Supporting Information

Tong et al. 10.1073/pnas.1413007111

SI Text

Assembly of *att* Site DNA Substrates and Recombination Conditions. All proteins were produced in *Escherichia coli* BL21(DE3) from expression plasmids under the control of the T7 promoter and using methods described previously (1): Int (2), Cre (3), CreCM2K201A (4), the Crn hybrid recombinases (2, 5), IHF (6), Xis, as Met-Gly-$(His)_{10}$-$(Ser)_2$-Gly-His-Ile-Glu-Gly-Arg-His-Xis (7), and Fis (8).

Oligonucleotides labeled with ^{32}P at the 5′ terminus were purified from the kinase reaction using the NORGEN-Biotek oligo clean-up kit (cat no. 34100) and were recovered in yields greater than 70%. Individual single DNA strands for *att*L were constructed from two oligos (~40–60 bases) and a single template strand overlapping the junction to be ligated; the strands comprising *att*R were constructed from three oligo and two templates. The unphosphorylated template strands were ~40 bases long and overlapped the junctions to be ligated by ~20 bp on each side of the junction such that that the calculated melting temperature of each overlap was ~50 °C in 10 mM Tris (pH 8) and 50 mM NaCl.

The ratio of ligation substrates to templates was approximately equimolar, and the DNA concentrations were 10 μM (or higher) in 10 mM Tris (pH 8) and 50 mM NaCl. Annealing was carried out in a water bath that was slowly cooled from 95 °C to 100 °C to room temperature on the bench top and then further cooled to −20 °C. Following incubation with T4 DNA ligase (NEB) at 18 °C for 4 h, reactions were placed at 0 °C overnight. The mixtures containing final full-length, single-stranded *att*-DNAs, 95–105 bases for *att*L and 135–145 bases for *att*R, were boiled for 2 min after the addition of 2× loading solution [98% (vol/vol) formamide] and then resolved on a denaturing polyacrylamide gel [7 M urea, 8–12% (wt/vol) polyacrylamide] at 25 V/cm for 90–105 min. The band of interest was excised from the gel and lightly pulverized using a 1-mL syringe, and the DNA was eluted with vigorous shaking in Tris EDTA at 37 °C for 4 h and then overnight at 25 °C. The eluent was separated from polyacrylamide by filtration through nylon floss, concentrated on Amicon filters (regenerated cellulose membrane, 3,000 molecular weight cut-off; cat no. UFC800396), and then purified using the NORGEN oligo clean-up kit. This procedure generates single-stranded *att*-DNAs in yields of 30–40%. To generate the final DNA duplexes (95–107 bp for *att*L and 135–145 bp for *att*R), complementary strands were annealed in 10 mM Tris (pH 8) and 50 mM NaCl at a DNA concentration of 5 and 0.4 μM for the unlabeled and radio-labeled DNAs, respectively. Supercoiled *att*P and long linear *att*B were made by annealing and ligating synthetic oligonucleotides into gapped circular DNAs constructed from circular single-strand phage DNAs, as described previously (1). The recombination buffer contained 25 mM Tris (pH 8.0), 75 mM NaCl, 5 mM EDTA, 2.5 mM DTT, 6 mM spermidine, 1.4 mg/mL BSA, and 0.25 mg/mL herring sperm DNA. In the cross-linking experiments, the amount of DTT was raised to 4.2 mM. In the integrative recombination reactions, herring sperm DNA was omitted because the large amount of vector DNA served as a carrier. Following the recombination reactions, 5× loading solution [15% (wt/vol) Ficoll in Tris-borate-EDTA buffer] was added to the mixture, which was then fractionated by electrophoresis as described below for each protocol.

For all Crn-promoted recombination reactions, a plasmid encoding the substrate to be labeled (*lot*B, *lot*L, or *lot*R) was linearized by digestion with HindIII, and the ends were filled in with the Klenow fragment of DNA polymerase (NEB) and [α-^{32}P]-deoxycytidine triphosphate. Recovery of the purified radio-labeled DNA was determined using agarose gel electrophoresis to determine the final concentration of substrate.

