ABSTRACT  The use of Cre recombinase to carry out conditional mutagenesis of transgenes and insert DNA cassettes into eukaryotic chromosomes is widespread. In addition to the numerous in vivo and in vitro applications that have been reported since Cre was first shown to function in yeast and mammalian cells nearly 30 years ago, the Cre–loxP system has also played an important role in understanding the mechanism of recombination by the tyrosine recombinase family of site-specific recombinases. The simplicity of this system, requiring only a single recombinase enzyme and short recombination sequences for robust activity in a variety of contexts, has been an important factor in both cases. This review discusses advances in the Cre recombinase field that have occurred over the past 12 years since the publication of Mobile DNA II. The focus is on those recent contributions that have provided new mechanistic insights into the reaction. Also discussed are modifications of Cre and/or the loxP sequence that have led to improvements in genome engineering applications.

INTRODUCTION
The use of Cre recombinase to carry out conditional mutagenesis of transgenes and insert DNA cassettes into eukaryotic chromosomes is widespread (1–9). Indeed, a PubMed search for “cre recombinase” in early 2014 returned over 4000 articles. In addition to the numerous in vivo and in vitro applications that have been reported since Cre was first shown to function in yeast and mammalian cells nearly 30 years ago (10, 11), the Cre–loxP system has also played an important role in understanding the mechanism of recombination by the tyrosine recombinase family of site-specific recombinases(12–14). The simplicity of this system, requiring only a single recombinase enzyme and short recombination sequences for robust activity in a variety of contexts (15), has been an important factor in both cases. Cre has also been used in experiments designed to understand the functions of other recombination systems (16–18).

In its physiological role for bacteriophage P1, Cre functions as an intramolecular resolvase. The lysogenic state of P1 normally does not involve integration into the host chromosome (19). Instead, the phage exists as a unit copy episome where Cre ensures faithful segregation by converting P1 chromosome dimers to monomers before cell division (20, 21). The reaction is carried out between 34-bp recombination sites called loxP, which are in a directly repeated orientation in the physiological substrates. Hence, Cre functions as a simple version of the Xer system, which performs a similar role for the bacterial chromosome (22).

The site-specific recombination reaction catalyzed by Cre is shown schematically in Fig. 1. Cre exists as a monomer in solution, even at high concentrations (23). However, the enzyme binds cooperatively and with high affinity to the loxP DNA sequence, resulting in a dimer of Cre subunits bound to each recombining site (24–27). Two Cre-bound loxP sites associate to form a synaptic complex (23, 28, 29), within which strand exchange is catalyzed using a mechanism thought to be shared by all of the tyrosine recombinases (13, 14).

The basic elements of the mechanism shown in Fig. 1 were generally accepted when Mobile DNA II was
published in 2002 (30). However, several important questions remained unanswered, some aspects of the mechanism had been challenged, and new questions would soon be raised. In particular, it was not clear if a specific synaptic complex was favored to initiate recombination and what the basis for this bias might be. The energetics of synapsis were largely unexplored. The roles and importance of many residues in the active site of Cre and of other tyrosine recombinases were also poorly understood.

In addition to these mechanistic questions, the field of researchers seeking to improve the use of Cre in transgenic mice and other organisms had their own sets of questions and goals. Among these were: Can the loxP site be modified to affect the directionality of the reaction and favor integration? Can Cre variants be engineered to recognize and recombine very different loxP sequences? Can more highly regulated versions of Cre be engineered?

In this review, I discuss advances in the Cre recombinase field that have occurred over the past 12 years, since the publication of Mobile DNA II. Readers are referred to earlier reviews for discussions relating to fundamental structural and mechanistic aspects of the Cre–loxP system (12, 13). Here, the focus will be on those recent contributions that have provided new mechanistic insights into the reaction, including modifications of Cre (Table 1) and/or loxP (Table 2) that have led or may lead to useful improvements in genome engineering applications. Progress in strategies of regulating Cre expression, where the focus has been on promoter selection and viral mechanisms of gene delivery, are not discussed because they are less related to the mechanistic aspects of Cre–loxP recombination.
WORKING WITH Cre

Part of the reason that Cre has proven to be a useful model system for biochemical and structural studies is the relative ease of expression and purification of the enzyme. Two useful expression and purification strategies have been described. One involves purification of an intein–Cre fusion on chitin beads followed by cleavage and release of Cre from the intein-bound resin (31). The second involves purification of native Cre using ion exchange and hydroxyapatite chromatography, resulting in a single band on silver-stained sodium dodecyl sulfate–polyacrylamide gel electrophoresis (32). Both are rapid, with milligram quantities of pure enzyme per liter of bacterial culture. Purified Cre can be highly concentrated under appropriate conditions and can be stored at −80°C for extended periods (32). A variety of affinity tags have also been employed to facilitate purification of Cre fusions.

Assays of Cre activity are straightforward. Cre binds with high affinity to loxP sites, with apparent Kd<1 nM in standard electrophoretic mobility shift assays with loxP-containing restriction fragments or PCR products (26). In vitro recombination assays can be carried out quantitatively in a number of reaction formats, with product formation occurring within minutes. Rapid quantitative evaluation of the activity of Cre mutants in Escherichia coli has also been described, using a single-copy reporter (33).

For in vivo experiments making use of Cre to perform deletions, inversions, and exchanges of loxP-flanked DNA elements, many resources are available in the form of plasmids, strains, and mice expressing Cre in various

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Selected loxP variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>loxP</td>
<td>Sequence</td>
</tr>
<tr>
<td>loxP</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxS1</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxS2</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxS11</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxFAS</td>
<td>AcAATACGTGTATA-accctt-CATACGAACTTGAT</td>
</tr>
<tr>
<td>lox2272</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>lox517</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>lox66</td>
<td>tacgtTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>lox71</td>
<td>ATACGTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>lox72</td>
<td>tacgtTCTTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxH</td>
<td>ATATCTCTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxM</td>
<td>ATATCTCTGATATGCTATACGAACTTGAT</td>
</tr>
<tr>
<td>loxLTR</td>
<td>AcAATACGTGTATA-accctt-CATACGAACTTGAT</td>
</tr>
</tbody>
</table>
tissue-specific and inducible forms. The reader is referred to several recent reviews for application-specific information (34–38).

