotides [5,8].

Serine integrases are large multidomain proteins, typically 400 amino acids or more. Their membership of the serine recombinase 'supergroup' is established by the sequence and structural similarity of a ~150-amino acid domain at the integrase N-terminus to the catalytic domains of well known 'small' serine recombinases such as the Tn*3* and γδ resolvases [3,5,17]. In the small serine recombinases, the N-terminal catalytic domain is immediately followed by a small DNA-binding C-terminal domain (CTD) that confers sequence specificity. The much larger serine integrase CTD remained relatively poorly understood until 2013, when the Van Duyne group published structures of this part of phage A118 integrase bound to a DNA half-site [18••]. These remarkably informative structures revealed that the integrase CTD in fact comprises three domains, two of which are involved in DNA sequence recognition. A third domain with a coiled-coil structure is proposed to play an essential role in synapsis of two sites bound by integrase dimers. Its different positions in complexes with *attP*, *attB*, *attL* and *attR* could account for specificity of recombination between *attP* and *attB* in the absence of RDF. Binding of RDF to integrase is proposed to alter the positioning of the coiled-coil domains and thereby favour *attL* × *attR* recombination [18••,19–21].

Serine integrase diversity

A large number of putative serine integrases can be identified from DNA sequence databases, along with other large serine recombinases that are not associated with phages [5,22,23••]. In order to reconstitute a complete integration system (*e.g.*, for a biotechnology application), one must also identify the recombination sites and the RDF. At present only about 20 serine integrase-based systems have been characterized in substantial depth, and even fewer (<10) have a known RDF [5].

Applications of serine integrases

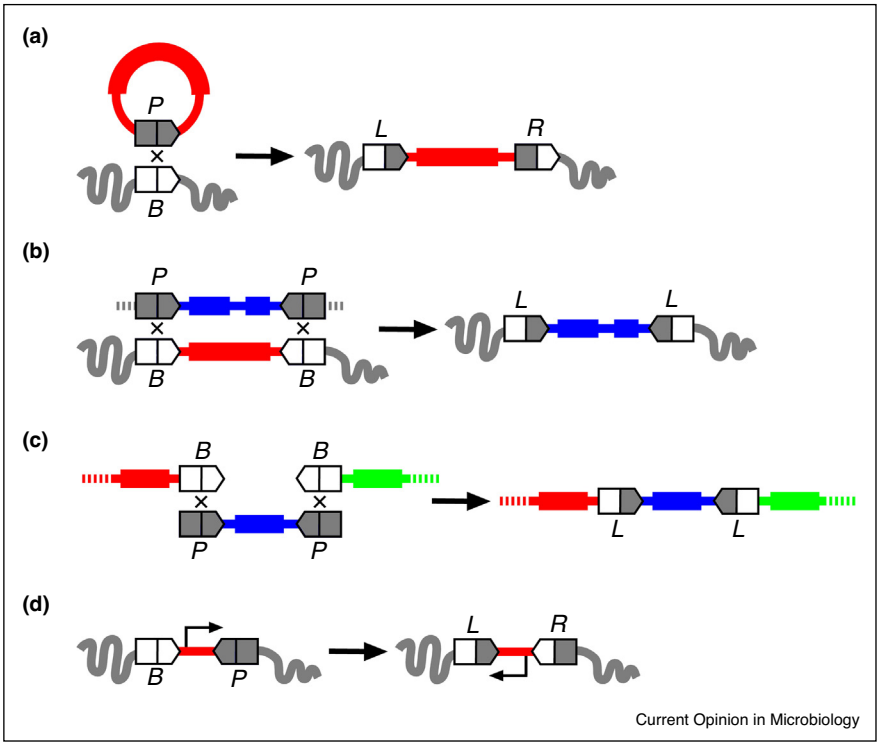
Very soon after the discovery of the first serine integrase systems, they began to be put to use. Early applications were for site-specific transgene integration, at first in bacteria; but it was soon shown that serine integrases would function efficiently in many organisms including multicellular eukaryotes such as flies and humans. Their use has now become widespread in numerous organisms, and applications have become more sophisticated, often involving rearrangements by multiple recombinases [5,8,9]. Recently, there has been a flood of interest in uses of serine integrases for synthetic biology applications; for assembly of long arrays of genes or other DNA fragments, creation of genetic switches, circuits and logic systems, and for permanent genomic recording of cellular events [8,9,23••,24–28,29•,30•,31,32••,33]. The diagrams in Figure 3 illustrate these types of processes. Although other site-specific recombination systems share features of precision and efficiency with the serine integrases, the unique selling point of the latter systems is their one-way directional nature; the products of integrase-mediated *attP* × *attB* recombination are typically inert to any further reactions (unless RDF is present).