Crn-promoted recombination reactions were carried out at 25 °C in 5mM EDTA pH 8, 6mM spermidine, 25 mM Tris, pH 8, 0.5mg/mL BSA, 2.5mM DTT, 30nM NaCl and optimized concentrations of IHF and Xis (for excision). Reactions were terminated by the addition of SDS-containing loading solution, heated for 10 min at 65 °C, analyzed by electrophoresis through 1.2% (wt/vol) agarose gels in Tris-acetate-EDTA buffer, and visualized on a phosphorimager.

Cross-Linking Experiments The cystamine-containing oligonucleotides were made by dissolving the dried pellet of a 35- to 65-base oligonucleotide containing a single O^6-phenyl-2′-deoxyinosine (*O*6-phenyl-dI) modification (Operon, HPLC purified, 1-μmol scale) in 300 μL of 5 M cystamine, converted from cystamine dihydrochloride (97%; ACROS Organics) as previously described (9–11). The reaction was incubated overnight at 65 °C, quenched with 180 μL of 8.74 M acetic acid, washed three times with 4 mL of ddH_2O using a 3K molecular weight cut-off filter (Amicon), and purified by ethanol precipitation at 4 °C overnight, as previously described (9–11). The product oligonucleotide containing a single N^6-(6-amino-3,4-dithiahexyl)-deoxyadenosine (cystamine-dA) was dissolved in ddH_2O to make a 100-μM solution.

To make the cysteine substitutions, we started with our standard overproducing plasmid encoding an Int protein in which the two surface-accessible cystamine residues had been changed to serine:C25S (12), named pMER10, to which was added C197S, named pMER15 or pWJTb (this work). This Int protein does not make any appreciable disulfide cross-links with the cystamine-substituted oligonucleotides or *att* sites. Single Cys substitutions were introduced into this strain by site-directed mutagenesis using the primer extension method (Quick Change II; Agilent Tech). To test the cross-linking efficiency of the cystamine label at each of the candidate positions, we synthesized 36- or 58-bp oligonucleotides containing an *O*6-phenyl-dI modification, which was converted to cystamine as described above. Following incubation of the indicated cystamine-substituted oligonucleotide with the indicated Cys-substituted Int (over a range of Int concentrations), the amount of covalent disulfide formed was assayed by SDS gel electrophoresis (*Materials and Methods*). The results of these assays, along with the recombination proficiency of each Cys-substituted Int, are summarized in Tables S1 and S2.

In every recombination reaction a ^{32}P label was incorporated at the 5′ terminus of the P or P′ arm carrying the cystamine label. Before, or during, each recombination reaction, the DNA was treated with 4.2 mM DTT to reduce the cystamine to a free thiol. Excisive recombination reactions containing 46 nM unlabeled *att*L or *att*R and 19 nM ^{32}P-labeled *att*L or *att*R were carried out at 25 °C for 2 h in recombination buffer in the presence of 1.4 μM Int, 1 μM IHF, 2.4 μM Xis, and 0.75 μM Fis (while the presence of Fis enhanced the efficiency of recombination in those cases where Xis was limiting, it had no effect on the patterns of Int bridging). For the integrative reactions, Xis, Fis, and herring sperm DNA were omitted, and *att*P and *att*B were present at 30 and 750 nM, respectively.

Following the 2-h recombination reaction, disulfide bond formation was promoted by adding 0.3 μmol hydrogen peroxide per 10-μL reaction, and the mixture was incubated for another 40 min. The excess hydrogen peroxide was quenched by addition of 0.4 μL catalase (MP Biomedicals; ≥1,000 U/mL) per 10-μL reaction and incubation at 37 °C for 15 min. Loading solution (5×) containing

15% (wt/vol) Ficoll and 1% SDS in Tris-Borate-EDTA buffer was added, and the reactions were fractionated by electrophoresis in 6% or 8% (wt/vol) SDS polyacrylamide gels at 12–19 V/cm for 4–8 h.

