*IN VIVO* INTEGRON INTEGRASE (INTI1) ABUNDANCE AFFECTS THE TYPE AND AMOUNT OF RECOMBINATION PRODUCTS DIFFERENTIALLY DURING THE GROWTH CYCLE

by

JULIA ELAINE SERVATIUS SHEARER

(Under the Direction of Anne O. Summers)

ABSTRACT

The integron integrase IntI1 mediates the transfer of antibiotic resistance gene cassettes between plasmid-carried integrons by two mechanisms, integration and excision or the formation and resolution of cointegrates. Cointegrates can be isolated from cells expressing integrase under the natural promoter $P_{int}$, but cassette capture has never been experimentally observed without excess integrase. Historically, integrase activity has been measured after conjugation or transformation of its recombinant products to a recipient cell. We constructed an integron capture vector pICV8 that acquires and donates antibiotic resistance gene cassettes in strains that overexpress IntI1 and detected recombination products in cells by PCR without conjugative transfer of recombinant plasmids. Qualitative PCR revealed that pICV8 readily forms cointegrates *in vivo* in strains expressing integrase under its natural promoter on a low copy number plasmid, though cassette capture products are not detected without excess integrase. Using dilution PCR, we report the first measurements of the time-

dependence of formation of integron intracellular products. Integrase-mediated recombination products were quantified over time using dilution PCR in low and high integrase strains. Only cointegrates were detectable in strains expressing integrase under a natural promoter, and the simple acquisition of gene cassettes could only be seen in strains overexpressing integrase. Recombination products in both the low and high integrase strains increased in late log to early stationary phase. *attC* x *attI* recombination has been reported to occur 10 times more frequently than *attI* x *attI* crossovers, but, unexpectedly, we detected more *attI* x *attI* than *attC* x *attI* recombination products in log phase for both strains. However, the high integrase strain showed increased *attC* recombination in stationary phase, consistent with the earlier observations on integrase crossover preferences. Thus, the intracellular abundance of IntI1 affects the amount and type of recombination events, especially toward the end of log phase.

INDEX WORDS: Integron, integrase, IntI1, *attI*, *attC*, recombination, antibiotic resistance, growth cycle, dilution PCR

# *IN VIVO* INTEGRON INTEGRASE (INTI1) ABUNDANCE AFFECTS THE TYPE AND AMOUNT OF RECOMBINATION PRODUCTS DIFFERENTIALLY DURING THE GROWTH CYCLE

by

JULIA ELAINE SERVATIUS SHEARER

B. S., University of Georgia, 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2007

# *IN VIVO* INTEGRON INTEGRASE (INTI1) ABUNDANCE AFFECTS THE TYPE AND AMOUNT OF RECOMBINATION PRODUCTS DIFFERENTIALLY DURING THE GROWTH CYCLE

by

JULIA ELAINE SERVATIUS SHEARER

Major Professor: Anne O. Summers

Committee: Anna Karls
Robert Maier
Eric Stabb

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2007

ACKNOWLEDGEMENTS

I would like to thank my major professor, Anne O. Summers, for her guidance and insight and my committee members for their knowledge and suggestions. I would also like to acknowledge all of my labmates, past and present, for their invaluable aid and give thanks to my husband and family for their unending support.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

## Introduction and Literature Review

## INTRODUCTION

Conservative site-specific recombination is a “cut and paste” mechanism of recombination utilizing short, specific DNA recognition sequences for insertion, excision (including resolution), and/or inversion reactions without the use of high-energy cofactors or DNA replication. This strict definition of site-specific recombination excludes some reactions that have been termed site-specific, such as DDE family-mediated transpositions and movement mediated by the IS*110*/IS*492* and Piv families of recombinases (66, 68, 79). Transposition, such as that mediated by the DDE superfamily of transposases, is similar to site-specific recombination in that both recombine via transesterification mechanisms that show some site-specificity without requiring ATP, but, unlike the two-step transesterifcation of site-specific recombinases, transposases use a one-step transesterification that requires DNA synthesis/repair mechanisms to complete the recombination reaction (66, 68, 79). Site-specific recombinases in nature catalyze diverse reactions such as the resolution of DNA dimers, inversion of DNA segments for phase variation, and integration of phage genomes, and the vast majority can be separated into two groups, the serine and tyrosine recombinases, named for the conserved amino acid that covalently binds DNA during the recombination reaction (see ref (68) for review).

Though serine and tyrosine recombinases share neither sequence homology nor specific reaction mechanism, they catalyze the same types of reactions using DNA recognition sites of similar appearance (68, 69, 176, 178). Recombinases of both families are characterized by their individual DNA recognition sequences, which are sites for binding and crossover and usually mediate recombination directionality. These DNA sites are referred to as core sites (Fig. 1.1 A) and usually consist of a palindromic sequence of nucleotides, with inverted repeats surrounding a central region of varying length (68, 69, 176). The inverted repeat regions are called half-sites, and each half site typically binds one subunit of recombinase. Directionality of a core site is usually determined by the specific sequence of the crossover region, but it can also be affected by the inverted repeats being imperfect, as is seen in *att* site recombination by integron integrases (68, 185). For site-specific recombination, the type of reaction that occurs largely depends on the arrangement of recognition sites on the DNA molecules. An inversion reaction (Fig 1.1 B), such as that involved in the switching of *Salmonella* flagellar antigen expression mediated by the serine invertase Hin (91), occurs between recombination sites arranged in inverse orientation on the same DNA molecule (66, 176). Integration, such as of phage genomes by integrases of both the serine and tyrosine families, occurs by recombination between core sites located on separate DNA molecules (Fig. 1.1 C). In phage integrases, a phage attachment site *attP* and a bacterial attachment site *attB* cause a directional insertion and produce *attL* and *attR* sites on the junctions of the recombination product (69). Excision is the reverse reaction of integration and usually requires an excision factor (Xis) in addition to the

recombinase. Excision of phage DNA occurs when *attL* and *attR* recombine to form an *attB* in the bacterial DNA sequence and *attP* in the phage DNA sequence (69, 176). Resolution is another type of excision reaction involving DNA recombination sites on opposite sides of a single DNA molecule arranged in direct repeat, such as those occurring in chromosomal or plasmid dimers and plasmid cointegrates. The γδ resolvase (68, 168, 176) and Cre (68, 200) are well-studied resolvases of the serine and tyrosine families, respectively.

Both serine and tyrosine recombinases are ubiquitous in eubacteria, widely present in archea, and more and more are being discovered in eukaryotes (68). In addition, several prokaryotic site-specific recombinases have been experimentally shown to function in eukaryotic cells (21, 69, 86, 93, 176), including the tyrosine recombinases Cre (169) and lambda integrase (31), and the serine phage integrases from ϕC31 and R4 found in *Streptomyces parvulus* (136, 195). The site-specific recombinases can mediate precise insertion and/or excision of DNA into sequence-specific locations, either engineered recognition sites or natural pseudo-sites, within eukaryotic genomes (21). Any DNA segment could theoretically be inserted into engineered recognition sites, and integrating DNA into naturally present pseudo-sites has possibilities for gene therapy treatments, showing that site-specific recombination, widely known to be vital in nature, may also become important for medical applications (69, 137-139).

The integron integrases form a unique and distinct family within the tyrosine recombinases, and are best explained after first describing site-specific recombinases in

general. This review first characterizes the serine family of site-specific recombinases, giving specific examples that demonstrate the regulation of site-specific recombination using accessory components (host cofactors or additional recombinase subunits). Second, I review the tyrosine family and describe three prototypical tyrosine recombinases (Cre, XerCD, and lambda integrase) that exemplify (respectively) the mechanism, versatility, and complex regulation within the tyrosine family. Third, I review integrons and their site-specific integrases by describing the importance of integrons in the dissemination of antibiotic resistance genes and characterizing the major classes of integron integrases and their heterogeneous DNA recombination sites.

**SERINE RECOMBINASES**

The serine recombinases are distinguished by the use of a serine nucleophile to attack the recombining DNA. The family separates into two groups, the smaller resolvase/invertases and transposases and the large integrases, based on size, function, and catalytic domain phylogenetic analysis (68, 176). The group of small serine recombinases, less than 200 amino acids in size, can be separated into two subgroups, the resolvases/invertases and the IS*607*-like transposases, by organization of protein domains (Fig. 1.2 A, B). The IS*607*-like transposases, including IS*607* from *Helicobacter pylori* and IS*1535* from *Mycobacterium tuberculosis*, are unusual in that their DNA-binding domain is N-terminal to the catalytic domain (Fig. 1.2 B), backwards from the other serine recombinases (94, 176). In addition, the resolvase/invertase traditional subgroup of serine recombinases, including well-known members like Hin

invertase from *Salmonella* and the transposon γδ resolvase from *E. coli*, form a tightly clustered group when their catalytic protein sequences are aligned (Fig. 1.3), as do those of the second subgroup, the IS*607*-like transposases (69, 176). The second group of serine recombinases is composed of the large phage integrases (Fig. 1.2 C). These serine integrases are between 400 and 800 amino acids in size and are mostly found in Gram-positive bacteria. They include the phage integrases from φC31 and R4 found in *Streptomyces parvulus* and TP901 from *Lactococcus lactis* (69, 176). Unlike the subgroups of the serine recombinase group, the large phage integrases do not cluster closely together from alignment of their catalytic domains (Fig. 1.3), however they contain additional conserved domains of unknown function (68, 69, 176) (Fig. 1.2 C).

**Serine recombinase reaction mechanism**

The serine recombinase mechanism for recombination has been determined mainly from studying the γδ resolvase, one of the smaller serine recombinases of the resolvase/invertase subgroup. Recombination occurs by a four-strand cleavage and rejoining mechanism (see Fig. 1.4 and ref (68) for illustration). The recombinase binds the DNA as an asymmetric dimer and cleaves in a staggered manner, the bottom strand cleavage 2 bp from the top strand cleavage, on both molecules of DNA whereby the active serine nucleophile binds the recessed 5' end of the DNA, leaving a 3' hydroxyl. Ligations occur after a 180° rotation of the complex, necessary to bring together the cleaved DNA, initially located on the outsides of the recombinase synapse (Fig. 1.4).

There is no Holliday junction intermediate in this process, though the precise synaptic structure is unknown (168, 176, 178). The 2-bp crossover sequence must be identical for efficient recombination and can provide directionality to the reaction (68). The large serine integrases are presumed to recombine in a similar manner, though they show little catalytic core sequence amino acid identity to their smaller relatives, or to each other, and their DNA crossover sequences vary considerably in size (69, 176).

**Small serine recombinases**

The smaller serine recombinases are less than 200 amino acids in length and are composed of two domains: the N-terminal catalytic domain of approximately 145 amino acids, which is also responsible for protein-protein interactions, and the DNA-binding C-terminal domain (68, 176) (Fig. 1.2 A, B). The catalytic domain encodes the active site residues of the recombinase and binds the 3' end of the DNA after cleavage and is also responsible for protein-protein interactions. The C-terminal domain encodes a helix-turn-helix motif thought to solely function in binding DNA (2, 68). The majority of known serine recombinases are in the resolvase/invertase subgroup, the most-studied being Hin and Gin invertases and the Tn3 and γδ resolvases. The resolvase members of the subfamily do not require accessory factors but tend to have large and complex DNA recognition sites, while the invertases use simple DNA core sites plus an enhancer sequence and corresponding host cofactor (68, 176). The IS*607*-like transposase proteins are similar to the resolvase/invertases in sequence structure except that the domains are reversed (Fig. 1.2 B), with the DNA-binding domain in the N-terminus

(176). These recombinases are called transposases since they excise and integrate IS elements, but they are members of the serine site-specific family by sequence and mechanism and are not at all related to the DDE transposases (68, 176). The IS*607*-like recombinases are unusual in that they do not require an excision factor (94, 176). IS*607* DNA insertion sites are also unusual in showing little sequence conservation, only the 2-bp crossover (GG) is necessary, and there appears to be no directionality in insertion (94, 176).

**Large serine integrases**

The large serine integrases are from greater than 400 to nearly 800 amino acids, and only show limited similarity to their smaller relatives in the N-terminal domain and no similarity in the C-terminus (69). The N-terminal catalytic domain shares conserved residues with the other serine recombinases, but the serine integrases do not form a tightly clustered group on alignment of their catalytic domains (Fig. 1.3), meaning they are a more divergent group. The large integrases also do not encode a DNA-binding domain with a helix-turn-helix motif, but they do have two additional domains of unknown function (68, 69, 176). The largest of these domains is approximately 220 amino acids in length and is located C-terminal to the catalytic domain of the large integrases (Fig. 1.2 C). So far, at least 10 conserved residues have been identified in the central domain, including three cysteines, two of which are required for $\phi$C31 integrase recombination (175, 176). In most large serine integrases, the central domain is followed by a non-conserved stretch of about 125 amino acids (Fig. 1.2 C). The

remainder of the protein consists of the short conserved domain of unknown function, a 30-amino acid sequence, and a C-terminus of varying length and sequence, from four to 200 amino acids (176). The serine phage integrases do not appear to require host accessory proteins like IHF, the integration host factor used by many recombinases such as lambda integrase (69, 101). The integrases from ϕC31 and TP901 both function *in vitro* and in mammalian cells without the addition of any cofactors (69, 70, 186, 193, 195). The large serine phage integrases appear to bind *attB*, *attP*, *attL*, and *attR* equally well, but only the integration reaction between *attB* and *attP* is seen *in vitro* (68, 175). However, an excisionase has been identified for TP901 integrase, though not yet for ϕC31, suggesting that a Xis protein may be necessary for an excision reaction mediated by a large serine integrase (24, 69).

**Serine recombinases with altered reaction specificity**

The elements that form recombination synaptic complexes are thought to determine the reaction specificity of site-specific recombinases. The recombinases themselves can likely catalyze intra- and intermolecular recombination between directly repeated or inversely repeated sites, meaning excision, inversion, and integration. The accessory DNA sequences (7, 95, 168, 176) and protein cofactors (95, 176) are thought to restrict recombination to a particular preferred reaction, intramolecular recombination between inversely oriented sites for inversion, as one example (95, 176). This theory is best demonstrated by the serine invertase/resolvases, in particular the Gin invertase and Tn*3* resolvase. These recombinases have been shown to complete only one type

of reaction *in vivo* and *in vitro*, but mutants have been isolated that can mediate all types of site-specific recombination reactions (7, 95, 176).

**Gin invertase.** The serine invertase Gin increases the host range of phage Mu by inversion of DNA called the G segment. The recombination site, *gix*, is a core site (Fig. 1.1 A, B) composed of two 12-bp inverted repeats for Gin binding surrounding an asymmetrical 2-bp crossover (96, 97). Adjacent to *gix* is a 60 bp enhancer sequence containing two 15-bp Fis binding sites (96, 99). The enhancer sequence position and orientation relative to *gix* are not specific for recombination, meaning the enhancer can be on either side of the core site at a variable distance. The inversion occurs between two inversely oriented *gix* sites on a single molecule, only on negatively supercoiled DNA substrate, and requires the host factor Fis, which binds to the enhancer sequence. Fis plus the enhancer is thought to stabilize the synaptic complex and both bend and unwind the DNA to promote cleavage for recombination (92, 96, 97). Fis-independent Gin mutants were isolated that had lost their recombination specificity; they could mediate recombination inter- as well as intramolecularly between direct and inverted repeats on linear, nicked, and supercoiled DNA substrates (95). The mutations are thought to have affected protein-protein interactions between subunits of Gin since the altered residues were not located within the active sites for cleavage and strand exchange. Fis bound to the DNA enhancer was not needed to position the DNA because the new Gin-Gin interaction was able to do so on its own, but these new interactions were more flexible than the usual synapse and allowed for the occurrence

of other types of recombination reactions (95). Thus, it is Fis bound to the DNA enhancer that restricts Gin-mediated recombination to the inversion reaction (95, 170, 177).

**Tn*3* resolvase.** Experiments with the Tn*3* resolvase have also yielded mutants similar to the Gin mutants. The Tn*3* resolvase does not use any host factors; instead, it has a complex 114 bp DNA recognition site, *res* (Fig. 1.5). The *res* site is composed of three subsites, each a core-type site that binds two resolvase subunits (67, 178). The cleavage and strand exchange reaction occurs only at subsite I, and the resolvase subunits bound to subsites II and III are thought to act in a manner similar to Fis bound to enhancer DNA with Gin. Only the resolvase subunits bound to each subsite I are involved in cleavage, and the other eight subunits bound to subsites II and III are involved in stabilizing the synapse, correctly orientating the DNA, and restricting the recombination to a resolution-type reaction (7, 168, 178). The Tn*3* resolvase mutants isolated could mediate recombination between a full *res* site and a subsite I, and one double mutant catalyzed a recombination between two subsite I (7). These mutants showed reduced substrate specificity like the Gin mutants (95, 170, 177); they could catalyze inversion as well as resolution with linear and supercoiled DNA substrates (7). The Gin invertase and Tn*3* resolvase mutant studies clearly support the idea that the accessory components of site-specific recombination reactions can determine substrate specificities and direct the type of reaction; thus, it is possible that all site-specific

recombinases could potentially mediate integration, excision (including resolution), and inversion reactions.

## TYROSINE RECOMBINASES

The tyrosine family presently encompasses over 300 recombinases, defined as such by sequence similarity (69). It is often referred to as the lambda integrase family, due to its most well known member. Tyrosine recombinases are only related to the serine recominases by function, not by sequence similarity (6, 66, 69). The catalytic domains of tyrosine recombinases are similar to two other enzyme groups: the type Ib topoisomerases that nick and religate DNA in order to remove supercoils, and the telomere resolvases that maintain DNA hairpins in linear replicons (28, 51, 68). A few examples of the functions of tyrosine recombinases serve in nature are the resolution of chromosomal and plasmid dimers, the inversion of DNA for phase variation, and the integration and excision of phage genomes (69). Tyrosine recombinase proteins usually contain two domains. The domain responsible for DNA core site recognition and binding is located in the N-terminal region (Fig. 1.6), though some tyrosine recombinases such as lambda integrase (Fig. 1.6 C) encode an additional DNA-binding domain N-terminal to the core site binding region (1, 53, 101, 207). The C-terminus encodes the catalytic domain, and residues of this domain are also involved in binding DNA at the site of recombination, such as the active tyrosine residue. The N-terminal core-binding domains are not conserved in sequence, but the crystal structures of tyrosine recombinases (Cre, XerD, and lambda integrase) have shown similarities in

protein folding (1). The tyrosine family can be separated into subgroups based on function and amino acid sequence alignment of their C-terminal catalytic domains (54, 132). However, unlike the serine family, the tyrosine recombinases do not cluster tightly together, forming more than seven subgroups with nearly half of the proteins left ungrouped (54, 132). The most divergent subclass includes the yeast invertases related to the FLP invertase of the 2 μm circle from *Saccharomyces cerevisiae*. These eukaryotic invertases are distant enough from the other tyrosine recombinases that they must be removed from alignment in order to generate a C-terminal domain evolutionary tree (54, 132). Other groupings of the tyrosine recombinases include the lambda subfamily of integrases, the P2 and P4 integrase subfamilies, the Rci shufflons, the Fim invertases, and the Xer subfamily of resolvases (54). The C-terminal catalytic domains of the tyrosine recombinase proteins share a conserved RKHRH pentad in addition to the active tyrosine (Fig. 1.6). The RHR and the tyrosine are considered the most important residues in the recombination active site since mutations lead to recombination-deficient proteins (57, 81, 115, 132, 140, 206). The first histidine is not completely conserved throughout the family, though present in more than 90% of tyrosine recombinases, and a tryptophan is sometimes present in place of the second histidine, like in Cre (69) (Fig. 1.6).

