

Microbial site-specific recombinases as gene editing tools

Abstract

Commonly used microbial site-specific recombinases used as editing and gene therapy tools in eukaryotic genomes suffer the need to pre-insert their recombination sites, or else need to undergo protein evolution to recognize endogenous pseudo sites. These difficulties can be bypassed using the lambdaoid coliphage HK022 site-specific integrase recombinase that owing to its smaller integration site and sequence promiscuity can utilize active endogenous target sites.

Keywords: site-specific recombination, gene-editing, RMCE, gene therapy

Volume 1 Issue 1 - 2015

Ezra Yagil

Department of Biochemistry and Molecular Biology, Tel-Aviv University, Israel

Correspondence: Ezra Yagil, Dept. of Biochemistry and Molecular Biology, Tel-Aviv University, Tel-Aviv 6997801, Israel, Tel +972546804181, Fax +97236406834, Email yagile@tauex.tau.ac.il

Received: June 22, 2015 | Published: August 13, 2015

Abbreviations: bp, base pairs; CRISPR, clustered regularly interspaced short palindromic repeats; DSB, double-stranded dna break; GOI, gene of interest; gRNA, guide RNA; IHF, integration host factor; O, overlap; RMCE, recombinase-mediated cassette exchange; RS, recombination site; TALEN, transcription activator-like effectors; Xis, excisionase; ZFN, zinc finger nuclease

Introduction

Site-specific recombinases catalyze recombination reactions between DNA pairs each carrying an identical but specific short sequence known as site-specific recombination site (RS). They differ in this regard from homologous recombinases that recombine practically any pair of homologous DNA sequences usually much longer than the RSs. Early attempts to use native homologous recombinases to accomplish eukaryotic gene manipulations proved inefficient. However, microbial site-specific recombination systems proved more useful for these purposes, initially when Brian Sauer introduced the Cre-lox system of *Bacteriophage P1* into mammalian cells¹ which was followed by the implementation of yeast's Flp-*frt* system. The RSs of these respective recombinases (lox and *frt*, respectively) are each 34 base pairs (bp) long and each comprises two complete or partial inverted repeats of 13bp (Figure 1a, b, capital letters) that flank a central polar sequence of 8bp (Figure 1a, b, small letters), known as the overlap (O). The sequences that include the inverted repeats (blue arrows in Figure 1) serve as binding sites for the relevant recombinase and O is the site of the DNA exchange reaction. Such RSs are usually absent in the genomes of higher eukaryotes. Therefore, Cre-lox or Flp-*frt* mediated genetic manipulations necessitate the introduction of the suitable RSs at specific genomic sites. Subsequently, a plasmid-borne gene of interest (GOI) abutted by a second compatible RS and a selection marker (usually an antibiotic resistant gene, Figure 2A) is used to selectively insert the GOI into the genome-inserted RS (Figure 2B) in a reaction catalyzed by the relevant recombinase that is provided in trans. As a result the GOI is site-specifically inserted along with the superfluous selection marker and extra plasmid DNA (Figure 2C). A major disadvantage of this procedure is that the inserted plasmid, flanked with two tandemly-oriented RSs (Figure 2C, broken arrows), serves as a substrate for the thermodynamically favorable reverse excision reaction (long vertical arrow in Figure 2) probably

because of their cis configuration (both relatively adjacent on the same DNA molecule). Hence, it was easier to create gene knockouts when two RSs were inserted in tandem on each site of the gene to be excised.^{2,3} Another popular microbial site specific gene editing tool is the Integrase system (Int-att) of the *Streptomyces* phage ϕ C31. In contrast to Cre-lox and Flp-*frt* the two attRSs of ϕ C31-Int (attB and attP) are not completely identical, 34 and 39 respectively (Figure 1c).⁴ attP x attB recombination catalyzed by its integrase recombinase (Int- ϕ C31) results in the recombinant attL x attR sites that flank the insertion; however, in the mammalian milieu, they are not reversible thereby allowing stable integration. Moreover, mammalian genomes carry some partially mismatched but active native attP sites (pseudo attP secondary sites) that were used as successful gene editing and gene therapy targets.⁵