Characterization of the Holliday Junction Complex. The Holliday junction (HJ) complex was visualized and characterized in polyacrylamide gels under native and denaturing conditions (Fig. S1). In native gels, the HJ complex migrates slower than any of the *att* site substrate or product complexes. In the presence of SDS (or ethidium bromide or on extraction with phenol/chloroform), the naked HJ DNA gives rise to a band that migrates slower than any of the *att* site DNAs. When the trapped HJ complexes are analyzed in urea gels, the top strands (TS) are seen to have undergone a single-strand exchange between the substrate partners, but the bottom strands (BS) remain parental, as expected from the known order of strand exchanges (Fig. S1*A*). In contrast, recombination reactions between partners not designed to trap HJs go through a complete reaction involving exchange of both top and bottom strands (Fig S3*B*). The formation of the HJ complex is dependent on all of the required recombination proteins, and in excisive reactions, their formation is stimulated by Fis when Xis is limiting. The HJ complex contains four monomers of λ-Integrase, as demonstrated by a quantitative assay using Cy3-labeled λ-Int and a ^{32}P-radiolabeled half-*att* suicide substrate as a standard for one Int per DNA molecule (Fig. S1*C*) (13). The isolated protein–DNA HJ complex is stable in polyacrylamide gels and in solution for more than 12 h at room temperature or 4 °C (Fig. S1*D*).

1. Matovina M, Seah N, Hamilton T, Warren D, Landy A (2010) Stoichiometric incorporation of base substitutions at specific sites in supercoiled DNA and supercoiled recombination intermediates. *Nucleic Acids Res* 38(18):e175.
2. Tirumalai RS, Kwon HJ, Cardente EH, Ellenberger T, Landy A (1998) Recognition of core-type DNA sites by λ integrase. *J Mol Biol* 279(3):513–527.
3. Hoess R, Abremski K, Irwin S, Kendall M, Mack A (1990) DNA specificity of the Cre recombinase resides in the 25 kDa carboxyl domain of the protein. *J Mol Biol* 216(4): 873–882.
4. Ghosh K, Guo F, Van Duyne GD (2007) Synapsis of loxP sites by Cre recombinase. *J Biol Chem* 282(33):24004–24016.
5. Warren D, Laxmikanthan G, Landy A (2008) A chimeric Cre recombinase with regulated directionality. *Proc Natl Acad Sci USA* 105(47):18278–18283.
6. Nash HA, Robertson CA, Flamm E, Weisberg RA, Miller HI (1987) Overproduction of Escherichia coli integration host factor, a protein with nonidentical subunits. *J Bacteriol* 169(9):4124–4127.
7. Warren D, Lee SY, Landy A (2005) Mutations in the amino-terminal domain of λ-integrase have differential effects on integrative and excisive recombination. *Mol Microbiol* 55(4):1104–1112.
8. Pan CQ, et al. (1994) Structure of the Escherichia coli Fis-DNA complex probed by protein conjugated with 1,10-phenanthroline copper(I) complex. *Proc Natl Acad Sci USA* 91(5):1721–1725.
9. Ferentz A, Verdine G (1991) Disulfide cross-linked oligonucleotides. *J Am Chem Soc* 113:4000–4002.
10. Ferentz A, Keating T, Verdine G (1993) Synthesis and characterization of disulfide cross-linked oligonucleotides. *J Am Chem Soc* 115:9006–9014.
11. Stanojevic D, Verdine GL (1995) Deconstruction of GCN4/GCRE into a monomeric peptide-DNA complex. *Nat Struct Biol* 2(6):450–457.
12. Tirumalai RS, Pargellis CA, Landy A (1996) Identification and characterization of the N-ethylmaleimide-sensitive site in lambda-integrase. *J Biol Chem* 271(47): 29599–29604.
13. Nunes-Düby SE, Matsumoto L, Landy A (1989) Half-att site substrates reveal the homology independence and minimal protein requirements for productive synapsis in λ excisive recombination. *Cell* 59(1):197–206.