**BINDING OF Cre TO loxP**

Cre is a two-domain protein, with a 130-residue N-terminal domain (NTD) closely linked to a 211-residue C-terminal domain (CTD; Fig. 2A). The 34-bp loxP site contains two 14-bp recombinase-binding elements (RBEs) that are arranged as nearly perfect inverted repeats on opposite sides of an asymmetric 6-bp crossover region (Fig. 2B). When Cre binds to an RBE in the loxP site, the two protein domains encircle the DNA duplex, forming a “C-shaped clamp” that forms extensive minor and major groove contacts. Two key binding elements are the αJ helix from the CTD that binds in the major groove near the center of the loxP half-site and the αB/αD helices from the NTD that straddle the major groove near the start of the crossover region (Fig. 2B). Cre binds to the first RBE with nanomolar affinity and the second Cre subunit then binds with subnanomolar affinity (26, 27). The basis for cooperative binding was evident from the first Cre–DNA crystal structure, where extensive contacts between the two Cre subunits bound to loxP were observed (39) (Fig. 2B).

The Cre–DNA interface is now well-documented in structural terms from several high resolution Cre–DNA crystals structures (23, 33, 39–43) and there is good general agreement between the structural models and both loxP mutagenesis studies and the DNA-binding properties of Cre mutants (44–46). The Cre–loxP interface is complex, with numerous water-mediated interactions but relatively few direct contacts between protein side chains and bases in the major groove. Developing a simple scheme for understanding Cre–loxP binding specificity has therefore been challenging. The most conspicuous polar interactions are made by Arg259, which forms two hydrogen bonds to the C·G base-pairs at positions ±10 (Fig. 3), and Gln90, which hydrogen-bonds to the A·T base-pairs at positions ±10.

![FIGURE 2](https://doi.org/10.1128/microbiolspec.MDNA3-0014-2014.f2)

**FIGURE 2**  Cre binding to loxP. (A) Cre bound to a loxP half-site. The two domains of Cre form a “C-shaped clamp” that wraps around the DNA duplex. Helices B and D from the N-terminal domain (NTD) straddle the major groove from one face and helix J sits in the major groove from the opposite face. An asterisk marks the point of insertion of several basic residue side chains (not shown) into the minor groove at the end of the site. (B) Cre bound to the loxP site. The loxP site is composed of two 14-bp recombinase-binding elements (RBEs) arranged as inverted repeats around an asymmetric 6-bp crossover region. The RBEs differ only in the base pairs adjacent to the crossover region. Arrows indicate the cleavage sites. Two primary sources of Cre–Cre interactions on the loxP site are evident: the two NTDs interact via helices A and E and the two C-terminal domains interact where helix-N from one subunit is buried in a hydrophobic pocket of the adjacent subunit. The scissile phosphates are drawn as yellow spheres in (A) and (B) and Tyr324 as yellow sticks in (A). doi:10.1128/microbiolspec.MDNA3-0014-2014.f2
base-pairs at positions ±5. Of the two, Arg259 plays crucial roles in affinity and specificity, whereas Gln90 plays more minor roles (44).

An example of the complex nature of the Cre–DNA interface was illustrated in a study by Rüfer & Sauer, who changed the sequence of the T·A base-pairs at positions ±11 and ±12 near the centers of the RBEs (47). Cre makes van der Waals contacts to the 5-methyl groups of these base-pairs through residues near the amino-terminus of helix αJ (Fig. 3). Recombination decreased by a factor of 10⁵ as a result of the substitutions, indicating that substantial disruption of the network of Cre–DNA interactions in this region of the major groove had occurred.

A selection scheme was used to identify changes in Cre that would restore activity, leading to the identification of Glu262 as a key residue that must be altered to permit recognition of the mutant loxP site. Curiously, Glu262 is located on the αJ helix, but some distance away from loxP residues 11 and 12 (Fig. 3). This residue makes a water-mediated interaction to the C9 base and hydrogen-bonds directly to the phosphate backbone, requiring that a proton be shared between the two acidic groups. Substitution of Glu262 by glutamine or several alternative residues allows Cre to bind and recombine the altered loxP sites, presumably by allowing more flexibility in the positioning of αJ in the major groove. The restored function comes at the expense of a significant relaxation in binding specificity at the 11 and 12 positions, as the Cre mutants also recombine the wild-type loxP site with high efficiency (47).

A related set of experiments was carried out by Baldwin and colleagues, who investigated in vitro recombination of a loxP site containing symmetric C→A substitutions at positions ±10 (48). These substitutions abolish the key hydrogen-bonding interaction formed by Arg259, leading to loss of in vivo recombination (45). Cre was shown to bind and recombine this substrate weakly in vitro and a Cre–loxP mutant crystal structure indicated that Glu262 and Glu266 were tethering the displaced Arg259 side chain via salt bridges. Interestingly, a Cre mutant where both glutamates were replaced with glutamine not only bound more efficiently to the loxP mutant, but showed an increase rate of cleavage (48). Thus, Cre–DNA interactions far from the loxP cleavage sites can have significant effects on catalysis. A subsequent study on alternative substitutions at loxP position 10 resulted in even more complex behavior, indicating that the effects of Cre–loxP recognition can be manifested at multiple stages in the reaction pathway (49).
Most Cre–DNA binding studies have relied on electrophoretic mobility shift assay (EMSA)-based experiments (25, 26). While these assays perform well in this system, it is difficult to draw conclusions about the effects of divalent ions, polyamines, etc. on binding affinity and cooperativity. Sauer and colleagues have studied the Cre–loxP interaction using surface plasmon resonance (SPR) and found a substantial increase in binding cooperativity when 10 mM MgCl₂ or 5 mM spermidine is present (27). A much smaller effect had been previously observed from EMSAs (26). The increase in binding cooperativity was suggested to result from a change in DNA conformation and/or bending, since Mg²⁺ and spermidine are known to facilitate such changes. Bending of the loxP site is expected to promote interactions between Cre subunits bound to the RBEs (Fig. 2B) (40, 50).

