Title

Spatially Directed Assembly of a Heterotetrameric Cre-Lox Synapse Restricts Recombination Specificity

Journal

Journal of Molecular Biology, 378(3)

ISSN

0022-2836

Authors

Gelato, Kathy A
Martin, Shelley S
Liu, Patty H
et al.

Publication Date

2008-05-01

DOI

10.1016/j.jmb.2008.02.058

Peer reviewed

doi:10.1016/j.jmb.2008.02.058 *J. Mol. Biol.* (2008) **378**, 653–665

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Spatially Directed Assembly of a Heterotetrameric Cre-Lox Synapse Restricts Recombination Specificity

Kathy A. Gelato[1], Shelley S. Martin[2], Patty H. Liu[2], April A. Saunders[2] and Enoch P. Baldwin[2,3]*

[1]*Biochemistry and Molecular Biology Graduate Group, University of California, Davis, Davis, CA 95616, USA*

[2]*Department of Molecular and Cellular Biology, University of California, Davis, Davis, CA 95616, USA*

[3]*Chemistry Department, University of California, Davis, Davis, CA 95616, USA*

Received 29 November 2007; received in revised form 13 February 2008; accepted 25 February 2008
Available online
4 March 2008

Edited by J. Karn

The pseudo-fourfold homotetrameric synapse formed by Cre protein and target DNA restricts site-specific recombination to sequences containing dyad-symmetric Cre-binding repeats. Mixtures of engineered altered-specificity Cre monomers can form heterotetramers that recombine nonidentical asymmetric sequences, allowing greater flexibility for target site selection in the genome of interest. However, the variety of tetramers allowed by random subunit association increases the chances of unintended reactivity at nontarget sites. This problem can be circumvented by specifying a unique spatial arrangement of heterotetramer subunits. By reconfiguring intersubunit protein–protein contacts, we directed the assembly of two different Cre monomers, each having a distinct DNA sequence specificity, in an alternating (ABAB) configuration. This designed heterotetramer preferentially recombined a particular pair of asymmetric Lox sites over other pairs, whereas a mixture of freely associating subunits showed little bias. Alone, the engineered monomers had reduced reactivity towards both dyad-symmetric and asymmetric sites. Specificity arose because the organization of Cre-binding repeats of the preferred substrate matched the programmed arrangement of the subunits in the heterotetrameric synapse. When this "spatial matching" principle is applied, Cre-mediated recombination can be directed to asymmetric DNA sequences with greater fidelity.

Keywords: Cre-LoxP site-specific recombination; protein engineering; orthogonal protein–protein interface; library screening/selection; directed subunit assembly

**Corresponding author.* Department of Molecular and Cellular Biology and the Chemistry Department, University of California, Davis, Davis, CA 95616, USA. E-mail address: epbaldwin@ucdavis.edu.

Present addresses: K. A. Gelato, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany; P. H. Liu, Department of Neurological Surgery, University of California, San Francisco, 13 Parnassus Avenue, San Francisco, CA 94143, USA; A. A. Saunders, Department of Molecular and Integrative Neurosciences, The Scripps Research Institute, La Jolla, CA 92037, USA.

Abbreviations used: HJ, Holliday junction; CTH, C-terminal helix; CTD, C-terminal domain; YSSR, tyrosine site-specific recombinase; CreWT, wild-type Cre recombinase with a Met-His_6 N-terminal tag fused to Ser2; CreALSHG, CreWT with five substitutions, Ile174Ala, Thr258Leu, Arg259Ser, Glu262His, and Glu266Gly; CreAAF, CreWT with three substitutions, Met299Ala, Val304Ala, and Ala334Phe; CreAA, CreWT split-interface mutant with two substitutions, Met299Ala and Val304Ala; ALSHG-F, CreALSHG split-interface mutant with an additional substitution, Ala334Phe; LoxP, natural Cre recombinase recognition site; LoxM7, LoxP variant that is the preferred substrate of CreALSHG, containing T7C, C8T, and G9A left arm 13-bp repeat and the dyad-related C26T, G27A, and A28G right arm 13-bp repeat substitutions; LoxPM7, chimeric Lox site with a LoxP left arm 13-bp repeat and a LoxM7 right arm 13-bp repeat; LoxM7P, chimeric Lox site with a LoxM7 left arm 13-bp repeat and a LoxP right arm 13-bp repeat.

Introduction

Controlled protein oligomerization underlies specificity in many biological processes. Multisubunit complex assembly provides for precise targeting of binding or enzymatic activities as well as regulation of those activities through allosteric interactions and proximity effects. For sequence-specific binding proteins and enzymes, oligomerization of nucleic-acid-interacting domains increases recognition site size and target specificity. The larger sites reduce the chance occurrences of recognition sequences elsewhere in the host genome, thereby minimizing detrimental off-target effects. For homo-oligomeric proteins, this increased specificity requires no extra protein-coding capacity but limits recognition to repeated elements. Hetero-oligomeric assemblies can expand the range of possible target sequences by combinatorial mixing of differently specific partners, as is the case with homeodomain, bZIP coiled-coil, and bHLH transcription factors.[1,2] However, the increased number of potential recognition sequences necessarily reduces the uniqueness of binding interactions. Specificity is achieved by a variety of mechanisms including regulated protein expression and localization, or stabilization of individual subunit combinations.[3–6] Here, we enforced pairing of two variant tyrosine site-specific recombinase (YSSR) subunits in a defined spatial arrangement in order to direct recombination to specific asymmetric substrates, thus eliminating the requirement for repeated recognition sites while retaining a narrow specificity.

YSSRs efficiently induce crossovers between 20- and 40-bp DNA target sequences. In nature, the resulting integrations, excisions, inversions, and translocations are employed in chromosome and plasmid segregation, plasmid amplification, virus integration, and gene regulation.[7–10] In biotechnology, YSSRs have been harnessed to manipulate chromosome structure in living cells and organisms to generate controlled gene deletions, integrate transgenes, and excise viral DNA.[11–16] However, the high sequence specificity of the naturally occurring YSSRs and the relaxed specificity of engineered versions remain an obstacle for the widespread use in applications requiring both flexible recognition potential and extreme precision, such as gene therapy.[17] One critical limitation is the homotetrameric nature of recombination complexes, which enforces recombination to occurring between nearly identical, dyad-symmetric DNA sequences.

One well-studied YSSR, bacteriophage P1 Cre protein, carries out site-specific recombination at 34-bp LoxP sites.[18–21] Cre-LoxP recombination is widely utilized for genome manipulations because of its robust activity and simple requirements. Numerous biochemical, biophysical, and structural investigations have generated paradigms for YSSR recombination mechanism and function.[22,23]

Recombination begins by assembly of a synaptic complex containing four Cre monomer subunits and two LoxP sites (Fig. 1a).[24] As is typical for YSSRs, LoxP sites contain an inverted dyad of Cre-binding 13-bp repeats, separated by an asymmetric intervening sequence, the 8-bp spacer (Fig. 1b).[19] In the complex, "crossing-over" recombination results from two pairs of single-strand exchanges, effected by reversible cleavages and strand swaps within the 8-bp spacers (Fig. 1a and b). The first exchange forms a Holliday junction (HJ) intermediate and the second exchange forms the products.[21] The order and progression of strand exchanges are directed by interplay between 8-bp spacer asymmetry, an associated DNA bend, and differentiation of Cre monomers into cleaving and noncleaving conformations. This aspect of Cre function has been extensively studied and reviewed elsewhere.[27–30]

The two domains of a Cre monomer comprise a clamp that surrounds the 13-bp repeat, creating an extensive interface of protein–DNA contacts.[24] Monomers are recruited to each LoxP 13-bp repeat and the active tetrameric recombination complex is subsequently assembled through a pseudo-fourfold cyclic arrangement of Cre–Cre interactions (Fig. 1c). A crucial intersubunit contact is a domain swap in which the C-terminal helix (CTH, residues 333–340) packs against the C-terminal domain (CTD, residues 131–326) of an adjacent monomer.

The high fidelity of wild-type Cre for the 13-bp repeat restricts the range of available recombination targets. Efficient reactions are only realized with close matches to LoxP and even single base changes in the 13-bp repeats can nearly abolish efficient function.[31–33] This limitation has been circumvented by directed evolution of Cre variants with altered DNA specificity.[16,25,32,34] For example, a quintuple mutant CreALSHG ["C2(+/−) #4" from the work of Santoro and Schultz[25]], prefers to recombine the LoxP variant, LoxM7. LoxM7 contains three base-pair substitutions at a key protein–DNA interface in the 13-bp repeat (Fig. 1b), and is not recombined by wild-type Cre (CreWT).

The homotetrameric Cre-LoxP synapse exerts further substrate restrictions (Fig. 2a, i). The symmetric arrangement of protein subunits and their DNA-binding surfaces matches the arrangement of 13-bp repeats of the identical LoxP sites in the complex. As a result, CreWT cannot recombine chimeric Lox sites containing both LoxP and LoxM7 13-bp repeats[26] (Fig. 1b), since two of the 13-bp repeats do not match the subunit specificities.

Two strategies have been employed to bypass the symmetry restriction of YSSRs. The first strategy utilizes relaxed-specificity recombinases to simultaneously recognize substantially different 13-bp repeats.[16,25,26,35] While this is a convenient approach, such promiscuity could lead to off-target recombination that would be unacceptable for high-fidelity applications, such as gene therapy. The second strategy uses combinations of altered-specificity Cre mutants.[26] A four-variant mixture would permit recombination between two nonidentical asymmetric targets with the only requirement being identity in the central six base pairs that are swapped (Fig. 2a, ii). However, such a mixture can assemble into 70

unique tetramers, which decreases fidelity by expanding the range of utilizable sequences. Further, currently available evolved specificity variants may not have high selectivity. For example, CreALSHG achieves 80% of the product levels of CreWT in LoxP integration reactions *in vitro* and shows only a 2.1-fold affinity preference for LoxM7 (Table 1). Thus, for mixtures, the many combinations of four different moderate-fidelity subunits may lead to partial or full recombination reactions at nontarget chromosomal sites. This concern is validated by the cytotoxicity of heterodimeric zinc finger nuclease chimeras used to introduce programmed double-strand breaks,[37,38] which presumably resulted from unintended DNA cleavages from the homodimers.