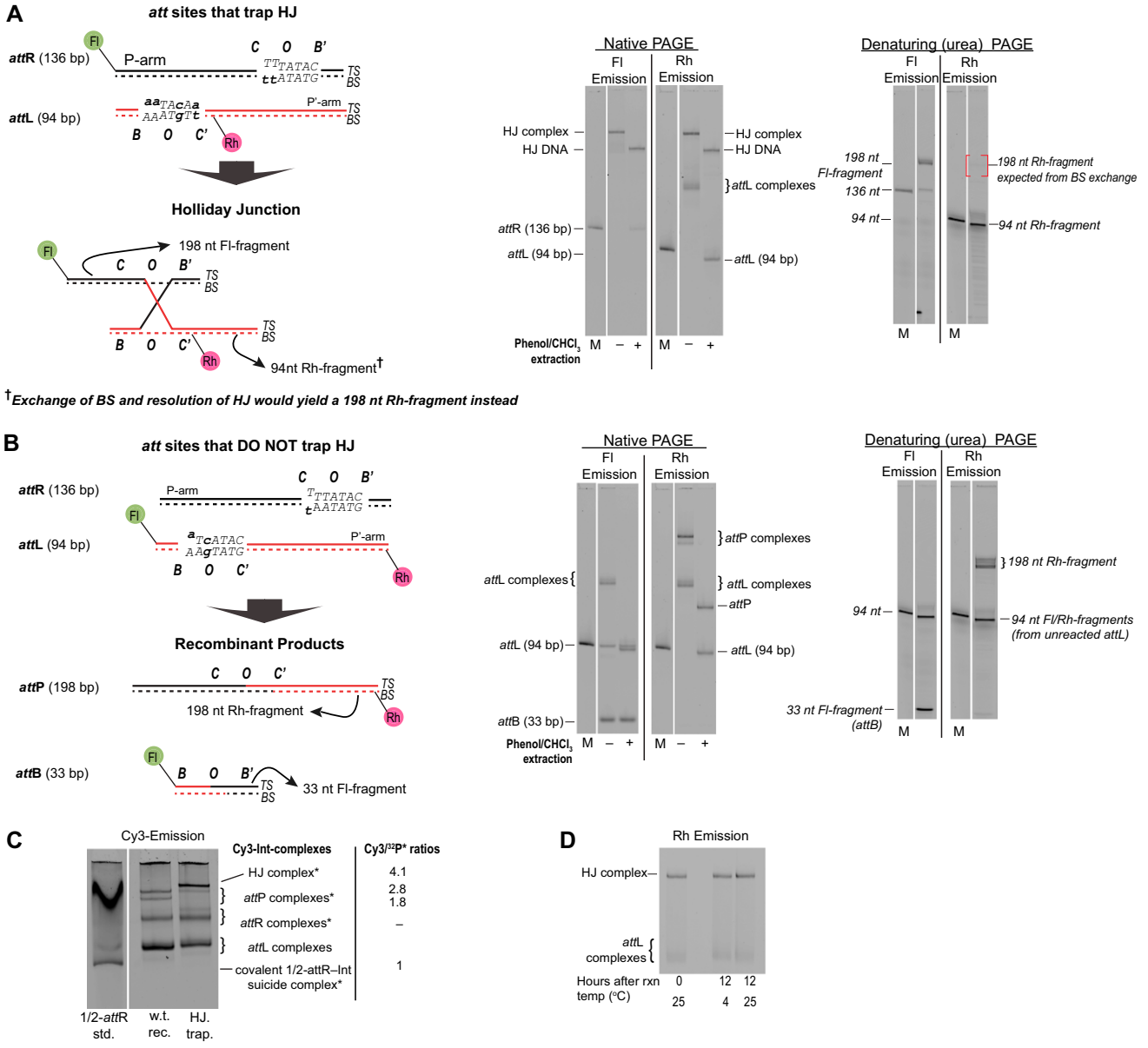
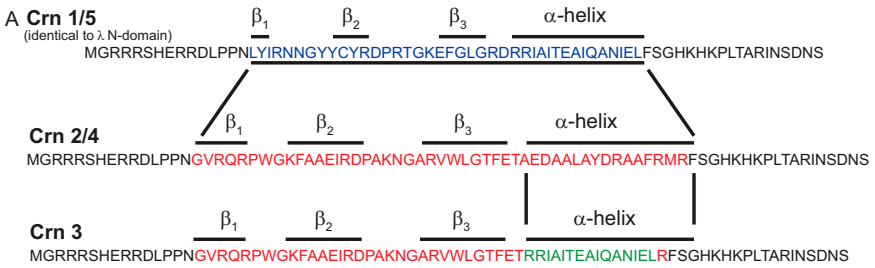


