The Legacy of Nat Sternberg: The Genesis of Cre-*lox* Technology

Michael Yarmolinsky[1] and Ronald Hoess[2]

[1]Laboratory of Biochemistry and Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892; email: myarmo@helix.nih.gov

[2]Chadds Ford, Pennsylvania 19317; email: hoessr@earthlink.net

Annu. Rev. Virol. 2015. 2:25–40

The *Annual Review of Virology* is online at virology.annualreviews.org

This article's doi: 10.1146/annurev-virology-100114-054930

Keywords

bacteriophage P1, plasmid, site-specific recombination, chromosome segregation, genetic engineering

Abstract

Cre-*lox* of bacteriophage P1 has become one of the most widely used tools for genetic engineering in eukaryotes. The origins of this tool date to more than 30 years ago when Nat L. Sternberg discovered the recombinase, Cre, and its specific locus of crossover, *lox*, while studying the maintenance of bacteriophage P1 as a stable plasmid. Recombinations mediated by Cre assist in cyclization of the DNA of infecting phage and in resolution of prophage multimers created by generalized recombination. Early in vitro work demonstrated that, although it shares similarities with the well-characterized bacteriophage λ integration, Cre-*lox* is in many ways far simpler in its requirements for carrying out recombination. These features would prove critical for its development as a powerful and versatile tool in genetic engineering. We review the history of the discovery and characterization of Cre-*lox* and touch upon the present direction of Cre-*lox* research.

INTRODUCTION

Biology has entered a new era, an era in which we can manipulate genes at will and do so even during the course of development. In the accomplishment of this technological feat, a major role has been played by site-specific recombination mediated by the recombinase, Cre, and specific DNA sites, *lox*.

Here we narrate the story of the discovery and initial characterization of Cre-*lox*. We mix science and our personal reminiscences of Nat L. Sternberg (1942–1995), the scientist whose insights and efforts made Cre-*lox* an invaluable genetic tool (**Figure 1**).

We were associated with Nat in different ways. M.Y. recruited Nat to work on bacteriophage P1, the organism that turned out to be the source of both Cre and *lox*. R.H. was recruited by Nat to study their interactions.

FROM HERESY TO RESPECTABILITY

Several historical circumstances delayed the discovery of temperate bacteriophages and made the study of the temperate bacteriophage P1 highly unlikely. The scientist-adventurer Félix d'Herelle, who, in 1917, gave the name *bacteriophage* to infectious particles responsible for killing bacteria, assumed that bacteriophages, whatever their nature, are invariably lethal. D'Herelle's assumption that all phages are virulent (as opposed to temperate) relegated to obscurity a phenomenon discovered in 1925 known now as lysogeny. Lysogeny is the hereditary power of certain bacteria to perpetuate a noninfectious entity (*Anlage* or prophage) able, on occasion, to produce bacteriophages and release them by cell lysis (1).

For a quarter of a century lysogeny was deemed artifactual, the fantasy of incompetent experimentalists whose sloppy technique resulted in bacterial cultures that carried along a viral contamination that only occasionally resulted in a productive infection, a "carrier state." André Lwoff, in his 1965 Nobel lecture, stated the situation with characteristic wit: "In 1949, the new school of American virologists, to which virology owes so much, condemned lysogeny. . . . Like those wisps of cloud that a breath of wind dispels, the problem was blown away from the temple of science and a smell of sulfur was left floating in the air. Lysogeny had become a heresy" (2).

Figure 1

Nat L. Sternberg (1942–1995). The undated photograph was provided by his widow, Ilene Harfenist Sternberg. It was the only satisfactory image of her self-effacing husband that she had.

When, in 1949, Giuseppe Bertani joined the laboratory of Salvador Luria, a leader of the new school of American virologists, and asked to study lysogeny, Luria was initially reluctant to give his blessing. Bertani attributed that reluctance to the wording of Luria's grant (3), but lysogeny's sulfurous smell must have had something to do with it, too.

Whatever his misgivings, Luria did agree to Bertani's request and obtained from Joshua and Esther Lederberg two lysogenic strains of *Escherichia coli*: K-12 (λ) and the strain Lisbonne-Carrère (Li). Bertani began studying both lysogens, but discontinued studying λ lysogeny at the urging of the Lederbergs, as Esther Lederberg had published on λ only informally (in 1950). Bertani, along with the Lederbergs, was the first to embark on the study of lysogeny on the American continent.

E. coli Li had been known to yield phages of various plaque sizes. That these might represent different phages had not been anticipated. In 1951 Bertani showed that the different types of plaques were made by the growth of phages of three immunologically distinct types. He named them P1, P2, and P3.

Understandably, Bertani concentrated his attention on P2, as it made plaques large enough to select mutant phages differing in plaque morphology. He did not entirely stop working on P1, although he did abandon P3. (Its resurrection in 1976 by Harvard Professor Nancy Kleckner was nipped in the bud by the Cambridge City Council, fearing the risk to her students of studying an organism requiring P3-level containment.) The P1*kc* from which all laboratory strains of P1 have been derived is a mutant isolated by Bertani that grows on both *E. coli* C and *E. coli* K-12. Like the original P1 (which was known to grow only on *Shigella dysenteriae*), its plaques on *E. coli* are tiny. This feature made it an uncongenial phage to study.

P1, like P3, might well have faded from view, but was saved from oblivion in 1954 by the initiative of Ed Lennox, a physicist turned biologist who had joined Luria's lab. He asked whether Bertani's phages could transduce (transfer genetic markers from a donor to a recipient host). The remarkable capacity of *Salmonella* phage P22 to transduce closely linked genetic markers had been demonstrated in 1951 by Norton Zinder and Joshua Lederberg. Lennox showed that P2 lacks that capacity, but P1 has it in spades (4). P1 is capacious enough to transduce prophages such as λ along with associated genetic markers. P1 promptly became a workhorse of genetic manipulation in bacteria and the phage itself an object of interest (5, 6). The DNA extracted from plaque-forming particles, characterized in the 1960s by Hideo Ikeda and Jun-ichi Tomizawa, was shown to be double-stranded, linear, terminally redundant, and cyclically permuted (7).