For YSSRs, if a unique heterotetrameric subunit arrangement is enforced, then recombination can be restricted to two pairs of substrates, related by circular permutations of their different recombinase-binding repeats (Fig. 2a, iii). Within this pair, the spatial arrangement of repeats would match that of the subunits in the heterotetramer, whereas noncognate sites in which repeats are not "spatially matched" would not be efficiently incorporated into recombination synapses. Enforcing assembly of a single arrangement increases the effective site size thereby reducing off-target effects. Applying a similar principle, an engineered heterodimeric intron homing endonuclease preferentially cleaved

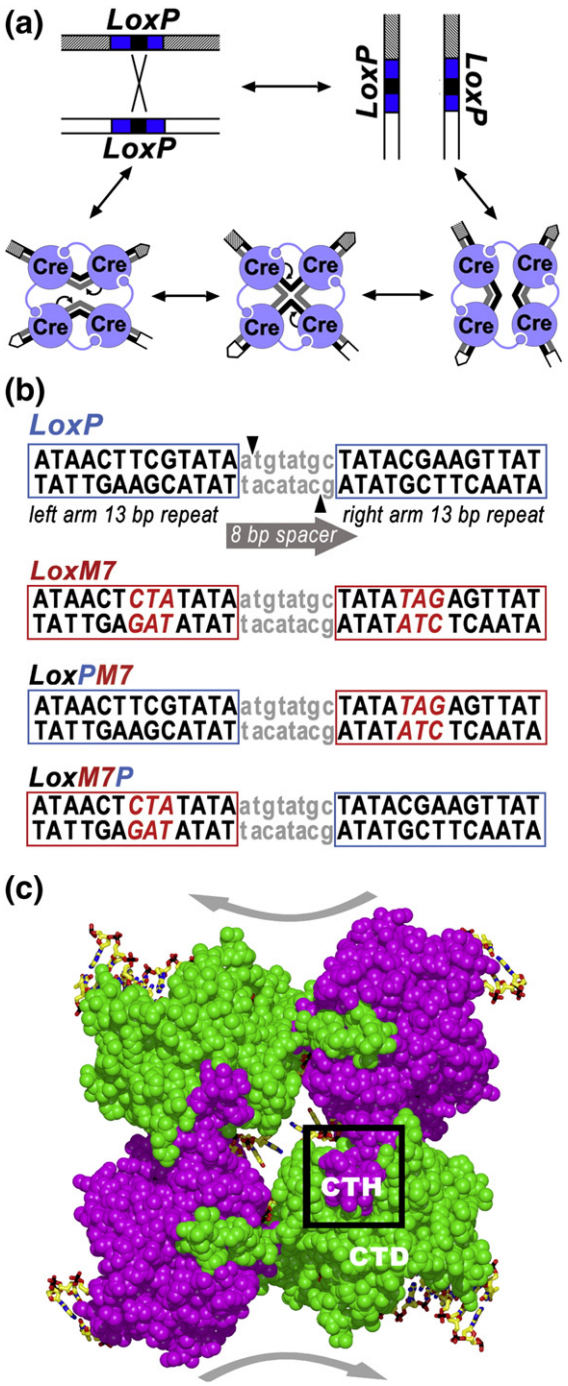
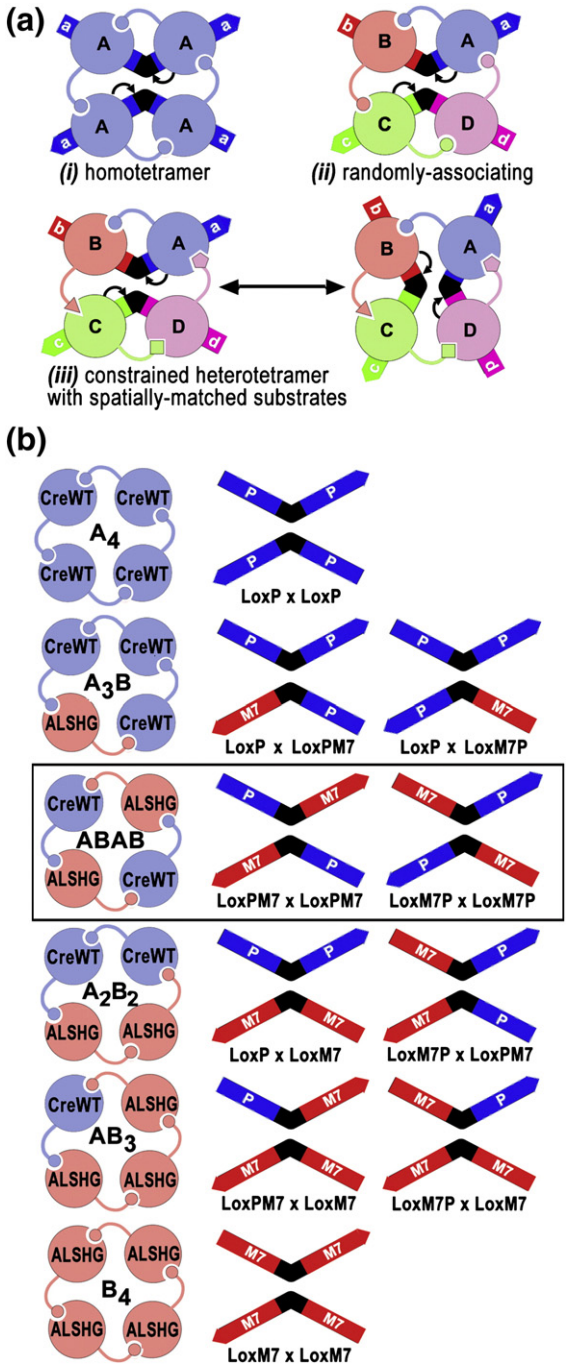


Fig. 1. Cre-LoxP recombination. (a) YSSR recombination structural mechanism.[22] Cre monomers are depicted as circles, and the two LoxP DNA strands are explicitly shown. LoxP directionality, as conventionalized in (b), is indicated by the pointed end. The critical CTH–CTD contact is depicted by the "ball-and-socket" connection between subunits (see c). The CTH is indicated by the small filled circle, the linker peptide is indicated by the connecting line, and the CTH docking site is indicated by the circular void. After complex assembly (left complex), two consecutive single-strand exchanges are carried out through cleavage/ligation events within the 8-bp spacer. The intervening 3′-phosphotyrosine covalent protein-DNA intermediates are not shown. The first exchange creates an HJ intermediate (center complex). The second strand exchange resolves the HJ to form the products, essentially as a reversal of the first exchange. (b) LoxP: 34-bp LoxP sites consist of a dyad of "left arm" and "right arm" Cre-binding 13-bp repeats (blue boxes) flanking an asymmetric 8-bp "spacer" that contains the sites of DNA cleavage and ligation (black arrowheads). The convention for site directionality is indicated by the gray arrow as pointing from left arm to right arm, and all subsequent figures follow this convention, as indicated by the arrowheads. The LoxM7 variant contains three base substitutions in each of the 13-bp repeats (red boxes). In all subsequent figures, blue indicates a LoxP 13-bp repeat and red indicates a LoxM7 13-bp repeat. The CreALSHG variant, obtained through directed evolution ["C2(±) #4" in Ref. 25], prefers to recombine LoxM7 compared to LoxP, whereas CreWT does not recombine LoxM7. Chimeric sites contain one LoxP and LoxM7 13-bp repeat each. LoxPM7 contains a LoxP left arm repeat and a LoxM7 right arm repeat, while LoxM7P contains a LoxM7 right arm repeat and a LoxP left arm repeat. CreWT also does not recombine these sites.[26] (c) The Cre-Lox recombination complex. The complex (PDB code 1CRX),[24] viewed from the CTD face, contains four Cre molecules and two Lox sites. With respect to the 8-bp spacer, the directionality of LoxP sites in the complex is anti-parallel (gray arrows; see b for the convention). This orientation is required for correct base-pairing of the 6 bp of exchanged DNA in products. Cre subunits are differentiated into alternating "cleaving" (green) and "noncleaving" (magenta) conformations, which enforces the pairwise nature, order, and regioselectivity of the exchanges.[24] The four Cre proteins associate through extensive protein–protein contacts that direct complex assembly. A critical contact involves a domain-swap of the CTH, which binds a pocket in the CTD of the adjacent monomer (box).

chimeric DNA substrates over the dyad-symmetric ones with high fidelity,[39] using a naturally occurring tandem fusion of the two nuclease DNA-binding domains to assure heterospecificity. However, this strategy cannot be readily applied to Cre because the N- and C-termini of recombinase monomers are on opposite sides of the recombination complex.

As "proof-of-concept" for the spatial matching principle in Cre-Lox recombination, we directed the assembly of a unique heterotetramer using engineered "orthogonal" heterospecific protein–protein interfaces[40–43]. We specifically positioned two different Cre monomers in a geometry that promoted efficient recombination between a pair of chimeric Lox sites that spatially matched the subunit arrangement, while restricting reactivity towards other pairs.



Results

Heterotetramer design

From the six unique synapses that result from random assortment of two Cre subunits having different DNA specificities, we selectively assembled an ABAB-type heterotetramer (Fig. 2b, left). For our two Cre proteins, we chose CreWT and CreALSHG[25] because both are well-characterized using *in vitro* recombination assays, alone and in combination.[26,44] For CreALSHG, the five substitutions confer specificity for LoxM7, mediated structurally by altered direct and water-bridged contacts.[44] However, it displays significant activity against LoxP, albeit with a 4.4-fold higher $S_{0.5}$ value and a sixfold slower rate at saturation compared to CreWT (Table 1).

Two different 13-bp repeats yield 10 unique Lox substrate pairs (Fig. 2b, right). The ABAB heterotetramer would promote recombination between identical Lox sites containing both 13-bp repeats, but not other substrate combinations. This arrangement requires only two orthogonal Cre–Cre interfaces to impose the alternating subunit positioning (Fig. 3a), whereas other arrangements would require four interfaces. Each Cre monomer in the ABAB heterotetramer contains reciprocal orthogonal CTH–CTD combinations to prevent self-oligomerization and enforce heterospecificity (Fig. 3b). The CreWT/CreALSHG heterotetramer would be expected to

Fig. 2. Heterotetrameric recombinases. Complexes are depicted as in Fig. 1a, except that the DNA duplexes are represented as a single-colored bar. Each color corresponds to a different protein-binding repeat sequence, with corresponding recombinase monomer specificities indicated by the lighter shaded circles. Site directionality is indicated by the pointed end of DNA. (a) A homotetrameric recombinase (i) will promote recombination between sites with identical or nearly identical inverted repeats, because the 13-bp repeat arrangement matches the homotetramer pseudo-fourfold symmetry. To allow recombination between dissimilar asymmetric sites, heterotetramers can be created by mixing four different specificity variants (ii), but up to 70 substrate pairs can be recognized. Mutually exclusive orthogonal interfaces direct heterotetramer assembly to a single defined subunit arrangement (iii). This complex will preferentially recombine pairs of sites whose repeats can match this arrangement. For example, an ABCD tetramer can perform ba × dc (left) or bc × da crosses, (right), but not ba × cd crosses. (b) With two different recombinase specificities, indicated as "CreWT" or "ALSHG" (see Fig. 1), six unique tetramers are possible (left side). In each, the DNA-binding surfaces match the positioning of one or two of the 10 possible unique substrate pairs arising from different arrangements of two distinct 13-bp repeats (right side). The ABAB tetramer should be specific for recombining identical chimeric sites (boxed).