The SPR study raises an important issue regarding the available structural models for understanding Cre–loxP binding cooperativity. Most binding experiments monitor association of the first, then the second Cre subunits binding to isolated loxP sites. The sites are either maintained at very low concentration (EMSA) or are tethered to the surface of a chip (SPR) under conditions where synaptic complex formation is disfavored. On the other hand, all published structural models of Cre–DNA complexes are of tetrameric assemblies in precleavage, cleaved, or Holliday junction states of the reaction. The interface formed between adjacent Cre subunits is facilitated by the bending that occurs in the loxP sites in these complexes and the extensive protein–protein interaction surface formed upon synapsis is therefore likely to play a role in stabilizing the observed DNA deformations. This raises the possibility that Cre-bound loxP sites may have somewhat different properties when initially formed compared with the conformations observed in tetrameric complexes. The weaker than expected binding observed in Cre–loxP cyclic permutation EMSA experiments would be consistent with this interpretation (50).

SYNOPSIS OF loxP SITES AND THE ORDER OF STRAND EXCHANGE

The recombination mechanism illustrated in Fig. 1 requires that the Cre-bound loxP sites associate in an antiparallel fashion, generating a two-fold symmetric complex within which sequential strand exchange takes place. The inherent symmetry, or “antiparallel nature” of the synaptic complex is evident in the topological outcome of experiments even when symmetrized sites are used (51). An important feature of a symmetric, antiparallel synaptic complex is that the bending direction of loxP and the location of the bend within the crossover region are identical in the two associating sites. An asymmetric, parallel arrangement of sites can also be rationalized, although such synaptic complexes would not be expected to undergo efficient strand exchange. For a parallel synaptic complex, the directions (and most likely the locations) of the DNA bends would differ in the two interacting loxP sites. If the energies associated with forming and associating loxP sites with opposite bend directions were equal, a 1:2:1 distribution of anti-TS: parallel:anti-BS synaptic complexes would be expected, where anti-TS and anti-BS refer to configurations poised to exchange the top strands or bottom strands, respectively (i.e. TS-Synaptic and BS-Synaptic in Fig. 1).

The question of whether a particular antiparallel synaptic complex forms to begin recombination is closely related to the question of whether the top or bottom strands of loxP are preferentially exchanged first during recombination. The order of strand exchange question has been particularly confusing in the Cre-loxP system because of conflicting reports in the literature. Early work by Hoess and colleagues demonstrated a clear preference for bottom-strand exchange based on isolation and analysis of the HJ intermediate (52). Later reports concluded that the top strands were exchanged first (42, 53), but one of these claims was reversed when it was discovered that the use of a Cre H289A active site mutant was responsible for a change in directionality of the enzyme (54). Biochemical studies by the Sadowski laboratory provided evidence in support of the original findings of Hoess and colleagues and identified those sequence elements in the loxP site that are primarily responsible for bottom strand exchange preference (55, 56).

The question of why Cre displays this bias in the reaction pathway was further addressed using 5′-bridging phosphorothioate-containing DNA substrates, a tool that had previously been used to study the role of homology in the λ-integrate system (57). Cre will cleave substrates where the 5′-bridging oxygen of the scissile phosphate has been replaced by sulfur, but the subsequent ligation reaction is sufficiently slow that the reaction can be considered irreversible. Combined with fluorescence resonance energy transfer (FRET) experiments with labeled loxP sites and inactive Cre mutants, it was shown that preferential formation of the bottom strand synaptic complex is responsible for establishing the order of strand exchange in the Cre system (58). As discussed below, single molecule FRET (smFRET) experiments later provided additional evidence in support of this idea (59).
An explanation for the bias in Cre-loxP synapsis was provided by a synaptic complex crystal structure containing wild-type loxP sites and the inactive Cre K201A mutant (23). In that structure, the bend in loxP compresses the minor groove at the start of the crossover region in the left half-site (Fig. 1, BS-Synapase). This deformation is more readily accommodated at the T-G step in the left half-site of loxP than it is at the C-A step in the right half-site, resulting in a synaptic complex where the left half-sites are preferentially kinked. Consequently, the Cre subunits bound to the right half-sites are activated for cleavage of the loxP bottom strands, which are geometrically positioned for exchange to generate the HJ intermediate (12). These reaction steps are illustrated in the top half of Fig. 1.

**ENERGETICS OF SYNAPSIS**

The ability of Cre to carry out efficient recombination in a variety of cellular and topological contexts is likely related to its ability to form very stable synaptic complexes. The energetics of this process have been studied using sedimentation equilibrium ultracentrifugation (23). Using either wild-type Cre or the inactive K201A mutant, Cre-bound loxP sites associate with $K_d = 10 \text{ nM}$, with only minor effects from divalent ions. An interesting observation from these studies was that synapsis is pH-dependent, with affinity dropping off sharply as the pH is increased above 8.5. This finding explained why Cre-loxP synaptic complexes could be identified on native gels, but quantitative experiments using electrophoresis titrations gave inconsistent results. Use of a modified electrophoresis buffer at pH 7.5 led to the development of a native gel electrophoresis assay that reproduced the findings from analytical ultracentrifugation experiments.