Fig. S1. Characterization of the HJ complexes. (*A*, *Left*) A schematic representation of *att*L and *att*R substrates that recombine to trap the HJ complex, with the *att*R top strand (TS) labeled with fluorescein (Fl, green ball) and the *att*L bottom strand (BS) labeled with TAMRA (Rh, red ball). (*Right*) Analysis of the reaction products by PAGE under native and denaturing conditions. The fluorescein (Fl) and TAMRA (Rh) emission were collected using a Typhoon laser scanner. In the native gel, proteins were removed from one aliquot of the HJ complex by extraction with phenol/chloroform; the naked HJ DNA and the HJ complex are distinguished from each other by their different electrophoretic mobilities. In the denaturing urea gel, the presence of the 198 nt Fl-labeled single strands indicates the recombining *att* sites have exchanged top strands. The presence of the 94-nt Rh-labeled single strands coupled with the absence of a 198-nt Rh-labeled single strands indicates the bottom strands remain in their parental state. Size markers (M) are *att* sites labeled with one or two (slower mobility) fluorescent dyes. The gel lanes in *A* and *B* have been selected from a single gel containing several different experiments. (*B*) The schematic representation and gels illustrating recombination between fluorescently labeled *att* sites that give rise to recombinant products *attP* and *attB* (both DNA strands have been exchanged). (*C*) The number of λ-Integrase monomers present in the HJ complex and in the recombinant *attP* complexes were quantified using ^{32}P-radiolabeled DNA and Cy3-labeled λ-Int (C25S Int labeled with Cy3-maleimide at the surface-accessible Cys197 was a gift from Jeffrey Mumm, Thermo Fisher Scientific, Carlsbad, CA.). A single monomer of Cy3-labeled λ-Int was covalently trapped using a ^{32}P-radiolabeled, half-*attR* suicide substrate (13) and the covalent complex was isolated by KCl-SDS precipitation (leftmost lane of the gel). Because the same ^{32}P-radiolabeled oligonucleotide was used to prepare the half-*attR* suicide substrate and *attR* recombination substrates, the^{32}P-counts and Cy3 intensity in the former could be used to quantify the ratio of DNA to Int in the product attP and HJ complexes (second and third lanes, respectively). The ratio of Cy3 to ^{32}P counts indicates the HJ complex contains four monomers of λ-Int; the larger of the two *attP* complexes contains three Ints, whereas the smaller contains two. The gel lanes have been selected from a single gel containing several different experiments. (*D*) The HJ complex was subjected to native PAGE immediately, or 12 h, after the reaction was completed. The reactions that were analyzed 12 h after completion were stored either at 4 or 25 °C, as indicated.



B

Name	Function	Description
Crn 1	Chimeric Cre	λ-amino terminal DNA binding domain fused to Cre
Crn 2	Unique specificity excisive briding enzyme	AtERF DNA binding domain fused to CreCM2 core domain
Crn 3	Unique specificity excisive briding enzyme	AtERF DNA binding domain in which the α-helix has been replaced with the amino-terminal α-helix from λ-Integrase fused to Cre CM2 core domain
Crn 4	C-domain proof of principle	AtERF DNA binding domain replaces λ-amino terminal domain in Crn1
Crn 5	N-domain proof of principle	λ-amino terminal DNA binding domain fused to CreCM2