Table 1. Complex assembly and kinetic parameters for Cre-Lox recombination reactions

		Complex assembly			Substrate turnover				
Cre protein(s)	DNA substrates	Extent[a] (%)	$S_{0.5}$ (nM)	Hill (α)	Extent[a] (%)	A	$k^1\times10^3$ (s^{-1})	$k^2\times10^3$ (s^{-1})	$t_{1/2}$ (s)
CreWT	LoxP×LoxP	47±1[2]	71±5	2.8±0.1	60±6[b]	0.42±0.02	4.5±1.3	32±3	48
	LoxM7×LoM7	NR[c] at 2400 nM complex							
CreALSHG	LoxM7×LoxM7	49±1[2]	150±16	3.7±0.5					
	LoxP×LoxP	38±8[4]	313±10	3.6±0.9	20±4[2]	0.66±0.03	1.1±0.2	9.7±0.1	300
CreAAF	LoxP×LoxP	45±4[4]	154±17	2.1±0.2	45±3[2]	0.84±0.01	1.5±0.7	22±4	346
CreWT+CreALSHG	LoxPM7×LoxPM7	54±5[4]	132±17	3.9±0.7	58±1[2]	0.59±0.07	5.3±0.1	23±1	67
	LoxM7P×LoxPM7	54±2[2]	144±18	4.3±0.4	56±2[2]	0.63±0.03	1.4±0.1	14±2	198
CreAAF+CreALSHG	LoxP×LoxP	21±8[3]	266±22	4.8±2.1					
CreAA+ALSHG-F	LoxM7P×LoxM7P	53±5[5]	179±19	3.6±0.6	65±7[4]	0.51±0.13	4.7±0.6	27±6	56
	LoxM7P×LoxPM7	<3% turnover at 1200 nM complex							
	LoxP×LoxM7	9±2[3]	178±9	5.9±1.3	10±8[2]	0.73±0.03	0.93±0.03	33±2	402
	LoxM7×LoxM7	NR at 360 nM complex							
	LoxP×LoxP	Reaction in qualitative assay comparable to LoxM7P×LoxM7P							
ALSHG-F	LoxM7P×LoxM7P	NR at 360 nM complex							
	LoxM7P×LoxPM7	NR at 360 nM complex							
	LoxP×LoxM7	NR at 360 nM complex							
	LoxM7×LoxM7	NR at 360 nM complex							
	LoxP×LoxP	NR at 360 nM complex							
CreAA	LoxM7P×LoxM7P	NR at 360 nM complex							
	LoxM7P×LoxPM7	NR at 360 nM complex							
	LoxP×LoxM7	NR at 360 nM complex							
	LoxM7×LoxM7	NR at 360 nM complex							
	LoxP×LoxP	Some products in qualitative assay at >240 nM complex[d]							

Complex assembly and kinetic parameters were calculated as described in Materials and Methods and are given as the averaged values from independent fits of replicate experiments±SD ($n>2$) or SE ($n=2$).

[a] Superscripts denote number of independent replicates.

[b] From Ref. 36.

[c] NR, no detectable reaction in the qualitative assay.

[d] Fig. 5a (middle panel).

efficiently perform LoxM7P×LoxM7P or LoxPM7×LoxPM7 reactions but not LoxM7×LoxP or LoxPM7×LoxM7P reactions.

Identification of an orthogonal Cre–Cre interface

Heterospecifically oligomerizing CreWT and CreALSHG subunits were obtained in three steps: (1) isolation of CreWT variants containing alternative functional Cre–Cre interfaces via library screening; (2) identification of one variant, CreAAF, which inefficiently forms synapses with CreWT; and (3) creation of CreWT- and CreALSHG-derived "split-interface" variants, CreAA and ALSHG-F, which contain noncomplimenting pairs of wild-type and orthogonal mutant interface surfaces.

To carry out the first step, we introduced amino-acid variation at the CTH–CTD interface, where side chains of CTH residues Ala334, Met335, and Leu338 pack against those of CTD Met299, Ala302, and Val304 (Fig. 3c). Formation of this contact is required for both tetramer assembly and cleavage activity.[30] We created an expression library of 8000 residue combinations by specifying all 20 amino acids at three positions that form a contiguous cluster, 299, 304, and 334, in which the large Met and Val side chains pack against the small Ala side chain. The library was transformed into a selection host in which cell survival is promoted by Cre-mediated excision of a conditional lethal marker. This method was used for both selection and efficiency screening of individual clones (see Materials and Methods). From 24 unique sequenced gene variants, five purified candidate proteins were assayed qualitatively for recombination activity *in vitro* (see Materials and Methods). The CreAAF variant contained three substitutions, Met299Ala, Val304Ala, and Ala334Phe, and had robust LoxP recombination activity. The pattern of amino acid replacements, with reciprocal large-to-small and small-to-large substitutions (Fig. 3d), suggested that CTH–CTD contacts with CreWT would be sterically disfavored. The variant alanine CTD substitutions in conjunction with wild-type Ala334 would be expected to create a destabilizing cavity, whereas the phenylalanine substituted for CTH Ala334 is too large to efficiently pack against wild-type Met299 and Val304 without rearrangements of the CTD interface.

Recombination competences were scored from the dependences of product levels on 2:1 Cre-Lox complex concentrations using a quantitative integration assay,[29,33,36] in which Cre-mediated recombination between a synthetic 34-bp Lox site and labeled 220-bp Lox-containing restriction fragment generates 113- and 141-bp products (Fig. 4a; see Materials and Methods). In LoxP×LoxP reactions (Fig. 4b), the CreAAF-containing complex yielded a sigmoidal dependence and a saturating level of product formation that were similar to that of the CreWT complex (44% *versus* 42% turnover), with a twofold higher $S_{0.5}$ value (154 nM *versus* 71 nM) (Fig. 4c, Table 1). The reduction in apparent affinity could

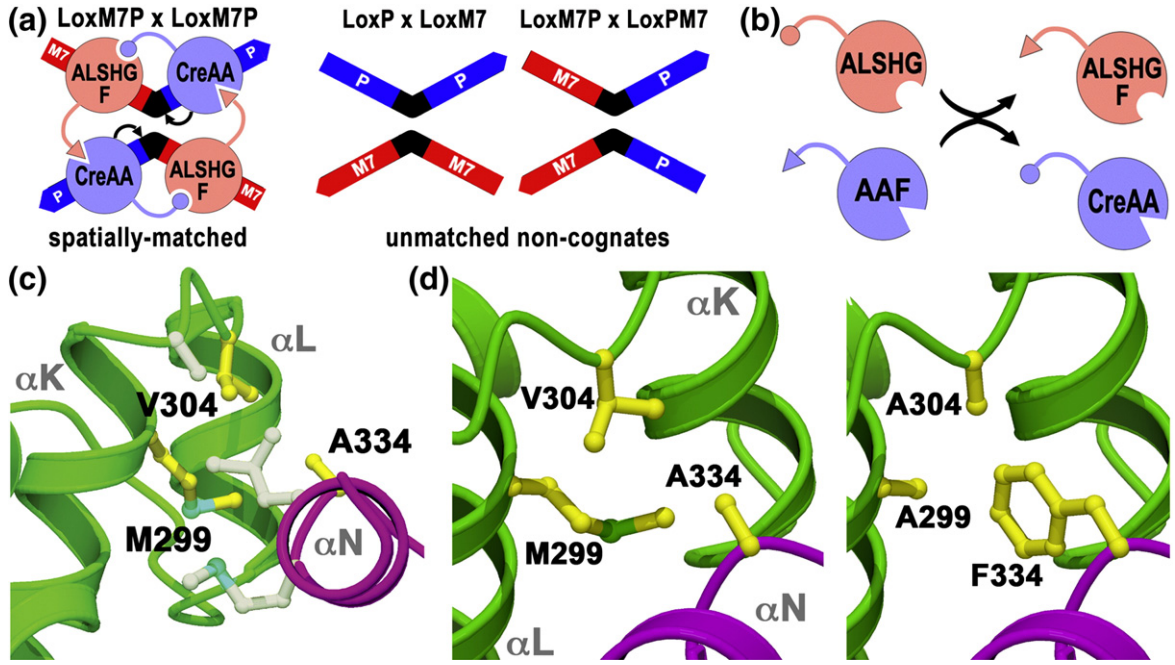


Fig. 3. Construction of the designed Cre heterotetramer. (a) "Spatial matching" of recombinase and substrates. The alternating ABAB arrangement of CreWT and CreALSHG subunits can be enforced by using two orthogonal interface surfaces, thereby favoring recombination of substrate pairs containing alternating 13-bp repeat arrangements (left) over unmatched arrangements (middle and right). (b) "Split interface" variants. CreAA and ALSHG-F contain complimentary combinations of orthogonal CreAAF and CreWT interfaces with CreWT and CreALSHG DNA specificities. The Ala334Phe substitution is indicated by the triangular CTH representation and the Met299Ala/Val304Ala CTD substitutions are indicated by the wedge-shaped indentation, while the corresponding CreWT interfaces are indicated by the circle and circular indentation. (c) CTH/CTD Cre–Cre interface. In the CTD (chain B of 1CRX, green), residues Met299, Ala302, and Val304 form a hydrophobic docking surface for residues Ala334, Met335, and Leu338 from the CTH (A chain, purple) from the adjacent subunit. Positions 299, 304 and 334 are the focus of this study and are highlighted in yellow. (d) Interface residues targeted for engineering. CTD residues Met299 and Val304 form a shallow pocket that accommodates CTH-residue Ala334 from the adjacent subunit (left panel). A hypothetical unminimized model of the "size-switch" CreAAF variant is shown in the right panel. The Ala299 and Ala304 side-chain truncations form a larger pocket to accommodate the larger Phe334. This volume swap creates alternate interface surfaces that interact unfavorably with CreWT surfaces. Contacts between CreAAF and CreWT subunits are disfavored because cavity formation or steric clashes would result.

result from smaller contribution of protein–protein interactions to complex assembly equilibria,[30,45] suggesting that although fully functional in terms of product production, the CreAAF CTH–CTD interface is not as avid as in the CreWT complex.