Similar synapsis experiments using symmetrized loxP sites provided additional insight into the preference for formation of a BS synaptic complex. The loxS1 site contains two right half-sites of loxP (Table 2) and a Cre-loxS1 crystal structure showed that bending occurs by negative roll at the C-A step, with resulting compression of the minor groove at GC-rich positions ±4 and ±3 (40). This Cre-bound site associates with $K_d \sim 200 \text{ nM}$, about 20-fold weaker than wild-type loxP. The loxS2 site contains two left half-sites (Table 2) and most likely bends with negative roll at the T-G step as observed in the synaptic complex structure containing wild-type loxP. Compression of the minor groove involves A-T base pairs in this case. The Cre-bound loxS2 site associates with $K_d \sim 5 \text{ nM}$, an affinity that is two-fold higher than for loxP. Hence, the bias towards the BS synaptic complex can be viewed most simply as an energetic preference for bending in the left, AT-rich loxP half-site.

The Cre K201A mutant allowed for direct analysis of synapsis in the absence of cleavage and strand exchange. The tight association measured for Cre K201A-loxP synapsis is close to that predicted from kinetic modeling studies using wild-type Cre (26), suggesting that loss of Lys201 does not strongly bias the results. As discussed below, Lys201 plays an essential role in catalysis by activating the 5′-hydroxyl leaving group during phosphoryl transfer (58, 60). In general, however, synapsis of loxP sites is coupled to other functional elements in the Cre–DNA tetramer. For example, The Cre Y324F and R173K active site mutants are both inactive for strand cleavage (46, 61) but they are also both partially defective in synapsis (23). Arg173 is involved in bending the loxP site and Tyr324 plays a role in structural organization of an important region of the CTD, perhaps explaining in part why these residues are required for efficient synapsis. An important consequence of these findings for Cre applications is that use of the Cre Y324F mutant as a negative control will not only eliminate the cleavage and recombination functions of Cre, but will also diminish any structural effects imposed by Cre-mediated synapsis of loxP sites in the target substrates.

Crystal structures of synaptic complexes have provided a detailed molecular understanding of the interactions responsible for stabilizing this intermediate (23, 40, 53). An extensive network of contacts between Cre subunits is formed, burying a substantial amount of solvent-accessible surface (Fig. 4). A number of synapsis-deficient Cre mutants have been identified (23, 29, 62), with the same complexity in interpretation as observed for the Cre–DNA interface. Some synapsis-deficient mutants are readily interpretable (e.g., Cre A36V), whereas others are involved more indirectly (e.g., Y324F).

Levine and coworkers have devised an experimental scheme to monitor synapsis and recombination of Cre-loxP complexes in real time using FRET (63). Using both intermolecular and intramolecular substrates containing donor/acceptor fluorophores, kinetic constants for a simplified model of the recombination reaction could be extracted. In general, the results are consistent with those obtained using gel-based analysis of reaction kinetics (26). An additional component of this study was the use of Cre-mediated synapsis/recombination kinetics as a tool to monitor the probability of DNA loop closure.
ALTERNATIVE MODELS OF THE SYNAPTIC COMPLEX

Most experiments and structural models involving Cre–loxP synaptic complexes have been interpreted in the context of the initial tetrameric assembly that forms immediately before cleavage and strand exchange to form the HJ intermediate. However, the product of top strand cleavage and strand exchange that is formed during resolution of the HJ intermediate is also a synaptic complex (i.e., TS-Synapse in Fig. 1). Kinetic models of the Cre reaction and single molecule experiments have both indicated that the product complex is exceptionally stable (26, 59), implying that synaptic complexes formed by wild-type Cre are likely to be dominated by product, rather than substrate loxP sites. The similar synapsis $K_d$ values obtained for wild-type Cre (presumed stable product synapse) and the K201A mutant (stable substrate synapse) suggest that the two synaptic complexes have similar dissociation rates, an idea that was supported by single-molecule studies (64).

There is currently no structural model that has been identified as a product synaptic complex and consequently it is not yet clear what type of DNA-bending is accommodated by the loxP sites at the end of the reaction. One formal possibility is that loxP is bent in the right half-site, in a manner similar to that observed in the Cre–loxS1 synaptic complex structure. As it is clear from centrifugation and electrophoresis experiments that synaptic complexes containing a sharp bend in the right half-site are not particularly stable, this model would be difficult to reconcile with the established stability of the product complex.

An alternative explanation may be found in a unique synaptic complex structure containing nonbridging phosphorothioate substitutions in the loxP sites (53). In this structure, the loxP sites are in an antiparallel top-strand cleavage configuration, but bending is not localized to the right half-site. Instead, a positive roll is distributed across the left half-site where the minor groove is opened towards the synapsed partner (shown as TS-synapse in Fig. 1). This configuration is a strong candidate for the product synaptic complex, where an energetically disfavored right half-site bend is replaced by an alternative left half-site bend that uses positive, rather than negative, roll to bend the site. Further investigations will be required to verify this assignment, which implies a strong structural asymmetry in the overall Cre–loxP reaction pathway.

These findings raise some intriguing questions about the Cre reaction. For example, if the product synaptic complex is so stable, why does it not form preferentially from Cre-bound loxP sites and either run the reaction in reverse or hold the sites together for an extended period in an unproductive state? A kinetic barrier to forming

![Figure 4](https://doi.org/10.1128/microbiolspec.MDNA3-0014-2014.f4)
the product complex directly (i.e., without strand exchange) could be part of the answer, but the molecular basis for such a barrier is currently not clear.

An additional question relating to the synaptic complex involves the degree to which the Cre-bound loxP sites are co-planar. Topological studies of the integration and excision reactions catalyzed by λ-integrase suggested that the active recombination complex has an intrinsic chirality (65). Related experiments with Flp and Cre recombinase were interpreted to imply that this chirality may be shared by all λ-integrase family members. The simplest explanation for a chirality consistent with this proposal would be an out of plane rotation of the two loxP sites during recombination, resulting in a right-handed crossing of the sites (65, 66). There are currently no structural models to support this idea because all tyrosine recombinase crystal structures of synaptic complexes and HJ intermediates feature core half-sites that are nearly co-planar.