What controls the frequency with which an infecting phage reduces to a prophage and the sporadic entry of the prophage into a lytic cycle? One of the several biological insights in Jacob and Monod's landmark paper of 1961 was the recognition that the vegetative viral functions are repressed in the prophage at the level of the gene by repressor proteins, and in the same way that inducible enzymes are repressed by their cognate repressors (8). Phage mutants with thermosensitive repressors soon became the "wild type" in laboratories studying lysogeny. These repressors are responsible for the immunity of lysogens to superinfection by a phage of the identical kind.

Where in the cell are prophages located? Already in her informal 1950 report, Esther Lederberg indicated that λ prophage was somehow linked to genetic markers of its host, but the nature of the linkage was unclear. Subsequent studies of bacterial sexuality in the early 1950s clarified this linkage. Lambda prophage was mapped to a unique chromosomal site and P2 prophage to more than one, but P1 prophage resisted attempts at assigning it a chromosomal location.

How does a prophage without a fixed chromosomal address replicate? The possibility that P1 prophages are passively replicated by transient integration into the host chromosome was

considered, but was conclusively eliminated in 1968 (7). Unlikely as it seems, passive replication of a short DNA segment that transiently integrates into the *E. coli* chromosome at the terminus region was found to mimic maintenance as a plasmid (9). The responsible site-specific recombinase was probably XerC. P1 prophage is an independent replication unit: a plasmid, and, as it turned out, a plasmid of low copy number and high stability. These features imply that P1, as a prophage, possesses the wherewithal to control its copy number within narrow limits and to distribute at least one plasmid copy to each nascent daughter cell. No other bacteriophages were known at the time to possess these capabilities.

The study of these fundamental processes is greatly facilitated by the existence of a simple procedure to amplify and purify the plasmid DNA. One has only to induce viral replication and extract the DNA from harvested bacteriophage particles. P1 seemed uniquely suited for studying the basic capabilities of plasmids. The phage, however, was largely unexplored.

PARIS

Nat Sternberg and I (M.Y.) arrived in Paris in the same year, 1970. Nat came on a postdoctoral fellowship to work in the laboratory of François Gros in a newly established Centre National de la Recherche Scientifique (CNRS) laboratory of molecular biology, I to set up my own laboratory on the floor below. The institute was located on the new campus of the Pierre and Marie Curie University. I had expressed an interest in a position in Paris, some six years previously, when a postdoctoral fellow at the Pasteur Institute in the laboratory of François Jacob; my request had not been forgotten.

Nat and I had considerable similarities in our cultural and scientific backgrounds. Both of us had grown up in New York City and possessed a shared interest in phage-host relationships. Interactions between λ and *E. coli* had been the subject of my own research in the United States both before my fellowship at the Pasteur Institute (escape synthesis, i.e., derepression of host genes mediated by prophage induction) and after (integration/excision). Nat had earned his doctorate in 1969 with Sewell Champe at Purdue, studying the head proteins of the virulent T-even phages, and in Paris began a study that introduced him to the temperate phage λ.

Whereas my laboratory in Paris was spacious and tranquil, Nat found himself jockeying for workspace in a noisy, overcrowded room where sudden shortages of media were not uncommon. Undaunted, Nat embarked on an ambitious project to select and characterize *E. coli* mutants with mutations affecting the capacity to synthesize head proteins of λ phage. He periodically reported on his progress in lucid presentations that casually displayed a remarkable productivity. When I asked Nat how he acquired his capacity for work with unflagging concentration, he referred to his years as a (reluctant) student in a rabbinical academy (yeshiva). In yeshiva, he said, you learn *Sitzfleisch*, the stamina to persevere. *Sitzfleisch* may have been part of it, but Nat's extraordinary talents also resided elsewhere.

When, in 1972, Nat's fellowship was coming to a close and he was preparing to return to the United States, he told me of his dissatisfaction with the opportunities available to him. About that time I learned that Robert Weisberg, whose thriving laboratory at the National Institutes of Health (NIH) was devoted to λ, had a post to fill and was planning to choose from among candidates at an upcoming phage meeting. Confident that Weisberg could do no better than offer the post to Nat, I took the initiative to telephone him and persuade him to interview Nat. In 1972 a long-distance call from overseas had gravitas. Nat, at his own expense, made the trip for the (successful) interview. For the next four years he held the position of staff fellow in the Weisberg lab, publishing a number of papers, five of them with a colleague in the lab, Lynn Enquist, the future founding editor of this journal.

In the same four-year period my group in Paris pursued genetic studies of P1 replication and immunity. Immunity maintenance being central to prophage stability, its study seemed a natural place to begin exploring the biology of P1.

In September of 1974 I visited the United States and let it be known to certain colleagues at NIH that I would be interested in returning to the fold. Nat learned of my interest and wrote to me in October to express his interest in becoming a member of my laboratory, should I succeed in establishing one at NIH. A position for Nat in my lab became a top priority.

FREDERICK

In 1976, I (M.Y.) accepted an offer to direct the Molecular Genetics Section in a new Laboratory of Molecular Biology in Frederick, Maryland, funded by the National Cancer Institute. The laboratory was located in Fort Detrick in space liberated by the closure of facilities for biological warfare research. Dedicated lobbying by Harvard Professor Matthew Meselson was instrumental in that important decision. Some reminders remained of Fort Detrick's sinister past, but there were compensations: considerable freedom and adequate funding.

Nat Sternberg became the first member of my group. He was of immeasurable practical help in setting up the lab prior to my arrival and the arrival of Stuart Austin (from my lab in Paris) and Robert Yuan (from Werner Arber's lab in Basel).

Nat was persuaded that P1 was, as he put it, "just ripe for the picking." Nat's first contribution to P1 biology, made while still in Weisberg's laboratory, was to construct a P1 library in a λ vector, thus making P1 genes much more amenable to individual study (10). Lambda, Nat recognized, provided possibilities for DNA manipulation that were not obtainable with the small plasmid vectors with which, a few years earlier, the gene cloning revolution began. He took advantage of features of the λ DNA packaging process to dissect the P1 DNA fragments. The P1 library proved useful in studying a variety of aspects of P1 biology, including replication, immunity, incompatibility, and site-specific recombination. It is this last topic that constitutes the subject of his first publication on P1 from Frederick, in 1978.

This article (of which Nat was sole author) is brief, but information-rich, and gives a foretaste of much that followed (11). In it he defined a site-specific recombination as the explanation for an unexpected linearity of the recombination map of P1. As noted earlier, P1 phage particles contain DNA that is cyclically permuted and terminally redundant. Recombination between P1 phages was expected to produce a circular, not a linear, map, but previous reports by June Scott and by Donald and Jean Walker indicated that the map was, indeed, linear. Nat proposed that the simplest explanation for this result is that P1 DNA contains a hotspot for genetic recombination located in the region corresponding to the ends of the map.