We next assayed the exclusion of the CreWT CTH–CTD interface by CreAAF using an *in vitro* interference assay (Fig. 4d). Concentration dependences of substrate turnover in a chimeric substrate recombination reaction, LoxPM7 × LoxPM7, were compared for equimolar mixtures of CreWT + CreALSHG or CreAAF + CreALSHG. Interface incompatibility was indicated by lower efficiency in CreAAF + CreALSHG reactions (Fig. 4e). Less substrate was converted to products or HJs (21% *versus* 54%), and higher complex concentrations were required to achieve half-maximum turnover ($S_{0.5}$ values, 266 nM *versus* 132 nM) (Table 1). The lack of complete interference may have resulted, in part, from productive binding to the LoxP 13-bp repeat by a CreALSHG homotetramer. As mentioned earlier, CreALSHG has significant LoxP × LoxP recombination activity, with an $S_{0.5}$ value of 313 nM and turningover 38% of substrate at saturation (Table 1). CreALSHG concentration dependence of recombination activity for chimeric substrates is intermediate between those for LoxP and LoxM7 (data not shown).

Construction and geometric specificity of the engineered ABAB heterotetramer

The CTH and CTD CreAAF substitutions were separated and combined with the complimentary wild-type interfaces (Fig. 3b, see Materials and Methods for details) to give CreAA, which contains a wild-type CTH with Met299Ala and Val304Ala CTD substitutions, and ALSHG-F, which contains a wild-type CTD and Ala334Phe CTH substitutions. When bound to their symmetric DNA substrates, these split-interface variants should be disfavored in forming homo-oligomeric contacts. Indeed, ALSHG-F/LoxM7 × LoxM7 reactions showed no LoxM7 recombination product at 540 nM complex, and CreAA required fourfold greater concentrations to achieve CreWT levels of substrate turnover in LoxP × LoxP reactions (Fig. 5a). As expected, CreAA

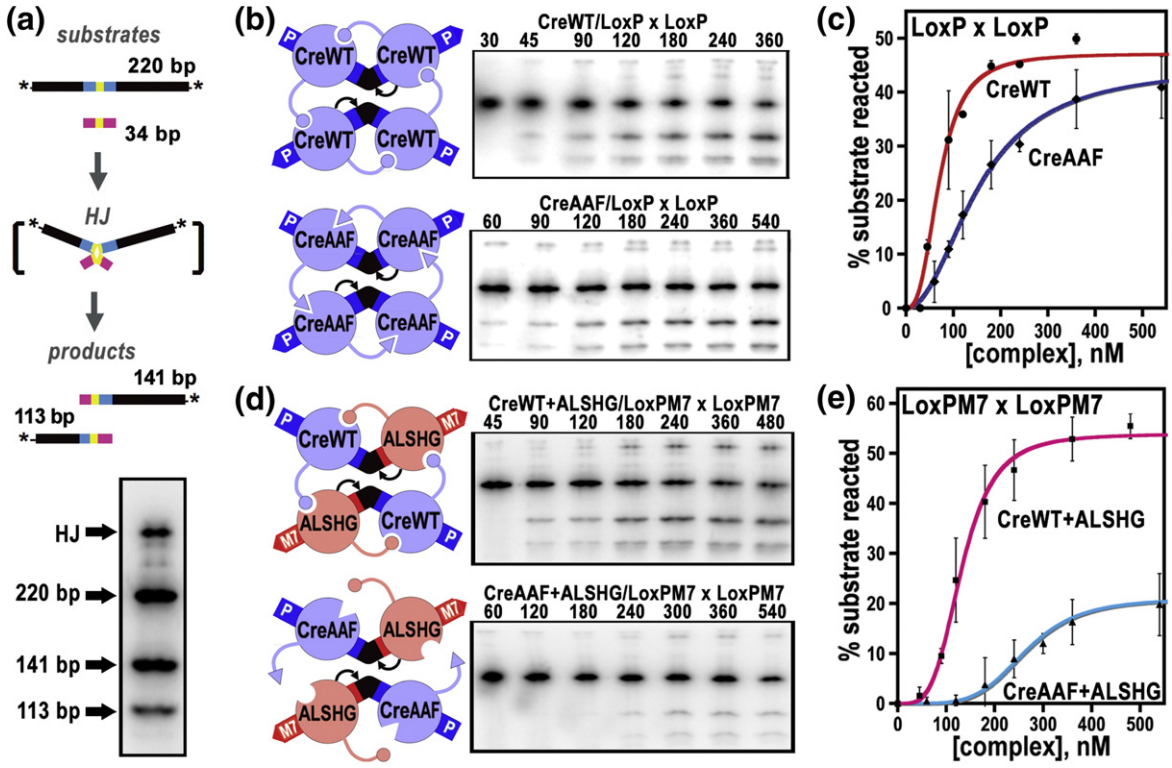


Fig. 4. Viability and selectivity of the engineered CreAAF interface. (a) Cre-Lox recombination assay. 13-bp repeats are indicated in cyan (220 bp) and magenta (34 bp), while the 8-bp spacers are indicated in yellow. Recombination swaps the 13-bp repeats. ^{32}P-5′-labelled 220-bp Lox-containing restriction fragments (*) were reacted with Cre proteins and 34-bp synthetic Lox (2:1 Cre monomers:Lox site) yielding 141- and 113-bp products and HJ intermediates. DNA components were separated by electrophoresis through SDS-PAGE gels, then visualized and quantified by phosphorimaging. (b) CreAAF is competent for LoxP × LoxP recombination. The concentration-dependences of active complex assembly as indicated by substrate turnover levels were compared for CreWT and CreAAF reactions. The concentrations of 2:1 Cre:LoxP complexes are indicated above each lane. The relevant complexes are diagrammed on the left of the corresponding phosphorimages. CTD and CTH interface surfaces are indicated as described in Fig. 3b. (c) Quantification of LoxP × LoxP titration reactions. Averaged, normalized measurements from two to four independent titration experiments and their standard deviations are shown by data points and error bars. The isotherms were generated from averaged fit Hill binding parameters given in Table 1 (Ref. 33). CreAAF reactions achieved comparable maximal levels to CreWT, but required twofold higher complex concentrations to achieve 50% maximum turnover. (d) CreAAF discriminates against wild-type CTH–CTD interfaces. In an "interference" assay, CreAAF and CreALSHG were recruited to adjacent positions on chimeric LoxPM7 sites (bottom left) and their ability to form active complexes was assessed, compared to a heterotetramer containing only wild-type interfaces (CreWT + CreALSHG/LoxPM7, upper left). Reactions were performed as described in (a) and (b). (e) Quantification of the interference assay. Titration data from (d) were treated as in (c). Data points and error bars depict the normalized averages and their standard deviations for three or four experiments. The isotherms are calculated from the averaged Hill parameters given in Table 1. CreAAF is less efficient at forming heterodimers with CreALSHG on chimeric LoxPM7 sites compared to CreWT, as evidenced by an increased $S_{0.5}$ value and reduced substrate turnover level compared to CreWT + CreALSHG reactions.

or ALSHG-F alone also produced no observable products in either LoxPM7 × LoxPM7 or LoxM7-P × LoxM7P reactions at concentrations up to 360 nM (Fig. 5b, left and middle panels). On the other hand, CreAA and ALSHG-F together readily recombined the expected pairs of chimeric substrates, LoxPM7 × LoxPM7 or LoxM7P × LoxM7P (Fig. 5b, right panels). Interestingly, the recombination efficiency of the two pairs is not the same, and LoxM7P × LoxM7P reactions achieve higher turnover levels than LoxPM7 × LoxPM7 reactions. In our quantitative analyses below, we used the more efficient LoxM7P × LoxM7P reactions.

Using the quantitative integration assay, we tested the ability of the CreAA + ALSHG-F heterotetramer to distinguish between chimeric substrates with cognate and noncognate arrangements of LoxP and LoxM7 13-bp repeats. In LoxM7P × LoxM7P complexes, the cyclic arrangement of repeats spatially matched the ABAB subunit alternation (Figs. 3a and 6a, left), resulting in efficient LoxM7P × LoxM7P recombination, turning over 54% of substrate with an $S_{0.5}$ value of 179 nM (Fig. 6a and e, upper panel; Table 1). By comparison, a 1:1 mixture of CreWT and ALSHG recombined various chimeric Lox combinations with similar efficiency. For example, in LoxPM7 × LoxPM7 and LoxPM7 × LoxM7P reactions, 54% of substrate is turned over with $S_{0.5}$ values of 130–145 nM (Figs. 4e, 6d, and 6e (*upper panel*), Table 1). In the qualitative assay, LoxM7P × LoxM7P and LoxP × LoxM7 reactions behaved comparably (data not shown).

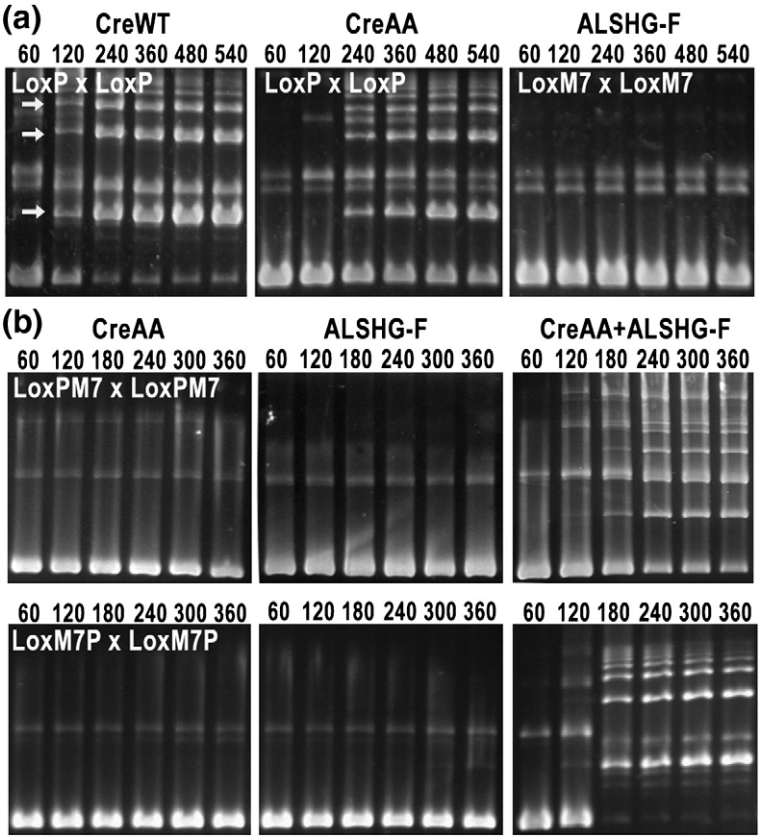


Fig. 5. Assessment of engineered monomer heterospecificity using a qualitative assay. (a–c) Varying concentrations of 2:1 Cre-Lox complex, indicated above each lane in nanomolars, were reacted with ~15 nM supercoiled Lox-containing pLITMUS plasmid for 16 h[36] (see Materials and Methods), electrophoresed through 1.2% agarose gels and stained with ethidium bromide. The major products are a complex mixture of linear monomers and multimers (white arrows, upper left panel) as well as catenanes and plasmid topoisomers. Single-site restriction enzyme digests yield the expected pair of product bands and a single reactant band (not shown). (a) CreAA/LoxP×LoxP (middle) and ALSHG-F/LoxM7×LoxM7 (right) compared to CreWT/LoxP×LoxP (left) reactions. The lack of complete conversion of supercoiled substrate suggests a greatly reduced recombination and/or topoisomerase activity. ALSHG-F was completely inactive in these assays. (b) CreAA (left) or ALSHG-F (middle) alone exhibited no recombination activity in chimeric substrate LoxPM7×LoxPM7 (upper panels) or LoxM7P×LoxM7P (lower panels) reactions, but the 1:1 CreAA+ALSHG-F mixture (right) recombined both substrates. However, LoxPM7×LoxPM7 reactions were much less efficient.