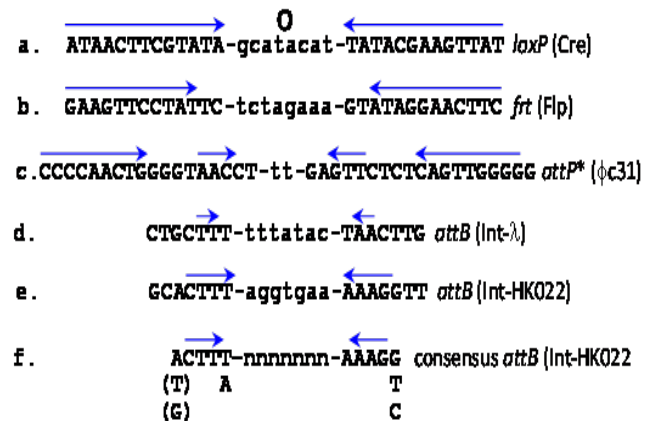


Figure 1 Recombination sites (RSs) of various microbial site-specific recombinases. The binding sites of the recombinase are shown in capital letters and the overlap sequences in small letters. O, overlap sequence. Arrows indicate inverted repeats. The bases below the consensus sequence (f) can replace the one in the first line; those in parentheses can do so only if the right side binding site maintains its AAAGG sequence. *in the case of Int- ϕ C31 (in contrast to Int- λ and Int-HK022) its attP, and not attB, is the genomic site in the eukarya.⁵

The insertion problem of entire plasmids was partially solved with the implementation of a dual site-specific recombination reaction

known as the “recombinase-mediated cassette exchange” (RMCE) reaction (Figure 3), first implemented in Jürgen Bode’s laboratory using the FLP-*frt* system.⁶ Here, a chromosomal sequence flanked by two incompatible RSs (RS1 and RS2, Figure 3B) is exchanged with a plasmid-borne sequence (Figure 3A) likewise flanked by RS1 and RS2 (Figure 3C). It is essential the RS1 and RS2 be incompatible to avoid recombination between them. If only one recombinase catalyzes the RMCE reaction it must recognize both compatible recombining RS pairs (at least one pair is mutated). Alternatively each compatible RS pair can match a different recombinase, in which case both recombinases must be supplied.^{2,7} A Cre-catalyzed RMCE reaction that swapped fragments over 100kb long was accomplished in mice.⁸ The need to pre-insert Cre and FLP RSs (*loxP* and *frt*, respectively) was partially overcome by the search and use of native secondary RSs with some mismatches that are inactive with their wild type recombinases. Rendering a recombinase specific for such secondary sites entails a lengthy directed protein evolution process. Several such modified recombinases able to mediate gene-editing reactions with endogenous secondary RSs were created in this manner.⁹⁻¹⁴ Several other microbial site-specific recombinases that are less popular in eukaryotic gene manipulations are listed in.¹⁵

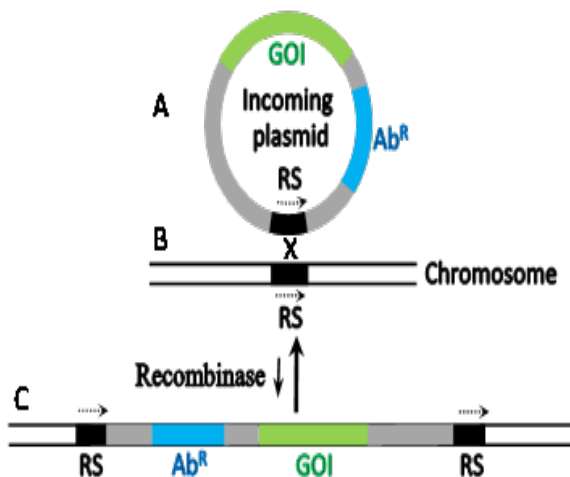


Figure 2 Chromosomal insertion of a gene of interest (GOI) by a site-specific recombination reaction and its more favorable reverse excision. AbR-antibiotic resistance gene, dotted arrows- polar direction of the overlap (O) sequence.

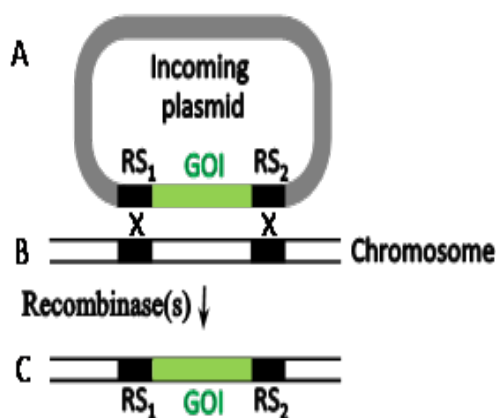


Figure 3 Recombinase-mediated cassette exchange reaction that swaps a DNA fragment between an incoming plasmid and the chromosome by a dual site-specific recombination reaction.