The perfect integrase for these applications would be a stable protein, in cells and/or *in vitro*, which is robustly active under varying conditions. Integrase-promoted recombination should be fast and highly directional (complete conversion of substrates to recombinants, with no back reaction). Furthermore, recombination should be site-specific (no off-target activity) and precise (no DNA damage or mutation of the sites). For some applications we would like to use several integrases together, each integrase acting exclusively on its cognate sites. Ideally there should be no crosstalk between these integrases, either at the DNA or the protein level, and each RDF should act strictly on its cognate integrase. The serine integrases characterized to date seem to have these desirable properties to varying extents. For future advanced applications of these systems, it will be essential to understand the idiosyncrasies of each integrase and to optimize them for the intended uses. Some integrases may be more suitable than others for specific applications, and we might be able to modify system parts to improve performance.

Finding more serine integrases

Until recently, serine integrase-based recombination systems were identified and characterized fairly sporadically as their coding sequences turned up in newly sequenced phage genomes. However, the recent upsurge in interest has led to systematic attempts to find more, diverse serine integration systems [23••]. Coding sequences for serine integrases are quite easy to find in databases because of the similarity of their N-terminal, catalytic domains (see above). However, the identification and delimitation of the recombination (*att*) sites for a novel integrase, and particularly the identification of its RDF, can be much more difficult. The most successful strategy for finding the *att* sites has been to sequence prophage from bacterial genomic DNA; *attL* and *attR* should be at the bacterial-phage DNA junctions, and from these the *attP* and *attB* sequences can be deduced [5,23••,34,35]. With the

Figure 3



Example applications of integrase-promoted recombination in synthetic biology and biotechnology. **(a)** Integration of a circular transgene-bearing DNA molecule by recombination with a chromosomal site. **(b)** Cassette exchange. **(c)** Assembly of multiple DNA fragments into a linear array. **(d)** An inversion switch. The grey squiggles represent flanking chromosomal DNA, and dashed lines represent flanking DNA that is not retained in the recombinant of interest. The bent black arrow in (d) represents a hypothetical promoter within the invertible DNA segment whose orientation might cause alternative expression states of flanking genes, for example as part of a genetic circuit. The orientation of an invertible segment can also represent a binary digit (0 or 1) in a genetic logic or memory system. For more details, see Ref. [9].

integrase and *att* sites, it can be demonstrated that the system is active and is not a defunct relic from a once-active phage. Only a few RDFs have been characterized [5]. Unlike the integrases, there is no clear sequence conservation among RDFs; although some closely related integrases have related RDFs, others vary dramatically in length and sequence [34,36]. Also, the RDF can be a protein with additional functions [37]. Nor is there a clear conservation of location of the RDF gene in the phage genome.

Site-specificity

Studies to date *in vivo* and *in vitro* indicate that serine integrases are typically very specific for recombination activity at their cognate sites [5]. Some envisaged applications which require extremely high specificity have been shown to be viable; for example, targeted transgene integration at a unique site in the genomic DNA of human cells, where there are literally billions of potential off-target sites [38]. However, serine integrases do show low levels of off-target activity at 'pseudosite' sequences that are similar to the natural *att* sites. For example, considerable numbers of mammalian genomic pseudosites for the widely used ϕC31 integrase have been identified [39•,40]. Site specificity might also become an issue in designed systems involving multiple integrases (see 'Orthogonality' section below).