Consistent with this hypothesis, he observed efficient recombination within a cloned small (6-kb) fragment of P1 DNA that, by marker rescue experiments, spans the ends of the genetic map. The recombination was detected as recombination between λ markers flanking the cloned DNA. Moreover, several other cloned P1 fragments, including some much larger, did not support efficient recombination. Thus, the absence of genetic linkage between the flanking markers could not be due to their separation by a large unmarked region. Nor could it be due to a nonspecific recombinase, because Nat showed the recombination to be independent of bacterial and λ genes essential for recombination. He postulated that the active fragment carried a site or sites at which a site-specific recombinase acts, that the recombinase is probably of P1 origin, and, if so, that it must be encoded within the same 6-kb P1 DNA fragment.

Nat named the site *lox*, for locus of crossover (x), simultaneously honoring a favorite traditional food of American Jewry. The choice of a name for the recombinase required a sense of its function.

The name did not appear in print until 1981, when Nat had isolated an amber mutation in the structural gene (12). The name Cre, on which Nat settled, is an abbreviation of cyclization recombinase and an anagram of Rec. It did not escape us that as lox, the smoked salmon, is commonly served over cream cheese, the proximity of the mutant cre_{am} gene to *lox*, the DNA site, has a certain logic.

In his 1978 publication on the P1 genetic map, Nat had already noted that experiments then in progress strongly suggested a role for *lox*-mediated recombination in cyclization of the phage DNA, an essential step in lysogenization as a plasmid. Host-mediated recombination between the terminally redundant regions (encompassing roughly 10% of the genome) can accomplish the task. If the duplicated regions in a significant number of phage particles were to include *lox*, then the DNA in those phages would be capable of cyclizing in the absence of the host recombination functions; in that case P1 should be able to lysogenize a *recA E. coli* at a modest but significant frequency. That it can do so had been shown in 1971 by Toshiya Takano (13).

Although P1 resembles λ in possessing a site-specific recombinase and a cognate recombination site, the prophage is maintained as a plasmid in *E. coli*, rather than, like λ, as part of the *E. coli* chromosome. By application of appropriate selection pressure, one can isolate rare individual *E. coli* strains (including *recA* mutants deficient in generalized recombination) in which P1 is carried as a stable integrated prophage.

At the time that Nat began studying P1, such bacterial strains had recently been described and the recombination site or sites within P1 shown to be situated at the ends of the P1 genetic map (14). Nat suspected that integration had occurred by Cre-mediated recombination, albeit highly inefficient, of cyclized P1 DNA into a bacterial pseudo-*lox* site. He showed this to be the case and named the putative pseudo-*lox* site *loxB*. He named the P1 *lox* site *loxP* and the hybrid sites at the ends of the integrated DNA *loxL* and *loxR*. Comparisons of these sites for their efficacy in recombination with a *loxP* site revealed that P × L and P × R are much more efficient than P × B. Whereas these anomalous recombinations have no apparent significance in the physiology of P1 itself, their study has proved to be of significant value for understanding the sequence of events in the recombination process.

The extensive terminal redundancy of P1 DNA in phage particles is a consequence of how the viral DNA is packaged from the linear polymer (concatemer) that is generated by rolling circle replication. Packaging into phage heads begins at a defined locus and proceeds for about four or five headfuls, each headful corresponding to about 1.1 genomes of DNA. The resulting redundant DNA ends in each phage particle enable cyclization in the infected cell to occur by bacterial recombination functions and make Cre-mediated recombination of duplicated *lox* sites largely superfluous. But Nat must have been correct in assigning cyclization as a significant function of Cre. His subsequent studies revealed that the locus on the DNA from which packaging begins (about 5 kb from *lox*) and the direction of packaging (toward *lox*) are such that the first DNA to be packaged in a phage head has a *lox* site centrally located in the roughly 10-kb region of terminal redundancy (15, 16). Selection pressure for an assist to cyclization must have placed it there.

Nat did not ignore other possible functions of Cre-mediated recombination and in his initial publication suggested a role in the formation of concatemers, the requisite substrate of headful packaging. Based on evidence that, under sufficient selection pressure, P1 integrated into the *E. coli* chromosome can be selected (14), Nat further suggested a likely role of P × B recombination in transducing particle generation. Whereas concatemer formation is important in phage production, concatemer resolution to monomers is important for plasmid maintenance. P1 prophage is stable despite its low copy number because of an active segregation mechanism. If the units to be segregated are covalently joined, plasmid loss at cell division will ensue. In 1981, Stuart Austin, Marcia Ziese, and Nat Sternberg documented this additional role for Cre-*lox* (17).

My (R.H.) introduction to Nat Sternberg came on a bleak November day in 1978. I was visiting the Frederick Cancer Research Center to interview for a position. The barrack-like buildings connected by elevated steam pipes and the six-story bricked-up building referred to as the anthrax tower added to the gloominess of the place. My last interview that day was with Nat. In the laboratory that he shared with Michael Yarmolinsky were crammed four shaking water baths, buckets bristling with dirty pipettes, and stacks of Petri dishes. Nat pulled up a couple of chairs for us, grabbed a legal pad, and proceeded to describe his projects. As I discovered then, talking and simultaneously drawing was Nat's expository style. Describing his work on the P1 repressor, he quickly filled the page, tore it off the pad, and wrote on the other side before using the next page. Near the end of the half hour he began describing the remarkable progress he had made on the site-specific recombination mediated by Cre-*lox* that he had found in P1. Having depleted the pad, Nat continued writing and drawing on both sides of the cardboard backing (**Figure 2**). It particularly interested me because I was just finishing a postdoctoral fellowship in the laboratory of Arthur Landy, studying the λ attachment site and the genes involved in λ site-specific recombination.