**Tyrosine recombinase reaction mechanism**

The tyrosine recombinase reaction mechanism differs from the serine recombinase mechanism in that two strands of DNA are exchanged and ligated to form

a Holliday junction intermediate prior to cleavage of the second pair of strands (Fig. 1.7). Dimers of the recombinases bind to the DNA recognition sites and cleave one strand of each DNA molecule with the 3' end bound to the protein by a phosphotyrosine linkage, leaving a 5' hydroxyl. The 5' hydroxyl then attacks the 3' phosphotyrosine linkage of the opposite DNA molecule, resulting in the formation of a Holliday junction intermediate. The recombinase attacking the complimentary DNA strands in the same process leads to resolution of the Holliday junction (66, 68, 131, 200). The tyrosine recombinases also differ from the serine family in structure of the recombination synapse. For the serine recombinases, the DNA is located on the outsides of the recombination complex, requiring a large rotation of the entire complex prior to ligation (66, 69, 168, 178, 200), but the DNA in a tyrosine recombinase reaction remains in the middle of the complex, requiring less movement of protein subunits to complete the reaction (66, 69, 131, 200). The tyrosine recombinase DNA recognition sites are most often core-type sites (Fig. 1.1 A) containing inverted repeats surrounding a 6 to 8 bp crossover region (Fig. 1.8), rather than the 2-bp crossover common to many serine recombinases (66, 69). Crossover between DNA recognition sites for the tyrosine family members often occurs between two simple core-type sites, such as for Cre, but, like with the serine family, some reactions require more complex DNA sites including accessory binding regions for extra recombinase subunits and/or protein cofactors (66, 69, 200). Recombinases such as Cre use only two simple core sites (*loxP*) for recombination (Fig. 1.1 A, Fig. 1.8 A) (200), but more complex family members like lambda integrase use a simple *attB* site in combination with a more complex *attP* site (Fig 1.8 C) that encodes accessory regions

for binding both protein cofactors and extra molecules of integrase (67, 69, 101). Recombinases like the XerCD resolvase can recombine more than one type of DNA recognition site (Fig. 1.8 B), with only some sites requiring the binding of cofactors (43, 44, 63). The DNA binding sites for tyrosine recombinases function like the sites of the serine family in directing and restricting recombination reactions. The wide range of tyrosine recombinase-mediated reactions is best demonstrated by three prototypical recombinases: the simple resolvase Cre, the multi-functional XerCD system, and the complex lambda integrase.

**Prototypical tyrosine recombinases**

**Cre resolvase.** Cre is arguably the most simple and well-studied tyrosine recombinase, though it does not branch into a particular subfamily with C-terminal phylogenetic analysis (54). It is a resolvase 343 amino acids in length (Fig. 1.6 A) which, in order to ensure stable inheritance of the P1 prophage, resolves multimers that can occur by homologous recombination (200). Though the purpose of Cre is to maintain stability of P1, high concentrations of Cre *in vivo* have been shown to actually destabilize the prophage, likely due to multimer formation by intermolecular recombination (144). Thus, the natural function of the recombination mechanism is strongly affected by Cre concentration—i.e. high Cre concentrations can overcome the natural regulation that restricts the recombinase to intramolecular recombination (144).

The DNA sites where Cre recombination occurs are called *loxP* (Fig. 1.8 A). They are simple core sites 34 bp in length, consisting of a pair of 13 bp perfect inverted

repeats, also known as recombination binding sites (RBEs), surrounding an 8 bp asymmetric spacer region. One Cre subunit binds to each RBE, and cleavage occurs on either end of the crossover region, the six base pairs in the center of the spacer region (200). The asymmetric spacer region sequence directs the order of strand cleavage, especially the base pairs in between the crossover region and the RBEs. In the traditional representation of a *loxP* site (Fig. 1.8 A), the A/T pair is shown on the left and the C/G on the right side of the *loxP* sequence. In that case, the Cre preferentially cleaves the bottom strand first, in between the crossover nucleotides and the G. Then, Cre shows a bias to cleave second between the crossover region and the A on the top strand to resolve the Holliday junction intermediate (85, 103, 104). Cre cleavage preference can be reversed if the A/T is replaced by G/C and vice versa. Additionally, Cre cleavage adjacent to the A residue is more efficient than is cleavage adjacent to the G residue. The more efficient reaction occurs second, resolving the Holliday junction intermediate, instead of initiating recombination. The less efficient cleavage is thought to occur first so that the more robust reaction can then bias the resolution to produce more recombination products rather than starting substrate (104).

The traditional mechanism used to illustrate Cre recombination is called the "branch migration" model (see ref (200) for review). In this model, a recombination synapse is formed by two Cre-*loxP* complexes. One Cre subunit on each DNA molecule cleaves the DNA and forms a covalent 3'-phosphotyrosine bond, leaving a 5' hydroxyl. Ligation occurs when the 5'-hydoxyl attacks the 3'-phosphotyrosine linkage of the opposite DNA molecule to form a Holliday junction intemediate. A branch migration

of the DNA-protein complex allows the other two Cre subunits to similarly attack the DNA for resolution of the Holliday junction intermediate and completion of the recombination reaction. In the 1990s, however, a new model was adopted to describe the tyrosine recombinase reaction. In the "strand-swapping isomerization" model, 2-3 bp of the cleaved DNA melts from its own complimentary strand to anneal to the complimentary strand of the opposite DNA molecule in preparation for strand exchange (Fig. 1.7), and an isomerization step occurs rather than branch migration prior to Holliday junction resolution (68, 131). Subsequent experiments in which Cre-*loxP* structures were examined greatly support this newer model for Cre recombination. This model explains the importance for crossover region sequence identity between recombination sites, and describes how the protein-DNA complex could easily alter conformation to allow strand exchange and position the unbroken DNA strands and second pair of Cre subunits to complete the reaction without major changes in quaternary structure (72, 200).

Three Cre-*loxP* crystal structures have been resolved that support the "strand-swapping isomerization" model (62, 71, 72, 200). The first Cre-*loxP* structure was of Cre bound to a suicide DNA substrate (72). A symmetrical *loxP* site was constructed containing nicks on the other side of what would become the 5'-hydroxyl-containing cytidine after Cre-mediated cleavage (Fig. 1.8 A). The nick caused release of the cytidine after DNA cleavage, so the complex lacked both the 5'-hydroxyl and the length of DNA necessary to complete strand exchange (72, 200). The structure revealed the covalent linkage of two Cre subunits to the scissile phosphates by 3'-phosphotyrosines,

as expected, and the 5' DNA, where the cytidines would have been, had separated from their complementary DNA strand. Models based on the crystal structure supported the isomerization model by revealing that the unpaired 5' DNA would have been in position to bind to the complementary strand of the opposite DNA molecule, had the missing 5' cytidine been present (72, 200). Three of the six crossover base pairs are exchanged in formation of the Holliday junction, while the other three are exchanged in the resolution of the junction (Fig. 1.7), all without much movement by the protein, so the quaternary structure of the complex remains unaffected. Once the Cre-DNA synapse forms, the cleavage of one DNA strand over the other would be preferred due to the positioning of the DNA. As stated above, Cre tends to preferentially cleave the "G-C" end of the loxP site first (Fig. 1.8 A), so the synaptic configuration that positions the DNA for that cleavage may be favored (85, 103, 104). The other two crystallized Cre-*loxP* structures were of Cre-bound Holliday junction intermediates and give further understanding of the isomerization step in the recombination model (62, 72, 200). One was constructed using 8 strands of DNA that assembled as a symmetrical *loxP* site forming a Holliday junction with four nicks, so branch migration could theoretically occur (62, 200). The second Cre-HJ complex had an artificially fixed branch point, but the resulting two crystal structures are essentially superimposable, suggesting that the artificial changes did not cause alteration of the natural complex (62, 200). In addition, other than the positioning of the DNA within the synaptic complex, all three Cre-DNA structures are basically the same. There are both N-terminal and C-terminal protein-protein interactions of Cre subunits. The N-terminal domains between the two Cre subunits at

each core site interact in a nearly four-fold symmetrical conformation that does not change with strand exchange; therefore, the N-terminal protein-protein interactions are thought to lend structural stability to the synaptic complex (200). A C-terminal helix, called “helix-N”, associates with a C-terminal “pocket” in the adjacent subunit to form a four subunit cyclical exchange (200). The four DNA arms in the synaptic complex form two acute and obtuse angles. The strands activated for cleavage and crossover are called the crossing strands, while the complementary strands not currently involved are referred to as the continuous strands (Fig. 1.7). During the isomerization step in recombination, the DNA backbone twists such that the continuous strands become the crossover strands and vice versa, leading to the second half of the recombination reaction, resolution of the Holliday junction intermediate (Fig. 1.7). This twist allows the full recombination reaction to occur without large changes in quaternary structure—such as the breaking and re-forming of protein-protein interactions that would be necessary for a branch migration of this complex (200). The Cre-*loxP* reaction provides the best and most complete model for recombination by a tyrosine site-specific recombinase, but the differences in sequence and structure between the many tyrosine recombinases and their DNA recognition sites indicate that their recombination mechanisms are unlikely to be completely identical to Cre-*loxP* (66, 200).

**XerCD recombinase.** XerCD is a tyrosine recombinase that resolves chromosomal or plasmid dimers that occur by homologous recombination in order to ensure proper segregation during cell division. While the XerCD from *Escherichia coli* is the best

studied, XerCD homologs have been found in almost all bacteria, as well as in at least one archeal species, and are highly conserved among the Enterobacteriaceae (107, 187). The *E. coli* XerC and XerD (Fig. 1.6 B) are both 298 amino acids in length and share 37% amino acid identity (20, 45, 187).

*XerCD as a resolvase*. There are three types of DNA recognition sites for XerCD in *E. coli*: *dif*, *cer*, and *psi* (Fig. 1.8 B). The *dif* sites are on the chromosome for resolution of chromosomal dimers (19, 20), *cer* sites are located on ColE1-type plasmids (20, 45), and *psi* sites are on pSC101 plasmid derivatives (46). ColE1 and pSC101 plasmids have evolved to take advantage of the existing XerCD machinery to resolve multimers, thereby ensuring stable plasmid inheritance (189). Recombination at *cer* and *psi* is restricted to sites in direct repeat on the same DNA molecule (Fig. 1.1 C), so that multimers are resolved rather than created, and requires accessory proteins (Fig. 1.8 B) as well as supercoiled DNA substrates (43, 44, 63). PepA, an aminopeptidase and DNA-binding protein, is required for XerCD recombination at both *cer* and *psi* sites and is important for the physical structure of the recombination synapse (5, 44). ArgR, an arginine repressor, enhances the XerCD-*cer* reaction, but XerCD-*psi* uses ArcA, an anaerobic repressor (3, 43). ArgR and ArcA are thought to provide less vital structural support for the recombination complex. ArgR is essential for recombination at *cer in vitro* unless an excess of PepA is added, due to cooperative binding of ArgR and PepA, and ArcA acts similarly for XerCD-*psi* (3, 43, 44).

Instead of requiring intramolecular supercoiled substrates and accessory factors that bind and wind the DNA to prevent dimer formation rather than resolution, XerCD-*dif* recombination is restricted to the time and place of cell division, partly by FtsK (8). FtsK is necessary for both XerCD-mediated chromosomal dimer resolution and chromosome segregation and is localized to the septum at cell division (179, 199, 203). Both intra- and intermolecular recombination are possible between *dif* sites on supercoiled, linear, or open circle substrates (8, 20). Although no specific accessory sequences are essential, the chromosomal site must be present within a 30 kb area of the terminus region in *E. coli*, and the sequences surrounding the *dif* site need to have opposite polarity (145, 192). This means that recombination does not occur when the DNA on one side of the *dif* site is inverted experimentally, unless the DNA on the opposite site of *dif* is also inverted (145). XerCD-*dif* recombination is also linked to homologous recombination in that no chromosomal dimer resolution is detected *in vivo* in a strain lacking RecA; thus, XerCD-mediated recombination at *dif* only occurs in the presence of a chromosomal dimer (153).

All three types of DNA recognition sites are arranged as a core sequence consisting of 11 bp for XerC and XerD binding surrounding a 6-8 bp spacer (6 bp for *dif* and *psi* and 8 bp for *cer*) (Fig. 1.8 B), but a 150-200 bp region upstream of the XerC binding site is also necessary for *cer* and *psi* recombination for interaction with the accessory proteins (Fig. 1.8 B) (63, 107, 199). Unlike *loxP* and most other site-specific recombinase DNA recognition sites, the core sequence is not an inverted repeat since XerC and XerD have different recognition sequences (Fig. 1.8 B) (63). XerCD, like Cre,

shows a preference in the order of strand cleavage during recombination. The strand with the XerC binding site 5' to the XerD binding site is designated the top strand (Fig. 1.8 B). In *dif* site recombination, FtsK promotes XerD-mediated cleavage of the bottom strand to form the Holliday junction intermediate; then, XerC resolves the intermediate by cutting the top strand (8, 199).

For *cer* and *psi* site recombination, XerC has been reported to cleave the top strand first (43, 44, 122), but this assertion is not fully supported by data in the literature. *In vitro* studies with *psi* have shown that a wild-type XerD plus a catalytically inactive XerC mutant can form and resolve Holliday junction intermediates to produce the original substrates, but a wild-type XerC plus a XerD mutant can only form Holliday junction intermediates and not resolve them (44). Thus, XerD can cleave first and second under certain conditions at *psi* sites, but XerC is less flexible. It is possible that XerC is only able to produce dead-end intermediates when cleaving first. Another *in vitro* study demonstrated that wild-type XerD with a XerC mutant was unable to resolve a synthesized *psi* site Holliday junction intermediate, while XerC plus a mutant XerD could resolve that same intermediate (78). This supports the idea that XerC mediates Holliday junction resolution rather than formation for *psi* sites, as it does in *dif* site recombination.

In *cer* site recombination, XerD-mediated bottom strand cleavage has not been observed *in vitro*, and XerC-mediated cleavage produces Holliday junction intermediates that do not resolve, not even to form the original substrates (44). *In vivo* experiments using an excess of XerC showed *cer* resolution products as well as XerC-mediated

Holliday junction intermediates. One possible explanation for this result is that XerD-mediated Holliday junction intermediates were cleaved by XerC to form the resolution products, and that the excess of XerC also caused XerC-mediated intermediates to accumulate that XerD could not be resolve (122). It has been hypothesized that other *E. coli* resolvases and/or DNA replication are necessary to resolve XerC-*cer* Holliday junction intermediates *in vivo* (44, 122), but XerD cleavage may simply need to occur first in *cer* site recombination, as it does in *dif* site recombination. There may be another accessory factor required for the complete *cer* recombination reaction, however. Recombination between *dif* sites in the absence of FtsK yields Holliday junction intermediates formed by XerC cleavage that can only resolve to form the original substrates (13). Similarly, XerC has been shown to resolve Holliday junctions with hybrid *cer* sites *in vitro* without any accessory factors, but only to form what would have been the original substrates rather than the recombinant products (5). This also is similar to Cre-*loxP* recombination in that XerC cleavage is more robust than XerD cleavage, like *loxP* cleavage adjacent to the A versus the G residue; thus, if XerD cleavage required an accessory factor that was not present, XerC cleavage would seem to occur first *in vitro* (104). Whether XerC cleavage occurs first or second in *cer* and *psi* site recombination, there may be an unidentified accessory factor necessary for XerD cleavage; though it is not the same factor required for Xer-*dif* because recombination at *cer* and *psi* does not require FtsK *in vivo* (153). Since it is widely thought that resolution of a Holliday junction intermediate to form either the original substrates or the recombinant products depends on a structural change in the intermediate, it is possible

that an unknown cofactor is required to effect that conformational change (131, 200). It will be very interesting to see the intricacies of XerCD recombination at *cer* and *psi* further revealed in the future.

*XerCD as an integrase*. Recently, XerCD, like the Gin invertase mutants, has been shown to mediate intermolecular recombination as an integrase (7, 88, 97, 199). CTXϕ is an integrated phage found in strains of *Vibrio cholerae* and carries the cholera enterotoxin, the major toxin responsible for causing disease (49, 199). Like ColE1 and pSC101 plasmid derivatives, CTXϕ has taken advantage of the XerCD system to mediate its own transfer. XerCD-mediated irreversible integration of CTXϕ has been shown to occur at *dif1*, one of two *dif* sites found in *Vibrio cholerae* (88, 199). The CTXϕ phage genome is a circular single strand of DNA, with two putative *attP*-type sites located 90 bp apart. These sites can bind each other to form a functional double-stranded *attP* site to which XerC and XerD can bind and form a recombination synapse with the *dif1* site on chromosome 1 of *V. cholerae* (199). XerC-mediated cleavage then forms a Holliday junction intermediate that cannot resolve, as resolution would result in linearization of the chromosome, and host-mediated DNA replication completes the integration reaction (199). XerD is necessary for CTXϕ integration but does not need to be catalytically active, suggesting a structural role, and, unlike a typical XerCD-*dif* recombination reaction, CTXϕ integration does not require FtsK since XerD cleavage does not occur (88, 199). A synthetic single-stranded CTXϕ *attP* was successfully recombined with a *dif1* site *in vitro* using *E. coli* XerCD. Recombination in *E. coli* was

demonstrated with both *E. coli* and *V. cholerae* XerCD using a plasmid-encoded *dif1* and an *attP* introduced by a conjugative plasmid incapable of replication in the strain (199). The CTXϕ integration demonstrates an advantageous strategy for phage and other mobile elements to insert into genomes using the ubiquitous XerCD system, aided by the high similarity among the XerCD proteins in different bacterial species. Two prophages found in pathogenic *E. coli* isolates contain possible *attP* sequences similar to CTXϕ, and it is likely that more examples will be found in the near future (199).