C

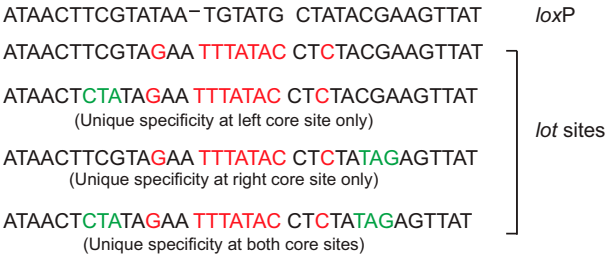


Fig. S2. Construction of the Crn recombinases. (*A*) Amino acid sequences of the NTDs of the indicated Crn recombinases. Crn1 has the NTD of λ Int [this is also the NTD of the previously described chimeric Cre (1)]. The 47-residue DNA binding domain of the AtERF protein (red) (2, 3) was used to replace the analogous 40-residue NTD present in Crn1 (blue) to create Crn2 and Crn4. Proteins containing this specificity bind to the G box consensus sequence, AGCCGCC (2), whereas the proteins containing the DNA binding domain of λ Int, Crn1, and Crn5 bind to the λ arm-type consensus sequence (MRGTCACTAT). Because the DNA binding domain of AtERF lacks an Xis binding site, the 15-residue α-helix of lambda integrase, which interacts with Xis (green), was used to replace the analogous 15-residue α-helix of the AtERF DNA binding domain (residues 47–66 of Crn2) to generate Crn3. (*B*) Composition of the five Crn recombinases. Crn1 and Crn4 have the core binding specificity of Cre. Crn2, Crn3, and Crn5 all contain the core-type DNA binding specificity of the Cre mutant CreCM2, which binds to modified *lox*P sites known as *lox*M7 (ATAACTCTATATA) (4, 5). This Cre mutant, which differs from WT Cre at five residues (I74A, T258L, R259S, E262H, and E266G), is unable to recombine canonical *lox*P sites. Crn1, 2, and 3 are described in the paper; Crn4 and Crn5 were constructed for the purposes of the experiment described in Fig. S3. (*C*) Core regions sequences of *lox*P and the four *lot* sites. The first *lot* sequence is the previously described core region used with the chimeric Cre (Crn1) and containing two Cre core sites (1). The following three lot sites have incorporated the unique M7 specificity (green) in one or both core sites, as indicated. The P and P′ arms of these sites are not shown (main text).

1. Warren D, Laxmikanthan G, Landy A (2008) A chimeric Cre recombinase with regulated directionality. *Proc Natl Acad Sci USA* 105(47):18278–18283.
2. Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M (1998) A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J* 17(18):5484–5496.
3. Wojciak JM, Sarkar D, Landy A, Clubb RT (2002) Arm-site binding by lambda -integrase: Solution structure and functional characterization of its amino-terminal domain. *Proc Natl Acad Sci USA* 99(6):3434–3439.
4. Saraf-Levy T, et al. (2006) Site-specific recombination of asymmetric lox sites mediated by a heterotetrameric Cre recombinase complex. *Bioorg Med Chem* 14(9):3081–3089.
5. Santoro SW, Schultz PG (2002) Directed evolution of the site specificity of Cre recombinase. *Proc Natl Acad Sci USA* 99(7):4185–4190.