In contrast to the spatially matched substrates, the CreAA+ALSHG-F heterotetramer discriminates against LoxP×LoxM7 and LoxM7P×LoxPM7 combinations. These substrate pairs require identical subunits to occupy adjacent positions in the complex (Fig. 6b and c, left). LoxPM7×LoxM7P reactions yielded less than 3% turnover at up to 1200 nM complex (Fig. 6b and e, upper panel; Table 1). In LoxP×LoxM7 reactions, more substrate (~9%) was turned over at saturation, with an $S_{0.5}$ value that was similar to those of the LoxM7P×LoxM7P reactions (Fig. 6c and e, upper panel, and Table 1). However, 70% of reacted substrates accumulate as HJ intermediates, compared to 15% for LoxM7P×LoxM7P reactions (data not shown). In other words, CreAA-ALSHG-F exhibits ~17-fold discrimination in product formation between LoxM7P×LoxM7P and LoxP×LoxM7 reactions, similar to the discrimination against LoxPM7×LoxM7P.

In addition to product yield, the CreAA+ALSHG-F heterotetramer also discriminated kinetically against LoxP×LoxM7 (Fig. 6e, lower panel) with a 10-fold lower initial rate compared to LoxM7P×LoxM7P (Table 1). A maximum turnover difference of 12.5-fold is achieved after 3 min (Fig. 6e, lower panel, inset). The CreAA+ALSHG-F reaction rate was essentially identical to that for CreWT+CreALSHG in LoxPM7×LoxPM7 reactions (Table 1), but was 2.5-fold faster than the mixture in LoxPM7×LoxM7P reactions (Table 1; Fig. 6e, lower panel).

Discussion

Functionally useful matches to the large YSSR target sequences are improbable even in large genomes, and the naturally occurring mouse and human Lox-related sequences do not support efficient recombination by wild-type Cre.[34,46] Altered-specificity variants can provide access to these, as well as novel symmetric[25,34] or asymmetric sites.[16,35] Use of a single relaxed-specificity recombinase to release the symmetry requirement imbues flexibility in target site selection, but such enzymes may be unsuitably promiscuous, leading to DNA cleavage or strand exchange at undesirable sites. Even wild-type Cre has some latitude in the 13-bp repeats it will bind and reportedly cleaves murine BAC sequences relatively frequently in bacteria.[47] If CreWT can recombine 10 different 13-bp variants with reasonable efficiency, which is a somewhat conservative estimate,[31] then there are ~5000 potential substrate pairs, increasing to ~80,000 if twenty 13-bp repeats are recognized. Heterotetramers can also alleviate symmetry and identity requirements[26] but, again, target fidelity is a potential problem. Seventy tetramers, composed of randomly assorting monomers with different non-overlapping DNA specificities, as estimated above for CreWT, can recognize at least 1.2 million potential substrate pairs. Thus, mixed heterotetramers could exhibit significant off-target reactivity, particularly if the lower fidelity level exhibited by

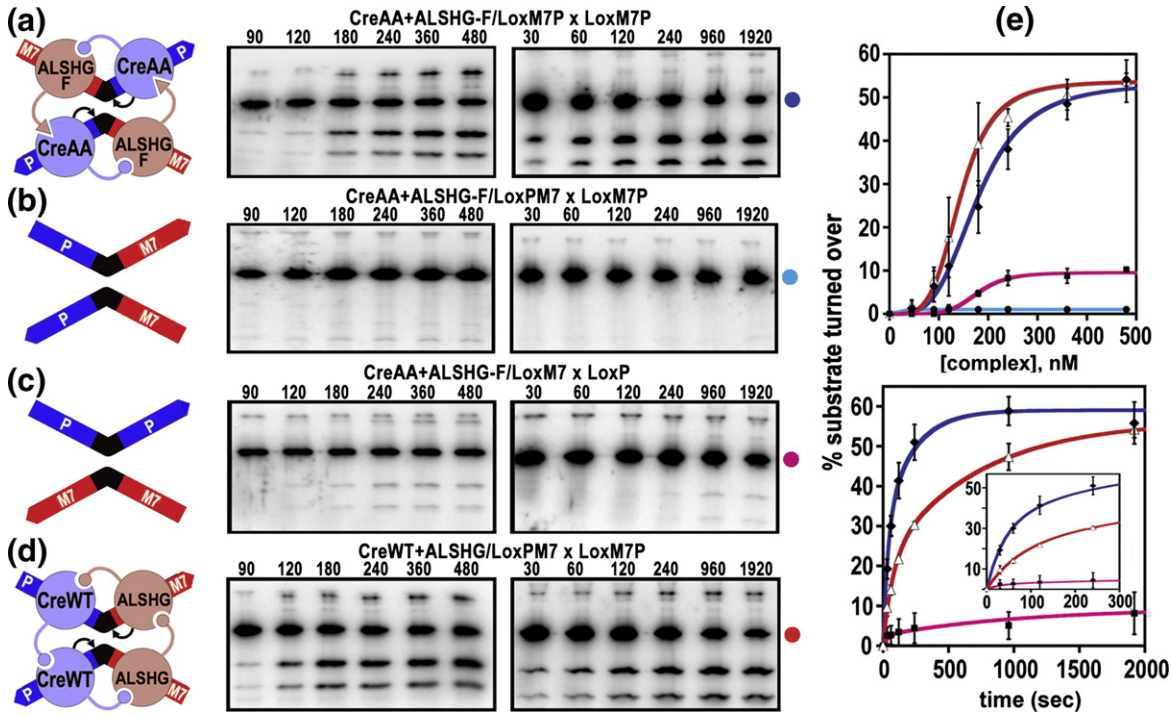


Fig. 6. The constrained ABAB heterotetramer distinguishes between different arrangements of LoxP and LoxM7 13-bp repeats. (a–d) The relevant complexes or substrate pairs are diagrammed on the left. Example phosphorimages of complex assembly titrations (left gel panel) and time course experiments (right gel panel) are shown. Titration reactions were carried out as described in Fig. 4 and Materials and Methods. Complex concentrations in nanomolars are given above the lanes. For time course reactions (1200 nM complex), aliquots were removed at the indicated times, given in seconds, quenched, and electrophoresed. The 220-bp fragment is indicated by the colored dot. The reactions are equimolar CreAA + ALSHG-F with (a) LoxM7P × LoxM7P; (b) LoxM7P × LoxPM7; and (c) LoxP × LoxM7. Similar data for the control reaction of the unconstrained heterotetramer CreWT + CreALSHG in a LoxPM7 × LoxM7P cross are given in (d). (e) The concentration (upper panel) or time dependences (lower panel and inset) of substrate turnover are given for each protein–DNA combination in (a–d): CreAA + ALSHG-F/LoxM7P × LoxM7P (filled diamonds, blue line); CreAA + ALSHG-F/LoxPM7 × LoxM7P (filled circles, cyan line); CreAA + ALSHG-F/LoxP × LoxM7 (filled squares, magenta line); and CreWT + ALSHG/LoxPM7 × LoxM7P (open triangles, red line). Average, normalized data points and their standard deviations are indicated, except for CreAA + ALSHG-F/LoxPM7 × LoxM7P, which turned over less than 3% substrate at 1200 nM complex. The curve colors correspond to the reactions indicated by dots at the right of (a–d). Complex assembly titration curves were derived from averaged Hill parameters from two to five experiments, given in Table 1. The time course progress curves were calculated from a biphasic kinetic model[33] using averaged parameters from two to four experiments, given in Table 1. The inset shows the expanded time course from 0 to 300 s.

CreALSHG *in vitro* is typical of evolved recombinase monomers.

Spatially constrained heterotetramers offer a solution by greatly reducing the number of possible substrates through a specified subunit geometry. For zinc finger nuclease chimeras, nontarget cleavage *in vivo* was suppressed by enforced heterodimerization using redesigned nuclease dimer interfaces,[48,49] indicating that simply limiting homodimerization results in useful improvements in target specificity. By comparison, the larger reduction in recombinase oligomerization complexity should effect even greater enhancement.

In this preliminary work, we provide the first evidence that constrained YSSR assembly, using reconfigured protein–protein interfaces, can direct recombination to particular asymmetric substrates. The CreAA-ALSHG-F heterotetramer, with its alternating subunit structure, facilely recombined a substrate pair with a cognate alternating arrangement of LoxM7 and LoxP 13-bp repeats and had low reactivity for arrangements with adjacent identical repeats. In contrast, the freely associating CreWT + CreALSHG mixture, which can assemble into six distinct tetramers, did not distinguish between these crosses. Together, the combination of binding and kinetic effects yielded an overall discrimination factor at least of 50- to 170-fold between spatially matched and unmatched substrates. An unexpected observation was that the CreAA/ALSHG-F heterotetramer had selectivity for LoxM7P × LoxM7P over LoxPM7 × LoxPM7 reactions, with threefold greater recombination levels for the preferred substrate (K. Gelato and E. Baldwin, unpublished data). This additional specificity level is the topic of a separate investigation, which will be addressed in a later manuscript.