**Lambda integrase**. The lambda integrase is considered the most complex tyrosine recombinase because of its multiple regulation mechanisms and use of several accessory factors, best demonstrated by the multipart *attP* recombination site (Fig. 1.8C). While the less complex tyrosine recombinases such as Cre and FLP, an invertase from the 2 μm circle found in the yeast *Saccharomyces cerevisiae*, each use two almost identical core sites (Fig. 1.1 A) of less then 50 bp in size for recombination, the integrase members of the family require larger and more diverse DNA recombination sites (69). The *attB* simple site is a core site composed of two inverted repeats that surround a 6-8 bp overlap sequence (Fig. 1.8 C). The overlap region shares sequence identity with *attP* where the DNA crossover occurs, as in Cre-*loxP* recombination, but the *attP* sites (Fig. 1.8 C) of the integrases are longer and more complex than the *attB* sites (69). The *attP* sites have, besides a simple core-type site, regions for binding accessory factors, like those of *cer* and *psi* sites for XerCD (43, 44,

63), and additional molecules of integrase (Fig. 1.8 C), somewhat similar to the *res* sites of resolvases from the serine family (67, 69, 178).

Lambda integrase, responsible for integrating (and excising) (Fig. 1.1 C) bacteriophage lambda DNA into a host's genome, is considered the most complex tyrosine recombinase and, unlike Cre, uses several cofactors for recombination (69, 159, 210). Recombinase accessory factors can aid recombination in several different ways. They can direct the binding of integrase to the DNA or bind the recombinase to induce a conformational change that allows binding of the DNA substrate. Accessory proteins can also bend a DNA substrate to promote formation of a recombination synapse (63). Therefore, an accessory factor may bind the recombinase only, the DNA only, or may be involved in both protein-protein and protein-DNA interactions, and the lambda integrase uses all of these strategies (69, 101, 159, 210).

The lambda integrase protein is 356 amino acids in length and consists of three domains (Fig 1.6C) (1, 9, 69). The N-terminal DNA-binding domain and C-terminal catalytic domain similar to Cre and XerD are present, but lambda integrase also possesses an additional N-terminal domain for binding accessory DNA sites, located 70 to 150 bp away from the core site in *attP* (Fig. 1.8 C) (1, 53, 101, 207). The central domain is similar in structure but not in sequence to the N-terminal domains of both Cre and XerD (1). The lambda Int N-terminal region has a high binding affinity for the arm-type binding sites on either end of *attP*. The accessory factor IHF bends the DNA to promote the lower-affinity binding of lambda Int to the core site (69, 125, 159); then, the arm-type binding of the integrase to the DNA physically regulates recombination by

structurally preventing much movement of the complex and provides support (125, 201). The arm binding is necessary for recombination and directs strand exchange order to favor product formation, thereby promoting a complete recombination reaction rather than the reforming of the initial substrates (18, 150, 201). Several cofactors aid in excision by binding *attR*, such as the host factors IHF (101, 125, 159) and Fis, the host-encoded factor for inversion stimulation also used by the serine invertase Gin (92, 159), and the Xis excisionase (69, 101, 210). Both *attR* and *attL* encode binding sites for IHF as well as integrase arm-type sites (Fig. 1.8 C) (9, 69). More regulatory mechanisms designed to influence recombination toward integration or excision of bacteriophage lambda include transcriptional controls of the integrase and host-mediated degradation of phage-encoded proteins, such as cII and Xis (10, 105). The relative amounts of lambda Int and Xis in a cell affect the recombination reaction choice of integration or excision (105). The ATP-dependent proteases Lon and FtsH in *E. coli* have been shown to degrade Xis *in vivo*. Excess Xis inhibits the integration of lambda DNA; therefore, Xis is degraded to promote integration (105). The controls on recombinase (and Xis) expression added to the multiple strategies in using various accessory factors plus the multipart DNA recognition sites show the amazing complexity used to regulate lambda integrase-mediated recombination.

### *Cis* vs. *trans* cleavage

A controversial debate regarding tyrosine recombinases, especially lambda integrase, involves whether a recombinase cleaves core site DNA *in cis* or *in trans*.

FLP, the yeast invertase, has been shown to cleave in *trans*. This means that the active tyrosine nucleophile that attacks the DNA is part of the recombinase subunit bound adjacent to the monomer including the R-H-R conserved residues in the recombination synpase (30, 102). Most data have suggested cleavage *in cis* for the other tyrosine recombinases, reflecting the phylogenetic distance between the FLP-type invertases and the rest of the family. Successful alignment of the tyrosine recombinase catalytic domains has required the exclusion of the yeast invertases, and analysis of the DNA-binding domain protein folding patterns and sequences revealed no similarity between the yeast and the prokaryotic or archeal N-terminal domains, suggesting that the FLP-like invertases are only distant relatives of the rest of the tyrosine recombinases (54, 190). Structural and biochemical studies have suggested that Cre works *in cis* study (66, 72, 200), though one study (172) claims evidence for *in trans* cleavage. XerCD experimental evidence indicates cleavage *in cis;* studies combining a XerC active tyrosine mutant with a XerC mutated at one of the conserved arginine residues could not perform recombination *in vitro*, since these important residues need to be present on the same protein for XerC to cleave (5). Also, the crystal structure of XerD supports an *in cis* cleavage mechanism since the active tyrosine is located close to the other active site residues, unlike in FLP (187). There is evidence for both cleavage *in cis* and *in trans* for lambda integrase (80, 133). The crystal structure of lambda integrase not bound to DNA suggests cleavage *in trans* because the active tyrosine is not near the other active site residues and significant conformational change would be necessary before cleavage *in cis* could occur (1, 100). However, structures of lambda integrase

bound to DNA illustrate an alternative form of the protein that would allow cleavage *in cis* and indicate that the 8 C-terminal residues of the protein can alter the position of the catalytic tyrosine. This protein tail distances the tyrosine from the cleavage site in a catalytically inactive form of the protein, but interacts with an adjacent integrase *in trans* in an active integrase molecule, freeing the tyrosine to associate with the R-H-R residues in the active site to attack the DNA (1). Presently, the FLP-like yeast invertases are the only tyrosine recombinases that definitively cleave *in trans*. However, whether a recombination reaction takes place *in cis* or *in trans* does not itself necessarily indicate a catalytic difference (5). For example, if lambda integrase can truly cleave both *in cis* and *in trans,* it simply represents another regulation mechanism for site-specific recombinases. For example, the tyrosine may not be in position to attack until a conformational change, such as a bend in the DNA or in the protein itself, is induced by an accessory factor. All members of a recombination synapse may need to be present in a particular orientation for attack *in cis*, which seems to be the case for lambda integrase (1). Or, as is for FLP, one subunit may leave the cleavage site unblocked to allow an adjacent subunit to lend its tyrosine *in trans* because the correct conformation cannot occur *in cis* with that substrate (5, 158). Evidence for cleavage *in cis* or *in trans* should not be considered completely definitive because the complex and varied nature of site-specific recombinases has shown that any aspect of a recombination reaction may be used for regulation.

## INTEGRONS AND THEIR INTEGRASES

Integrons are naturally occurring genetic elements found on plasmids and on the chromosomes of bacteria. They contain a site-specific recombination system mediated by an integrase (IntI) of the tyrosine recombinase family where gene cassettes, small mobile DNA elements, are excised or inserted at a recombination site, *attI*. Integrons were first discovered in multiresistant strains of Enterobacteriacea (180). Gene cassettes often encode antibiotic resistance, and integrons were first distinguished as a major mechanism through which these resistances are disseminated throughout bacterial populations by horizontal gene transfer (119). More recently, however, hundreds of gene cassettes have been detected in environmental DNA, most encoding proteins of unknown function (183), and new integrases have been identified on the chromosomes of dozens of diverse bacterial species (118). This shows that a large reservoir of integron gene cassettes exists in the environment and that integrons carry much more than just antibiotic resistance genes in nature (124, 183).

### IntI similarities/differences to other tyrosine recombinases

Integron integrases form a distinct family within the tyrosine recombinases, most closely related to the XerCD family (41, 163). Integron integrases catalyze several site-specific recombination reactions including intermolecular integration (Fig 1.1 C), such as cointegrate formation and gene cassette integration, and intramolecular excision, such as cointegrate resolution and gene cassette excision (38). Integron integrase-mediated inversion (Fig. 1.1 B) has not been observed, likely due to the unique DNA recognitions

sites controlling recombination directionality (i.e. the recombination sites are only found in direct repeat). The integron integrases, like other tyrosine recombinases, require the RKHRH pentad and the conserved catalytic tyrosine for recombination (Fig. 1.6 D). Mutation of these conserved residues results in an integrase unable to catalyze excision or integration (40, 65, 123). Integron integrases contain a unique IntI region within their catalytic domain (Fig. 1.6 D) that encodes conserved residues also required for recombination (123, 128). The DNA recombination sites, *attI* and *attC*, of integron integrases are complex and very diverse. *attC* sites in particular vary greatly in size (57-141 bp) and sequence but are recognized by multiple IntI proteins (41, 118, 154). Integron integrases can recombine *att* sites with little to no homology in the crossover region of their core sites (Fig. 1.1 A), when even small differences in this region strongly reduce recombination mediated by other tyrosine recombinases (68, 101, 104). In addition, there is only evidence for cleavage at one position within *attI* and *attC* sites, indicating either a single strand exchange or cleavage of both DNA strands at the same position (185). Integron integrases seem to have low recombination activity compared to other tyrosine recombinases because little recombination is observed experimentally without using excess IntI (36, 116). In addition, no accessory proteins aiding IntI-mediated recombination have yet been identified. Because of the many different recombination reactions catalyzed by integron integrases and the complex DNA recognition sites, it is likely that these integrases have multiple regulatory mechanisms similar to those of lambda integrase and/or XerCD, but little is presently known about

how integron integrase-mediated recombination is regulated, largely due to a lack of successful *in vitro* experiments (82, 118).

**General characteristics of integrons**

Integrons are composed of an integrase gene, *intI*, an *attI* integrase recognition site, and gene cassettes (Fig. 1.9 A). The gene cassettes are small mobile DNA elements that consist of a recombination site (*attC*) and an open reading frame, often encoding antibiotic resistance (73, 76, 116, 180). The cassettes are transcribed by a common promoter $P_c$ located just upstream of the *attI* site, though only precisely identified in class 1 integrons, and the integrase is divergently transcribed from the gene cassettes (Fig. 1.9 A) (35, 180). The non-mobile integron regions, the integrase gene and *attI* site, are known as the 5' conserved segment (or conserved sequence). Class 1 integrons also have a non-mobile 3' conserved sequence (Fig. 1.9 B), usually containing at least three genes, *sulI*, *qacEΔ1*, and *orf5* (76).

Integrase-mediated recombination effects gene cassette movement between integrons by two different mechanisms: excision and integration (Fig. 1.10) or the formation and resolution of cointegrates (Fig. 1.11). In excision and integration, gene cassettes are excised into covalently closed circles and then inserted into the *attI* site of a different integron in the same cell (36, 37). This process is called "cassette capture" (Fig. 1.10). The second mechanism, the formation and resolution of cointegrates, occurs between plasmid-carried integrons, and the integrase mediates RecA-independent recombination between the donor and recipient plasmids at the *attI* and/or

*attC* sites to form a cointegrate plasmid (Fig. 1.11). Unlike cointegrates formed by replicative transposition (79), there is no replication involved in integron integrase-mediated cointegrate formation. The cointegrate plasmid can then be resolved by the integrase to create two separate plasmids, either of which can contain the cassette(s) adjacent to the insertion site (Fig. 1.11) (38, 82, 116, 157). The cassette capture products resulting from the two different mechanisms are identical in sequence; therefore, it cannot be determined which mechanism was used to generate a product (38).

Integrons are classified by their integrase sequence, with >98% amino acid identity among integrases within a class. Although over 20 integron classes have been identified, only 5 have been well characterized (39, 118, 128). Integrons are also separated into types based on location: within a mobile element or on a chromosome. Integrons located within mobile elements are sometimes referred to as MRIs (multi-resistance integrons) because these integrons are most associated with antibiotic resistance (15, 163). This term is somewhat misleading since not all mobile element-associated integrons carry multiple antibiotic resistances (and chromosomal integrons can carry antibiotic resistances, as well), but they are all found on a plasmid and/or within a transposon. Thus, I refer to these mobile element-associated integrons as MIs. The other type of integron is located on a chromosome and not within a mobile element. These integrons are often referred to as superintegrons (SIs), but this term is controversial (75). SIs are distinguished by encoding over 20 gene cassettes with closely related *attC* sequences (118). Some chromosomal integrons not located on

transposons encode no gene cassettes (106) or only a few gene cassettes with heterogeneous *attC* sites (52); therefore, they are neither SIs nor MIs. In addition, it is not yet known whether some identified chromosomal integrases are located within a mobile element. Thus, I refer to chromosomal integrons as CIs, unless they are known to encode over 20 gene cassettes (SIs) or are known to be within a mobile element (MIs). Most MIs are separated into class 1, 2, and 3 integrons (Fig. 1.9 B-D). The prototypical SI is the integron encoding the integrase VchIntIA and nearly 200 gene cassettes within the small chromosome of *Vibrio cholerae* (Fig. 1.9 E) (120), and the integron on the *Shewanella oneidensis* chromosome encoding three heterogeneous *attC* sites is an example of a CI (Fig. 1.9 F) (52).

**Dissemination of antibiotic resistance**

Class 1 MIs, encoding the integrase IntI1, carry multiple antibiotic resistances in the widest range of different bacterial species and environments of any other integron class by far (55, 118). It is the presence of these multiple antibiotic resistances that led to the initial discovery of integrons (26, 76, 180).

Bacterial resistance to antibiotics has become a major challenge in healthcare around the globe. Acquisition of antibiotic resistance genes by horizontal gene transfer has been implicated as the major cause of dissemination of these resistances throughout bacterial populations (48, 119). Once antibiotic resistance genes are introduced into a population, they are not easily lost, even in absence of selection (119, 165). One example of the persistence of these genes is the prevalence of

chloramphenicol resistance in groundwater enteric isolates after a decade of little to no use of the antibiotic (165). Integrons are one major genetic system through which multigene antibiotic resistance can be transferred. In fact, integrons are the most prevalent genetic system through which multiple resistances are spread at one time in the Enterobacteriaceae (119). Integrons were first named in 1989 (180), and over 75 different integron-associated antibiotic resistance gene cassettes had already been identified by 1999 (119). Integrons carry gene cassette arrays encoding multiple antibiotic resistances and are associated with multiresistant Enterobacteriaceae isolated from hospitals and the environment (61, 109, 112, 121, 126, 128, 142, 166, 196, 211).

**Multiresistant clinical strains carry integrons.** Clinical Enterobacteriaceae isolates from around the world have been tested for the presence of multiple antibiotic resistances and the presence of IntI1 from class 1 integrons. A statistically significant correlation has been found between multidrug resistance and the presence of IntI1 in clinical isolates (109). In one study, sulfamethoxazole was the most prevalent resistance, present in approximately 90% of class 1 integrase-containing strains (110). This is unsurprising, since a non-mobile *sul1* gene is usually found associated with the 3' CS of class 1 integrons (Fig. 1.9 B). Gene cassettes most often seen in the class 1 integrons found in clinical strains are those conferring resistance to older aminoglycosides like kanamycin and streptomycin, demonstrating the persistence of these genes (29, 108, 167, 205). However, resistances to newer antimicrobials are being transferred by integrons, as well. *qnr* genes encoding quinolone resistance have

been found associated with class 1 integrons in *Klebsiella pneumoniae* and *E. coli* isolates (89, 197).

**Resistance gene cassettes and integrases in environmental samples.** PCR of environmental DNA has identified integrases and antibiotic resistance gene cassettes associated with integrons in diverse environments such as in soil and in the fecal matter of livestock and companion animals (61, 126, 128, 183). Although originally discovered within Gram-negative Enterobacteriaceae, studies have identified integrons within both low-GC and high-GC Gram-positive bacteria and show that Gram-positive bacteria may represent a larger reservoir for integrons than Gram-negative bacteria in certain ecosystems (33, 126, 127). In addition, more and more chromosomal integrons, some encoding hundreds of types of gene cassettes including some for antibiotic resistance, are being identified within a wide range of different bacterial species, from *Vibrio fischeri* to *Geobacter metallireducens* and *Treponema denticola* (118). Thus, integrons are more widespread than previously imagined and may carry antibiotic resistances between diverse bacterial species.

**Integron resistance gene cassettes prior to widespread use of antibiotics.** Antibiotic resistance genes were present within integrons prior to widespread use of antibiotics. Isolated in Japan in the 1950s, plasmid NR1 (R100) encodes the Tn*21* class 1 integron, In2 (Fig. 1.9 B), containing an *aadA1* cassette that confers streptomycin/spectinomycin resistance as well as *sul1* (113, 143, 204). The plasmid

pSa was isolated in the 1960s and encodes *sul1* and two gene cassettes, *aacA4* and *aadA2*, for kanamycin/gentamycin and streptomycin/spectinomycin resistance, respectively (143, 204). Several other plasmids were isolated during the 1960s and early 1970s containing class 1 integrons, such as R388 and R46 (26, 77, 143). Sulfonamides were used for approximately 10 years prior to use of penicillin; therefore, selection for integrons encoding *sul1* would have occurred during that time (143). Gene cassettes encoding antibiotic resistance homologs have even been discovered in a chromosomal integron within a strain of *Vibrio metschnikovii* isolated in 1888 (162), suggesting that some antibiotic resistance genes could originally have had other functions (160). The presence of integrons prior to use of antibiotics in medicine emphasizes that integrons are not solely important for disseminating antibiotic resistances.