Precision

One might expect that serine integrases will have evolved to carry out very precise recombination in the phage's host cells, leaving no residual DNA damage or mutations, and this property is observed in *in vitro* experiments. However, evidence is accumulating that recombination can be less precise when integrases act in cells that are not their natural phage hosts; sequencing of recombinants has revealed various types of damage, including mutations at the centres of the *att* sites (rendering them refractory to further reactions) and deletions at the sites which may extend into flanking DNA [38]. The biochemical causes of this damage remain to be established; possibilities include effects of chromatin and DNA-binding proteins, or interruption of recombination by cellular events such as transcription and replication. Some integrases may be less prone to cause damage than others [41,42].

Speed, efficiency and directionality

Current data on the speed of integrase-mediated recombination *in vivo* are patchy, but recombination can be complete within quite short times (hours) in some systems [23••]. *In vitro*, reactions can be fast, with half-times for conversion of substrates to products of a few minutes, but some systems appear to recombine faster than others [4•,13•,43•,44].

Recombination efficiency and directionality are subject to numerous factors specific to the experimental system being studied, including the nature of the substrate (intramolecular or intermolecular reactions, distance between sites in the same molecule, *etc.*), stoichiometry of the recombining sites, and cellular environment [5,11•,14,45]. For accurate comparisons it is best to study recombination *in vitro*, where it is apparent that some integrases can give higher levels of conversion to products than others [46–49]. Mathematical modelling of the kinetics of ϕC31 integrase-mediated recombination implies that incomplete (~80%) conversion of an *attP* × *attB* plasmid substrate to recombinant products *in vitro* is an intrinsic feature of that system linked to the mechanism of directionality [43•,50]. This modelling predicts that directionality might be enhanceable by directed evolution and/or engineering of the integrase protein, or modifications of the recombination sites.

Orthogonality

If multiple serine integrases are to be used together, they should be orthogonal; that is, an integrase should not recognize or recombine sites cognate to the other integrases, nor should it interact directly with the other integrase proteins such that it interferes with their activity, nor (ideally) should it interact with the RDFs of the other integrases. Some integrase pairs that have been analysed in detail apparently fulfil these requirements (*e.g.*, ϕC31 and Bxb1 integrases; [51]), whereas others do not, and display 'crosstalk' [23••,36,48].

Conclusions and future directions

We already know of the existence of large numbers of distinct serine integrases. One obvious future challenge is to establish one or more sets of high-performance, orthogonal integrases with their sites and RDFs, as toolboxes that can reliably be utilized together by anyone with an envisaged application. This programme will involve assessment of each integrase for its conformity with the criteria highlighted above. So far, up to 11 integrases have been used successfully together in one system, albeit with some evidence of crosstalk [23••]. In the longer run, as well as characterizing new natural integrases it may also prove valuable to investigate ways of improving the integrases that we have, for example by directed evolution approaches [52] or by systematic protein engineering (*e.g.*, by attachment of modifier domains [53] or modification of active site residues). Another long-term goal which has already seen some preliminary effort is to engineer the specificity of existing serine integrases so that they act orthogonally on new target sequences [22,54]. Making such engineered systems fully orthogonal is a big challenge because it involves changes to protein–protein interactions as well as protein-DNA recognition (see Figure 2). All efforts to subvert the natural serine integrase systems for our own purposes are deeply dependent on advances in our understanding of their molecular structures and functions, where, despite the progress summarized above, there is still much to learn.

Acknowledgements

I am very grateful to Femi Olorunniji and Sean Colloms for critical reading of the manuscript. My research is supported by the Biotechnology and Biosciences Research Council, grant BB/K003356/1.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Craig NL: **A moveable feast: an introduction to mobile DNA**. In *Mobile DNA III*. Edited by Craig NL, Chandler M, Gellert M, Lambowitz A, Rice PA, Sandmeyer SB. Washington DC: ASM Press; 2015:3-39.

2. Boeke JD: **Putting mobile DNA to work: the toolbox**. In *Mobile DNA II*. Edited by Craig NL, Craigie R, Gellert M, Lambowitz AM. Washington DC: ASM Press; 2002:24-37.

3. Stark WM: **The serine recombinases**. In *Mobile DNA III*. Edited by Craig NL, Chandler M, Gellert M, Lambowitz A, Rice PA, Sandmeyer SB. Washington DC: ASM Press; 2015:73-89.