**REQUIREMENTS FOR loxP HOMOLOGY**

Efficient Cre–loxP recombination requires that the central six base-pairs comprising the crossover region of loxP be identical in the two recombining sites (67). Even single base-pair heterologies can severely reduce recombination, as demonstrated in studies where the loxP spacer region has been systematically modified (68, 69). The requirement for identical crossover sequences was originally interpreted in the context of branch migration of the Holliday intermediate, but structural models of Cre–loxP HJs have indicated that branch migration is unlikely to play a role in this system (42, 43, 70). Instead, sequence identity is required for efficient strand swapping and ligation in both the HJ-forming and HJ-resolving steps in the pathway (71). In molecular terms, formation of canonical base-pairs during strand exchange is thought to be required to establish a geometry that promotes efficient ligation and formation of stable intermediates or products. It is worth noting, however, that crossover sequence identity is not required for all tyrosine recombinases (72).

An important implication of loxP homology requirements is that sites with distinct crossover sequences can coexist in DNA substrates without efficient deletion or inversion of the intervening sequences by Cre. This is the basis of the recombinase-mediated cassette exchange reaction, which in its simplest form uses pairs of sites containing incompatible crossover sequences (discussed below). An example of an incompatible site is loxFAS (Table 2), which performs nearly as well as the loxP site (73). Systematic studies of the loxP crossover region and the flanking base pairs at the RBE borders indicate that changes to this central region of loxP all result in some loss of recombination efficiency (68, 69).

It is also now clear that the requirement for sequence identity between the crossovers of recombining sites is not absolute. Under some in vivo conditions, Cre will carry out recombination between incompatible sites and will carry out deletions for sites arranged as inverted repeats at a measurable frequency (74). In some cases, the sequences of the recombinant products cannot be explained by Cre-mediated strand exchange of heterologous crossovers, suggesting that cellular factors may play a role in a subset of these noncanonical rearrangements (75).

**CATALYSIS OF CLEAVAGE AND STRAND EXCHANGE**

The Cre active site contains seven residues that are conserved to varying degrees among the tyrosine recombinases (Fig. 5). The structure of a vanadate transition state mimic of the phosphoryl transfer reaction, combined with saturation mutagenesis of each active site residue, has provided insight into the catalytic steps involved in Cre–loxP recombination (33, 76). When combined with high-resolution structures of the precleavage (23, 40) and covalent intermediates (39) of the reaction, the positions and hydrogen-bonding patterns of the catalytic residues can be followed through the catalytic steps. These interaction networks are illustrated in Fig. 5.

The Cre active site residues can be divided into three sets. In the first set, Tyr324 and Lys201 are highly conserved, play unique catalytic roles, and no substitutions for either residue results in significant activity. Tyr324 is the nucleophile in the cleavage reaction and Lys201 is the general acid primarily responsible for activating the O5’ leaving group.

In the second set, Arg173 and Arg292 are also highly conserved and play important roles in stabilizing the negative charge of the phosphate in the transition states of the cleavage and ligation reactions. Both residues make double hydrogen bonds to oxygen atoms using their side-chain guanidino groups and likely play dual roles in catalysis. Arg173 hydrogen bonds to O1P and to O5’ of the scissile phosphate and biochemical evidence suggest that the O5’ contact contributes to leaving group activation during cleavage (77). The R173K mutant is inactive in vitro and only weakly active in vivo, but the R173H mutant shows strong activity both in vitro and
in vivo, indicating that transition state stabilization at O1P is likely to be the primary catalytic role. Arg292 hydrogen bonds to O2P and to the Tyr324 hydroxyl oxygen in the transition state. This suggests that a secondary role of Arg292 may involve stabilizing the phenolate oxygen that is the leaving group during ligation.

In the third group of residues, His289 and Glu276 are moderately conserved, but Trp315 is more often histidine in the tyrosine recombinase family. Like Cre, the Flp recombinase uses a tryptophan in this position of the active site and biochemical studies have indicated that the primary role of this residue is structural, rather than catalytic (78, 79). Indeed, the Cre W315F mutant has higher activity than W315H, indicating a similar functional role in the Cre system (33).

Neither His289 nor Glu176 are required for catalysis, with several alternative residues showing moderate levels of recombination activity in each case. His289 can hydrogen-bond to O1P and to the Tyr324 hydroxyl oxygen, suggesting a dual stabilization role similar to, but less important than, that seen for Arg292. Glu176 plays a structural role, buttressing the positions of the Arg173 side chain and the main chains of both Arg173 and Lys201.

The role of electrostatic stabilization in Cre catalysis has been probed using methylphosphonate (MeP)-containing loxP substrates, where one of the nonbridging oxygen atoms is replaced by a methyl group (80). Both wild-type Cre and Cre R292A will efficiently cleave loxP sites containing a scissile MeP, in agreement with

**FIGURE 5** The Cre active site. Hydrogen-bonding interactions between the seven active site residues and the scissile phosphate are shown for the precleavage intermediate, transition state mimic, and covalent intermediate, based on crystal structures of the corresponding complexes. Glu176 forms additional hydrogen bonds to the backbone amides of both Arg173 and Lys201 in each case, which are not shown. The 5'-hydroxyl group is not present in the covalent intermediate structure, but hydrogen bonding to Lys201/Arg173 is likely (not shown). The cleavage reaction proceeds from left to right, whereas the ligation reaction proceeds right to left in this scheme. Note that O1 and O2 of the transition state are labeled in the opposite sense to that used in Gibb et al. (33). doi:10.1128/Microbiolspec.MDNA3-0014-2014.f5
a role of electrostatic stabilization by this residue. The Cre active site also protects the MeP-containing active site from both phosphodiester hydrolysis and from hydrolysis of the 3'-phosphotyrosyl linkage, a property that distinguishes Cre from the related Flp recombination and TopIB systems. Subsequent studies revealed that Arg173 is also not required for cleavage of MeP-containing sites and even the R173A, R292A double mutant is competent for cleavage (81). In the ligation reaction, R292A is competent, R173A is weakly active, and the double mutant is inactive, providing additional support for the role of Arg173 in the activation of the 5'-hydroxyl oxygen.