**Spread of integrons and cassettes by horizontal gene transfer.** The dissemination of integrons is accomplished by horizontal gene transfer. Horizontal gene transfer is thought to be the main mechanism through which bacteria evolve (87, 134) and is how pathogens can acquire virulence factors such as toxins, as well as antibiotic resistances (Frost et al 2005). In fact, approximately 25% of most bacterial genomes are thought to have originated from a different cell (124, 134). MIs, particularly class 1 integrons, have multiple methods for horizontal transfer besides simple gene cassette transfer between integrons (Figs. 1.10, 1.11). 1. Most class 1 integrons are defective transposons thought able to move when transposition genes are supplied *in trans* (Fig. 1.9 B) (25,

180). In addition, many integrons, not just those of class 1, are located on larger mobile transposons (80), like Tn*7* (Fig. 1.9 C) (83) and Tn*21* (Fig. 1.9 B) (113). Tn*21*-like class 1 integrons and Tn*7*-like class 2 integrons are widely distributed in nature (83, 126, 211). 2. MIs, whether or not located on a larger transposon, are usually on large conjugative plasmids that can move between bacterial cells (22, 76). A 3.8 kb class 1 integron located on a plasmid in *Corynebacterium glutamicum* was found to be identical, with the exception of two base pairs, to an integron carried on a plasmid from *Pseudomonas aeruginosa*, demonstrating the possibility of wide-range horizontal transfer of an integron between distantly related organisms (127). 3. Homologous recombination is likely to play a large part in the exchange of arrays of gene cassettes when more than one integron is present in a cell (141). This would occur most often between class 1 integrons since they usually share long lengths of sequence at the 5' and 3' ends, but there is sufficient homology for this to occur with other classes of MIs (or sometimes between different classes) (141). An R388-like plasmid was isolated that contained a class 1 integron differing from In3 on R388 by encoding a completely different cassette array and a different version of the $P_c$ promoter (Fig. 1.9 A, B). This integron likely formed by homologous recombination between In3 and a second integron at the 5' and 3' conserved sequences to form this R388-like plasmid—which can spread its new antibiotic resistance gene cassette array to other bacterial cells by conjugation (see number 2 above) (141). With so many mechanisms for horizontal gene transfer, it is not surprising that class 1 integrons, and other MIs, have been found in diverse environments and bacterial species.

**Observing cassette acquisition *in vivo* requires excess integrase**

All integrons can theoretically capture gene cassettes by either the integration and excision mechanism (Fig. 1.10) or by the formation and resolution of cointegrates (Fig. 1.11), but chromosomal integrons are unlikely to form cointegrates because chromosomal dimers are deleterious to bacteria (107). Though excision/integration is likely to be highly preferred by chromosomal integrons, it is difficult to determine which mechanism plasmid-carried integrons favor since the final products of the two mechanisms are identical. The excision/integration mechanism is stated in the literature to be the major mechanism by which cassette acquisition occurs (38, 76) (the cointegrate formation/resolution mechanism is often not even mentioned), but this conclusion seems to be based on only one experiment using excess integrase (38). In fact, excess integrase has always been used in experiments observing cassette capture.

The initial integron recombination studies were conducted by Martinez and de la Cruz on class 1 integrons from Tn*21* and plasmid R388 (116, 117). They showed that pACYC184 derivatives carrying the integron region from Tn*21* would recombine with R388 to form cointegrate plasmids. They then determined that the cointegration frequency greatly increased when integrase was supplied in *trans* on a third plasmid (116). This began the use of vectors like pSU2056, a pUC9-based plasmid encoding *intI1* under a lac promoter, to increase integrase-mediated recombination in the laboratory (116). The excess integrase allows excision/integration and cointegrate formation/resolution to readily occur, while no integration products or resolved

cointegrates are detected without excess integrase (36) (and see Chapters 2, 3). In fact, neither excision nor integration of gene cassettes have been detected without excess integrase, but cointegrates have been observed to form when integrase is expressed under the natural $P_{int}$ promoter (38, 73, 116, 117).

Collis and Hall examined a plasmid derivative of pSa that encoded an integron with *aacA4* and *aadA2* gene cassettes in a *recA*$^-$ *E. coli* strain also containing pSU2056 to provide excess integrase *in trans*; cointegrates seen in their study were formed of two identical plasmids (38). They concluded that cointegrate formation and resolution (Fig. 1.11) could not be responsible for most of the excision of *aadA2* observed, because that mechanism should result in an equal number of plasmids carrying a duplicate *aadA2* (38). However, since excess integrase was present and since integration and excision would also have been occurring in the cells, other factors could have affected the ratio of cointegrates to the number of plasmids carrying a duplicate *aadA2* cassette. 1. The use of excess integrase could have increased the preference for the excision/integration mechanism in recombination, meaning more excision of *aadA2*. Excess recombinase *in vivo* can affect reaction preferences and cause unexpected outcomes for tyrosine recombinases (105, 122, 144). High concentrations of lambda integrase cause an increase in integration versus excision (105), and high concentrations of Cre cause an increase in intermolecular versus intramolecular recombination (144). 2. Having directly repeated gene cassettes could have affected recombination mechanism choice. For example, an integron containing two *aadA2* cassettes in tandem with identical *attC* sites may induce excision of the extra cassette. In integrons containing gene cassette

duplicates, the two cassettes are not usually found directly adjacent to one another but are separated by one or more cassettes (16, 36, 73, 76, 152, 182). In the same Collis and Hall study, plasmids encoding two *aadA2* cassettes in tandem were detected by PCR, but the attempts to isolate the plasmids were unsuccessful, suggesting they were rare (38). 3. Using only one plasmid for recombination could affect the number of cointegrates; cointegrates may resolve more quickly when the cointegrate formed is actually a plasmid dimer. Several mechanisms exist to resolve dimers because they promote plasmid instability (188).

The above studies conclusively show that only cointegrates (and not cassette capture products) are detected without excess integrase and that all integrase-mediated products (cointegrates, excisions, and integrations) increase (i.e. are readily observed) with the addition of excess integrase. The preference for cointegrate formation by plasmid-carried class 1 integrons when integrase is expressed under the natural $P_{int}$ promoter may indicate that this mechanism is important in nature. Since cassette capture has only been observed using excess integrase, all of the experiments discussed in this review (for all classes of integrase) use overexpressed integrase.

**Importance of cointegrates in nature**

Integrase-mediated cointegrate formation may be a recombination mechanism preferred by plasmids because cointegrates are important in nature for the evolution of plasmids and as a strategy to outcompete other plasmids in a population (191). Plasmid incompatibility is when two plasmids are unable to be stably maintained within

the same cell (130, 191). Incompatibility is related to the replicon of a plasmid and can be altered by mutating the origin of replication region. However, mutations within the replication origin most often result in a large increase in plasmid copy number, causing a slower growth rate and eventual death of the host line, which is outcompeted by strains containing the non-mutated plasmid (191). There are more than 50 recognized incompatibility groups (59). How did these different groups evolve when replication region mutations are deleterious? Gene duplication is a way for genes to evolve without loss of function, and plasmid incompatibility is thought to evolve in a similar way (135, 191). For plasmid cointegrates, one replication region can evolve while the other performs the necessary replication functions until a stable new incompatibility group is generated; then, the two plasmids can separate, or remain as a cointegrate (191). Cointegrates can outcompete single-replicon plasmids in culture. Transfer of a plasmid into a cell already containing a plasmid with the same incompatibility group results in loss of the original plasmid, but a single replicon plasmid cannot outcompete a double replicon cointegrate (84, 191). Cointegrates are abundant in nature; one study of several large plasmids isolated from fecal *E. coli* strains found that 37% of the 43 plasmids encoding identifiable incompatibility groups were incompatible with more than one group (64). Several virulence factor-containing plasmids isolated from pathogens are also incompatible with more than one group (47, 58, 191). In addition, most isolated plasmids are not tested for incompatibility; therefore, it is likely that many multi-replicon plasmids have not been identified as cointegrates (58, 191).

**Integron DNA recombination sites**

Integron integrases mediate recombination at *attI* and *attC* recombination sites (Fig. 1.10, 1.11), and, in rare instances, at secondary sites (42, 82, 157). A secondary recombination site is a site not located within an integron that shares some sequence identity with an *att* site. Insertion at a secondary site is irreversible, since integrase-mediated excision requires a recombination site at both ends of a gene cassette (42, 157). Like the DNA sites of other tyrosine recombinases, integron recombination sites all contain a core-type site (also called a simple site) with inverted repeats surrounding a central region (Fig. 1.1 A, 1.12, 1.13). Very few bases are conserved between different integron *att* sites, however. The right side of the core site inverted repeat encodes a consensus sequence of GTTRRRY, though only the GTT is absolutely conserved. Crossover of the DNA during integrase-mediated recombination occurs between the conserved G and TT nucleotides, and/or between the AA and C bases on the complementary strand, (Fig. 1.12, 1.13) for *attC* x *attC*, *attI* x *attC*, and *attI* x secondary site recombination (82, 185). In very rare cases, recombination crossover has been observed adjacent to the left side inverted repeat of the simple site, at the expected second cleavage position for a tyrosine recombinase (Fig. 1.13) (185). However, base pair mismatches within the sequence in between the inverted repeats, which would be the crossover region for a typical tyrosine recombinase (Fig. 1.1 A), do not affect recombination frequency. In one study, recombination did not decrease when the sequence identity of the putative crossover region was reduced from a perfect match to a 5 base pair mismatch (185). Identical crossover region sequences are essential for

efficient recombination mediated by other tyrosine recombinases (68, 101, 104). This suggests that the integron integrase either mediates a single strand exchange or that the DNA is cleaved at the same position on both strands (185).

***attI.*** *attI* sites consist of a single core-type site with inverted repeats for IntI binding plus two directly repeated IntI binding sites, DR1 and DR2 (Fig. 1.12 A). This orientation of binding sites with a set of inverted repeats and two direct repeats resembles the TnpI integrase binding site of Tn*5401* (14, 40, 156). The strongest IntI1 binding site, by footprint and deletion analysis, is the DR1 site (40, 65). Purified IntI1 binds to a double-stranded *attI1* site (56) but only weakly binds *attI3*, and IntI3 binding to *attI1* could not be detected (41). Thus, integron integrases show a strong preference for their own *attI* sites. This is not surprising since the *attI* sites of different integron classes share little sequence identity (Fig. 1.12 B) (41, 83). One study reported only 2% activity for IntI1 recombination using *attI2* or *attI3* sites with an *attC* site, when compared to recombination using *attI1* (74). Another study confirmed that IntI1 could recombine using *attI2* at very low frequency, but IntI2 recombination using *attI1* was not detected (83).

***attC.*** *attC* sites consist of two core-type simple sites surrounding a region of varying length (Fig. 1.13 A). They were initially called 59 base elements (26), though they range in size from 57-141 (118, 154). The double inverted repeats of *attC*s are palindromic and could form a cruciform structure in supercoiled DNA (Fig. 1.13 B) (56, 82, 90). The

sequences of different *attC* sites are heterogeneous, suggesting that structure rather than sequence plays a large role in *attC* recognition by the integrase, especially since different classes of integrase recognize several sizes and sequences of *attC* sites (41, 52, 90). Though IntI1 binds double-stranded *attI* sites *in vitro*, purified IntI1 preferentially binds the bottom strand of *attC*s (56, 90), though one study shows IntI1 can bind to a double-stranded *attC* site using high concentrations of IntI1 (40). Since *attC* sites have the potential to form cruciform structures (Fig. 1.13 B), single-stranded oligos *in vitro* could form hairpins that imitate half of the cruciform structure (90). It is possible that double-stranded *attC* binding is inhibited *in vitro* because the DNA requires accessory proteins to form the correct secondary structure, and replication or transcription could also stabilize cruciform structures *in vivo* and influence IntI1-mediated recombination reactions (56). Alterations in the shape of *attC* could provide regulation *in vivo* by sensing transcription, replication, or DNA topology (90). A potential cruciform DNA site indicates that IntI1-*attC* recombination could resemble telomere resolution in the replication of linear chromosomes, which also uses cruciform DNA structures and tyrosine recombinase-like proteins (98).

The non-palindromic sequences within *attC* influence IntI1 strand binding preferences. Two bases of *attC*s (the imperfect bases of the inverted repeats) that would create bulges in any hairpins or cruciform structures formed (Fig. 1.13 B) are important in directing binding of IntI1 to the bottom strand (90). Gel mobility shift experiments show that IntI1 covalently binds the bottom strand of the *aadA1 attC* of Tn*21* (Fig. 1.9 B) by a phosphotyrosine linkage—a tyrosine mutant (IntI1Y312F) did not

bind covalently (90). IntI1 was also able to recombine a single-stranded *attC* site with a plasmid *attI* site *in vivo* in a suicide conjugation experiment (23), possibly by a mechanism similar to the XerCD-mediated integration of CTXϕ in *Vibrio cholerae* (199). A similar experiment using a double-stranded *attC* substrate was unsuccessful (23), but the integration of transformed double-stranded artificial gene cassettes has been shown *in vivo* (36). Thus, IntI1 may recombine double-stranded and single-stranded *attC* sites under certain conditions.

**IntI recombination site preferences.** Though integron integrases can recombine different combinations of *attI* and *attC* sites, they do show site crossover preferences. Recombination efficiency studies examined cointegrate formation using a plasmid construct containing a single recombination site, a second plasmid with both *attI* and *attC* sites, and a third plasmid providing excess IntI1 *in trans* (42). These studies revealed that *attI* x *attC* recombination is the preferred crossover, followed by *attC* x *attC*, with *attI* x *attI* reactions occurring least often (42). Not all *attC* sites are equally preferred in recombination, however, likely depending on the particular *attC* sequence (42). A preference for *attI* x *attC* recombination was also seen in IntI1-mediated insertion of artificial gene cassettes (36). Artificial gene cassettes were constructed by digesting DNA containing two identical gene cassettes in a row and ligating to form a circular cassette, and these non-replicating circularized cassettes were transformed into *E. coli* strains containing a plasmid carrying a recipient integron with *attI* and *attC* sites and an integrase-overexpressing plasmid (36). Transformants expressing the

resistance of the gene cassette were selected and restriction digests of the isolated plasmids showed that the artificial gene cassettes preferred insertion at an *attI* site over an *attC* site, meaning *attI* x *attC* recombination was favored over *attC* x *attC* recombination for the insertion reaction (36). Recombination of *attC*s with secondary sites and of *attI*s with secondary sites can also occur in cointegrate formation or insertion reactions but are even more rare than *attI* x *attI* reactions (42, 82, 157). Insertion of gene cassettes at secondary sites in the chromosome would be deleterious for a cell and is assumed to be restricted by tight expression of IntI1 and the preferential recognition of specific recombination sites by the integrase (82, 90).

Though different classes of integrases can recognize diverse *attC* sequences, they do show different preferences for particular *attC* sites in recombination. For example, a IntI2*179E mutant, a functional class 2 integrase, slightly prefers excision of a *sat* cassette in the second position over excision of the *aadA1* in the third position (13% to 10%), while IntI1 highly prefers excision of *aadA1* over *sat* (73% to 5.2%) from the same cassette array (Fig. 1.9 C) (82). Excision of both of those cassettes would include recombination at the *sat attC*, so IntI1 prefers recombination at the *aadA1 attC* over recombination at the *dfrA1 attC* (with *dfrA1* in the first cassette position) (Fig. 1.9 C) (82). Experiments comparing frequencies of VchIntIA-mediated deletion of artificial gene cassettes showed that VchIntIA recombined an *orfA attC* and an *ereA attC* poorly (10 to 100 times less frequently) when compared to an *aadA7 attC*, but IntI1 showed similar deletion frequencies for all three *attC* sites (15). In addition, IntI1 shows a strong preference for mediating recombination at the *orfA attC* in R388 over the *dfrB2 attC* site,

but IntI3 uses both *attC* sites in R388 equally well (39). These *attC* site preferences suggest that different classes of integrase differ in IntI-*attC* interactions.

**Characteristics of class 1, 2, and 3 MIs, the *V. cholerae* SI, and the *S. oneidensis* CI**

**Evolution of class 1 integrons.** Class 1 integrons were the first integron class to be discovered and are also the most abundant in nature (55, 118, 180). They are responsible for carrying multiple antibiotic resistance gene cassettes in clinical and environmental samples in Gram-negative and Gram-positive bacteria (61, 109, 112, 121, 126, 128, 142, 166, 196, 211). The nearly 100 gene cassettes found in class 1 integrons confer antibiotic resistance to all aminoglycosides, all β-lactams, macrolides, quinolones, rifampicin, trimethoprim, and more (55, 118). Class 1 integrons usually consist of 5' and 3' conserved sequences (CS) (Fig 1.9 B). The 5' CS includes the integrase gene *intI1* and the *attI1* site, and the 3' CS includes three nonmobile genes: *qacEΔ1* (resistance to quaternary ammonium compounds), *sul1* (sulfonamide resistance), and *orf5* (76). The abundance of and variability among class 1 integrons has provoked much discussion about their evolution. The relatedness between particular class 1 integrons can be difficult to determine because of the multiple ways integrons and gene cassettes can be transferred. As mentioned above, the presence of the conserved sequences on either end of most class 1 integrons (Fig. 1.9 B) suggests that homologous recombination is a likely mechanism for gaining and losing sequences between integrons (see above "Spread of integrons and gene cassettes by horizontal

gene transfer") (141). The *sul1* gene is found in the 3' end of the vast majority of class 1 integrons and is assumed to have been associated with integrons early in their evolution (143, 180). Codon usage analysis suggests that *intI1* and *sul1* likely evolved from closely related organisms but that *qacEΔ1* and mobile gene cassettes have more diverse origins (17). The Tn*402* integron (Fig. 1.9 B) resembles a possible ancestor for class 1 integrons; it does not contain *sul1*, but encodes a full mobile *qacE* cassette rather than *qacEΔ1*, and is the only integron capable of transposition on its own (17, 25, 151, 173). Tn*402* encodes two other cassettes upstream of *qacE*, *dfrB3* and *orfD*, and four *tni* transposition genes are located at the 3' end (Fig. 1.9 B). Other class 1 integrons have been identified as defective transposons, encoding IR elements at both ends and an incomplete set of transposition genes (25, 180). These integrons contain IS*1326* and a deletion in the transposition genes found in Tn*402* (Fig. 1.9 B). Since they still encode the IR sequences on either end, these integrons are thought to be capable of transposition if the transposition genes were supplied *in trans* (141). The integron In0 from pVS1 of *Pseudomonas aeruginosa* (17) represents a later version of class 1 integron than Tn*402* and could be an ancestor to the integron In2 of Tn*21*, requiring only the *aadA1* gene cassette at *attI1* and the acquisition of IS*1353* (Fig. 1.9 B) (25).

After the discovery of the *Vibrio cholerae* superintegron (120), dozens of integrases were identified on the chromosomes of different bacterial species (34, 60, 106, 118, 161, 198). Though it was already known that integrons predated the use of antibiotics, the prevalence of chromosomal integrons discovered in divergent species of

bacteria indicated that integrons are ancient structures, even more ancient than previously imagined (118, 163). Thus, class 1 integrons (and all MIs) likely evolved from chromosomal integrons that were incorporated into transposons or other mobile elements (118, 163, 184). Eventually, the transposon/integrons gained antibiotic resistance genes, like Tn*402* (Fig. 1.9 B), and then most integrons lost their transposition genes, like In0 and In2 (Fig. 1.9 B), as discussed above.