The CreAA-ALSHG-F heterotetramer superficially resembles a natural bipartite recombinase, *Escherichia coli* XerCD, which recognizes the dissimilar arms of asymmetric *dif* sites[50,51] with distinct subunits. Si-

milar to our synthetic analog, XerC or XerD alone exhibit little *dif* recombination activity,[52] because essential protein–protein interactions coordinate strand exchanges.[53] The CreAA-ALSHG-F heterotetramer exhibited preferential reactivity, but unlike XerCD, not high fidelity. As predicted, LoxM7-P×LoxPM7 recombination levels were very low, perhaps a consequence of assembling the expected heterotetramer in a synapse in which the substrate 8-bp spacers are misaligned in parallel, as opposed to the viable anti-parallel arrangement (Fig. 7a, left). Incompatibly placed asymmetric bends in the 8-bp spacer may preclude stable complex formation,[30] and the first strand exchange would generate high-energy HJs that would not appreciably accumulate or proceed to the second strand exchange (Fig. 7a, middle and right). The low but significant LoxP×LoxM7 reactivity was not predicted and can be explained by several scenarios. An ABAB heterotetramer would utilize the demonstrated ability of ALSHG-F to react with LoxP repeats, but require CreAA to function on the LoxM7 repeats (Fig. 7b, i). Alternatively, homodimer binding of CreAA to LoxP and ALSHG-F to LoxM7 would allow wild-type CTH–CTD interfaces to bridge the two dimer–DNA complexes but would place the nonfunctional Met299/Val304-Phe334 interface within the ALSHG-F homodimers (Fig. 7b, ii). Similarly, a heterotetramer containing three ALSHG-F subunits would place a functional interface on LoxP with promiscuous ALSHG-F occupying a LoxP 13-bp repeat, but at the cost of one bridging interaction (Fig. 7, iii). As a consequence, HJ accumulation in this reaction may result from inefficient second exchange by a "mismatched" LoxM7-bound CreAA subunit or an ALSHG-F/ALSHG-F homodimer.

The LoxP×LoxM7 reactivity suggests that disfavored Cre–DNA or Cre–Cre contacts can be overcome by the cooperative interactions in the complex. Higher-fidelity monomers would partially alleviate this problem. Illustrating the consequence of CreALSHG promiscuity, the CreAA+ALSHG-F heterotetramer performs LoxP×LoxP crosses more efficiently than CreAA alone (data not shown), but does not carry out LoxM7×LoxM7 crosses, likely because of high CreAA DNA specificity. Greater CreAA-ALSHG-F assembly fidelity would also enhance its substrate bias. For a tetramer, a 0.7-kcal/mol difference at each interface could result in up to 100-fold discrimination in assembly. The selected CreAAF interface exhibited a side-chain volume redistribution reminiscent of size-switch combinations in alternate T4 lysozyme cores isolated by similar methods.[54] Such size swaps afforded 0.9–1.3 kcal increased stability over the cavity-containing or overpacked single mutants, analogous to homodimeric interfaces of CreAA and ALSHG-F, but the apparent discrimination by the CreAAF interface is only an estimated 0.2–0.4 kcal/mol per contact. More exclusive monomer–monomer pairs can be obtained either by increasing the size of the engineered CTH–CTD interface or by additionally reengineering N-terminal domain intersubunit contacts.

Homodimer reactivities also suggest different discrimination mechanisms at CreAA/CreAA and ALSHG/ALSHG-F interfaces, with effects on catalysis as well as protein–protein affinity (Fig. 5a, middle *versus* right panels). The results suggest that the Met299/Val304-Phe334 CTD–CTH combination does not support strand exchange, which is somewhat surprising because cavity-creating substitutions are generally more destabilizing than over-

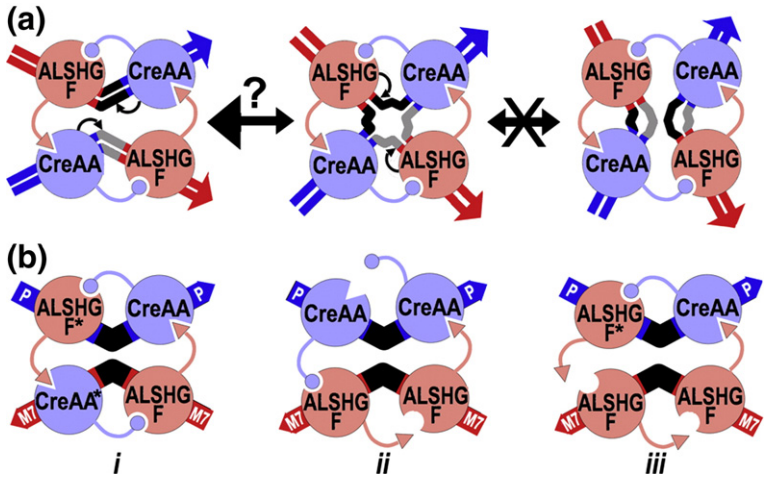


Fig. 7. Rationales for CreAA+ALSHG-F reactivities with substrate pairs that are not spatially matched. (a) Low reactivity of LoxM7P×LoxPM7 recombination. In Cre/LoxP×LoxP recombination complexes, anti-parallel arrangement of LoxP sites insures correct pairing for the central six base pairs in the HJ intermediate and product (see Fig. 1a). In CreAA+ALSHG-F/LoxM7P×LoxPM7 complexes, the spatially matched arrangement of 13-bp repeats places the Lox sites in parallel orientation, leading to four G-T/AC mismatches in the central six base pairs of the HJs and products, adding an 8–10 kcal/mol barrier to strand exchange. (b) Three possible scenarios for LoxM7×LoxP recombination. (i) Noncognate binding of the CreAA-ALSHG-F ABAB tetramer, with unmatched subunits marked with an asterisk (*). While ALSHG has reasonable LoxP recombination activity, CreWT has no measurable activity. (ii) Formation of an A_2B_2 heterotetramer by CreAA and ALSHG-F binding to LoxP and LoxM7 repeats, respectively. While the "bridging" interfaces between dimers are compatible, the intradimer interfaces are less effective. CreAA alone can support some LoxP recombination, suggesting that a CreAA/CreAA interface is partially functional, while ASLHG-F cannot recombine LoxM7, suggesting that ALSHG-F/ALSHG-F interface is not (Fig. 5a). (iii) In an AB_3 CreAA-ALSHG-F heterotetramer, all subunits are bound to 13-bp repeats that support some recombination, but there are only two functional interfaces, one bridging and one intradimer.

packing ones.[55] The overpacked interface might impede Tyr324 positioning for nucleophilic attack of the scissile phosphate, while the cavity-containing Ala299/Ala304-Ala334 interface may be less avid but also less structurally perturbed.

Besides the potential for increasing engineered recombinase fidelity, spatially directed assembly can also be used to investigate subunit interactions within the recombination synapse. Engineered interfaces in conjunction with DNA specificity variants can uniquely place additionally modified monomers at any of four unique positions in order to probe the mechanisms and spatial paths of cleavage coordination, assembly cooperativity, and allosteric communication.

Materials and Methods

Proteins

The following proteins were used in this study. CreWT, wild-type Cre recombinase with a Met-His_6 N-terminal tag fused to Ser2; CreALSHG, CreWT with five substitutions, Ile174Ala, Thr258Leu, Arg259Ser, Glu262His, and Glu266Gly; CreAAF, CreWT with three substitutions, Met299Ala, Val304Ala, Ala334Phe; CreAA, CreWT with two substitutions, Met299Ala, Val304Ala; ALSHG-F, CreALSHG with an additional substitution, Ala334Phe. His-tagged Cre proteins were expressed from pET28b(+)-derived constructs in BL21(DE3) cells (Novagen), purified, and stored as previously described.[56]

DNA

Plasmid DNA encoding CreALSHG was kindly provided by Steve Santoro from the Schultz Laboratory at The Scripps Research Institute, and the mutations were exchanged into the pET28-His6Cre vector.[44] Expression vectors for CreAAF and other variants were obtained as described below. Lox-containing pLITMUS plasmids (New England Biolabs), labeled restriction fragments, and synthetic Lox DNA substrates were also prepared as described previously.[36] Mutagenic primers were obtained from MWG Biotech (50 nmol scale) and phosphorylated with T4 kinase and ATP. The Lox sites used in this study are shown in Fig. 1b.

Heterospecifically oligomerizing CreAA and ALSHG-F proteins

CreAAF was obtained by *in vitro* selection for alternate CTH–CTD packing interfaces, which will be described in detail elsewhere. Briefly, a pET28b-His6Cre-based library was constructed in which randomized codons (XXG/C) were substituted for amino acid positions 299, 304, and 334. The library was transformed into an *E. coli* host harboring a chromosomal *rpsL20* streptomycin resistance allele, a DE3 T7 polymerase-expressing prophage (Novagen), and a Cm^r selection plasmid. The selection plasmid contains the wild-type rpsL gene, dominant for streptomycin sensitivity,[57] flanked by LoxP sites in direct orientation. This strain is Cm^r and Str^s. After transformation with the expression library, $Cm^r/Str^r/Kan^r$ cells expressing active Cre variants are selected on triple antibiotic plates, since the dominant Str^s gene is excised via recombinase activity. After DNA extraction and rescreening, 24 unique sequences were recovered from 38 clones. Unique clones were scored for recombination *in vivo* using the recovery efficiency of the Str^r phenotype. Five candidate proteins, including CreAAF, were purified and tested *in vitro* for overall activity and oligomerization heterospecificity (see Results).

The heterospecific split-interface variants CreAA and ALSHG-F were created by site-directed mutagenesis[58] of pET28b-His6Cre and pET28b-His6CreALSHG using the following oligonucleotides: Met299Ala/Val304Ala, GATCTCCGGTATTGAAgCTCCAGCcCGGGCCgcATCATCTCGCGCGGC, and Ala334Phe, GCGCACCATaaAaCCgGTTTCACT, respectively.

Integrative recombination activity assays

Recombination activity was assessed through Cre-mediated integration of a minimal 34-bp synthetic Lox duplex into a reporter substrate containing a single Lox site in a single turnover reaction.[36] Reactions are carried out at 21 °C in optimized Cre reaction buffer [300 mM lithium acetate, 20 mM Tris–acetate, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM DTT, pH 8.3]. Reactions (50–100 μL total) were initiated by mixing 0.9 volume reaction buffer and DNA with 0.1 volume of protein mixtures in 300 mM NaCl, 20 mM Tris–acetate, 1 mM EDTA, pH 8.3, quenched with 1 volume of 2× loading buffer (1× = 1% SDS, 20 mM DTT, 6% glycerol, 0.5 mg/mL proteinase K, and 0.05% bromphenol blue) and digested for 1 h at 37 °C prior to electrophoresis. Complex concentrations are expressed as the amounts of 2:1 Cre:Lox (e.g., 1 nM complex = 2 nM total protein + 1 nM synthetic 34-bp Lox site).