4. • Thorpe HM, Smith MCM: ***In vitro* site-specific integration of bacteriophage DNA catalysed by a recombinase of the resolvase/invertase family**. *Proc. Natl. Acad. Sci. U. S. A.* 1998, **95**:5505-5510.
This paper demonstrated ϕC31 integrase-catalysed recombination *in vitro* and characterized its substrate requirements.

5. Smith MCM: **Phage-encoded serine integrases and other large serine recombinases**. In *Mobile DNA III*. Edited by Craig NL, Chandler M, Gellert M, Lambowitz A, Rice PA, Sandmeyer SB. Washington DC: ASM Press; 2015:253-272.

6. 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**:2205-2216.

7. Landy A: **The λ integrase site-specific recombination pathway**. In *Mobile DNA III*. Edited by Craig NL, Chandler M, Gellert M, Lambowitz A, Rice PA, Sandmeyer SB. Washington DC: ASM Press; 2015:91-118.

8. Olorunniji FJ, Rosser SJ, Stark WM: **Site-specific recombinases: molecular machines for the genetic revolution**. *Biochem. J.* 2016, **473**:673-684.

9. Fogg PCM, Colloms SD, Rosser SJ, Stark WM, Smith MCM: **New applications for phage integrases**. *J. Mol. Biol.* 2014, **426**:2703-2716.

10. Katzen F: **Gateway recombinational cloning: a biological operating system**. *Exp. Opin. Drug Discov.* 2007, **2**:571-589.

11. • Olorunniji FJ, Buck DE, Colloms SD, McEwan AR, Smith MCM, Stark WM, Rosser SJ: **Gated rotation mechanism of site-specific recombination by ϕC31 integrase**. *Proc. Natl. Acad. Sci. U. S. A.* 2012, **109**:19661-19666.
This paper established that the mechanism of DNA strand exchange by serine integrases is similar to that of the 'small serine recombinases', involving subunit rotation.

12. Keenholtz RA, Grindley NDF, Hatfull GF, Marko JF: **Crossover-site sequence and DNA torsional stress control strand interchanges by the Bxb1 site-specific serine recombinase**. *Nucleic Acids Res.* 2016, **44**:8921-8932.

13. • 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. U. S. A.* 2008, **105**:3238-3243.
This paper (and references therein) helped to clarify the mechanisms used by serine integrases to select specific pairs of sites for recombination and ensure their correct alignment.

14. Smith MCM, Brown WRA, McEwan AR, Rowley PA: **Site-specific recombination by φC31 integrase and other large serine recombinases**. *Biochem. Soc. Trans.* 2010, **38**:388-394.

15. Khaleel T, Younger E, McEwan AR, Varghese AS, Smith MCM: **A phage protein that binds ϕC31 integrase to switch its directionality**. *Mol. Microbiol.* 2011, **80**:1450-1463.

16. Fan H-F, Hsieh T-s, Ma C-H, Jayaram M: **Single-molecule analysis of ϕC31 integrase-mediated site-specific recombination by tethered particle motion**. *Nucleic Acids Res.* 2016, **44**:10804-10823.

17. Yuan P, Gupta K, Van Duyne GD: **Tetrameric structure of a serine integrase catalytic domain**. *Structure* 2008, **16**:1275-1286.

18. •• Rutherford K, Yuan P, Perry K, Sharp R, Van Duyne GD: **Attachment site recognition and regulation of directionality by the serine integrases**. *Nucleic Acids Res.* 2013, **41**:8341-8356.
This landmark paper presents crystal structures of the C-terminal part of a serine integrase bound to DNA, and proposes a structure-based model for directionality of recombination.

19. Van Duyne GD, Rutherford K: **Large serine recombinase domain structure and attachment site binding**. *Crit. Rev. Biochem. Mol. Biol.* 2013, **48**:476-491.

20. Rutherford K, Van Duyne GD: **The ins and outs of serine integrase site-specific recombination**. *Curr. Opin. Struct. Biol.* 2014, **24**:125-131.

21. Mandali S, Gupta K, Dawson AR, Van Duyne GD, Johnson RC: **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.

22. Farruggio AP, Calos MP: **Serine integrase chimeras with activity in *E coli* and HeLa cells**. *Biol. Open* 2014, **3**:895-903.