**Most class 2 integrons encode a nonfunctional *intI2*.** Most class 2 integrons have been found in Tn*7*-related transposons (12, 16, 83). Tn*7* is widely distributed in nature, but the gene cassettes of class 2 integrons are much less diverse than class 1 integrons; less than 10 different gene cassettes have been found in class 2 integrons, meaning they usually carry the same few gene cassettes (12, 83). The class 2 integron in Tn*7* carries four gene cassettes, three of which encode for antibiotic resistance (Fig. 1.9 C). One Tn*7*-like class 2 integron discovered on a plasmid in a clinical *Escherichia coli* isolate contains an *ereA* cassette conferring resistance to erythromycin (16). Previously, *ereA* gene cassettes had only been detected within class 1 integrons from clinical isolates (29, 146, 147, 194). The *ereA* gene cassette encoded its own promoter, rare in a gene cassette (16). Class 2 integrons, including the *ereA*-containing class 2 integron, are unusual in that *intI2* encodes an internal stop codon; thus, they do not express a functional IntI2 protein (83). In experiments where the *intI2* stop codon was changed to encode glutamate, IntI2 was expressed and could mediate cassette excision (83). It was hypothesized that the gene cassettes of class 2 integrons could only be

exchanged in nature if IntI1 was present in the same cell to provide integrase *in trans* (83). The strain containing the class 2 integron with *ereA* also contained a class 1 integron, so IntI1 was likely present in the cell to mediate transfer of *ereA* (16). Recently, however, a class 2 integron containing a full *intI2* gene was found in two strains of *Providencia stuartii* isolated from beef cattle, encoding a glutamine instead of the stop codon seen in all other class 2 integrons (12). This integron contained four gene cassettes encoding six unknown open reading frames (orfs) not homologous to known antibiotic resistance genes (Fig. 1.9 C), instead of the cassettes found in Tn*7*-like integrons (12, 83). After this discovery, it seems likely that more class 2 integrons with functional integrases and diverse gene cassettes will be isolated in the future.

**Recombination site preferences of the class 3 integrase.** The first class 3 integron was identified in *Serratia marcescens* (4), and encodes $bla_{IMP-1}$ and *aacA4* gene cassettes (Fig. 1.9 D) conferring resistance to carbapenems and aminoglycosides, respectively. It contains a partial *tni* module and at least one IR sequence, similar to the Tn*402* class 1 integron, but the class 3 integron is reversed within the transposon when compared to Tn*402* (Fig. 1.9 B, D) (39). IntI3 is 59% identical to IntI1 and 46% identical to IntI2 (41). IntI3 has been shown to mediate cointegrate formation, integration, and excision (39, 41). For the class 3 integron, insertion of artificial circular gene cassettes occurs preferentially at *attI3*; thus, IntI3, like IntI1, prefers *attI* x *attC* recombination to *attC* x *attC* (39, 41). IntI3 mediates recombination at secondary sites at low frequency

like IntI1, but uses mainly “GTT” (over 75%) secondary sites while IntI1 uses “GTT” and “GAT” secondary sites equally well (39, 157).

Though not as prevalent as class 1 integrons, class 3 integrons have been found in varying species of bacteria across the globe. Class 3 integrons have been identified within two *Delftia* environmental isolates in Canada (208), and IntI3 has been detected by PCR in clinical isolates from Japan, including strains of *Pseudomonas aeruginosa, Alcaligenes xylosoxidans, Pseudomonas putida,* and *Klebsiella pneumoniae* (171, 174).

**The *Vibrio cholerae* SI and crystal structure of VchIntIA.** A superintegron (SI) describes a chromosomal integron not on a mobile element that encodes over 20 gene cassettes with similar *attC* sequences (118). The first superintegron (SI) was discovered on chromosome 2 of *Vibrio cholerae*, and differed from known integrons by size, content, and homogeneity of *attC* sequences (120). Since the discovery of the first SI, dozens of chromosomal integrons have been identified (34, 60, 106, 120, 161, 198) in diverse genera such as *Geobacter* and *Treponema*, leading to the theory that MIs evolved from CIs (see above). The *V. cholerae* SI (Fig. 1.9 E) encompasses 126 kb with 179 cassettes encoding approximately 215 orfs (118, 120, 163). The cassette genes are mainly unique to *V. cholerae*, unlike the widely distributed antibiotic resistance genes encoded by MIs, and encode for cellular functions such as metabolism and for virulence factors such as toxins (118). Unsurprisingly, several cassettes contain their own promoter sequences; although a promoter is a rare feature in MI gene cassettes, a single cassette promoter obviously could not transcribe 215 orfs (118).

Transcription of the majority of gene cassettes in the *V. cholerae* SI is positively regulated by HapR and RpoS and negatively regulated by RpoN; thus, gene cassette transcription is upregulated at high cell densities and in response to stress (209). The *attC* sites of the *V. cholerae* SI are 122-124 nucleotides in size, similar in size to the larger *attC*s of MI cassettes, but are over 85% identical in sequence, unlike the divergent sequences of MI *attC*s. Because of the close sequence similarity among the *V. cholerae* SI *attC*s, they were termed VCRs (for *V. cholerae* repeats) (11).

The integrase of the *V. cholerae* SI, VchIntIA (formerly called IntI4) (163), is 45% identical to IntI1 and IntI3 and 49% identical to IntI2 (41). VchIntIA is the only integron integrase for which a crystal structure has been resolved (114). The crystal structure is of VchIntIA bound to the bottom strand of a VCR-type *attC* site. VchIntIA is composed of two domains (Fig. 1.6 D), like most other tyrosine recombinases such as XerD (Fig. 1.6 B). The N-terminal DNA-binding domain structure resembles that of XerD (187) and lambda integrase (1), and the C-terminal catalytic domain includes the IntI region conserved in integron integrases (Fig. 1.6 D) (123). The VchIntIA crystal structure shows that all of the active site residues are from the same recombinase subunit, suggesting cleavage *in cis* (114). The crystal structure indicates a two-fold symmetric synapse with the single-stranded DNA substrate (MacDonald et al 2006), instead of the four-fold symmetry observed for other tyrosine recombinases (187, 200). The two-fold symmetry may inhibit isomerization of the recombination complex (Fig 1.7) in order to prevent Holliday junction resolution by a second cleavage and exchange (114) since a second strand exchange with a single-stranded DNA substrate would not give a desired

recombination product (199). Interactions between inactive (non-attacking) and active (attacking) recombinase subunits seen in the crystal structure may serve as a switch to differentiate *attC* and *attI* binding (114). This "molecular switch" (114) is an interesting structural mechanism to explain how an IntI protein can recombine such diverse DNA sites. Hopefully, a crystal structure of an IntI bound to an *attI* site will be resolved in the near future in order to show the *attI* conformation of the molecular switch.

**The *Shewanella oneidensis* CI encodes three types of *attC* sites.** A small chromosomal integron (CI) was discovered in *Shewanella oneidensis* MR-1 that exemplifies the ability of integrases to recombine several different types of recombination sites. This integron contains a new class of integrase, IntISon, that is 63% identical to IntI2 (41) and 45% identical to IntI1, IntI3, and VchIntIA (41, 52). This CI encodes three gene cassettes (Fig. 1.9 F), each with a different type of *attC* site; one *attC* site is unique (Drouin et al 2002), one is similar to the *attC*s found in the *Vibrio cholerae* SI (VCR-like) (32), and one is similar to *attC*s found associated with *aadA* and *aadB* cassettes from MIs (155). The first cassette with the unique *attC* site includes an insertion sequence with two genes, a recombinase and a transposase (52). This IS-containing cassette is also located in two other positions on the *S. oneidensis* chromosome (52). The second cassette with the VCR-like *attC* encodes an orf of unknown function homologous to *S. oneidensis* and *V. cholerae* chromosomal genes, and a repeat of this orf is found downstream of the integron encoded in the opposite direction (Fig. 1.9 F) (52). Surprisingly, the third cassette with the *aad*-like *attC* only

contains a noncoding repeat region (52). IntISon has been experimentally shown to mediate excision and integration reactions with a large variety of *attC* sequences commonly found in MI gene cassettes, as expected for an integron integrase—especially one naturally associated with a variety of *attC* sites (41, 52). The *attI* site, *attISon*, is most similar in sequence to *attI2*; in fact, IntISon can mediate recombination at *attI2*, but not at *attI1* or *attI3* (52).

**Expression of Gene Cassettes**

The gene cassettes of MIs are transcribed from a common promoter $P_c$ (Figs. 1.9 A, 1.14). Less than 10 MI cassettes have been found to encode their own promoter (16, 17, 148, 164, 181). In fact, the open reading frames in many gene cassettes start near the beginning of the cassette, without enough space for a promoter sequence (35). Since gene cassettes are transcribed by a common promoter, the cassettes located closer to the promoter have higher expression than those located further downstream (35, 111).

There are four versions of class 1 integron cassette promoters found in nature (Fig. 1.14). For $P_c$ (also known as $P_{ant}$), there is a strong version (in integrons from R388, pSa, and In58 from *P. aeruginosa*), a weak version (in the In0 integron from pVS1), and a hybrid version, the -10 region of the strong and the -35 of the weak, (in integrons from R46 and In111 from *E. coli*) (111, 149, 202). There is a second promoter called P2. The -10 and -35 regions of P2 are always present, but are usually only 14 bp apart, so they are not active as a promoter (Fig. 1.14 B). Some integrons encode 3 Gs

to make the spacing more efficient, and the promoter is then active. The active P2 promoter has always been found with the weak version of the $P_c$ promoter, not with the strong or hybrid versions, such as on the integron In2 of Tn*21* (Fig. 1.14B) (35, 111). The strong $P_c$ has the most activity (6.5 times stronger than the tac promoter (50), the weak $P_c$ plus P2 (with P2 providing about 90% of the activity) is second in activity, 3.2 times that of tac promoter (111). The hybrid $P_c$ is stronger than the weak $P_c$, but both are weaker than $P_{tac}$ (111). These results have been supported by testing the relative antibiotic resistance levels of cassettes under control of the different promoters (35). The resistance levels conferred by cassettes like *aadA2* are also dependent on the position of the cassette within an integron. Streptomycin resistance has been shown to decrease by approximately 4-fold with the movement of an *aadA2* cassette from the first to the second position from the strong $P_c$ promoter. Resistance decreased by as much as 20-fold, in one case, when *aadA2* occupied the third cassette position (35). This decrease in expression was affected by which particular cassettes were located upstream. This is likely due to which *attC* sites were upstream of the cassette since analysis of integron mRNA transcripts have shown that transcription tends to stop near the ends of cassettes. Thus, *attC* sites can act as transcriptional terminators, and some sites may be more efficient at stopping transcription than others (35).

Though the best-studied cassette promoters are those of class 1 integrons, phenotypic selection of antibiotic resistance cassettes encoded by class 2 and class 3 integrons shows that their gene cassettes are expressed (16, 39, 83). The Tn*7* class 2 integron likely has a relatively strong cassette promoter because all three cassette-

encoded antibiotic resistances are phenotypically expressed (16). By comparing the relative phenotypic resistance level conferred by the *aadA2* gene cassette in the first position, the class 3 $P_c$ promoter strength is within the range established for the different $P_c$ promoters of class 1 integrons (39).

Since the *V. cholerae* SI encodes over 150 gene cassettes, some of which with their own promoters, a $P_c$ promoter may not be necessary. It is known that the transcription of several *V. cholerae* SI gene cassettes is influenced by RpoS, RpoN, and HapR (209), but not all of the SI cassettes are expressed. An SI gene cassette encoding chloramphenicol resistance was acquired by a plasmid-carried class 1 integron when IntI1 was expressed *in trans* in *V. cholerae* (160). The chloramphenicol was not expressed when the gene cassette was within the *V. cholerae* SI but was expressed by the $P_c$ of the class 1 integron. This led to the proposal that SI gene cassettes are a storage method whereby selective pressure can cause integrase-mediated rearrangement of cassettes to allow for gene expression; otherwise, in order to reduce the selective burden of so many extra genes on the chromosome, most of the gene cassettes remain unexpressed (160).

**Gene cassette evolution**

The *attC* sites of MIs vary greatly in both size and sequence, indicating either a wide pool of possible sequences from which *attC* sites evolved or the development of diverse *attC*s from a few ancestors over a very long period of time (155). Chromosomal superintegrons contain gene cassettes whose *attC* sites are closely related in size and

sequence (163). This suggests that superintegrons could be a reservoir for MI gene cassettes and that the diverse MI *attC* sites resulted from acquiring cassettes from multiple bacterial species by horizontal gene transfer (118, 160, 163). The successful capture of several SI gene cassettes by a plasmid-carried class 1 integron in *Vibrio cholerae*, using excess IntI1 supplied *in trans*, supports the theory that MIs acquire cassettes from SIs (160). In addition, twelve MI gene cassettes have been identified that contain *attC* sites similar in sequence to the *attC*s of *Xanthomonas* and *Vibrio* SI cassettes (118, 163).

It is difficult to determine how gene cassettes were first assembled and from where their genes and *attC*s originated. Cassette-encoded genes are not just copies of known antibiotic resistance genes with downstream *attC* sequences; they are novel genes that form their own subfamilies within the families of known antibiotic resistance genes, making it difficult to identify the origin of a gene within an integron cassette (155). It has been hypothesized that gene cassettes were formed by reverse transcription of mRNA molecules (155). The *attC* sites could have already been present as transcription terminators or could have been added separately. There are examples of closely related genes having closely related *attC*s, examples of closely related genes having divergent *attC*s, and examples of unrelated genes having similar *attC*s (155). The former suggests that the genes and *attC*s of gene cassettes are assembled together, but the latter two suggest that gene cassettes are assembled separately (155). The discovery of group II introns adjacent to gene cassettes led to a new theory for the assembly of gene cassettes (27). RecA-mediated recombination between an intron

downstream from a gene and a second intron adjacent to an *attC* site could assemble a gene cassette, after the intron was spliced out and reverse transcription occurred (27). Possible intron-*attC* and gene-intron intermediates discovered in the *Nitrosomonas europaea* genome sequence support this theory (27).

**PCR detection of integrases and of the environmental gene cassette metagenome**

PCR was initially used to detect class 1 integrons in *Enterobacteriaceae* and pseudomonads by primers that anneal to the 5' and 3' conserved sequences (Fig. 1.9 B). The 5' CS primer binds to an area between the integrase coding region and *attI1*, and the 3' CS primer anneals to a sequence just upstream of *qacEΔ1* (17, 112). This PCR method can only detect class 1 integrons encoding these conserved regions surrounding a few cassettes, but that description includes the vast majority of identified class 1 integrons (141). Large percentages of *Enterobacteriaceae* and pseudomonad clinical isolates were shown to contain class 1 integrons with varying cassette arrays by this conserved segment PCR method (112). Less specific PCR primers have since been developed to target integrases and cassettes in environmental DNA. New integron classes have been identified in environmental DNA by PCR with primers annealing to a conserved sequence near the 3' end of the *intI* gene and to a conserved segment of *attC* sequence (Fig. 1.9 A). Thus, the PCR product includes most of the integrase gene, the *attI* site, and most of at least one gene cassette (128). The same method has also been used to isolate gene cassettes from environmental DNA, using

degenerate primers that anneal to *attC* sequences. Over 100 novel gene cassettes have been identified by this technique, revealing a huge reservoir of gene cassettes in the environment, referred to as the "gene cassette metagenome" (87, 183). There is some specificity to primers created to anneal to *attC* sequences, even when using degenerate primers, since *attC*s are so variable and share little sequence identity (Fig. 1.13). Even so, based on cassette PCR, a 50 $m^2$ soil sample is estimated to contain at least 2000 different gene cassettes (124). The fact that so many novel (and no known) sequences have been discovered using cassette PCR in spite of the potential primer specificity emphasizes the diversity of gene cassettes found in nature (129, 183). Some cassette sequences identified by PCR are homologous to antibiotic resistance genes, showing that the environmental gene cassette metagenome is a reservoir for antibiotic resistance gene cassettes (87, 124, 129).

**SUMMARY**

Although not related by sequence, serine and tyrosine site-specific recombinases perform similar functions in nature by mediating DNA inversion, integration, and excision reactions (68, 69, 79, 176). Site-specific recombinases of both families are able to potentially catalyze all three types of recombination reactions. Recombination is restricted to a preferred reaction by the precise assembly of the nucleoprotein complex forming the recombination synapse. This assembly is directed by the recombinase binding to specific DNA recognition sites, accessory factors and/or additional recombinase subunits binding to accessory DNA sites, and protein-protein interactions

between recombinase subunits and/or with accessory proteins. These recombination restrictions can be seen in the serine family by examining mutants of Gin invertase and Tn*3* resolvase. Gin invertase demonstrates regulating recombination by an accessory factor (96, 99), and Tn*3* resolvase restricts recombination by binding additional recombinase subunits (168, 178). Fis-independent Gin mutants (95, 96, 177) and Tn*3* mutants (7) both show reduced substrate specificity—they are able to mediate recombination between directly repeated or inverse sites on linear or supercoiled DNA.

Tyrosine recombinases, like the serine family, restrict recombination by recognizing specific complex DNA recombination sites, and those sites encode regions for binding accessory factors and/or additional subunits of recombinase that also regulate recombination. Some DNA recombination sites of tyrosine recombinases consist of only a core site, such as *loxP* (200), *dif* (20), and *attB* (101) (for Cre, XerCD, and lambda integrase, respectively) (Fig. 1.8). XerCD *cer* and *psi* sites are more complex, they have a core site plus sites for binding accessory factors (Fig. 1.8) (43, 122). Lambda integrase *attP*s are the most complex, containing a core site, integrase arm binding sites, and multiple accessory factor binding sites (Fig. 1.8 C) (1, 101, 207). Integron integrases have four integrase binding sites per *att* site (Figs. 1.12, 1.13) (40, 56, 65), but no accessory factor binding sites have yet been identified. Integron integrase *attI* sites consist of two directly repeated IntI binding sites plus a core site (40). Integron integrase *attC* sites are similar to *res* sites (Fig. 1.5) (like those of the Tn*3* resolvase of the serine family (67)) in having multiple core sites, though *res* sites contain three core-type sets of inverted repeats and *attC*s only have two (185), but crossover

typically occurs at only one particular core-type site for both *res* and *attC* sites (168, 185).