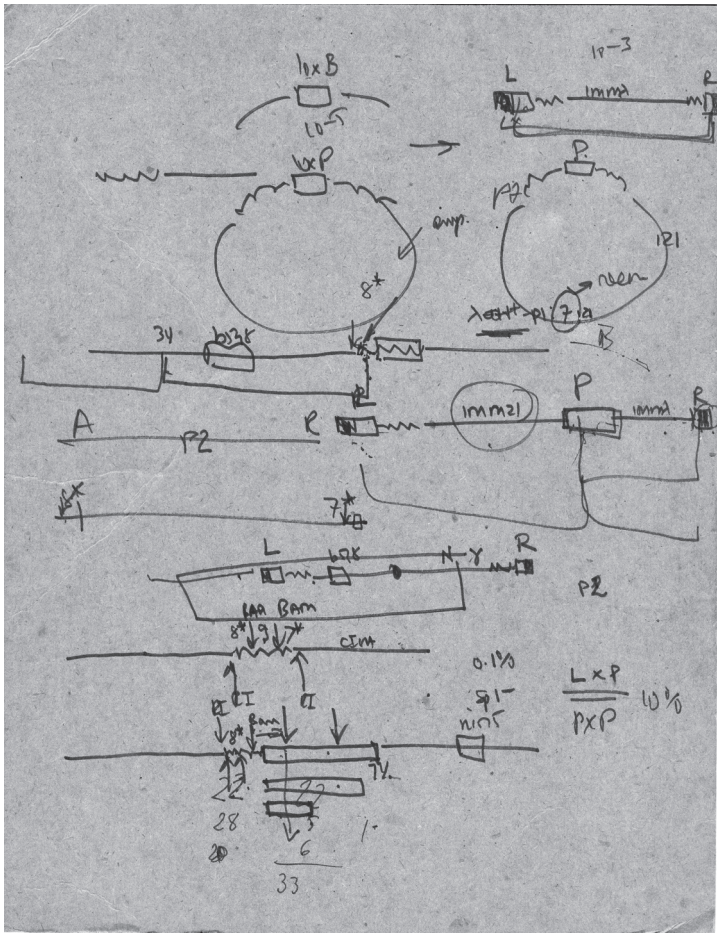

Figure 2

The back of the legal pad on which Nat Sternberg wrote during the interview with R.H., November 16, 1978.

When I accepted the offer to work at Frederick I had planned to continue working on λ, but not long after I arrived in the spring of 1979, Nat asked me rather casually whether I might be interested in helping out on the Cre-*lox* project. This was an opportunity too good to pass up, and within a few months nearly all my attention was directed at *lox*. Many of the features that have made Cre-*lox* a useful tool are best appreciated by comparison with what I will refer to as the λ paradigm for site-specific recombination. Lambda Int-*att* had been studied for over a decade prior to the discovery of Cre-*lox*, first by extensive genetic analysis, then, following development of an in vitro reaction, by biochemistry (18). By the late 1970s, many of the protein components had been purified and the sites of recombination (attachment sites, viral and bacterial) had been sequenced. The picture of recombination that emerged was complex. The recombinase (Int) and a number of accessory proteins first had to assemble into a nucleoprotein structure, which would then capture the bacterial attachment site. How recombination then took place was unclear, but a four-stranded, mobile (Holliday) junction seemed a likely intermediate. At that time λ Int-*att* represented the best-characterized site-specific recombination. Not surprisingly, Nat and I, along with Kenneth Abremski, who would later join us, were influenced by the λ paradigm, because each of us had worked on λ recombination. It influenced the design of our experiments and our expectations. Thus we were excited, even at the earliest stage, to uncover differences between Cre-*lox* and Int-*att*. More importantly, many of these differences would later translate into making Cre-*lox* the versatile genetic engineering tool that it is today.

By the end of 1979 we had determined the sequences of the various *lox* sites (19). The most striking feature was the presence of perfect 13-bp inverted repeats separated by an 8-bp spacer encompassing the crossover point. We supposed that the inverted repeats were potential binding sites for Cre. The minimal size of a functional *lox* site remained to be determined, but we assumed, by analogy with the λ attachment site, that additional DNA sequences beyond the inverted repeats were likely to be required for recombination.

In order to define the *loxP* site boundaries more precisely, we moved the putative *loxP*-containing region to a small plasmid convenient for further manipulation. It carried an ampicillin-resistance marker. Nat devised a simple pickup assay for testing *lox* function. The plasmid carrying the *lox* site to be tested was introduced into an *E. coli* strain expressing a functional *cre* gene. These cells were then infected with a λ containing a functional *loxP* site. If Cre-mediated recombination could occur between the two *lox* sites, then the resulting lysate should be able to transduce the ampicillin-resistance marker of the plasmid with high frequency (19).

In the original assay, the region of several hundred base pairs of P1 DNA flanking the functional *loxP* site in one partner of the recombination made it impossible to exclude its role in the process. Further experiments revealed that a plasmid containing two *loxP* sites, each less than 60 bp in length, could undergo efficient Cre-mediated recombination (20). Thus, unlike λ's attachment site of about 240 bp, the *loxP* site is quite small and therefore easily manipulated.

Early in 1981 Nat received permission to form his own group. However, because Yarmolinsky's group continued to study aspects of P1 biology and prokaryote research was deemed of only marginal relevance to cancer research, Nat was informed that he would have to switch his efforts to eukaryotes. For Nat, whose experience was entirely in the study of bacteria and their viruses, this was a daunting challenge. He accepted it, but had no intention of giving up his work on P1. Nat forged ahead on two fronts. He began studying recombination in mouse L cells, and he continued studying various aspects of P1 biology.

Nat's first hire was Ken Abremski, who had just finished postdoctoral studies at NIH on the λ protein Xis, which is required in conjunction with Int for excision of λ prophage from the bacterial chromosome (18, 21). Within months of his arrival, Ken developed an in vitro Cre-dependent recombination. One of the first tests of the in vitro reaction used two *loxP*-containing λ substrates

that Nat had constructed. Each substrate contained two *loxP* sites, one as direct repeats and the other as inverted repeats. Both substrates recombined efficiently in vivo and in vitro (20). This was in marked contrast to λ Int, which could recombine only directly repeated sites, and to the resolvases of certain of the small mobile genetic elements, such as that of γδ, which are able to catalyze only inversions (22).

Expression vectors were constructed that provided us with ample quantities of Cre protein. Ken made rapid progress in purifying Cre and determining the requirements for the recombination reaction. At almost every turn we found some difference when compared with λ Int recombination. Unlike λ Int recombination, no external energy source or additional accessory proteins were required for the reaction (23). Additionally, the different forms of the DNA substrate containing *loxP* sites, whether supercoiled, relaxed circle, or linear molecules, were all equally efficient substrates. Given the minimal requirements of the in vitro reaction along with the robustness of the recombination, working with Cre was as simple as doing a restriction enzyme digest.