A number of Cre active site substitutions lead to accumulation of the covalent phosphotyrosyl intermediate (e.g., R292K and H289Q) or the HJ intermediate (e.g., W315F and E176A) of the reaction (33). In many cases, one can rationalize these properties in terms of the transition state structure. For example, the observation that Cre R292K generates large amounts of covalent intermediate can be explained by the loss of hydrogen-bonding to the Tyr324 hydroxyl oxygen by this mutant and a consequent lowering of leaving group activation during the ligation step. This implies that Arg292 has a more important role in leaving group activation during ligation than in activation of the nucleophile during cleavage. The redundant roles and in some cases subtle mutagenesis effects of the less conserved Cre active site residues may reflect the evolutionary fine-tuning that has taken place to strike the correct balance between cleavage and ligation rates. Such a balance is presumably required to efficiently catalyze the sequential strand exchange reactions involved in site-specific recombination.

**INSIGHTS FROM SINGLE MOLECULE EXPERIMENTS**

Experimental approaches that allow monitoring of individual Cre-loxP synaptic complexes undergoing recombination are relatively new to the field, but have already led to several interesting findings. In the first such study, Fan employed tethered particle motion (TPM) to monitor Cre-mediated synapsis and recombination of loxP sites (64). In this technique, a DNA substrate containing loxP sites is fixed on one end to a surface and on the other end to a bead. The amplitude of brownian motion of the bead is monitored as a function of time and related to the extension of the DNA. Synapsis can be inferred from a shortening of the DNA length and recombination from persistence of the shortened DNA upon treatment with sodium dodecyl sulfate, when the loxP sites are arranged as direct repeats. Similar experiments using loxP sites arranged as inverted repeats can allow identification of HJ intermediates.

The TPM experiments have indicated the presence of two different rates of decay of Cre-loxP synaptic complexes: a faster rate that is associated with dissociation of the complex without recombination (~3×10⁻²/s) and a slower rate that is associated with dissociation of product loxP sites (~2×10⁻³/s). The slower rate reflects both the recombination process and the stability of the product complex. Interestingly, the K201A mutant also displayed this biphasic dissociation behavior with similar rate constants, despite its lack of ability to carry out recombination.

More recently, Pinkney et al. have combined tethered fluorophore motion (TFM) with single-molecule FRET (smFRET) to provide both a global view of the state of the DNA substrate and a local readout of the recombining loxP sites from donor and acceptor groups attached at strategic positions (59). The experiments supported previous work showing that Cre initiates recombination by exchanging the bottom strands and that the strand bias is caused by preferential formation of the anti-BS synaptic complex. These experiments also confirmed findings from the TPM study (64) that a significant fraction of synapsed loxP sites dissociate without undergoing recombination.

The additional information provided by smFRET measurements has led to a more detailed picture of the events unfolding during recombination compared with that obtained from TPM or TFM data alone. Two distinct FRET efficiency distributions were observed for synapsis of inactive loxP sites by wild-type Cre (those that dissociate without undergoing recombination), indicating that two distinct synaptic structures are formed. Interestingly, only the lower (major population) efficiency distribution is observed for the inactive K201A mutant, and only the higher (minor population) efficiency is observed for the synopsis-deficient A36V mutant. Modeling of the loosely tethered fluorophores onto templates based on Cre–DNA crystal structures suggests that the higher efficiency value corresponds to the distance expected from structural models of the synaptic complex, whereas the lower efficiency results in a distance that is ~10 Å longer than expected (59). The actual relationships between the observed FRET efficiencies and the underlying synaptic structures is not yet clear, but it does appear that there is still much to be learned about the structural and biochemical natures of the complexes that form on the recombination pathway.
DIRECTED EVOLUTION OF Cre TO RECOGNIZE ALTERNATIVE SITES

There has been considerable interest in the question of whether Cre can be modified to recognize and recombine alternative sequences, including naturally occurring sequences in mammalian and viral genomes. Buchholz and Stewart addressed this question by designing experiments to efficiently evolve Cre to recombine “loxH” sites through multiple cycles of PCR-based mutagenesis and selection in E. coli (82). The loxH site is from human chromosome 22 and differs from loxP by four changes in the half-site RBEs and a different crossover sequence (Table 2). The resulting evolved “Fre” recombinases were able to delete loxH-flanked gene segments in E. coli and in mammalian cells, with diminished activity on loxP-containing substrates.

Santoro and Schulz also used directed evolution to generate Cre variants with altered specificity (83). In this case, E. coli cells were sorted by fluorescence-activated cell sorting according to the recombination activities of their Cre-expressing plasmids, allowing both positive selection for recombination of the altered loxP sequence and negative selection against recombination of wild-type loxP sites. A series of modified loxP sites was chosen for these experiments based on disruption of recombination by wild-type Cre. This approach led to a Cre variant (called C2(α)#4) that efficiently recombines loxM7 sites containing three changes in each RBE (Table 2), but does not efficiently recombine loxP sites. For some altered loxP targets, active Cre variants could be identified by positive selection, but not by both positive and negative selection. This suggests that it may be difficult to change the specificity of Cre for some sequences without retention of significant activity on loxP (83).