Tyrosine recombinases direct recombination to preferentially cleave one strand of DNA first; the less efficient cleavage occurs first so that the more robust reaction resolves the Holliday junction intermediate, biasing the reaction to form reaction products rather than regenerate the starting substrates. This preferential strand cleavage is observed in Cre-*loxP* (85, 103, 104), XerCD-*dif* (8, 199), and lambda integrase recombination reactions (18, 150, 201). Integron integrase recombination differs from typical tyrosine recombination reactions in that cleavage only occurs at one position, indicating either a single-strand exchange or cleavage of both strands at the same position (185), or possibly a single-strand exchange for recombination at *attC* and double-stranded cleavage for *attI* x *attI* recombination. IntI1 preferentially binds double-stranded *attI* sites (40, 56, 65) and single-stranded *attC* sites (56, 90), though high concentrations of IntI1 can bind a double-stranded *attC* site, as well (40). IntI1 has been observed to mediate *attC* x *attI* integration reactions with single-stranded (23) and double-stranded substrates (36) under certain conditions. Other tyrosine recombinases can also mediate recombination using both double-stranded and single-stranded substrates. Besides resolving double-stranded plasmid and chromosomal dimers, XerCD mediates the integration of the single-stranded CTXϕ into chromosome 1 of *V. cholerae* (199).

Tyrosine recombinases show that recombination reaction choice can be affected by *in vivo* recombinase concentrations. High Cre concentrations can overcome the

regulation that usually restricts recombination to intramolecular reactions and allow intermolecular recombination, resulting in destabilization of the P1 prophage (144). Lambda integrase regulates the recombination choice of integration versus excision in part by the relative concentrations of integrase and excisionase (Xis) in a cell (105); thus, an excess of integrase promotes integration and inhibits excision. In addition, an excess of XerC *in vivo* can cause an accumulation of XerC-mediated Holliday junction intermediates instead of promoting complete recombination reactions (122). Integron integrases are greatly affected by intracellular concentrations *in vivo* because cointegrates are the only integrase-mediated products observed when *intI* is expressed under the natural promoter $P_{int}$ (38, 73, 116, 117). Cassette capture, requiring cointegrate resolution or integration of an excised cassette, is only observed when integrase is overexpressed (36, 38, 73, 116). This indicates that integron integrases have lower activity than other tyrosine recombinases and suggests that cointegrate formation may be a mechanism used more often in nature by plasmid-carried integrons than previously thought.

Integrons were first identified because they are often associated with multiple antibiotic resistance genes (180). Mobile element associated integrons (MIs), mainly class 1 integrons, are responsible for spreading antibiotic resistances among Gram-negative and Gram-positive bacteria in clinical and environmental settings (61, 109, 112, 121, 126, 128, 142, 166, 196, 211). MIs are thought to have evolved from chromosomal integrons (CIs) that became associated with a mobile element (118, 163, 184). Chromosomal integrons, including superintegrons (SIs) containing more than 20

gene cassettes, are also reservoirs of mobile cassettes by encoding genes for antibiotic resistances and virulence factors that can be acquired by MIs (118, 160). The MIs can then spread those gene cassettes by horizontal gene transfer, such as conjugation if they are located in a conjugative plasmid and/or transposition if they are located within a transposon (25, 76, 180). PCR of environmental DNA has revealed a metagenome of gene cassettes—showing that thousands of different mobile gene cassettes could be present in the soil bacteria of a relatively small area (183) and suggesting that large SIs, like the *V. cholerae* SI encoding over 175 gene cassettes (120), may be abundant in the environment (118).

Integron integrases show site preferences in recombination reactions—*attI* x *attC* reactions are favored over *attI* x *attI* and *attC* x *attC* reactions (42). Integrases prefer their own *attI* sites, and do not usually mediate recombination using the *attI* site of a different class of integrase (52, 74). Though the different classes of integron integrases can recognize many different types of *attC*s, the IntIs can favor one *attC* site over another in recombination, as well (15, 39, 83). The VchIntI crystal structure revealed a "molecular switch" in the protein structure that could change conformation based on its DNA substrate and may differentiate *attI* and *attC* binding (114). This molecular switch may help the integrase restrict recombination to desired reactions by promoting binding to either *attI* or *attC* sites.

Chapters 2 and 3 examine integron integrase-mediated recombination products in cells expressing integrase under the natural $P_{int}$ promoter and in cells overexpressing integrase. Recombination products are first qualitatively detected and then quantified

over time in broth cultures to determine whether these products simply accumulate over time or if variations are seen during culture growth, indicating a culture cycle-dependent regulation of IntI1-mediated recombination.

## REFERENCES


1. **Aihara, H., H. J. Kwon, S. E. Nunes-Duby, A. Landy, and T. Ellenberger.** 2003. A conformational switch controls the DNA cleavage activity of λ integrase. Molecular Cell **12:**187-198.
2. **Akopian, A., J. He, M. R. Boocock, and W. M. Stark.** 2003. Chimeric recombinases with designed DNA sequence recognition. PNAS **100:**8688-8691.
3. **Alén, C., D. J. Sherratt, and S. D. Colloms.** 1997. Direct interaction of aminopeptidase A with recombination site DNA in Xer site-specific recombination. EMBO J **16:**5188-97.
4. **Arakawa, Y., M. Murakami, K. Suzuki, H. Ito, R. Wacharotayankun, S. Ohsuka, N. Kato, and M. Ohta.** 1995. A novel integron-like element carrying the metallo-beta-lactamase gene *bla*$_{IMP}$. Antimicrob Agents Chemother **39:**1612-5.
5. **Arciszewska, L. K., and D. J. Sherratt.** 1995. Xer site-specific recombination *in vitro*. EMBO J **14:**2112-20.
6. **Argos, P., A. Landy, K. Abremski, J. B. Egan, E. Haggard-Ljungquist, R. H. Hoess, M. L. Kahn, B. Kalionis, S. V. Narayana, L. S. Pierson, 3rd, and et al.** 1986. The integrase family of site-specific recombinases: regional similarities and global diversity. Embo J **5:**433-40.

7. **Arnold, P. H., D. G. Blake, N. D. Grindley, M. R. Boocock, and W. M. Stark.** 1999. Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. EMBO J **18:**1407-14.

8. **Aussel, L., F.-X. Barre, M. Aroyo, A. Stasiak, A. Z. Stasiak, and D. Sherratt.** 2002. FtsK Is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. Cell **108:**195-205.

9. **Azaro, M. A., and A. Landy.** 2002. λ Integrase and the λ Int Family, p. 118-148. *In* N. L. Craig, R. Craigie, M. Gellert, A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D. C.

10. **Banuett, F., M. A. Hoyt, L. McFarlane, H. Echols, and I. Herskowitz.** 1986. *hflB*, a new *Escherichia coli* locus regulating lysogeny and the level of bacteriophage lambda cII protein. Journal of Molecular Biology **187:**213-224.

11. **Barker, A., C. A. Clark, and P. A. Manning.** 1994. Identification of VCR, a repeated sequence associated with a locus encoding a hemagglutinin in *Vibrio cholerae* O1. J Bacteriol **176:**5450-8.

12. **Barlow, R. S., and K. S. Gobius.** 2006. Diverse class 2 integrons in bacteria from beef cattle sources. J. Antimicrob. Chemother. **58:**1133-1138.

13. **Barre, F.-X., M. Aroyo, S. D. Colloms, A. Helfrich, F. Cornet, and D. J. Sherratt.** 2000. FtsK functions in the processing of a Holliday junction intermediate during bacterial chromosome segregation. Genes Dev. **14:**2976-2988.

14. **Baum, J. A.** 1995. TnpI recombinase: identification of sites within Tn*5401* required for TnpI binding and site-specific recombination. J Bacteriol **177:**4036-42.

15. **Biskri, L., M. Bouvier, A. M. Guerout, S. Boisnard, and D. Mazel.** 2005. Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. J Bacteriol **187:**1740-50.

16. **Biskri, L., and D. Mazel.** 2003. Erythromycin esterase gene *ere*(A) is located in a functional gene cassette in an unusual class 2 integron. Antimicrob. Agents Chemother. **47:**3326-31.

17. **Bissonnette, L., and P. H. Roy.** 1992. Characterization of In0 of *Pseudomonas aeruginosa* plasmid pVS1, an ancestor of integrons of multiresistance plasmids and transposons of gram-negative bacteria. J. Bacteriol. **174:**1248-57.

18. **Biswas, T., H. Aihara, M. Radman-Livaja, D. Filman, A. Landy, and T. Ellenberger.** 2005. A structural basis for allosteric control of DNA recombination by λ integrase. Nature **435:**1059-1066.

19. **Blakely, G., S. Colloms, G. May, M. Burke, and D. Sherratt.** 1991. *Escherichia coli* XerC recombinase is required for chromosomal segregation at cell division. New Biol **3:**789-98.

20. **Blakely, G., G. May, R. McCulloch, L. K. Arciszewska, M. Burke, S. T. Lovett, and D. J. Sherratt.** 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. Cell **75:**351-361.

21. **Bode, J., T. Schlake, M. Iber, D. Schubeler, J. Seibler, E. Snezhkov, and L. Nikolaev.** 2000. The transgeneticist's toolbox: novel methods for the targeted modification of eukaryotic genomes. Biol Chem **381:**801-13.

22. **Boucher, Y., M. Labbate, J. E. Koenig, and H. W. Stokes.** 2007. Integrons: mobilizable platforms that promote genetic diversity in bacteria. Trends Microbiol.

23. **Bouvier, M., G. Demarre, and D. Mazel.** 2005. Integron cassette insertion: a recombination process involving a folded single strand substrate. EMBO J **24:**4356-67.

24. **Breuner, A., L. Brondsted, and K. Hammer.** 1999. Novel organization of genes involved in prophage excision identified in the temperate lactococcal bacteriophage TP901-1. J. Bacteriol. **181:**7291-7297.

25. **Brown, H. J., H. W. Stokes, and R. M. Hall.** 1996. The integrons In0, In2, and In5 are defective transposon derivatives. J. Bacteriol. **178:**4429-37.

26. **Cameron, F. H., D. J. Groot Obbink, V. P. Ackerman, and R. M. Hall.** 1986. Nucleotide sequence of the AAD(2'') aminoglycoside adenylyltransferase determinant *aadB*. Evolutionary relationship of this region with those surrounding *aadA* in R538-1 and *dhfrII* in R388. Nucleic Acids Res **14:**8625-35.

27. **Centron, D., and P. H. Roy.** 2002. Presence of a group II ntron in a multiresistant *Serratia marcescens* strain that harbors three integrons and a novel gene fusion. Antimicrob. Agents Chemother. **46:**1402-1409.

28. **Chaconas, G., P. E. Stewart, K. Tilly, J. L. Bono, and P. Rosa.** 2001. Telomere resolution in the Lyme disease spirochete. Embo J **20:**3229-37.

29. **Chang, C. Y., L. L. Chang, Y. H. Chang, T. M. Lee, and S. F. Chang.** 2000. Characterisation of drug resistance gene cassettes associated with class 1 integrons in clinical isolates of *Escherichia coli* from Taiwan, ROC. J Med Microbiol **49:**1097-102.

30. **Chen, J.-W., J. Lee, and M. Jayaram.** 1992. DNA cleavage *in trans* by the active site tyrosine during Flp recombination: Switching protein partners before exchanging strands. Cell **69:**647-658.

31. **Christ, N., and P. Droge.** 2002. Genetic manipulation of mouse embryonic stem cells by mutant lambda integrase. Genesis **32:**203-208.

32. **Clark, C. A., L. Purins, P. Kaewrakon, and P. A. Manning.** 1997. VCR repetitive sequence elements in the *Vibrio cholerae* chromosome constitute a mega-integron. Mol Microbiol **26:**1137-8.

33. **Clark, N. C., O. Olsvik, J. M. Swenson, C. A. Spiegel, and F. C. Tenover.** 1999. Detection of a streptomycin/spectinomycin adenylyltransferase gene (*aadA*) in *Enterococcus faecalis*. Antimicrob. Agents Chemother. **43:**157-60.

34. **Coleman, N., S. Tetu, N. Wilson, and A. Holmes.** 2004. An unusual integron in *Treponema denticola*. Microbiology **150:**3524-6.

35. **Collis, C., and R. Hall.** 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. Antimicrob. Agents Chemother. **39:**155-162.

36. **Collis, C. M., G. Grammaticopoulos, J. Briton, H. W. Stokes, and R. M. Hall.** 1993. Site-specific insertion of gene cassettes into integrons. Mol. Microbiol. **9:**41-52.

37. **Collis, C. M., and R. M. Hall.** 1992. Gene cassettes from the insert region of integrons are excised as covalently closed circles. Mol. Microbiol. **6:**2875-85.

38. **Collis, C. M., and R. M. Hall.** 1992. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. J. Bacteriol. **174:**1574-85.

39. **Collis, C. M., M. J. Kim, S. R. Partridge, H. W. Stokes, and R. M. Hall.** 2002. Characterization of the class 3 integron and the site-specific recombination system it determines. J. Bacteriol. **184:**3017-26.

40. **Collis, C. M., M. J. Kim, H. W. Stokes, and R. M. Hall.** 1998. Binding of the purified integron DNA integrase IntI1 to integron- and cassette-associated recombination sites. Mol. Microbiol. **29:**477-90.

41. **Collis, C. M., M. J. Kim, H. W. Stokes, and R. M. Hall.** 2002. Integron-encoded IntI integrases preferentially recognize the adjacent cognate *attI* site in recombination with a 59-be site. Mol. Microbiol. **46:**1415-27.

42. **Collis, C. M., G. D. Recchia, M. J. Kim, H. W. Stokes, and R. M. Hall.** 2001. Efficiency of recombination reactions catalyzed by class 1 integron integrase IntI1. J. Bacteriol. **183:**2535-42.

43. **Colloms, S. D., C. Alén, and D. J. Sherratt.** 1998. The ArcA/ArcB two-component regulatory system of *Escherichia coli* is essential for Xer site-specific recombination at *psi*. Molecular Microbiology **28:**521-530.

44. **Colloms, S. D., R. McCulloch, K. Grant, L. Neilson, and D. J. Sherratt.** 1996. Xer-mediated site-specific recombination *in vitro*. EMBO J **15:**1172-81.

45. **Colloms, S. D., P. Sykora, G. Szatmari, and D. J. Sherratt.** 1990. Recombination at ColE1 *cer* requires the Escherichia coli *xerC* gene product, a member of the lambda integrase family of site-specific recombinases. J. Bacteriol. **172:**6973-6980.

46. **Cornet, F., I. Mortier, J. Patte, and J. M. Louarn.** 1994. Plasmid pSC101 harbors a recombination site, *psi*, which is able to resolve plasmid multimers and to substitute for the analogous chromosomal *Escherichia coli* site *dif*. J Bacteriol **176:**3188-95.

47. **Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas.** 1988. Identification and classification of bacterial plasmids. Microbiol Rev **52:**375-95.

48. **Davies, J.** 1994. Inactivation of antibiotics and the dissemination of resistance genes. Science **264:**375-82.

49. **Davis, B. M., H. H. Kimsey, A. V. Kane, and M. K. Waldor.** 2002. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO J **21:**4240-9.

50. **de Boer, H. A., L. J. Comstock, and M. Vasser.** 1983. The tac promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc Natl Acad Sci U S A **80:**21-5.

51. **Deneke, J., G. Ziegelin, R. Lurz, and E. Lanka.** 2000. The protelomerase of temperate *Escherichia coli* phage N15 has cleaving-joining activity. Proc Natl Acad Sci U S A **97:**7721-6.

52. **Drouin, F., J. Melancon, and P. H. Roy.** 2002. The IntI-like tyrosine recombinase of *Shewanella oneidensis* is active as an integron integrase. J. Bacteriol. **184:**1811-5.

53. **Enquist, L. W., A. Kikuchi, and R. A. Weisberg.** 1979. The role of lambda integrase in integration and excision. Cold Spring Harb Symp Quant Biol **43 Pt 2:**1115-20.

54. **Esposito, D., and J. J. Scocca.** 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucl. Acids Res. **25:**3605-3614.

55. **Fluit, A. C., and F. J. Schmitz.** 2004. Resistance integrons and super-integrons. Clinical Microbiology and Infection **10:**272-288.

56. **Francia, M. V., J. C. Zabala, F. de la Cruz, and J. M. Garcia Lobo.** 1999. The IntI1 integron integrase preferentially binds single-stranded DNA of the *attC* site. J. Bacteriol. **181:**6844-6849.

57. **Friesen, H., and P. D. Sadowski.** 1992. Mutagenesis of a conserved region of the gene encoding the FLP recombinase of *Saccharomyces cerevisiae*. A role for arginine 191 in binding and ligation. J Mol Biol **225:**313-26.

58. **Froehlich, B., J. Parkhill, M. Sanders, M. A. Quail, and J. R. Scott.** 2005. The pCoo plasmid of enterotoxigenic *Escherichia coli* is a mosaic cointegrate. J Bacteriol **187:**6509-16.

59. **Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.** 2005. MOBILE GENETIC ELEMENTS: THE AGENTS OF OPEN SOURCE EVOLUTION. Nature Reviews Microbiology **3:**722-732.

60. **Gillings, M. R., M. P. Holley, H. W. Stokes, and A. J. Holmes.** 2005. Integrons in *Xanthomonas*: a source of species genome diversity. Proc Natl Acad Sci U S A **102:**4419-24.

61. **Goldstein, C., M. D. Lee, S. Sanchez, C. Hudson, B. Phillips, B. Register, M. Grady, C. Liebert, A. O. Summers, D. G. White, and J. J. Maurer.** 2001. Incidence of class 1 and 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. Antimicrob. Agents Chemother. **45:**723-6.

62. **Gopaul, D. N., F. Guo, and G. D. Van Duyne.** 1998. Structure of the Holliday junction intermediate in Cre-*loxP* site-specific recombination. EMBO J **17:**4175-87.

63. **Gourlay, S. C., and S. D. Colloms.** 2004. Control of Cre recombination by regulatory elements from Xer recombination systems. Molecular Microbiology **52:**53-65.

64. **Grant, A. J., P. I. Bird, and J. Pittard.** 1980. Naturally occurring plasmids exhibiting incompatibility with members of incompatibility groups I and P. J Bacteriol **144:**758-65.

65. **Gravel, A., N. Messier, and P. H. Roy.** 1998. Point mutations in the integron integrase IntI1 that affect recombination and/or substrate recognition. J. Bacteriol. **180:**5437-5442.

66. **Grindley, N. D. F.** 1997. Site-specific recombination: Synapsis and strand exchange revealed. Current Biology **7:**R608-R612.

67. **Grindley, N. D. F., M. R. Lauth, R. G. Wells, R. J. Wityk, J. J. Salvo, and R. R. Reed.** 1982. Transposon-mediated site-specific recombination: Identification of three binding sites for resolvase at the *res* sites of γδ and Tn 3. Cell **30:**19-27.

68. **Grindley, N. D. F., K. L. Whiteson, and P. A. Rice.** 2006. MECHANISMS OF SITE-SPECIFIC RECOMBINATION. Annual Review of Biochemistry **75:**567-605.