WILMINGTON (DUPONT)

During our time in Frederick we were officially employed by Litton Industries, which ran our part of Fort Detrick under contract for the National Cancer Institute. This arrangement enabled the National Cancer Institute to respond rapidly to changing trends in cancer research, and as part of a prokaryotic research group, we were aware of our vulnerability when the contract came up for renewal in 1984. For Nat and the members of his burgeoning laboratory, rescue came via DuPont. A relatively new member of the Frederick Cancer Research Center, Mark Pearson, was offered a position in DuPont's Central Research and Development Department with the mandate to create a world-class molecular biology laboratory. Building on a well-established group of animal virologists, molecular geneticists, and chemists, Pearson rapidly assembled a department that included Nat's entire group as well as biophysicists with expertise in molecular structure.

Nat's group moved to DuPont in the spring of 1984. Unlike the dreary environs of Frederick, we found ourselves on a tree-lined campus with spacious, newly renovated labs. Funding for equipment and supplies was never an issue. By this time Ken and I were absorbed in working out the molecular details of Cre recombination. With purified Cre we were able to show by DNA footprinting that Cre does not contact DNA outside the 34 bp of the *loxP* site (24). Because Cre is able to carry out either excision or inversion depending on the orientation of the *loxP* sites, the footprinting results imply that the 8-bp spacer region imparts the directionality to the site. We demonstrated in vitro that Cre generates a 6-bp staggered cut within the spacer region (25) and made single-base substitutions within the spacer region to determine how they affected recombination. The results showed that homology of the spacer regions is required for recombination (26). Later these findings were exploited in genetic engineering with Cre and *lox*. Multiple *lox* sites, each addressable by the sequence of the spacer region, could be introduced into a cell, and only like sites would recombine with each other (27).

The products of recombination between directly repeated *lox* sites on a highly supercoiled substrate were found not to be entangled with each other, as might be expected. Instead, recombination yielded supercoiled molecules entirely free of each other or linked as simple catenanes, rings that are topologically, but not chemically, linked (20, 28). This result suggested that before the initiation of recombination, the recombination sites come together in a way that physically separates the domains between them, as might occur if the sites are brought together via a sliding of DNA upon itself. Note that the rings of a DNA catenane are readily unlinked by cellular topoisomerases.

One of Nat's early papers on Cre-*lox* looked at strand exchange during recombination of *lox* sites (29). A series of λ-P1 hybrid phages carrying *loxP* or *loxR* were crossed and the reassortment of outside markers scored. From these results it was noted that branch migration had frequently occurred following an initial crossover. More direct physical evidence would come later when we began isolating and characterizing various mutants of Cre. One particular mutant appeared partially deficient for recombination in vitro (30). When the products of the reaction were analyzed it was shown that Cre preferentially exchanged one set of strands but only partially exchanged the second set of strands, thereby generating Holliday structure intermediates among the products. To rule out that the Holliday intermediates were aberrant end products of the reaction, they were purified and shown to be resolved on further addition of Cre, thus demonstrating they were intermediates in the reaction.

In addition to the biochemical dissection of the Cre-*lox* interaction, the sequence of Cre itself was determined (31). Not long thereafter, an effort was made to align the sequence of Cre and other known recombinases, including Int of λ, P2, P22, P4, φ80, and φ186 (32). Despite the enormous sequence diversity among these recombinases, alignment by their C termini indicated a few small regions of homology with three residues (His, Arg, and Tyr) strictly conserved among all members. These residues were hypothesized to be involved in the catalysis of strand cleavage by the recombinases. A short time later a fourth conserved residue (Arg) was also identified (33). Three of these residues—Arg, His, and Arg—were shown by X-ray crystallography to form hydrogen bonds to the phosphate to which the conserved Tyr residue becomes transiently covalently bonded during cleavage, thus preserving the energy of the phosphodiester bond during the exchange (34). Over the years more than 100 potential recombinases have been identified by these alignments, indicating that Cre is a member of an enormously divergent recombinase family.

Nat was particularly interested in the regulation of Cre expression and how it is modulated in relation to the P1 life cycle. Three promoters of varying strengths were identified upstream of Cre (31). One of these promoters contains Dam methylation sites, and experiments showed that in the absence of Dam methylation Cre expression increases. This leads to the idea that upon infection, before newly replicated P1 DNA could be fully methylated, Cre expression would be elevated precisely when needed to cyclize the linear viral DNA.

Cre expression is also expected to be elevated during lytic growth, whether following infection or following induction of an established prophage. This elevated expression could, in principle, increase the burst size of phage particles by recruiting those P1 plasmids that have not initiated replication into P1 plasmids that are in the process of replication (35), or it could do so simply by forming concatemers, the substrate for DNA packaging into phage heads (11). P1 concatemers are formed during lytic growth by rolling circle (σ) replication, which, during the course of infection, takes over from bidirectional (θ) replication (36).

Whereas concatemers are important for producing viral particles, they (as well as circular oligomers) are disastrous for plasmid maintenance, because they are unpartitionable units. The problem for segregation created by circular chromosome dimers generated by interchromosome recombination was recognized as early as 1932 by Barbara McClintock [cited by Barre et al. (37)]. As Barre et al. observe, the problem appears to have been largely ignored prior to work on P1 segregation (17). During lysogenic growth of P1, plasmid oligomers are formed occasionally by RecA-mediated homologous recombination. Their resolution prior to cell division, a function of Cre-*lox*, is essential for prophage stability (17, 38). Ironically, problems arising from recombination events that are infrequent with respect to the cell cycle are remedied by raising the frequency of those events. Raising the recombination frequency without introducing new problems is accomplished by limiting the site of recombination to a highly localized region.

The Int of λ and Cre of P1 are similar in having more than a single role in the life of their respective phages, but they differ in the diversity of their substrates. Whereas Int acts on supercoiled circular DNA, the substrates of Cre include the linear DNA of infecting phages, the rolling circle DNA of phages in the lytic growth phase, and the supercoiled DNA of plasmid oligomers. The striking promiscuity of Cre, which is so seductive a feature of the enzyme for genetic engineers, correlates with the diversity of the substrates with which it has biologically significant interactions.