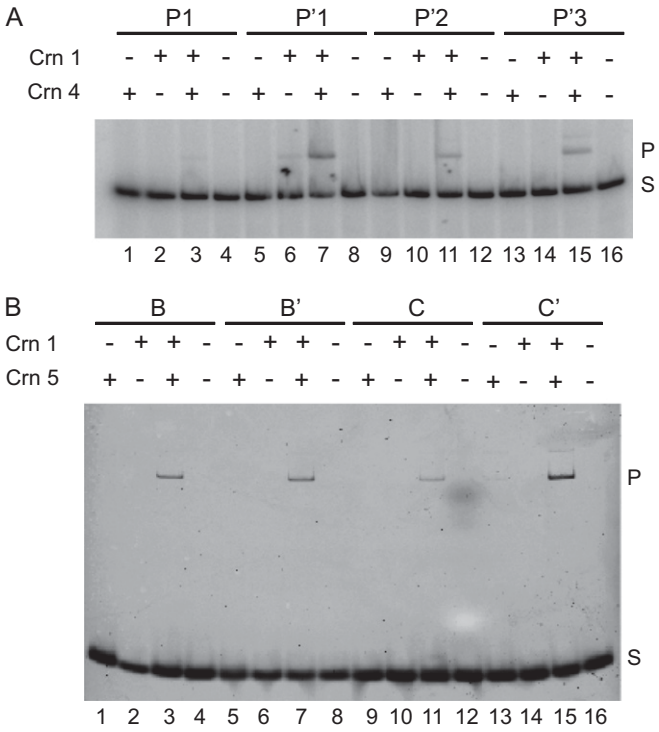


Fig. S3. Testing the specificity and functionality of the Crn NTDs and CTDs in integrative recombination. (*A*) To test the specificity and functionality of the NTDs alone, we used Crn1 and Crn4, whose NTDs are from λ and AtERF, respectively. Both recombinases have the same Cre CTD. The ability of a single GCC box-containing DNA sequence to replace each one of the λ arm-type sites was tested individually in four sets of recombination reactions. Each set of recombination reactions contained Crn1, Crn4, or both recombinases together. In each set of reactions, containing one GCC arm-type site and three λ arm-type sites, recombination is only seen when both proteins are present. The very low level of recombination for the P1 site (not seen here in the absence of significant over exposure) is discussed in *Results* and *Discussion*. (*B*) In analogous experiments to the test specificity and functionality of the CTDs alone, we used Crn1 and Crn5, whose CTDs are from canonical Cre and CreCM2, respectively. Both recombinases have the same λ NTD. The ability of a single M7 *lot* core-type sequence to replace each one of the Cre *lot* core-type sites was tested individually in four sets of recombination reactions. Each set of recombination reactions contained Crn1, Crn5, or both recombinases together. All of the reactions also contained a catalytically inactive (1) Cre-CM2(K201A) protein (lacking an NTD), which has the function of suppressing the low level of nonspecific binding of Crn1 to the lotM7 C′ site but does not interfere with the efficiency or specificity of recombination by the two competent (bridging) recombinases (see *Results*).

1. Ghosh K, Guo F, Van Duyne GD (2007) Synapsis of loxP sites by Cre recombinase. *J Biol Chem* 282(33):24004–24016.

Table S1. NTD recombination and arm site cross-linking

		Cross-linking assay																	
		A G G T C A C T A T C C A G T G A T																	
Int mutant	Recombination assay	1	1b	2	2b	3	3b	4	4b	5	5b	6	6b	7	7b	8	8b	9	9b
R19C	‡					++			++										
N20C	‡‡‡								++										
N21C	‡‡	++		++															
Y23C	‡‡‡			+		+													
C25S WT	‡‡								+										
R27C	‡‡											+					+		
R30C	NT	+		+															
G32C	‡‡‡																+		+
E34C	‡‡											++							

The 9-bp sequence of the P′1 arm-type site is shown (b indicates the bottom strand). The yield of the product in the integrative recombination assays is indicated relative to Int C25S (which has the same activity as WT Int) (‡‡), (‡) less than, and (‡‡‡) more than C25S, respectively. In the cross-linking assay, +, band corresponding to the covalently cross-linked complex was weaker than the unbound DNA across the entire range of Int concentrations tested; ++, cross-linked band was stronger than that of the unbound DNA at the optimal Int concentration (0.3–0.6 μM). The pair N20C–A(4b) was chosen (bold) as the best candidate. It was then tested at all arm sites and found to be suitable for the entire experiment.

Table S2. CTD recombination and core site Cross-linking

		Cross-linking assay*																	
		A C T A A G T T G T G A T T C A A[†] C																	
Int mutant	Recombination assay	1	1b	2	2b	3	3b	4	4b	5	5b	6	6b	7	7b	8	8b	9	9b
K95C	‡									++									
T96C	—							++											
N99C	—							++		++									
Y100C	—					NT		NT											
L142C	‡			++		++		++											
G283C	‡‡																++		

*C′ sequence was used.

[†]This base pair (T-A) was introduced at the B site because it was not present in WT.

The 9-bp sequence of the C′ core-type site is shown (b indicates the bottom strand). The results of the integrative recombination assays are indicated as follows: —, no recombination; ‡, no recombination in the absence of DTT and weak recombination in the presence of DTT; ‡‡, good recombination in the absence or presence of DTT. In the cross linking assay, +, band corresponding to the covalently cross-linked complex was weaker than the unbound DNA across the entire range of Int concentrations tested; ++, cross-linked band was stronger than that of the unbound DNA at the optimal Int concentration (0.3–0.6 μM). The integrative recombination results were confirmed in excisive recombination reactions. The only Int mutant showing good recombination (G283C) had only one candidate core site partner (A at 8b). This pair G283C–A(8b) (bold) was tested at the other core sites and proved suitable for the entire experiment. The A at this position was present in all core sites except B, where it is a T. Therefore, when the tether on B was required, the WT T:A base pair was changed to A:T.