The well-documented site-specific recombination system Int-*att* of the temperate coliphage λ (Figure 4) has been much less employed as a gene editing tool. Its bacterial *attB* RS is considerably shorter (21bp, Figure 1d, Figure 4B). It comprises two partially inverted 7bp Int-binding sites flanking a 7bp O region. However, the reason for its lingering behind is that *attB*’s 21bp phage partner RS (*attP*) is flanked by two longer arms (188 and 74bp, brown rectangles in Figure 4A) that carry additional indispensable binding sites for as well as binding sites for the indispensable host and phage-encoded accessory proteins IHF and Xis.^{16,17} As with ϕ C31 RSs, an *attB* x *attP* integration reaction leads to the formation of flanking recombinant *attL* and *attR* sites of the inserted DNA (and the prophage in nature, Figure 4C). A similar site-specific recombination mechanism characterizes the λ -related coliphage HK022.^{18,19} While both lambdoid Int systems were recruited to function as gene editing tools in mammalian cells, in contrast to the Int-*att* system of ϕ C31, their reverse *attL* x *attR* excision reactions are also functional in mammalian cells. However, only Int-HK022 is able to catalyze site-specific recombination reactions in the absence of the accessory proteins, albeit, at reduced efficiency.^{20,21} For λ to do so, an IHF-independent mutation of Int is used.²² The freedom from the accessory proteins rendered unnecessary the cognate *attP*~HK022 arms.²³ Moreover, it turned out that replacing the O sequence of *attB*-HK022 with a random 7bp sequence still supports efficient Int-mediated site-specific recombination as long as its cognate phage recombination site *attP* features an identical O sequence.²⁴ This property, combined with the relative shortness of *attB* allowed identifying native and active secondary *attB* sites that flank several human deleterious mutations, each with a different O sequence, and some even with minor modifications in their Int-binding sites. Owing to the difference in their O sequences renders the flanking secondary *attBs* incompatible with each other. Hence, such active endogenous sites that flank human hereditary mutations (Figure 5B) are potential RSs in RMCE reactions catalyzed by unmodified Int-HK022 thereby super seeding the need to insert any *attB* sites or alter the specificity of the recombinase by directed protein evolution. Some of these flanking *attBs* are currently exploited for performing two sequential RMCE reactions catalyzed by wild type Int-HK022 with the aim cure such mutations (Figure 5). The reason of the two RMCE reactions is for selection purposes and for the recovery of the native *attB* s that leads to “cleanly” cured cells that, except of the cured mutation, are devoid of any other genome DNA modification (Figure 5E).

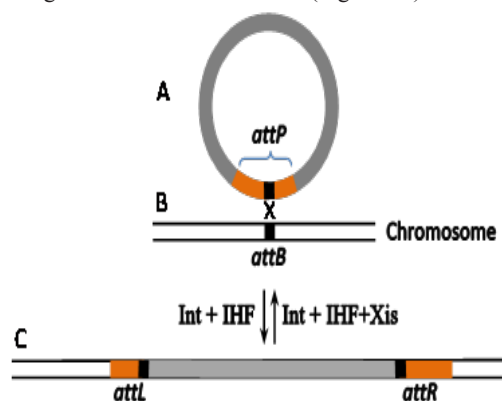


Figure 4 Site-specific recombination catalyzed by Int-HK022 and Int- λ . A- incoming plasmid or in nature the circularized infected phage DNA. B -chromosomal with an *attB* RS and in nature the *Escherichia coli* chromosome with its wild type native *attB* site. C- is the modified chromosome and in nature the lysogenic *E. coli* with the inserted prophage