Cre-*lox* COMMERCIALIZATION AND A FAUSTIAN BARGAIN

One member of Nat's group, Brian Sauer, upon moving to DuPont, began experiments to determine whether Cre-*lox* could function in eukaryotic cells. Initially that possibility seemed unlikely. Cre, a prokaryotic protein, would now be made in the cytoplasm and would need to find its way to the nucleus. Once in the nucleus, Cre would need to find *lox* sites while surrounded by orders of magnitude more DNA than in an *E. coli* cell. Despite initial concerns raised by these considerations, Brian constructed vectors for the expression of *cre* in yeast as well as *lox* substrates that could easily monitor Cre-mediated recombination. The substrates were cassettes with *lox* sites on each side of the yeast *LEU2* gene. These were then stably inserted into the yeast genome in two different chromosomal locations by homologous recombination. A separate vector was then introduced containing the *cre* gene under the control of the yeast galactose promoter. Induction with galactose resulted in *cre* expression and generation of leucine auxotrophs as a result of excision of *LEU2* via *lox* recombination (39). This was the basis for the first patent filing on Cre-*lox* functioning in eukaryotic cells. Initially the work received a muted response from the scientific community because functionally expressing a bacterial or phage protein in yeast was not considered that remarkable. Brian, undeterred, went on to show that Cre-*lox* could also function efficiently in mouse cell lines (40). A stable mouse cell line was established with *cre* expressed from an inducible metallothionein promoter. Various *lox* substrates were subsequently introduced into this cell line by transformation, and upon induction of the promoter, efficient *lox* recombination could be detected. It was the beginning of a revolution in genetic engineering. Late in 1990, Brian was granted a US patent for site-specific recombination of DNA in eukaryotic cells. It marked the commercialization of what was to become one of the most widely used tools in genetic engineering. Most companies, when issued a patent of this sort, continue working on the technology with the hopes of not only refining the tools but also broadening and strengthening their patent position. This was not to be the case, and by the close of 1990, research on Cre-*lox* at DuPont had ended.

In the fall of 1990 the molecular biology department had an off-site retreat organized by Mark Pearson. An after-dinner talk on the last evening was given by the head of DuPont Central Research and Development. The talk began innocuously enough, lauding our accomplishments as a department. It soon became apparent, when he discussed the DuPont bottom line, that Mark's vision of continuing a world-class molecular biology department was a mirage. Within a few short months a joint venture was announced between DuPont Medical Products and Merck. No longer was there a need for a freestanding molecular biology department at DuPont; we were now incorporated within a pharmaceutical company.

We had made our bargain and, after almost seven years, it was time to pay. A few of us managed to escape to the relative shelter of academia to continue basic research, but for most, this was not the case. The real tragedy was that Nat, because of a recurring cancer, was unable to leave. He was entering the last four years of his life, during which he endured almost continuous rounds of chemotherapy which left him bald and emaciated. Barely able to stand at his lab bench, Nat continued working. During this time, with little to no encouragement from anyone

in the company, he and a few postdoctoral students made significant contributions to the human genome project (and other genome projects) by developing P1 in vitro packaging for cloning large DNA molecules (41). The vectors included not only P1 packaging sites but also *lox* sites for cyclizing the large DNA once injected into *E. coli*. For replication of the clones, a P1 replicon was used to maintain low copy number, but a controllable lytic P1 replicon was also included to provide for amplification of clonal DNA. It was as if Nat was reconstructing P1 from first principles.

EPILOGUE

On June 29, 1995, three months before his death, a celebration of Nat's career was held at NIH as a gala Lambda Lunch (a long-established NIH interest group). A large contingent of laboratory members and friends from Wilmington as well as former colleagues from NIH attended. The day was marked by a number of talks relating to Nat's many contributions over the years. Along with the science there was mention of Nat's impatience to get on with his work, evident from his teeth-marks on hard-to-open restriction enzyme vials and the abnormal whine of unbalanced Eppendorf centrifuges. The highlight of the day was Nat's own talk, in which he described his training in the yeshiva, debating Talmudic questions, and then turned to his current interest in creating vectors for cloning large DNAs. His enthusiasm for science, even in those final days, remained unabated. It seems only fitting that his spark has been kept alive through the establishment by his former colleagues of the Nat Sternberg Thesis Prize, awarded annually for the best thesis in prokaryotic research. For those of us who had the good fortune to work alongside Nat, our lives were enriched by his example as scientist and friend.

In the past twenty years, the avenues of research into site-specific recombination that had been laid out by Nat and his coworkers have been extended in two directions, aided by high-tech advances in visualizing the choreography of proteins and DNA within living cells and in observing the staged performances of individual DNA molecules in vitro. Advances of no less importance have come from old-fashioned, low-tech bacterial genetics. One avenue of research leads to a deeper understanding of the cell cycle context of site-specific recombination. The other avenue, seemingly more relevant to Cre-*lox* technology, leads to an understanding of the successive molecular events in the recombination process. We cannot do justice here to recent developments in either of these areas, but cannot resist adding a few words about each.

Nat did not live to become aware of the important bacterial cell division protein FtsK and its involvement in a function of a bacterial recombinase, XerCD. Like Cre and Int, both components of this bacterial recombinase possess an essential Arg-His-Arg triad and the catalytic Tyr residue at their active sites (42). The existence of the *ftsK* gene (discovered in *E. coli* but since shown to be widely distributed among eubacteria) was reported only late in the year of Nat's death (43). The FtsK protein is a DNA translocase that reads the polarity of DNA based on its affinity for 8-bp sequences that are preferentially oriented toward the replication terminus (44). From its location in the nascent septum it rapidly reels in DNA, sorting sister terminus regions on either side of the septal ring. It efficiently displaces bound proteins in the process until it finds the region of the replication terminus (**Figure 3**). Within that region lies *dif*, the site of action of the XerCD recombinase. When the XerCD-*dif* complex reaches FtsK, the translocase activates XerD to generate Holliday junction intermediates that can then be resolved by XerC (45). In this way FtsK facilitates completion of chromosome segregation prior to cell division. If Cre and suitably positioned *lox* sites are substituted for XerCD and *dif* sites, Cre will satisfactorily resolve chromosome dimers (46). Dimer resolution by Cre-*lox* was shown to depend on both FtsK and the appropriate location of the *lox* sites. This action of FtsK on Cre introduces an intriguing new