23. •• Yang L, Nielsen AAK, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, Voigt CA: **Permanent genetic memory with >1-byte capacity**. *Nat. Methods* 2014, **11**:1261-1266.
The authors identified many novel serine integrases, and utilized 11 of them in construction of an intracellular DNA-based memory system.

24. Tomimatsu K, Kokura K, Nishida T, Yoshimura Y, Kazuki Y, Narita M, Oshimura M, Ohbayashi T: **Multiple expression cassette exchange via TP 901-1, R4, and Bxb1 integrase systems on a mouse artificial chromosome**. *FEBS Open Bio* 2017, **7**:306-317.

25. Weinberg BH, Hang Pham NT, Caraballo LD, Lozanoski T, Engel A, Bhatia S, Wong WW: **Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells**. *Nat. Biotechnol.* 2017, **35**:453-462.

26. Rubens JR, Selvaggio G, Lu TK: **Synthetic mixed-signal computation in living cells**. *Nat. Commun.* 2016, **7**:11658.

27. Hsiao V, Hori Y, Rothemund PWK, Murray RM: **A population-based temporal logic gate for timing and recording chemical events**. *Mol. Syst. Biol.* 2016, **12**:869.

28. Bonnet J, Subsoontorn P, Endy D: **Rewritable digital data storage in live cells via engineered control of recombination directionality**. *Proc. Natl. Acad. Sci. U. S. A.* 2012, **109**:8884-8889.

29. • Bonnet J, Yin P, Ortiz ME, Subsoontorn P, Endy D: **Amplifying genetic logic gates**. *Science* 2013, **340**:599-603.
This paper illustrates the use of Bxb1 and TP 901-1 serine integrases in 'genetic logic gates' within cells.

30. • Siuti P, Yazbek J, Lu TK: **Synthetic circuits integrating logic and memory in living cells**. *Nat. Biotechnol.* 2013, **31**:448-452.
These authors also constructed serine integrase-based genetic logic gates, including all 16 two-input Boolean logic operations.

31. Zhang L, Zhao G, Ding X: **Tandem assembly of the epothilone biosynthetic gene cluster by in vitro site-specific recombination**. *Sci. Rep.* 2011, **1**:141.

32. •• Colloms SD, Merrick CA, Olorunniji FJ, Stark WM, Smith MCM, Osbourn A, Keasling JD, Rosser SJ: **Rapid metabolic pathway assembly and modification using serine integrase site-specific recombination**. *Nucleic Acids Res.* 2014, **42**:e23.
The authors demonstrate that serine integrases can be efficient tools for creation of gene arrays for metabolic pathway assembly and other synthetic biology applications.

33. Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK: **Synthetic recombinase-based state machines in living cells**. *Science* 2016, **353** aad8559 (2016).

34. Fogg PCM, Haley JA, Stark WM, Smith MCM: **Genome integration and excision by a new *Streptomyces* bacteriophage, ϕJoe**. *Appl. Environ. Microbiol.* 2017, **83**:e02767-16.

35. Yoon B, Kim I, Nam J-A, Chang H-I, Ha CH: ***In vivo* and *in vitro* characterization of site-specific recombination of a novel serine integrase from the temperate phage EFC-1**. *Biochem. Biophys. Res. Commun.* 2016, **473**:336-341.

36. Zhang L, Zhu B, Dai R, Zhao G, Ding X: **Control of directionality in *Streptomyces* phage ϕBT1 integrase-mediated site-specific recombination**. *PLoS One* 2013, **8**:e80434.

37. Savinov A, Pan J, Ghosh P, Hatfull GF: **The Bxb1 gp47 recombination directionality factor is required not only for prophage excision, but also for phage DNA replication**. *Gene* 2012, **495**:42-48.

38. Calos MP: **Phage integrases for genome editing**. In *Genome Editing*. Edited by Cathomen T, Hirsch M, Porteus M. *Advances in Experimental Medicine and Biology* **895**: American Society of Gene and Cell Therapy; 2016: 81–95.