For qualitative assessments, the reporter substrate was ~15 nM supercoiled pLITMUS38(+) plasmid containing a 34-bp Lox inserted between the SnaBI and EcoRV sites.[44,36] The products were electrophoretically separated using a 1.2% agarose gel in Tris–acetate–EDTA buffer and visualized with ethidium bromide under UV light. Qualitative comparisons are made from CCD images of the gels, but rigorous quantitation is unfeasible.

For quantitative measurements, the reporter substrate was a 220-bp BamHI/StuI restriction fragment from the pLITMUS-Lox plasmid (2 or 10 nM), which was 5′-end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase.[36] Samples were quenched and electrophoresed through SDS/10% polyacrylamide gels. A Fuji Image plate was exposed to the dried gels and scanned using Molecular Dynamics Storm 860 image plate reader. Relative labeled DNA band intensities were quantified using ImageQuant. Reaction levels are taken as the amount of substrate that has been converted to 113- and 141-bp products and HJ intermediates ("substrate turnover").

Assessment of complex assembly competence

To assess the strengths of Cre–Cre and Cre-Lox interactions, substrate turnover was measured as a function of complex concentration by titrating reporter substrate with different Cre–synthetic Lox concentrations in 16-h-end-point reactions.[36] The complex concentrations typically ranged 15–500 nM (see figures for amounts). Complex assembly parameters were determined by fitting the amount of substrate turnover (v') and total complex concentration ([complex]) to the function $v' = (f \times [\text{complex}]^{\alpha}/(S_{0.5}{}^{\alpha} + [\text{complex}]^{\alpha})$.[36] The fit parameters were f, the

maximum amount of substrate turnover; $S_{0.5}$, the complex concentration when half the maximum product is produced (an apparent dissociation constant); and α, the apparent Hill coefficient (reported as the average of two to five independent experiments ± the standard deviation in Table 1). The $S_{0.5}$ value is a complex function of monomer DNA binding, dimerization, and tetramer assembly equilibria. The apparent α value likely reflects, in part, the degree of cooperativity arising from protein–protein interactions in complex assembly.

Single turnover kinetics

In time course reactions, 10 nM ^{32}P-labeled reporter was reacted with saturating complex (2400 nM Cre and 1200 nM Lox). Reactions were initiated by addition of complex and quenched after 30 to 1920 s. Rate parameters were determined by fitting the measured percent substrate reacted (v') and reaction time (t) to the function $v'(t) = f(1 - [Ae^{-k_1 t} + (1 - A)e^{-k_2 t}])$. The fit parameters were f, the percent of substrate turnover at $t = 8$; A and k_1, the amplitude and rate constant, respectively, for the reaction slow phase; k_2, the rate constant for the fast phase (reported as the average of two to four independent experiments ± the standard deviation in Table 1). The biphasic function is usually required to adequately fit the data,[36] suggesting that distinct fast- and slow-reacting complexes are formed early in the assembly process, but the data do not distinguish between a converging or sequential relationship between the two paths.

Acknowledgements

This work was funded by the National Institutes of Health grants RO1-GM63109 (E.B.) and T32-GM07377-25 (K.G.).

Supplementary Data

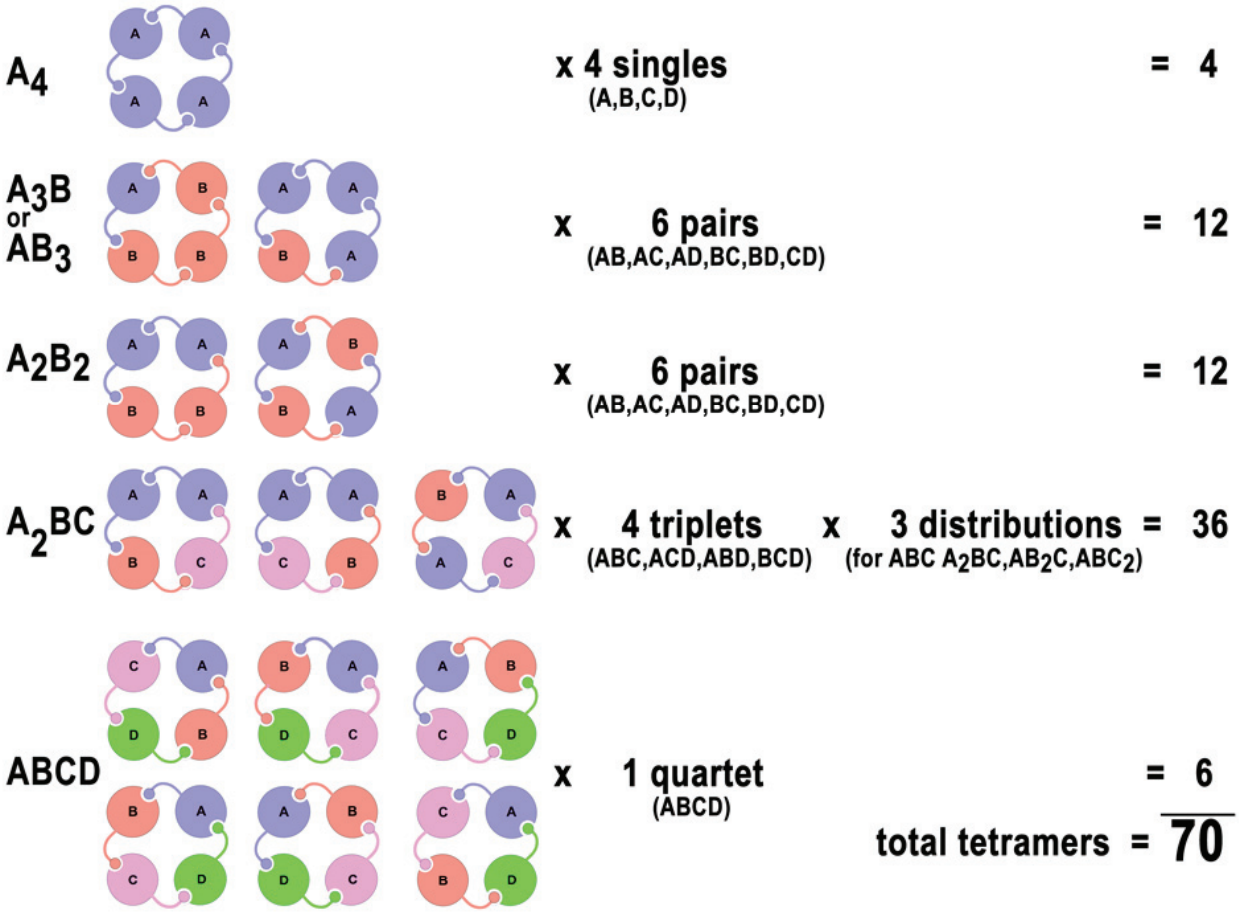
Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.02.058

References

1. Wolberger, C. (1999). Multiprotein–DNA complexes in transcriptional regulation. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 29–56.
2. Remenyi, A., Scholer, H. R. & Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol. Biol.* **11**, 812–815.
3. Calkhoven, C. F. & Ab, G. (1996). Multiple steps in the regulation of transcription-factor level and activity. *Biochem. J.* **317**, 329–342.
4. Chen, F. E. & Ghosh, G. (1999). Regulation of DNA binding by Rel/NF-kappaB transcription factors: structural views. *Oncogene*, **18**, 6845–6852.
5. Kewley, R. J., Whitelaw, M. L. & Chapman-Smith, A. (2004). The mammalian basic helix–loop–helix/PAS family of transcriptional regulators. *Int. J. Biochem. Cell Biol.* **36**, 189–204.
6. Mann, R. S. (1995). The specificity of homeotic gene function. *Bioessays*, **17**, 855–863.
7. Landy, A. (1993). Mechanistic and structural complexity in the site-specific recombination pathways of Int and FLP. *Curr. Opin. Genet. Dev.* **3**, 699–707.
8. Sadowski, P. D. (1993). Site-specific genetic recombination: hops, flips, and flops. *FASEB J.* **7**, 760–767.
9. Sherratt, D. J., Arciszewska, L. K., Blakely, G., Colloms, S., Grant, K., Leslie, N. & McCulloch, R. (1995). Site-specific recombination and circular chromosome segregation. *Philos. Trans. R. Soc. London, Ser. B*, **347**, 37–42.
10. Jayaram, M., Mehta, S., Uzri, D., Voziyanov, Y. & Velmurugan, S. (2004). Site-specific recombination and partitioning systems in the stable high copy propagation of the 2-micron yeast plasmid. *Prog. Nucleic. Acid Res. Mol. Biol.* **77**, 127–172.
11. Ow, D. W. (2002). Recombinase-directed plant transformation for the post-genomic era. *Plant Mol. Biol.* **48**, 183–200.
12. Branda, C. S. & Dymecki, S. M. (2004). Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell*, **6**, 7–28.
13. Jones, J. R., Shelton, K. D. & Magnuson, M. A. (2005). Strategies for the use of site-specific recombinases in genome engineering. *Methods Mol. Med.* **103**, 245–257.
14. Oberstein, A., Pare, A., Kaplan, L. & Small, S. (2005). Site-specific transgenesis by Cre-mediated recombination in *Drosophila*. *Nat. Methods*, **2**, 583–585.
15. Garcia-Otin, A. L. & Guillou, F. (2006). Mammalian genome targeting using site-specific recombinases. *Front. Biosci.* **11**, 1108–1136.
16. Sarkar, I., Hauber, I., Hauber, J. & Buchholz, F. (2007). HIV-1 proviral DNA excision using an evolved recombinase. *Science*, **316**, 1912–1915.
17. Schmidt-Supprian, M. & Rajewsky, K. (2007). Vagaries of conditional gene targeting. *Nat. Immunol.* **8**, 665–668.
18. Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M. & Hoess, R. (1981). Site-specific recombination and its role in the life cycle of bacteriophage P1. *Cold Spring Harbor Symp. Quant. Biol.* **45**, 297–309.
19. Hoess, R. H., Ziese, M. & Sternberg, N. (1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. *Proc. Natl Acad. Sci. USA*, **79**, 3398–3402.
20. Hoess, R. H. & Abremski, K. (1985). Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J. Mol. Biol.* **181**, 351–362.
21. Hoess, R., Wierzbicki, A. & Abremski, K. (1987). Isolation and characterization of intermediates in site-specific recombination. *Proc. Natl Acad. Sci. USA*, **84**, 6840–6844.
22. Van Duyne, G. D. (2001). A structural view of cre-loxp site-specific recombination. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 87–104.
23. Grindley, N. D., Whiteson, K. L. & Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* **75**, 567–605.
24. Guo, F., Gopaul, D. N. & van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40–46.
25. Santoro, S. W. & Schultz, P. G. (2002). Directed evolution of the site specificity of Cre recombinase. *Proc. Natl Acad. Sci. USA*, **99**, 4185–4190.
26. Saraf-Levy, T., Santoro, S. W., Volpin, H., Kushnirsky, T., Eyal, Y., Schultz, P. G. *et al.* (2006). Site-specific recombination of asymmetric lox sites mediated by a heterotetrameric Cre recombinase complex. *Bioorg. Med. Chem.* **14**, 3081–3089.