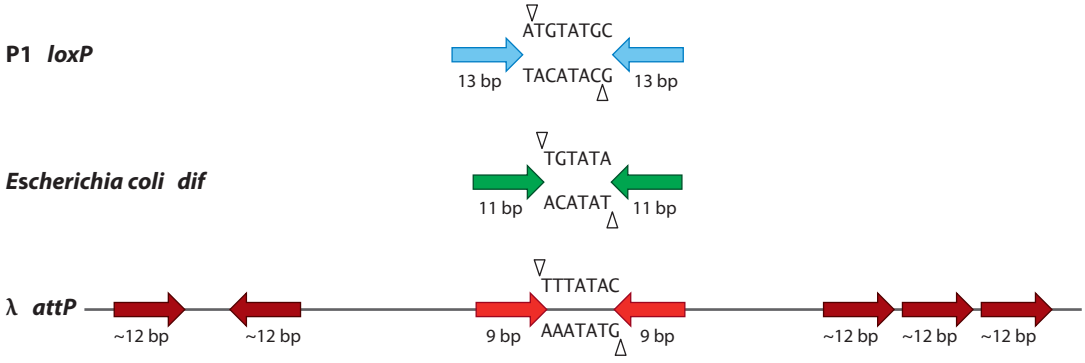

Figure 3

Comparison of the binding sites of recombinases Cre (19), XerCD (48), and Int (49). During recombination each recombinase generates staggered cleavages (*small triangles*) at the periphery of a spacer between elements of dyad symmetry of a core binding site. Int has additional arm recognition elements (*dark red arrows*) distributed over 240 bp. They are related (but not identical) and necessary for the assembly of an active recombination complex in which the DNA is bound to an Int tetramer and to host proteins IHF and Fis.

aspect of Cre-mediated recombination in vivo. The intimate molecular details of recombinase activation by FtsK have now been tracked by three single-molecule techniques (47).

Not long after Nat's death, Greg Van Duyne and his group solved the crystal structure of Cre-*lox* (34). The structure not only revealed the molecular architecture of Cre and the *loxP* site but also showed them in the process of recombining via a Holliday structure intermediate. This work was rapidly followed by reports of additional structures where altered *loxP* sites or mutant Cre proteins were cocrystallized (50). The combination of structural work with traditional biochemical analysis has resulted in a detailed picture of recombination. Cre binds to the *loxP* site and induces the DNA to bend. Formation of a synaptic structure begins a well-ordered series of steps in which cleavage and strand exchange first occur predominantly on the bottom strand to form the Holliday structure intermediate. Following isomerization of the Holliday structure, top-strand cleavage and strand exchange take place to complete the reaction. This picture has recently been bolstered by the new technique of single-molecule fluorescence in which individual molecules can be observed during the recombination process (51).

At the beginning of this review we noted the coincidence that λ and P1 had parallel roles in the rediscovery of lysogeny and then in the recognition of the important, but differing, roles of site-specific recombination in the life of these phages. Lambda's Int-mediated recombination is unidirectional and is under control by three accessory proteins, of which one (Xis) can reverse its direction. Int belongs to the class of regulated recombinases. P1's Cre-mediated recombination is bidirectional, requires no accessory proteins, and is unregulated. It belongs, along with XerCD, to the class of recombinases that are largely unregulated. In the course of this review, as in the course of characterizing Cre-*lox*, the contrast between Int and Cre and between *att*-λ and *loxP* has been a recurring theme. It therefore seems fitting that we close by considering their union, specifically the interaction of an Int-Cre chimera with chimeric binding sites. These were created in the laboratory of Arthur Landy (52). A small (76-residue) N-terminal fragment of Int was fused to Cre. The resulting chimeric enzyme recombines complex binding sites with all the regulated directionality and responses to accessory proteins that characterize Int-promoted recombination. The authors suggest that this construction, in its striking simplicity, may simulate the evolutionary interconversions responsible for the large variety of site-specific recombinases observed in the living world.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Richard D'Ari, Robert LaRossa, Greg Van Duyne, and David Sherratt for their critical reading of a draft of this article and François Cornet for enlightening comments on FtsK Orienting Polar Sequences (KOPS). This essay was supported by the Intramural Research Program of the National Cancer Institute's Center for Cancer Research at the National Institutes of Health.

LITERATURE CITED

1. Lwoff A. 1953. Lysogeny. *Bacteriol. Rev.* 17:269–337
2. Lwoff A. *Nobel Lecture: Interaction Among Virus, Cell, and Organism*. Stockholm: Nobel Media AB. **http://www.nobelprize.org/nobel_prizes/medicine/laureates/1965/lwoff-lecture.html**
3. Bertani G. 2004. Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *J. Bacteriol.* 186:595–600
4. Lennox ES. 1955. Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* 1:190–206
5. Yarmolinsky MB, Sternberg N. 1988. Bacteriophage P1. In *The Bacteriophages*, Vol. 1, ed. R Calendar, pp. 291–438. New York: Plenum
6. Łobocka MB, Rose DJ, Plunket G III, Rusin M, Samojedny A, et al. 2004. Genome of bacteriophage P1. *J. Bacteriol.* 186:7032–68
7. Ikeda H, Tomizawa J. 1968. Prophage P1, an extra-chromosomal replication unit. *Cold Spring Harb. Symp. Quant. Biol.* 33:791–98
8. Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3:315–56
9. Clerget M. 1991. Site-specific recombination promoted by a short DNA segment of plasmid R1 and by a homologous segment in the terminus region of the *Escherichia coli* chromosome. *New Biol.* 3:780–88
10. Sternberg N, Tiemeier D, Enquist L. 1977. In vitro packaging of a λ D*am* vector containing *Eco*RI DNA fragments of *Escherichia coli* and phage P1. *Gene* 1:255–80
11. Sternberg N. 1978. Demonstration and analysis of P1 site-specific recombination using λ-P1 hybrid phages constructed in vitro. *Cold Spring Harb. Symp. Quant. Biol.* 43:1143–46
12. Sternberg N, Hamilton D. 1981. Bacteriophage P1 site-specific recombination. I. Recombination between *loxP* sites. *J. Mol. Biol.* 150:467–86
13. Takano T. 1971. Bacterial mutant defective in plasmid formation: requirement for the lon^+ allele. *PNAS* 68:1469–73
14. Chesney RH, Vapnek D, Scott JR. 1978. Recombination between the plasmid prophages P1 and P7 and the *E. coli* chromosome. *Contrib. Microbiol. Immunol.* 6:78–88
15. Sternberg N, Coulby J. 1987. Recognition and cleavage of the bacteriophage P1 packaging site (*pac*). I. Differential processing of the cleaved ends in vivo. *J. Mol. Biol.* 194:453–68
16. Sternberg N, Coulby J. 1987. Recognition and cleavage of the bacteriophage P1 packaging site (*pac*). II. Functional limits of *pac* and location of *pac* cleavage termini. *J. Mol. Biol.* 194:469–83
17. Austin S, Zeise M, Sternberg N. 1981. A novel role for site-specific recombination in the maintenance of bacterial replicons. *Cell* 25:729–36
18. Nash HA. 1981. Integration and excision of bacteriophage λ: the mechanism of conservative site specific recombination. *Annu. Rev. Genet.* 15:143–67
19. Hoess RH, Ziese M, Sternberg N. 1982. P1 site-specific recombination: nucleotide sequence of the recombining sites. *PNAS* 79:3398–402
20. Abremski K, Hoess R, Sternberg N. 1983. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked productions following recombination. *Cell* 32:1301–11