39. • Chalberg TW, Portlock JL, Olivares EC, Thyagarajan B, Kirby PJ, Hillman RT, Hoelters J, Calos MP: **Integration specificity of phage ϕC31 integrase in the human genome**. *J. Mol. Biol.* 2006, **357**:28-48.
The authors carried out a detailed study of the specificity of ϕC31 integrase-mediated integration in human cells, identifying and ranking numerous pseudosites.

40. Qu L, Ma Q, Zhou Z, Ma H, Huang Y, Huang S, Zeng F, Zeng Y: **A profile of native integration sites used by ϕC31 integrase in the bovine genome**. *J. Genet. Genomics* 2012, **39**:217-224.

41. Xu Z, Thomas L, Davies B, Chalmers R, Smith M, Brown W: **Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome**. *BMC Biotechnol.* 2013, **13**:87.

42. Xu Z, Brown WRA: **Comparison and optimization of ten phage encoded serine integrases for genome engineering in *Saccharomyces cerevisiae***. *BMC Biotechnol.* 2016, **16**:13.

43. • Pokhilko A, Zhao J, Ebenhöh O, Smith MCM, Stark WM, Colloms SD: **The mechanism of ϕC31 integrase directionality: experimental analysis and computational modelling**. *Nucleic Acids Res.* 2016, **44**:7342-7360.
These authors present detailed kinetics data on ϕC31 integrase activity in vitro and use the data to construct a mathematical model explaining the directionality phenomenon.

44. Morita K, Yamamoto T, Fusada N, Komatsu M, Ikeda H, Hirano N, Takahashi H: ***In vitro* characterization of the site-specific recombination system based on actinophage TG1 integrase**. *Mol. Genet. Genomics* 2009, **282**:607-616.

45. Maucksch C, Aneja MK, Hennen E, Bohla A, Hoffmann F, Elfinger M, Rosenecker J, Rudolph C: **Cell type differences in activity of the *Streptomyces* bacteriophage ϕC31 integrase**. *Nucleic Acids Res.* 2008, **36**:5462-5471.

46. Bibb LA, Hancox MI, Hatfull GF: **Integration and excision by the large serine recombinase ϕRv1 integrase**. *Mol. Microbiol.* 2005, **55**:1896-1910.

47. Mandali S, Dhar G, Avliyakulov NK, Haykinson MJ, Johnson RC: **The site-specific integration reaction of *Listeria* phage A118 integrase, a serine recombinase**. *Mobile DNA* 2013, **4**:2.

48. Singh S, Rockenbach K, Dedrick RM, VanDemark AP, Hatfull GF: **Cross-talk between diverse serine integrases**. *J. Mol. Biol.* 2013, **426**:318-331.

49. Wang X, Tang B, Ye Y, Mao Y, Lei X, Zhao G, Ding X: **Bxb1 integrase serves as a highly efficient DNA recombinase in rapid metabolite pathway assembly**. *Acta Biochim. Biophys. Sin.* 2017, **49**:44-50.

50. Pokhilko A, Zhao J, Stark WM, Colloms SD, Ebenhöh O: **A simplified mathematical model of directional DNA site-specific recombination by serine integrases**. *J. R. Soc. Interface* 2017, **14**:20160618.

51. Zhu F, Gamboa M, Farruggio AP, Hippenmeyer S, Tasic B, Schüle B, Chen-Tsai Y, Calos MP: **DICE, an efficient system for iterative genomic editing in human pluripotent stem cells**. *Nucleic Acids Res.* 2014, **42**:e34.

52. Keravala A, Lee S, Thyagarajan B, Olivares EC, Gabrovsky VE, Woodard LE, Calos MP: **Mutational derivatives of PhiC31 integrase with increased efficiency and specificity**. *Mol. Ther.* 2009, **17**:112-120.

53. Sharma N, Moldt B, Dalsgaard T, Jensen TG, Mikkelsen JG: **Regulated gene insertion by steroid-induced ϕC31 integrase**. *Nucleic Acids Res.* 2008, **36**:e67.

54. Sclimenti CR, Thyagarajan B, Calos MP: **Directed evolution of a recombinase for improved genomic integration at a native human sequence**. *Nucleic Acids Res.* 2001, **29**:5044-5051.