INCREASING SPECIFICITY BY ENGINEERING Cre—Cre INTERFACES

The success in generating modified recombinases that recombine altered loxP sites led to the question of whether different Cre variants could function together on asymmetric sites where the RBEs are different. This idea was tested with sites containing RBE pairs from loxP and M7 and mixtures of wild-type Cre and the C2 variant (Tables 1 and 2) (86). Whereas neither recombinase alone will efficiently recombine hybrid loxP-M7 sites, the recombinase mixture is effective. These results indicate that dual specificity recombinase pairs such as are found in the Xer system could be useful in Cre applications. The ability to independently manipulate the binding properties of the two recombinases should be an attractive alternative to evolving a single enzyme that simultaneously recognizes both half-sites, as was done in the Tre example discussed in the previous section.
A complication with the two-recombinase approach is introduced by the high level of cooperative binding in the Cre-\textit{loxP} system. Cre binding to a “good” half-site can facilitate binding to a poor half-site through the intersubunit contacts responsible for cooperativity. Indeed, wild-type Cre will recombine the hybrid \textit{loxP}-M7 sites when present at elevated concentrations (86). When two high-affinity Cre variants are present, four distinct types of off-target sequences could in principle be recognized.

To address this issue, Baldwin and colleagues have engineered the interface formed between adjacent Cre subunits (87). They focused on the contacts formed between the C-terminal helix-N (CTH) of Cre and the corresponding binding pocket in the CTD and identified an alternative set of interacting residues that function efficiently with one another, but not with a wild-type Cre partner. Using the hybrid \textit{loxP}-M7 sites as a test system, Cre and the C2 variant were then modified so that a favorable CTH-CTD interaction would form only between Cre and C2 when bound to their corresponding RBEs. The result was an improvement in specificity towards the hybrid sites when both modified recombinases were present, illustrating that this approach may provide a useful contribution to the engineering of custom recombination systems.

Eroshenko and Church have recently identified Cre mutants that have improved specificity for \textit{loxP} sites by modifying intersubunit interactions (88). The authors provided a strong argument that a reduction in binding cooperativity should result in improved accuracy for binding to \textit{loxP} sites relative to off-target sites. The R32V and R32M substitutions were predicted to disrupt an intersubunit salt-bridge between helices (aA and uC) in the NTDs of Cre subunits bound to \textit{loxP}, thereby reducing cooperativity of binding. Although both mutants recombine less efficiently than wild-type Cre on \textit{loxP} sites, the difference in efficiency is even greater for off-target sites and for pseudo-\textit{loxP} sites in the \textit{E. coli} genome. Hence, an increase in recombination fidelity can be achieved at the expense of some recombination efficiency.

Arg32 is involved in a complex network of polar contacts in the interface between Cre subunits, involving Arg-Arg stacking, contributions from three acidic side chains and several tightly associated water molecules. The same network of interactions exists between Cre subunits that stabilize the synaptic complex, suggesting that the R32V and R32M mutants are likely to be deficient in this step of the pathway. The close structural relationship between binding cooperativity and synapsis in the Cre-\textit{loxP} system raises the intriguing question of whether mutants might be identified that primarily affect the former, with minimal effects on the latter. Such mutants might be valuable tools, following on the Eroshenko and Church results.

**TURNING Cre INTO AN INTEGRASE**

Cre readily catalyzes intermolecular recombination between \textit{loxP} sites, but the reverse intramolecular excision reaction is so efficient that it has been challenging to implement Cre as an effective integrase. This is not surprising, given that Cre normally functions as a resolvase in bacteriophage P1 biology. Two approaches have been successful in carrying out Cre-mediated integrations of DNA segments into chromosomes. The first is referred to as the left element (LE)/right element (RE) strategy, where the recombining sites each have one wild-type half-site and one mutant half-site (Fig. 6A) (89). The most widely used example is the \textit{lox66/lox71} combination, where the sites contain 5-bp changes in the left and right RBEs of \textit{loxP}, respectively (Table 2). The products of \textit{lox66 × lox71} recombination are \textit{loxP} and \textit{lox72} sites, where both half-sites of \textit{lox72} contain the RBE mutations. The diminished ability of \textit{lox72} to undergo recombination provides stability to the product of integration, and so some directionality to the reaction (89). Alternative LE/RE \textit{loxP} sites have been reported more recently, with much higher integration efficiencies (90). The LE/RE approach has also been used to bias Cre-mediated inversion reactions \textit{in vitro} (91, 92).

The second strategy is the recombinase-mediated cassette exchange (RMCE), which employs two incompatible \textit{loxP} variants flanking DNA cassettes on the donor and acceptor molecules (93–96). The sites differ in their crossover sequences and therefore will not recombine with one another with high efficiency. An initial recombination event between compatible sites results in integration of the donor into the acceptor and subsequent recombination between the second set of compatible sites results in a net exchange of DNA cassettes between the two substrates (Fig. 6B). A recent review of the RMCE reaction discusses the issues involved in optimizing the process, including a comparison of Cre versus Flp recombinases as RMCE enzymes (97).

Considerable effort has been devoted to testing which alternative crossover sequence to use for RMCE, since all crossover variants underperform the native \textit{loxP} sequence. The use of \textit{lox66/lox71} for one pair and \textit{lox2272} sites for the second pair has been reported to result in particularly stable products, even allowing the integration of a normally unstable Cre expression cassette in
mouse embryonic stem cells (98). In this case, the sites in the product cassette have incompatible crossovers, plus one of the sites (lox72) is poorly functional. On the other hand, RMCE using \( \text{loxP} \) sites as one pair and \( \text{loxFAS} \) sites as the second pair has a very high efficiency of integration (9). The \( \text{loxFAS} \) site contains a relatively efficient \( \text{loxP} \) crossover alternative (Table 2).

A number of variations of Cre-mediated RMCE have also been described. Dual RMCE refers to the use of two different recombinases, each with its own pair of sites (99, 100). Cre, Flp, and the \( \phi \text{C31} \) serine integrase have been used for this purpose, eliminating the need to identify incompatible crossover sites for a single recombinase that will resist cassette excision over extended time periods. Schemes for executing successive rounds of RMCE-based integration have also been reported, where each cassette exchange results in inactivation of one product site, but introduction of a new site for use in subsequent reactions (101, 102).