21. Abremski K, Gottesman S. 1981. Site-specific recombination: Xis-independent excisive recombination of bacteriophage lambda. *J. Mol. Biol.* 153:67–78
22. Reed RR. 1981. Transposon-mediated site-specific recombination: a defined in vitro system. *Cell* 25:713–19
23. Abremski K, Hoess R. 1984. Bacteriophage P1 site-specific recombination: purification and properties of the Cre recombinase protein. *J. Biol. Chem.* 259:1509–14
24. Hoess R, Abremski K. 1984. Interaction of the bacteriophage P1 recombinase Cre with the recombining site *loxP*. *PNAS* 81:1026–29
25. Hoess R, Abremski K. 1985. Mechanism of strand cleavage and exchange in the Cre-*lox* site-specific recombination system. *J. Mol. Biol.* 181:351–62
26. Hoess R, Wierzbicki A, Abremski K. 1986. The role of the *loxP* spacer region in P1 site-specific recombination. *Nucleic Acids Res.* 14:2287–300
27. Sauer B. 1996. Multiplex Cre/*lox* recombination permits selective site-specific DNA targeting to both a natural and an engineered site in the yeast genome. *Nucleic Acids Res.* 24:4608–13
28. Abremski K, Hoess R. 1985. Phage P1 Cre-*lox* site-specific recombination: effects of DNA supercoiling on catenation and knotting of recombinant products. *J. Mol. Biol.* 184:211–20
29. Sternberg N. 1981. Bacteriophage P1 site-specific recombination. III. Strand exchange during recombination at *lox* sites. *J. Mol. Biol.* 150:603–8
30. Hoess R, Wierzbicki A, Abremski K. 1987. Isolation and characterization of potential intermediates in site-specific recombination. *PNAS* 84:6840–44
31. Sternberg N, Sauer B, Hoess R, Abremski K. 1986. Bacteriophage P1 *cre* gene and its regulatory region: evidence for multiple promoters and regulation by DNA methylation. *J. Mol. Biol.* 187:197–212
32. Argos P, Landy A, Abremski K, Egan JB, Haggard-Ljungquist E, et al. 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J.* 5:433–40
33. Abremski K, Hoess R. 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Protein Eng.* 5:87–91
34. Guo F, Gopaul DN, Van Duyne GD. 1997. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* 389:40–46
35. Adams DE, Bliska JB, Cozzarelli NR. 1992. Cre-*lox* recombination in *Escherichia coli* cells: mechanistic differences from the in vitro reaction. *J. Mol. Biol.* 226:661–73
36. Cohen G. 1983. Electron microscopy study of early lytic replication forms of bacteriophage P1 DNA. *Virology* 131:159–70
37. Barre FX, Soballe B, Michel B, Aroyo M, Robertson M, et al. 2001. Circles: the replication-recombination-chromosome segregation connection. *PNAS* 98:8189–95
38. 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 Harb. Symp. Quant. Biol.* 45:297–309
39. Sauer B. 1987. Functional expression of the *cre-lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 7:2087–96
40. Sauer B, Henderson N. 1988. Site-specific recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *PNAS* 85:5166–70
41. Sternberg N. 1994. The P1 cloning system past and future. *Mamm. Genome* 5:397–404
42. Esposito D, Scocca JJ. 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucleic Acids Res.* 25:3605–14
43. Begg KJ, Dewar SJ, Donachie WD. 1995. A new *Escherichia coli* cell division gene, *ftsK*. *J. Bacteriol.* 177:6211–22
44. Bigot S, Sivanatan V, Possoz C, Barre FX, Cornet F. 2007. FtsK, a literate chromosome segregation machine. *Mol. Microbiol.* 64:1434–41
45. Grainge I, Lesterlin C, Sherratt DJ. 2011. Activation of XerCD-*dif* recombination by the FtsK DNA translocase. *Nucleic Acids Res.* 39:5140–48
46. Capiaux H, Lesterlin C, Pérals K, Louarn JM, Cornet F. 2002. A dual role for the FtsK protein in *Escherichia coli* chromosome segregation. *EMBO Rep.* 3:532–36
47. Zawadzki P, May PF, Baker RA, Pinkney JN, Kapanidis AN, et al. 2013. Conformational transitions during FtsK translocase activation of individual XerCD-*dif* recombination complexes. *PNAS* 110:17302–7

48. Carnoy C, Roten CA. 2009. The *dif*/Xer recombination systems in proteobacteria. *PLOS ONE* 4:e6531
49. Hsu PL, Ross W, Landy A. 1980. The λ phage *att* site: functional limits and interaction with Int protein. *Nature* 285:85–91
50. Van Duyne GD. 2001. A structural view of Cre-*loxP* site-specific recombination. *Annu. Rev. Biophys. Biomol. Struct.* 30:87–104
51. Pinkney JN, Zawadzki P, Mazauryic J, Arciszewska LK, Sherratt DJ, Kapanidis AN. 2012. Capturing reaction paths and intermediates in Cre-*loxP* recombination using single molecule fluorescence. *PNAS* 109:20871–76
52. Warren D, Laxmikanthan G, Landy A. 2008. A chimeric Cre recombinase with regulated directionality. *PNAS* 105:18278–83