69. **Groth, A. C., and M. P. Calos.** 2004. Phage Integrases: Biology and Applications. Journal of Molecular Biology **335:**667-678.

70. **Groth, A. C., E. C. Olivares, B. Thyagarajan, and M. P. Calos.** 2000. A phage integrase directs efficient site-specific integration in human cells. PNAS **97:**5995-6000.

71. **Guo, F., D. N. Gopaul, and G. D. Van Duyne.** 1999. Asymmetric DNA bending in the Cre-*loxP* site-specific recombination synapse. PNAS **96:**7143-7148.

72. **Guo, F., D. N. Gopaul, and G. D. Van Duyne.** 1997. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. Nature **389:**40-46.

73. **Hall, R. M., D. E. Brookes, and H. W. Stokes.** 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Mol. Microbiol. **5:**1941-59.

74. **Hall, R. M., C. M. Collis, M. J. Kim, S. R. Partridge, G. D. Recchia, and H. W. Stokes.** 1999. Mobile gene cassettes and integrons in evolution. Ann. N. Y. Acad. Sci. **870:**68-80.

75. **Hall, R. M., A. J. Holmes, P. H. Roy, and H. W. Stokes.** 2007. What are superintegrons? Nat Rev Micro **5**.

76. **Hall, R. M., and H. W. Stokes.** 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica **90:**115-32.

77. **Hall, R. M., and C. Vockler.** 1987. The region of the IncN plasmid R46 coding for resistance to beta-lactam antibiotics, streptomycin/spectinomycin and sulphonamides is closely related to antibiotic resistance segments found in IncW plasmids and in Tn*21*-like transposons. Nucleic Acids Res **15:**7491-501.

78. **Hallet, B., L. K. Arciszewska, and D. J. Sherratt.** 1999. Reciprocal control of catalysis by the tyrosine recombinases XerC and XerD: An enzymatic switch in site-specific recombination. Molecular Cell **4:**949-959.

79. **Hallet, B., and D. J. Sherratt.** 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. FEMS Microbiology Reviews **21:**157-178.

80. **Han, Y. W., R. I. Gumport, and J. F. Gardner.** 1993. Complementation of bacteriophage lambda integrase mutants: evidence for an intersubunit active site. EMBO J **12:**4577-84.

81. **Han, Y. W., R. I. Gumport, and J. F. Gardner.** 1994. Mapping the functional domains of bacteriophage lambda integrase protein. J Mol Biol **235:**908-25.

82. **Hansson, K., O. Skold, and L. Sundstrom.** 1997. Non-palindromic *attI* sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. **26:**441-53.

83. **Hansson, K., L. Sundstrom, A. Pelletier, and P. H. Roy.** 2002. IntI2 integron integrase in Tn7. J. Bacteriol. **184:**1712-1721.

84. **Hauman, J. H., R. W. Hedges, W. F. Coetzee, and J. N. Coetzee.** 1982. Plasmid R394 is a cointegrate. J Gen Microbiol **128:**2791-5.

85. **Hoess, R., A. Wierzbicki, and K. Abremski.** 1987. Isolation and characterization of intermediates in site-specific recombination. PNAS **84:**6840-6844.

86. **Hollis, R. P., S. M. Stoll, C. R. Sclimenti, J. Lin, Y. Chen-Tsai, and M. P. Calos.** 2003. Phage integrases for the construction and manipulation of transgenic mammals. Reprod Biol Endocrinol **1:**79.

87. **Holmes, A. J., M. R. Gillings, B. S. Nield, B. C. Mabbutt, K. M. Nevalainen, and H. W. Stokes.** 2003. The gene cassette metagenome is a basic resource for bacterial genome evolution. Environ. Microbiol. **5:**383-94.

88. **Huber, K. E., and M. K. Waldor.** 2002. Filamentous phage integration requires the host recombinases XerC and XerD. Nature **417:**656-659.

89. **Jacoby, G. A., K. E. Walsh, D. M. Mills, V. J. Walker, H. Oh, A. Robicsek, and D. C. Hooper.** 2006. *qnrB*, another plasmid-mediated gene for quinolone resistance. Antimicrob. Agents Chemother. **50:**1178-1182.

90. **Johansson, C., M. Kamali-Moghaddam, and L. Sundstrom.** 2004. Integron integrase binds to bulged hairpin DNA. Nucl. Acids Res. **32:**4033-4043.

91. **Johnson, R., and M. Bruist.** 1989. Intermediates in Hin-mediated DNA inversion: a role for Fis and the recombinational enhancer in the strand exchange reaction. EMBO J **8:**1581-90.

92. **Kahmann, R., F. Rudt, C. Koch, and G. Mertens.** 1985. G inversion in bacteriophage Mu DNA is stimulated by a site within the invertase gene and a host factor. Cell **41:**771-780.

93. **Keravala, A., A. Groth, S. Jarrahian, B. Thyagarajan, J. Hoyt, P. Kirby, and M. Calos.** 2006. A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. Molecular Genetics and Genomics **276:**135-146.

94. **Kersulyte, D., A. K. Mukhopadhyay, M. Shirai, T. Nakazawa, and D. E. Berg.** 2000. Functional organization and insertion specificity of IS*607*, a chimeric element of *Helicobacter pylori*. J. Bacteriol. **182:**5300-5308.

95. **Klippel, A., K. Cloppenborg, and R. Kahmann.** 1988. Isolation and characterization of unusual *gin* mutants. EMBO J **7:**3983-9.

96. **Klippel, A., R. Kanaar, R. Kahmann, and N. R. Cozzarelli.** 1993. Analysis of strand exchange and DNA binding of enhancer-independent Gin recombinase mutants. EMBO J **12:**1047-57.

97. **Klippel, A., G. Mertens, T. Patschinsky, and R. Kahmann.** 1988. The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA via a phosphoserine at amino acid position 9. Embo J **7:**1229-37.

98. **Kobryn, K., and G. Chaconas.** 2002. ResT, a telomere resolvase encoded by the Lyme disease spirochete. Mol Cell **9:**195-201.

99. **Koch, C., and R. Kahmann.** 1986. Purification and properties of the *Escherichia coli* host factor required for inversion of the G segment in bacteriophage Mu. J. Biol. Chem. **261:**15673-15678.

100. **Kwon, H. J., R. Tirumalai, A. Landy, and T. Ellenberger.** 1997. Flexibility in DNA recombination: Structure of the lambda integrase catalytic core. Science **276:**126-131.

101. **Landy, A.** 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. Annual Review of Biochemistry **58:**913-941.

102. **Lee, J., I. Whang, J. Lee, and M. Jayaram.** 1994. Directed protein replacement in recombination full sites reveals trans-horizontal DNA cleavage by Flp recombinase. Embo J **13:**5346-54.

103. **Lee, L., and P. D. Sadowski.** 2001. Directional resolution of synthetic Holliday structures by the Cre recombinase. J. Biol. Chem. **276:**31092-31098.

104. **Lee, L., and P. D. Sadowski.** 2003. Sequence of the *loxP* Site Determines the Order of Strand Exchange by the Cre Recombinase. Journal of Molecular Biology **326:**397-412.

105. **Leffers, G. G., Jr., and S. Gottesman.** 1998. Lambda Xis degradation *in vivo* by Lon and FtsH. J. Bacteriol. **180:**1573-1577.

106. **Leon, G., and P. H. Roy.** 2003. Excision and integration of cassettes by an integron integrase of *Nitrosomonas europaea*. J Bacteriol **185:**2036-41.

107. **Lesterlin, C., F.-X. Barre, and F. Cornet.** 2004. Genetic recombination and the cell cycle: what we have learned from chromosome dimers. Molecular Microbiology **54:**1151-1160.

108. **Leverstein-van Hall, M. A., A. T. Box, H. E. Blok, A. Paauw, A. C. Fluit, and J. Verhoef.** 2002. Evidence of extensive interspecies transfer of integron-mediated antimicrobial resistance genes among multidrug-resistant Enterobacteriaceae in a clinical setting. J Infect Dis **186:**49-56.

109. **Leverstein-van Hall, M. A., M. B. HE, T. D. AR, A. Paauw, A. C. Fluit, and J. Verhoef.** 2003. Multidrug resistance among Enterobacteriaceae is strongly associated with the presence of integrons and is independent of species or isolate origin. J. Infect. Dis. **187:**251-9.

110. **Leverstein-Van Hall, M. A., A. Paauw, A. T. Box, H. E. Blok, J. Verhoef, and A. C. Fluit.** 2002. Presence of integron-associated resistance in the community is widespread and contributes to multidrug resistance in the hospital. J Clin Microbiol **40:**3038-40.

111. **Lévesque, C., S. Brassard, J. Lapointe, and P. H. Roy.** 1994. Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. Gene **142:**49-54.

112. **Lévesque, C., L. Piche, C. Larose, and P. H. Roy.** 1995. PCR mapping of integrons reveals several novel combinations of resistance genes. Antimicrob. Agents Chemother. **39:**185-91.

113. **Liebert, C. A., R. M. Hall, and A. O. Summers.** 1999. Transposon Tn*21*, flagship of the floating genome. Microbiol. Mol. Biol. Rev. **63:**507-522.

114. **MacDonald, D., G. I. Demarre, M. Bouvier, D. Mazel, and D. N. Gopaul.** 2006. Structural basis for broad DNA-specificity in integron recombination. Nature **440:**1157-1162.

115. **MacWilliams, M. P., R. I. Gumport, and J. F. Gardner.** 1996. Genetic analysis of the bacteriophage lambda *attL* nucleoprotein complex. Genetics **143:**1069-79.

116. **Martinez, E., and F. de la Cruz.** 1990. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. **9:**1275-81.

117. **Martinez, E., and F. de la Cruz.** 1988. Transposon Tn*21* encodes a RecA-independent site-specific integration system. Mol. Gen. Genet. **211:**320-5.

118. **Mazel, D.** 2006. Integrons: agents of bacterial evolution. Nat Rev Micro **4:**608-620.

119. **Mazel, D., and J. Davies.** 1999. Antibiotic resistance in microbes. Cell Mol. Life Sci. **56:**742-54.

120. **Mazel, D., B. Dychinco, V. A. Webb, and J. Davies.** 1998. A distinctive class of integron in the *Vibrio cholerae* genome. Science **280:**605-8.

121. **Mazel, D., B. Dychinco, V. A. Webb, and J. Davies.** 2000. Antibiotic resistance in the ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob. Agents Chemother. **44:**1568-74.

122. **McCulloch, R., L. W. Coggins, S. D. Colloms, and D. J. Sherratt.** 1994. Xer-mediated site-specific recombination at *cer* generates Holliday junctions *in vivo*. EMBO J **13:**1844-55.

123. **Messier, N., and P. H. Roy.** 2001. Integron integrases possess a unique additional domain necessary for activity. J. Bacteriol. **183:**6699-6706.

124. **Michael, C. A.** 2004. Mobile gene cassettes: a fundamental resource for bacterial evolution. Am. Nat. **164:**1-12. Epub 2004 May 03.

125. **Moitoso de Vargas, L., S. Kim, and A. Landy.** 1989. DNA looping generated by DNA bending protein IHF and the two domains of lambda integrase. Science **244:**1457-61.

126. **Nandi, S., J. J. Maurer, C. Hofacre, and A. O. Summers.** 2004. Gram-positive bacteria are a major reservoir of Class 1 antibiotic resistance integrons in poultry litter. Proc. Natl. Acad. Sci. USA **101:**7118-7122.

127. **Nesvera, J., J. Hochmannova, and M. Patek.** 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. FEMS Microbiol. Lett. **169:**391-5.

128. **Nield, B. S., A. J. Holmes, M. R. Gillings, G. D. Recchia, B. C. Mabbutt, K. M. H. Nevalainen, and H. W. Stokes.** 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. **195:**59-65.

129. **Nield, B. S., R. D. Willows, A. E. Torda, M. R. Gillings, A. J. Holmes, K. M. Nevalainen, H. W. Stokes, and B. C. Mabbutt.** 2004. New enzymes from environmental cassette arrays: functional attributes of a phosphotransferase and an RNA-methyltransferase. Protein Sci **13:**1651-9.

130. **Novick, R. P., R. C. Clowes, S. N. Cohen, R. Curtiss, 3rd, N. Datta, and S. Falkow.** 1976. Uniform nomenclature for bacterial plasmids: a proposal. Bacteriol Rev **40:**168-89.

131. **Nunes-Duby, S. E., M. A. Azaro, and A. Landy.** 1995. Swapping DNA strands and sensing homology without branch migration in λ site-specific recombination. Current Biology **5:**139-148.

132. **Nunes-Duby, S. E., H. J. Kwon, R. S. Tirumalai, T. Ellenberger, and A. Landy.** 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. Nucl. Acids Res. **26:**391-406.

133. **Nunes-Duby, S. E., R. S. Tirumalai, L. Dorgai, E. Yagil, R. A. Weisberg, and A. Landy.** 1994. Lambda integrase cleaves DNA *in cis*. Embo J **13:**4421-30.

134. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299-304.

135. **Ohno, S.** 1970. Evolution by Gene Duplication. Springer-Verlag, New York.

136. **Olivares, E. C., R. P. Hollis, and M. P. Calos.** 2001. Phage R4 integrase mediates site-specific integration in human cells. Gene **278:**167-176.

137. **Ortiz-Urda, S., Q. Lin, C. L. Green, D. R. Keene, M. P. Marinkovich, and P. A. Khavari.** 2003. Injection of genetically engineered fibroblasts corrects regenerated human epidermolysis bullosa skin tissue. J. Clin. Invest. **111:**251-255.

138. **Ortiz-Urda, S., B. Thyagarajan, D. Keene, Q. Lin, M. Calos, and P. Khavari.** 2003. PhiC31 integrase-mediated nonviral genetic correction of junctional epidermolysis bullosa. Hum Gene Ther **14:**923-8.

139. **Ortiz-Urda, S., B. Thyagarajan, D. R. Keene, Q. Lin, M. Fang, M. P. Calos, and P. A. Khavari.** 2002. Stable nonviral genetic correction of inherited human skin disease. Nat Med **8:**1166-1170.

140. **Parsons, R. L., P. V. Prasad, R. M. Harshey, and M. Jayaram.** 1988. Step-arrest mutants of FLP recombinase: implications for the catalytic mechanism of DNA recombination. Mol Cell Biol **8:**3303-10.

141. **Partridge, S. R., H. J. Brown, and R. M. Hall.** 2002. Characterization and movement of the class 1 integron known as Tn*2521* and Tn*1405*. Antimicrob. Agents Chemother. **46:**1288-94.

142. **Partridge, S. R., C. M. Collis, and R. M. Hall.** 2002. Class 1 integron containing a new gene cassette, *aadA10*, associated with Tn*1404* from R151. Antimicrob. Agents Chemother. **46:**2400-8.

143. **Partridge, S. R., G. D. Recchia, H. W. Stokes, and R. M. Hall.** 2001. Family of class 1 integrons related to In4 from Tn*1696*. Antimicrob. Agents Chemother. **45:**3014-20.

144. **Paul, S., and D. Summers.** 2004. ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. Plasmid **52:**63-68.

145. **Perals, K., F. Cornet, Y. Merlet, I. Delon, and J.-M. Louarn.** 2000. Functional polarization of the Escherichia coli chromosome terminus: the dif site acts in chromosome dimer resolution only when located between long stretches of opposite polarity. Molecular Microbiology **36:**33-43.

146. **Peters, E. D., M. A. Leverstein-van Hall, A. T. Box, J. Verhoef, and A. C. Fluit.** 2001. Novel gene cassettes and integrons. Antimicrob Agents Chemother **45:**2961-4.

147. **Plante, I., D. Centron, and P. H. Roy.** 2003. An integron cassette encoding erythromycin esterase, *ere(A)*, from *Providencia stuartii*. J Antimicrob Chemother **51:**787-90.

148. **Ploy, M. C., P. Courvalin, and T. Lambert.** 1998. Characterization of In40 of *Enterobacter aerogenes* BM2688, a class 1 integron with two new gene cassettes, *cmlA2* and *qacF*. Antimicrob Agents Chemother **42:**2557-63.

149. **Poirel, L., T. Naas, D. Nicolas, L. Collet, S. Bellais, J.-D. Cavallo, and P. Nordmann.** 2000. Characterization of VIM-2, a carbapenem-hydrolyzing metallo-

beta -lactamase and its plasmid- and integron-borne gene from a *Pseudomonas aeruginosa* clinical isolate in France. Antimicrob. Agents Chemother. **44:**891-897.

150. **Radman-Livaja, M., C. Shaw, M. Azaro, T. Biswas, T. Ellenberger, and A. Landy.** 2003. Arm sequences contribute to the architecture and catalytic function of a lambda integrase-Holliday junction complex. Molecular Cell **11:**783-794.

151. **Radstrom, P., O. Skold, G. Swedberg, J. Flensburg, P. H. Roy, and L. Sundstrom.** 1994. Transposon Tn*5090* of plasmid R751, which carries an integron, is related to Tn7, Mu, and the retroelements. J Bacteriol **176:**3257-68.

152. **Ramirez, M. S., C. Quiroga, and D. Centron.** 2005. Novel rearrangement of a class 2 integron in two non-epidemiologically related isolates of *Acinetobacter baumannii*. Antimicrob Agents Chemother **49:**5179-81.

153. **Recchia, G. D., M. Aroyo, D. Wolf, G. Blakely, and D. J. Sherratt.** 1999. FtsK-dependent and -independent pathways of Xer site-specific recombination. EMBO J **18:**5724-34.

154. **Recchia, G. D., and R. M. Hall.** 1995. Gene cassettes: a new class of mobile element. Microbiology **141 ( Pt 12):**3015-27.

155. **Recchia, G. D., and R. M. Hall.** 1997. Origins of the mobile gene cassettes found in integrons. Trends Microbiol. **5:**389-94.

156. **Recchia, G. D., and D. J. Sherratt.** 2002. Gene Acquisition in Bacteria by Integron-Mediated Site-Specific Recombination, p. 162-176. *In* N. L. Craig, R. Craigie, M. Gellert, A. M. Lambowitz (ed.), Mobile DNA II. ASM Press, Washington, D. C.

157. **Recchia, G. D., H. W. Stokes, and R. M. Hall.** 1994. Characterisation of specific and secondary recombination sites recognised by the integron DNA integrase. Nucleic Acids Res. **22:**2071-8.

158. **Rice, P. A., and T. A. Steitz.** 1994. Model for a DNA-mediated synaptic complex suggested by crystal packing of gamma delta resolvase subunits. EMBO J **13:**1514-24.