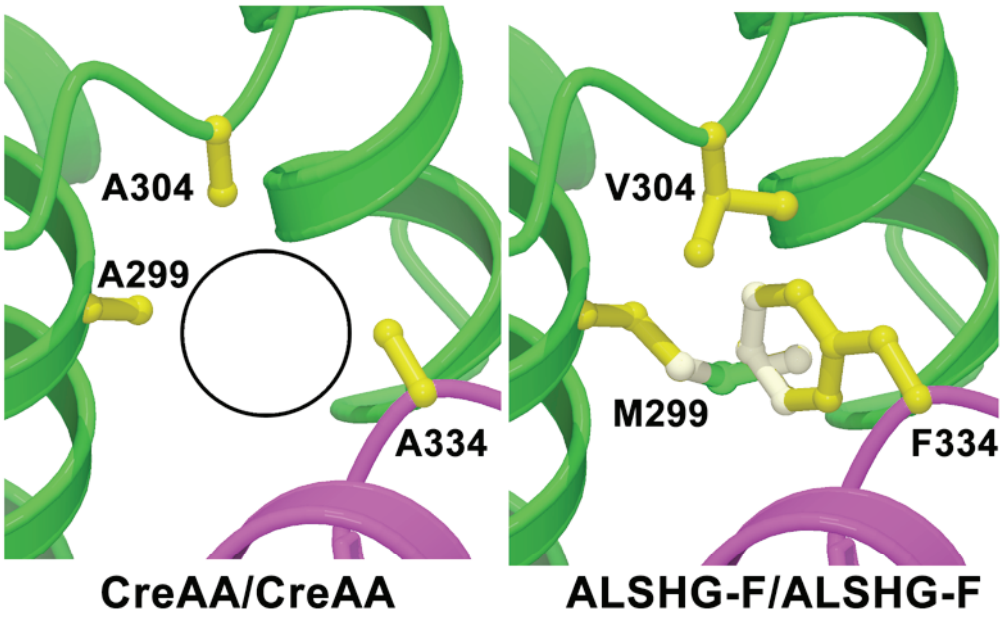
27. Lee, L. & Sadowski, P. D. (2003). Sequence of the loxP site determines the order of strand exchange by the Cre recombinase. *J. Mol. Biol.* **326**, 397–412.
28. Lee, L. & Sadowski, P. D. (2005). Strand selection by the tyrosine recombinases. In *Prog. Nucleic. Acid Res. Mol. Biol.* (Moldave, K., ed), vol. 80, pp. 1–42, Academic Press, New York.
29. Gelato, K. A., Martin, S. S. & Baldwin, E. P. (2005). Reversed DNA strand cleavage specificity in initiation of Cre-LoxP recombination induced by the His289Ala active-site substitution. *J. Mol. Biol.* **354**, 233–245.
30. Ghosh, K., Guo, F. & Van Duyne, G. D. (2007). Synapsis of loxP sites by Cre recombinase. *J. Biol. Chem.* **282**, 24004–24016.
31. Hartung, M. & Kisters-Woike, B. (1998). Cre mutants with altered DNA binding properties. *J. Biol. Chem.* **273**, 22884–22891.
32. Rufer, A. W. & Sauer, B. (2002). Non-contact positions impose site selectivity on Cre recombinase. *Nucleic Acids Res.* **30**, 2764–2771.
33. Gelato, K. A., Martin, S. S., Wong, S. & Baldwin, E. P. (2006). Multiple levels of affinity-dependent DNA discrimination in Cre-LoxP recombination. *Biochemistry*, **45**, 12216–12226.
34. Buchholz, F. & Stewart, A. F. (2001). Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nat. Biotechnol.* **19**, 1047–1052.
35. Bolusani, S., Ma, C. H., Paek, A., Konieczka, J. H., Jayaram, M. & Voziyanov, Y. (2006). Evolution of variants of yeast site-specific recombinase Flp that utilize native genomic sequences as recombination target sites. *Nucleic Acids Res.* **34**, 5259–5269.
36. Martin, S. S., Chu, V. C. & Baldwin, E. (2003). Modulation of the active complex assembly and turnover rate by protein–DNA interactions in Cre-LoxP recombination. *Biochemistry*, **42**, 6814–6826.
37. Porteus, M. H. & Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science*, **300**, 763.
38. Alwin, S., Gere, M. B., Guhl, E., Effertz, K., Barbas, C. F., 3rd, Segal, D. J. *et al.* (2005). Custom zinc-finger nucleases for use in human cells. *Mol. Ther.* **12**, 610–617.
39. Chevalier, B. S., Kortemme, T., Chadsey, M. S., Baker, D., Monnat, R. J. & Stoddard, B. L. (2002). Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell*, **10**, 895–905.
40. Nautiyal, S., Woolfson, D. N., King, D. S. & Alber, T. (1995). A designed heterotrimeric coiled coil. *Biochemistry*, **34**, 11645–11651.
41. Arndt, K. M., Pelletier, J. N., Muller, K. M., Alber, T., Michnick, S. W. & Pluckthun, A. (2000). A heterodimeric coiled-coil peptide pair selected *in vivo* from a designed library-versus-library ensemble. *J. Mol. Biol.* **295**, 627–639.
42. Ali, M. H., Taylor, C. M., Grigoryan, G., Allen, K. N., Imperiali, B. & Keating, A. E. (2005). Design of a heterospecific, tetrameric, 21-residue miniprotein with mixed alpha/beta structure. *Structure*, **13**, 225–234.
43. Kortemme, T., Joachimiak, L. A., Bullock, A. N., Schuler, A. D., Stoddard, B. L. & Baker, D. (2004). Computational redesign of protein–protein interaction specificity. *Nat. Struct. Mol. Biol.* **11**, 371–379.
44. Baldwin, E. P., Martin, S. S., Abel, J., Gelato, K. A., Kim, H., Schultz, P. G. & Santoro, S. W. (2003). A specificity switch in selected cre recombinase variants is mediated by macromolecular plasticity and water. *Chem. Biol.* **10**, 1085–1094.
45. Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R. & Stewart, A. F. (1998). Comparative kinetic analysis of FLP and cre recombinases: mathematical models for DNA binding and recombination. *J. Mol. Biol.* **284**, 363–384.
46. Thyagarajan, B., Guimaraes, M. J., Groth, A. C. & Calos, M. P. (2000). Mammalian genomes contain active recombinase recognition sites. *Gene*, **244**, 47–54.
47. Semprini, S., Troup, T. J., Kotelevtseva, N., King, K., Davis, J. R., Mullins, L. J. *et al.* (2007). Cryptic loxP sites in mammalian genomes: genome-wide distribution and relevance for the efficiency of BAC/PAC recombineering techniques. *Nucleic Acids Res.* **35**, 1402–1410.
48. Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I. *et al.* (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778–785.
49. Szczepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D. J. & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* **25**, 786–793.
50. Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T. & Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at dif and cer in *E. coli* K12. *Cell*, **75**, 351–361.
51. Hayes, F. & Sherratt, D. J. (1997). Recombinase binding specificity at the chromosome dimer resolution site dif of *Escherichia coli*. *J. Mol. Biol.* **266**, 525–537.
52. Blakely, G. W. & Sherratt, D. J. (1994). Interactions of the site-specific recombinases XerC and XerD with the recombination site dif. *Nucleic Acids Res.* **22**, 5613–5620.
53. Hallet, B., Arciszewska, L. K. & Sherratt, D. J. (1999). Reciprocal control of catalysis by the tyrosine recombinases XerC and XerD: an enzymatic switch in site-specific recombination. *Mol. Cell*, **4**, 949–959.
54. Baldwin, E., Xu, J., Hajiseyedjavadi, O., Baase, W. A. & Matthews, B. W. (1996). Thermodynamic and structural compensation in "size-switch" core repacking variants of bacteriophage T4 lysozyme. *J. Mol. Biol.* **259**, 542–559.
55. Baldwin, E. P. & Matthews, B. W. (1994). Core-packing constraints, hydrophobicity and protein design. *Curr. Opin. Biotechnol.* **5**, 396–402.
56. Woods, K. C., Martin, S. S., Chu, V. C. & Baldwin, E. P. (2001). Quasi-equivalence in site-specific recombinase structure and function: crystal structure and activity of trimeric Cre recombinase bound to a three-way Lox DNA junction. *J. Mol. Biol.* **313**, 49–69.
57. Russell, C. B. & Dahlquist, F. W. (1989). Exchange of chromosomal and plasmid alleles in *Escherichia coli* by selection for loss of a dominant antibiotic sensitivity marker. *J. Bacteriol.* **171**, 2614–2618.
58. Kunkel, T. A., Bebenek, K. & McClary, J. (1991). Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol.* **204**, 125–139.

Figure S1. Enumeration of unique Cre heterotetramers.



The complete collection of 70 unique Cre tetramers composed of four different types of subunits, after correcting for four-fold symmetry. Because the tetramer four-fold axis is directional, the enantiomorphs (clockwise ABCD vs DCBA) are not identical.

Figure S2. Models of non-cognate CreAA/CreAA and ALSHG-F/ALSHG-F interfaces.



Un-minimized models of "incompatible" homomeric CTH-CTD interfaces for CreAA and ALSHG-F proteins, based on the 1CRX coordinates (Guo *et al., Nature 389,* 40-46 (1997)). The CreAA/CreAA contact, in which residues 299, 304 and 334 are Ala, would be expected to contain a destabilizing cavity. The ALSHG-F/ALSHG-F interface consists of bulky hydrophobic CTD Met299 and Val304 residues which might occlude the site for CTD Phe334 burial (white-coloured atoms are less than 2.5 Å distant in the model). Alternatively, the CTH may reorient to avoid a steric clash, and interfere with proper positioning of catalytic Tyr324.