**RECONSTITUTION OF ACTIVE Cre FROM FRAGMENTS**

The regulation of timing and location of Cre activity in transgenic animals has played a crucial role in the design and execution of experiments where Cre is used as a genetic switch. The use of tissue-specific and ligand-activated promoters to regulate Cre expression is widespread, with a variety of Cre-expressing mice now available (34, 35, 103). However, the basal rates and stochastic nature of expression from many promoters have been associated with cell toxicity and lack of optimal regulation (104–108). An additional level of regulation has been achieved in recent years by requiring that two Cre fragments assemble to form an active enzyme.

One approach to the “split Cre” method involves generating fusions of Cre fragments to domains that can be induced by ligands to associate and promote Cre activity (Fig. 7A). The first example of this type used
FKBP fused to residues 1 to 59 of Cre and FRB fused to residues 60 to 341 (109). The small molecule rapamycin induces heterodimerization of FKBP and FRB (110), leading to reconstitution of an active Cre subunit. This method of regulation has been demonstrated to function in vitro (109) as well as in mice (111).

Residues 1 to 59 of Cre include the first two α-helices of the NTD. This region is important for both cooperative binding to loxP (Fig. 2B) and for forming a stable synaptic complex (Fig. 4). Cre A36V disrupts these interactions and as a consequence this mutant is poorly active (23, 29, 59). The Cre 60–341 fragment may resemble Cre A36V in some respects, but with an even stronger phenotype. It is therefore tempting to speculate that synapsis of loxP sites is what is primarily being regulated in this split Cre system.

A subsequent study implemented the same Cre fragments, but fused them to the self-associating coiled-coil region from yeast GCN4 and expressed the two fragments under different promoters (112). The fragments spontaneously assemble in vivo to form an active enzyme, but with a more complex level of regulation than is possible using Cre expressed from a single promoter.

In an alternative approach, overlapping fragments of Cre were identified that could assemble without the aid of dimerization domains and could be independently regulated with different promoters. The first such implementation used Cre fragments 1 to 196 and 181 to 341...
341, where the overlap is located within the Cre CTD (Fig. 7B) (113). Subsequent studies modified the system to provide different modes of regulation. In the first case, one fragment was expressed in the target cells and the other was added to the cell culture as a purified protein (114). In the second study, the fragments were fused to heterodimeric leucine zipper domains, increasing the efficiency of the original system (115).

**DESTABILIZED Cre**

The most recent innovation in the quest to regulate Cre activity involves expression of destabilized forms of Cre (116). In this study, an E. coli dihydrofolate reductase mutant was fused to Cre (DD-Cre), causing it to be rapidly degraded by the proteosome (117). Addition of trimethoprim stabilizes the dihydrofolate reductase mutant, allowing Cre to escape degradation and carry out recombination of loxP-flanked alleles. As trimethoprim can cross the blood–brain barrier, this method of regulation allows both conditional mutagenesis in the central nervous system and a much reduced basal level of Cre activity.

**Cre TRANSDUCTION**

It was demonstrated in 2001 that recombination can be achieved in mammalian cells by transduction of a purified His$_6$-NLS-Cre-F fusion protein (118). In this construct, NLS refers to the simian virus 40 nuclear localization sequence and F refers to the 12-residue membrane translocation sequence from the Kaposi fibroblast growth factor FGF-4 (119). This finding implied that one could deliver a one-time pulse of Cre to target cells and avoid the undesirable side-effects resulting from transient transfections or low levels of constitutive expression.

In studies appearing several months later, it was further reported that Cre can be effectively transduced with a variety of membrane translocation sequence-like tags, including NLS-Cre (120–122). Indeed, even unmodified Cre can be used to transduce mammalian cells with moderate efficiency when present at slightly elevated concentrations (120). Optimization of various combinations of tags resulted in identification of His$_6$-TAT-NLS-Cre as a particularly effective construct, resulting in rapid recombination at micromolar Cre concentrations (121). In this case, TAT is an 11-residue peptide from HIV-TAT that is known to promote the transduction of heterologous proteins across mammalian cell membranes (123). This idea was later applied to human embryonic stem cells and human embryonic stem progeny cells. Addition of His$_6$-TAT-NLS-Cre fusion to the culture medium at micromolar concentrations resulted in recombination in ~100% of cells containing a loxP-flanked reporter, underscoring the usefulness of this approach (124).

The demonstration that purified Cre can be delivered to growing cells in a manner that results in nearly quantitative recombination of genomic loxP sites raised a new question of whether chemically modified versions of Cre with even more complex modes of regulation could be used. Deiters and coworkers have answered this question with the development of a light-activated version of Cre (125). Using unnatural amino acid incorporation methodologies in E. coli, Tyr324 was replaced by o-nitrobenzyl tyrosine (ONBY), effectively blocking the active site nucleophile and rendering Cre inactive. Upon irradiation with 365 nm UVA light, the nitrobenzyl group was efficiently removed, resulting in active Cre and recombination of loxP sites. When ONBY-Cre was transduced into mammalian cells, Cre could not only be activated by UVA light irradiation, but spatial control of Cre activation could be demonstrated by limiting the activating light to a specific region of a monolayer of cells.

**FUTURE PROSPECTS**

It would be reasonable to conclude that we now have a good mechanistic understanding of the Cre-loxP recombination reaction and that the application of Cre-loxP technologies is well-established in many areas of the biological sciences. However, it is also clear from investigations over the past decade that we do not yet know the answers to all the questions posed at the start of this chapter. Understanding and changing the sequence specificity of Cre remains extremely challenging and a complete structural and biochemical understanding of all the steps in the recombination pathway is still lacking. The available experimental data indicate that Cre and the loxP sequence have been fine-tuned by evolution, where even single changes to nucleotides in the crossover region can profoundly affect recombination efficiency. In many cases, we still do not understand the origins of these effects. At the time that *Mobile DNA II* was being written, it would have been difficult to predict all of the advances in mechanistic understanding and improvements in regulating Cre activity that would be made in the coming years. We should therefore anticipate at least a few more surprises in the years to come.
REFERENCES