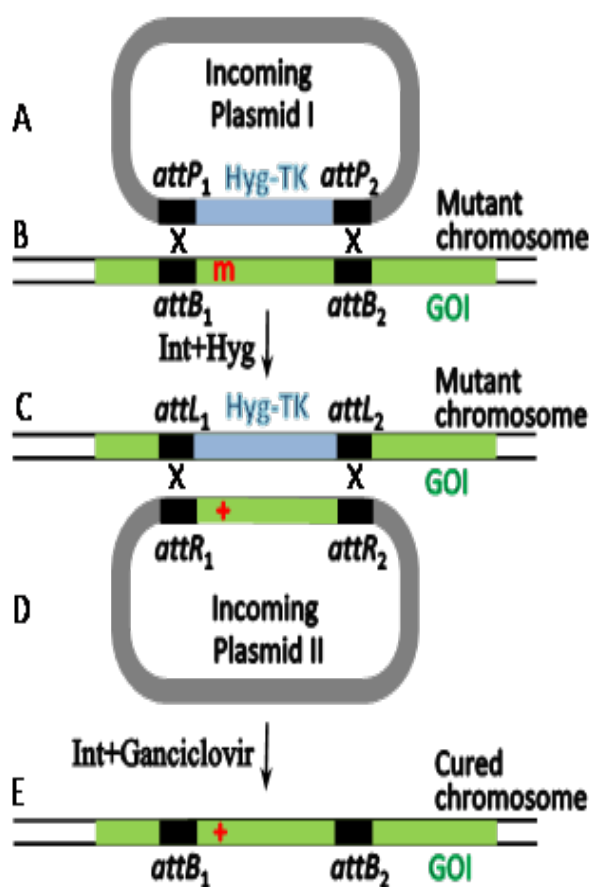


Figure 5 Gene curing by two sequential Int-HK022 catalyzed RMCE reactions. First reaction: An Incoming plasmid (A) of the first reaction whose Hyg-TK fused gene³¹ flanked by two relevant attP sites is co-transformed with an Int-expressing plasmid into cells (B) that carry a mutation (m) flanked by two endogenous active secondary attB sites within a gene of interest (GOI). Recombinant cells (C) are selected with hygromycin. Second reaction: The recombinant Hyg-TK cells (C) are co-transfected with the Int-expressing plasmid and with a plasmid (D) that carries the exchanged GOI fragment but with the corrected wild type allele (+), flanked by two relevant attR sites. Transformed cells are treated with Ganciclovir that selects against cells expressing thymidine-kinase (TK).³¹ Surviving cells are expected to be those whose Int activity has catalyzed the exchange of the Hyg-TK gene with a normal GOI fragment by a dual attL × attR reaction yielding the cured GOI flanked with its original native attB sites (E).

An analysis of several potential human native secondary attB-HK022 sites that were either active or inactive, along with minor mismatches in the Int binding sites that flank a random O sequence lead to a consensus sequence of a 15bp (with some flexibility) active attB-HK022 site that maintained its CTT...AAG inverted repeat (Figure 1f).²⁴ In λ , minor changes in its O sequence likewise did not affect recombination provided identical in both recombining att partners.^{25,26} Whether in λ 's att sites as well as in loxP and frt an extended randomness in the O sequence of their recombining RS pairs allows recombination was not yet reported.

Recent gene editing approaches based on the recruitment of the endogenous homologous recombination repair system in response to double-stranded DNA break (DSB) seem to supplant those based on the microbial site-specific recombination systems. Key to the success of the homologous recombination repair systems is the ability to

introduce DSBs at any desired chromosomal loci. This feat has been originally accomplished using sequence-specific DNA-bound Zinc-finger double-stranded DNA nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) fused to a blunt-end restriction enzyme motif (usually FokI). More recently, these engineered DNA-bound nuclease proteins are replaced by site-specific DSB induction mediated by engineered derivatives of the bacterial adaptive immunity CRISPR/Cas system.^{27,28} However, to accomplish productive RMCE reactions these systems still associate with site-specific recombinases. E.g. the CRISPR/Cas system was already used to insert and perform RMCE reactions in *Drosophila* based on the activity of Int- ϕ C31 system²⁹ and a seed-mouse line was constructed that uses the Cre-lox based RMCE reactions in the mouse Rosa26 locus.^{30,31} Namely, these RMCE constructs still use pre-inserted RSs of the prokaryotic site-specific recombinases and they are not necessarily located in their native genome location. The proposed int-HK022 RMCE reactions, where possible, have the spatial advantage of using native RS sequences to cleanly correct deleterious mutations within their native loci (Figure 5).

Acknowledgements

Gabi Kaufmann's valuable comments and corrections helped improved the manuscript. Work done in our laboratory was supported by GIF, the German Israeli Foundation for Scientific Research and Development (Grant 1062/2008) the Israel Science Foundation (Grant 702/11) and the US-Israel Binational Science Foundation Jerusalem, Israel (Grant 2003394).

Conflict of interest

The author declares no conflict of interest.

References

1. Sauer B, Henderson N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A*. 1988;85(14):5166–5170.
2. Wirth D, Gama-Norton L, Riemer P, et al. Road to precision: recombinase-based targeting technologies for genome engineering. *Curr Opin Biotechnol*. 2007;18(5):411–419.
3. Branda CS, Dymecki SM. Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev Cell*. 2004;6(1):7–28.
4. Calos MP. The phiC31 integrase system for gene therapy. *Curr Gene Ther*. 2006;6(6):633–645.
5. Chavez CL, Calos MP. Therapeutic applications of the phiC31 integrase system. *Current Gene Therapy*. 2011;11(5):375–381.
6. Schlake T, Bode J. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry*. 1994;33(43):12746–12751.
7. Turan S, Zehe C, Kuehle J, et al. Recombinase-mediated cassette exchange (RMCE) – a rapidly-expanding toolbox for targeted genomic modifications. *Gene*. 2013;515(1):1–27.
8. Wallace HAC, Marques-Kranc F, Richardson M, et al. Manipulating the mouse genome to engineer precise functional synthetic replacements with human sequence. *Cell*. 2007;128(1):197–209.
9. Abi-Ghanem J, Chusainov J, Karimova M, et al. Engineering of a target site-specific recombinase by a combined evolution- and structure-guided approach. *Nucleic Acids Res*. 2013;41(4):2394–2403.

10. Buchholz F, Stewart AF. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat Biotechnol.* 2001;19(11):1047–1052.
11. Sarkar I, Hauber I, Hauber J, et al. HIV-1 proviral DNA excision using an evolved recombinase. *Science.* 2007;316(5833):1912–1915.
12. Bolusani S, Ma CH, Paek A, et al. Evolution of variants of yeast site-specific recombinase Flp that utilize native genomic sequences as recombination target sites. *Nucleic Acids Res.* 2006;34(18):5259–5269.
13. Gaj T, Sirk SJ, Barbas CF. Expanding the scope of site-specific recombinases for genetic and metabolic engineering. *Biotechnol Bioeng.* 2014;111(1):1–15.
14. Shah R, Li F, Voziyanova E, et al. Target-specific variants of Flp recombinase mediate genome engineering reactions in mammalian cells. *FEBS J.* 2015;282(17):3323–3333.
15. Groth, Calos MP. Phage integrases: Biology and applications. *J Mol Biol.* 2004;335(3):667–678.
16. Azaro MA, Landy A. λ Integrase and the λ Int Family. In: Craig NL, Craigie R, Gellert M, Lambowitz A, editors. *Mobile DNA II*. USA: ASM Press; 2002. p. 118–148.
17. Biswas T, Aihara H, Radman-Livaja M, et al. A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature.* 2005;435(7045):1059–1066.
18. Weisberg RA, Gottesmann ME, Hendrix R, et al. Family values in the age of genomics: comparative analyses of temperate bacteriophage HK022. *Annu Rev Genet.* 1999;33:565–602.
19. Yagil E, Dolev S, Oberto J, et al. Determinants of site-specific recombination in the lambdaoid coliphage HK022. An evolutionary change in specificity. *J Mol Biol.* 1989;207(4):695–717.
20. Harel-Levy, Goltsman J, Tuby CN, et al. Human genomic site-specific recombination catalyzed by coliphage HK022 integrase. *J Biotechnol.* 2008;134(1–2):46–54.
21. Malchin N, Goltsman J, Dabool L, et al. Optimization of coliphage HK022 Integrase activity in human cells. *Gene.* 2009;437(1–2):9–13.
22. Corona T, Bao QY, Christ N, et al. Activation of site-specific DNA integration in human cells by a single chain integration host factor. *Nucleic Acids Res.* 2003;31(17):5140–5148.
23. Malchin N, Tuby CN, Yagil E, et al. Arm site independence of coliphage HK022 integrase in human cells. *Mol Genet Genomics.* 2011;285(5):403–413.
24. Kolot M, Malchin N, Elias A, et al. Site promiscuity of coliphage HK022 integrase as a tool for gene therapy. *Gene Ther.* 2015;22(7):521–527.
25. Bauer CE, Gardner J, Gumport RI. Extent of sequence homology required for bacteriophage lambda site-specific recombination. *J Mol Biol.* 1985;181(2):187–197.
26. Weisberg RA, Enquist LW, Foeller C, et al. Role for DNA homology in site-specific recombination. The isolation and characterization of a site affinity mutant of coliphage lambda. *J Mol Biol.* 1983;170(2):319–342.
27. Gaj T, Gersbach CA, Barbas CF III. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 2013;31(7):397–405.
28. Kim H, Kim JS. A guide to genome engineering with programmable nucleases. *Nat Rev Genet.* 2014;15(5):321–334.
29. Zhang X, Koolhaas WH, Schnorrer F. A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in *Drosophila*. *G3 (Bethesda).* 2014;4(12):2409–2418.
30. Quadros RM, Harms DW, Ohtsuka M, et al. Insertion of sequences at the original provirus integration site of mouse ROSA26 locus using the CRISPR/Cas9 system. *FEBS Open Bio.* 2015;5:191–197.
31. Lupton SD, Brunton LL, Kalberg VA, et al. Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. *Mol Cell Biol.* 1991;11(6):3374–3378.