159. **Robertson, C. A., and H. A. Nash.** 1988. Bending of the bacteriophage lambda attachment site by *Escherichia coli* integration host factor. J. Biol. Chem. **263:**3554-3557.

160. **Rowe-Magnus, D. A., A.-M. Guerout, and D. Mazel.** 2002. Bacterial resistance evolution by recruitment of super-integron gene cassettes. Molecular Microbiology **43:**1657-1669.

161. **Rowe-Magnus, D. A., A. M. Guerout, L. Biskri, P. Bouige, and D. Mazel.** 2003. Comparative analysis of superintegrons: engineering extensive genetic diversity in the Vibrionaceae. Genome Res **13:**428-42.

162. **Rowe-Magnus, D. A., A. M. Guerout, and D. Mazel.** 1999. Super-integrons. Res Microbiol **150:**641-51.

163. **Rowe-Magnus, D. A., A. M. Guerout, P. Ploncard, B. Dychinco, J. Davies, and D. Mazel.** 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. Proc. Natl. Acad. Sci. USA **98:**652-7.

164. **Rowe-Magnus, D. A., and D. Mazel.** 2002. The role of integrons in antibiotic resistance gene capture. Int J Med Microbiol **292:**115-25.

165. **Salyers, A. A., and C. F. Amabile-Cuevas.** 1997. Why are antibiotic resistance genes so resistant to elimination? Antimicrob. Agents Chemother. **41:**2321-5.

166. **Sandvang, D.** 1999. Novel streptomycin and spectinomycin resistance gene as a gene cassette within a class 1 integron isolated from *Escherichia coli*. Antimicrob. Agents Chemother. **43:**3036-8.

167. **Sandvang, D., and F. M. Aarestrup.** 2000. Characterization of aminoglycoside resistance genes and class 1 integrons in porcine and bovine gentamicin-resistant *Escherichia coli*. Microb Drug Resist **6:**19-27.

168. **Sarkis, G. J., L. L. Murley, A. E. Leschziner, M. R. Boocock, W. M. Stark, and N. D. F. Grindley.** 2001. A model for the gamma delta resolvase synaptic complex. Molecular Cell **8:**623-631.

169. **Sauer, B.** 1998. Inducible gene targeting in mice using the Cre/*lox* system. Methods **14:**381-392.

170. **Schneider, F., M. Schwikardi, G. Muskhelishvili, and P. Droge.** 2000. A DNA-binding domain swap converts the invertase *gin* into a resolvase. Journal of Molecular Biology **295:**767-775.

171. **Senda, K., Y. Arakawa, S. Ichiyama, K. Nakashima, H. Ito, S. Ohsuka, K. Shimokata, N. Kato, and M. Ohta.** 1996. PCR detection of metallo-beta-lactamase gene ($bla_{IMP}$) in gram-negative rods resistant to broad-spectrum beta-lactams. J Clin Microbiol **34:**2909-13.

172. **Shaikh, A. C., and P. D. Sadowski.** 1997. The Cre recombinase cleaves the *lox* site *in trans*. J. Biol. Chem. **272:**5695-5702.

173. **Shapiro, J. A., and P. Sporn.** 1977. Tn*402*: a new transposable element determining trimethoprim resistance that inserts in bacteriophage lambda. J Bacteriol **129:**1632-5.

174. **Shibata, N., Y. Doi, K. Yamane, T. Yagi, H. Kurokawa, K. Shibayama, H. Kato, K. Kai, and Y. Arakawa.** 2003. PCR typing of genetic determinants for metallo-beta-lactamases and integrases carried by gram-negative bacteria isolated in Japan, with focus on the class 3 integron. J Clin Microbiol **41:**5407-13.

175. **Smith, M. C. A., R. Till, K. Brady, P. Soultanas, H. Thorpe, and M. C. M. Smith.** 2004. Synapsis and DNA cleavage in ΦC31 integrase-mediated site-specific recombination. Nucl. Acids Res. **32:**2607-2617.

176. **Smith, M. C. M., and H. M. Thorpe.** 2002. Diversity in the serine recombinases. Molecular Microbiology **44:**299-307.

177. **Spaeny-Dekking, L., E. Schlicher, K. Franken, P. van de Putte, and N. Goosen.** 1995. Gin mutants that can be suppressed by a Fis-independent mutation. J Bacteriol **177:**222-8.

178. **Stark, W. M., D. J. Sherratt, and M. R. Boocock.** 1989. Site-specific recombination by Tn3 resolvase: Topological changes in the forward and reverse reactions. Cell **58:**779-790.

179. **Steiner, W., G. Liu, W. D. Donachie, and P. Kuempel.** 1999. The cytoplasmic domain of FtsK protein is required for resolution of chromosome dimers. Molecular Microbiology **31:**579-583.

180. **Stokes, H. W., and R. M. Hall.** 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol. Microbiol. **3:**1669-83.

181. **Stokes, H. W., and R. M. Hall.** 1991. Sequence analysis of the inducible chloramphenicol resistance determinant in the Tn*1696* integron suggests regulation by translational attenuation. Plasmid **26:**10-9.

182. **Stokes, H. W., and R. M. Hall.** 1992. The integron In1 in plasmid R46 includes two copies of the *oxa2* gene cassette. Plasmid **28:**225-34.

183. **Stokes, H. W., A. J. Holmes, B. S. Nield, M. P. Holley, K. M. Nevalainen, B. C. Mabbutt, and M. R. Gillings.** 2001. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. Appl. Environ. Microbiol. **67:**5240-6.

184. **Stokes, H. W., C. L. Nesbo, M. Holley, M. I. Bahl, M. R. Gillings, and Y. Boucher.** 2006. Class 1 integrons potentially predating the association with Tn*402*-like transposition genes are present in a sediment microbial community. J. Bacteriol. **188:**5722-5730.

185. **Stokes, H. W., D. B. O'Gorman, G. D. Recchia, M. Parsekhian, and R. M. Hall.** 1997. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. Mol Microbiol **26:**731-45.

186. **Stoll, S. M., D. S. Ginsburg, and M. P. Calos.** 2002. Phage TP901-1 site-specific integrase functions in human cells. J. Bacteriol. **184:**3657-3663.

187. **Subramanya, H. S., L. K. Arciszewska, R. A. Baker, L. E. Bird, D. J. Sherratt, and D. B. Wigley.** 1997. Crystal structure of the site-specific recombinase, XerD. EMBO J **16:**5178-87.

188. **Summers, D. K., C. W. Beton, and H. L. Withers.** 1993. Multicopy plasmid instability: the dimer catastrophe hypothesis. Mol Microbiol **8:**1031-8.

189. **Summers, D. K., and D. J. Sherratt.** 1984. Multimerization of high copy number plasmids causes instability: CoIE1 encodes a determinant essential for plasmid monomerization and stability. Cell **36:**1097-103.

190. **Swalla, B. M., R. I. Gumport, and J. F. Gardner.** 2003. Conservation of structure and function among tyrosine recombinases: homology-based modeling of the lambda integrase core-binding domain. Nucl. Acids Res. **31:**805-818.

191. **Sykora, P.** 1992. Macroevolution of plasmids: a model for plasmid speciation. J Theor Biol **159:**53-65.

192. **Tecklenburg, M., A. Naumer, O. Nagappan, and P. Kuempel.** 1995. The *dif* resolvase locus of the *Escherichia coli* chromosome can be replaced by a 33-bp sequence, but function depends on location. Proc Natl Acad Sci U S A **92:**1352-6.

193. **Thorpe, H. M., and M. C. M. Smith.** 1998. *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. PNAS **95:**5505-5510.

194. **Thungapathra, M., Amita, K. K. Sinha, S. R. Chaudhuri, P. Garg, T. Ramamurthy, G. B. Nair, and A. Ghosh.** 2002. Occurrence of antibiotic resistance gene cassettes *aac(6')-Ib*, *dfrA5*, *dfrA12*, and *ereA2* in class I integrons in non-O1, non-O139 *Vibrio cholerae* strains in India. Antimicrob Agents Chemother **46:**2948-55.

195. **Thyagarajan, B., E. C. Olivares, R. P. Hollis, D. S. Ginsburg, and M. P. Calos.** 2001. Site-specific genomic integration in mammalian cells mediated by phage ΦC31 integrase. Mol. Cell. Biol. **21:**3926-3934.

196. **Tosini, F., P. Visca, I. Luzzi, A. M. Dionisi, C. Pezzella, A. Petrucca, and A. Carattoli.** 1998. Class 1 integron-borne multiple-antibiotic resistance carried by IncFI and IncL/M plasmids in *Salmonella enterica* serotype Typhimurium. Antimicrob. Agents Chemother. **42:**3053-8.

197. **Tran, J. H., and G. A. Jacoby.** 2002. Mechanism of plasmid-mediated quinolone resistance. Proceedings of the National Academy of Sciences **99:**5638-5642.

198. **Vaisvila, R., R. D. Morgan, J. Posfai, and E. A. Raleigh.** 2001. Discovery and distribution of super-integrons among Pseudomonads. Molecular Microbiology **42:**587-601.

199. **Val, M.-E., M. Bouvier, J. Campos, D. Sherratt, F. Cornet, D. Mazel, and F.-X. Barre.** 2005. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. Molecular Cell **19:**559-566.

200. **Van Duyne, G. D.** 2001. A STRUCTURAL VIEW OF Cre-*loxP* SITE-SPECIFIC RECOMBINATION. Annual Review of Biophysics and Biomolecular Structure **30:**87-104.

201. **Van Duyne, G. D.** 2005. Lambda Integrase: Armed for Recombination. Current Biology **15:**R658-R660.

202. **Vourli, S., L. S. Tzouvelekis, E. Tzelepi, E. Lebessi, N. J. Legakis, and V. Miriagou.** 2003. Characterization of In111, a class 1 integron that carries the extended-spectrum β-lactamase gene $bla_{IBC-1}$. FEMS Microbiology Letters **225:**149-153.

203. **Wang, L., and J. Lutkenhaus.** 1998. FtsK is an essential cell division protein that is localized to the septum and induced as part of the SOS response. Molecular Microbiology **29:**731-740.

204. **Watanabe, T., C. Furuse, and S. Sakaizumi.** 1968. Transduction of Various R Factors by Phage P1 in *Escherichia coli* and by Phage P22 in *Salmonella typhimurium*. J. Bacteriol. **96:**1791-1795.

205. **White, P. A., C. J. McIver, and W. D. Rawlinson.** 2001. Integrons and gene cassettes in the Enterobacteriaceae. Antimicrob Agents Chemother **45:**2658-61.

206. **Wierzbicki, A., M. Kendall, K. Abremski, and R. Hoess.** 1987. A mutational analysis of the bacteriophage P1 recombinase Cre. J Mol Biol **195:**785-94.

207. **Wojciak, J. M., D. Sarkar, A. Landy, and R. T. Clubb.** 2002. Arm-site binding by lambda -integrase: Solution structure and functional characterization of its amino-terminal domain. PNAS **99:**3434-3439.

208. **Xu, H., J. Davies, and V. Miao.** 2007. Molecular characterization of class 3 integrons from *Delftia* spp. J. Bacteriol.**:**JB.00348-07.

209. **Yildiz, F. H., X. S. Liu, A. Heydorn, and G. K. Schoolnik.** 2004. Molecular analysis of rugosity in a *Vibrio cholerae* O1 El Tor phase variant. Molecular Microbiology **53:**497-515.

210. **Yin, S., W. Bushman, and A. Landy.** 1985. Interaction of the lambda site-specific recombination protein Xis with attachment site DNA. PNAS **82:**1040-1044.

211. **Zuhlsdorf, M. T., and B. Wiedemann.** 1992. Tn*21*-specific structures in gram-negative bacteria from clinical isolates. Antimicrob. Agents Chemother. **36:**1915-21.

**Figure 1.1.** Basic site-specific recombination reactions. **A.** Close-up of a basic recombination core site. Each individual DNA recombination site is composed of inverted repeats (purple arrows) surrounding a crossover region of variable size (orange line). The nucleotides show an example of an inverted repeat sequence. **B.** Inversion. Intramolecular recombination between sites of opposite orientation flanking a segment of DNA on a single molecule. **C.** Integration and excision. Intermolecular recombination results in integration (or cointegrate formation if the green and black DNA molecules are both plasmids). Excision, or resolution, is the reverse reaction, recombination between sites in direct repeat on a single DNA molecule. For phage integrases, the blue arrow represents *attP*, the red arrow is *attB*, and the left and right red/blue arrows are *attL* and *attR*, respectively.

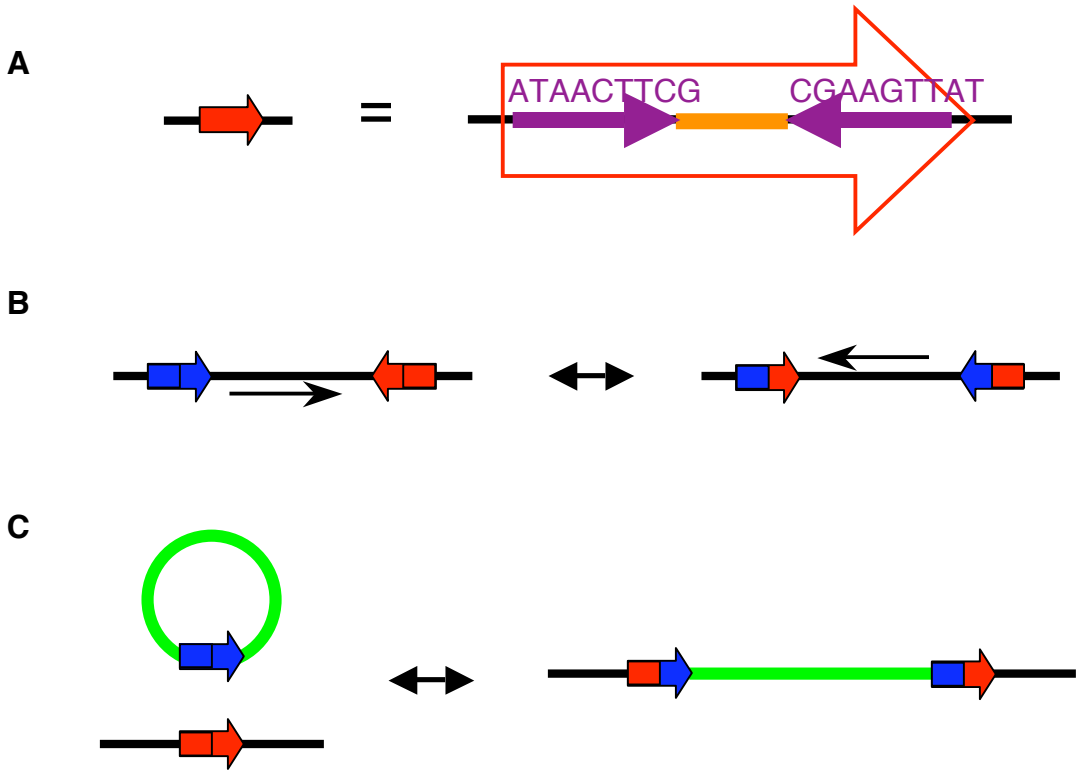

A
=
ATAACTTCG
CGAAGTTAT
B
C

**Figure 1.2.** Protein domains of serine recombinases. The catalytic domains are shown in green, and DNA-binding domains are blue. Orange and purple represent conserved domains of unknown function (176). **A.** The resolvase/invertases. **B.** The IS607-like transposases. **C.** The large serine integrases.

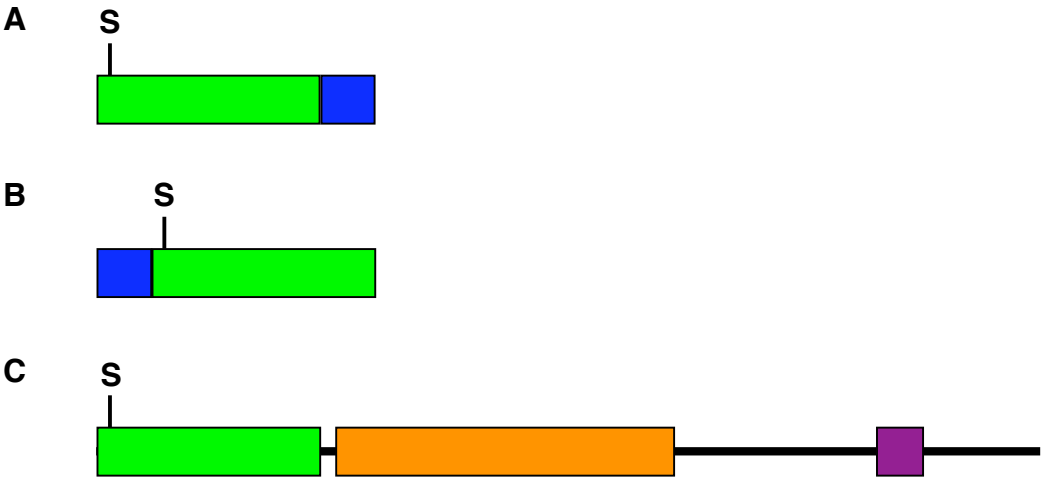

A
S
B
S
C
S

**Figure 1.3.** Phylogenetic tree of the catalytic domains of twelve serine recombinases. Alignment of the catalytic domains was performed by ClustalW using MEGA 4.0 (http://www.megasoftware.net). The tree is drawn to scale, with evolutionary distance as number of amino acid substitutions per site. Numbers are the percentage of replicate trees in which the associated protein sequences clustered together in the bootstrap test. The small serine recombinases are shaded: resolvase/invertases are blue and IS*607*-like transposases are pink. The unshaded proteins are large serine integtrases. The resolvase/invertases group together, as do the IS*607*-like transposases. The large integrases do not form their own cluster and are just as likely to appear more related to one of the other groups as to each other (176). The recombinases are: Hin=invertase, *Salmonella typhimurium* (NP_461699); Gin=invertase, bacteriophage Mu (NP_050655); Tn3=resolvase, *Escherichia coli* (YP_190221); γδ=resolvase,Tn*1000* on F plasmid (NP_061388); TP901=integrase, *Lactococcus lactis* phage (NP_112664); TnpX=integrase, *Clostridium difficile* Tn*4451* (AAF66226); IS1535=IS*607*-like, Mycobacterium tuberculosis (NP_335381); IS607=transposase, *Helicobacter pylori* (YP_626943); ISC1913=IS*607*-like, *Solfolobus sulfataricus*; MJ0014=IS*607*-like, *Methanococcus jannaschii* (Q60329); R4=integrase, *Streptomyces* phage (Q37839); ϕC31=integrase, *Streptomyces* phage (NP_047974).

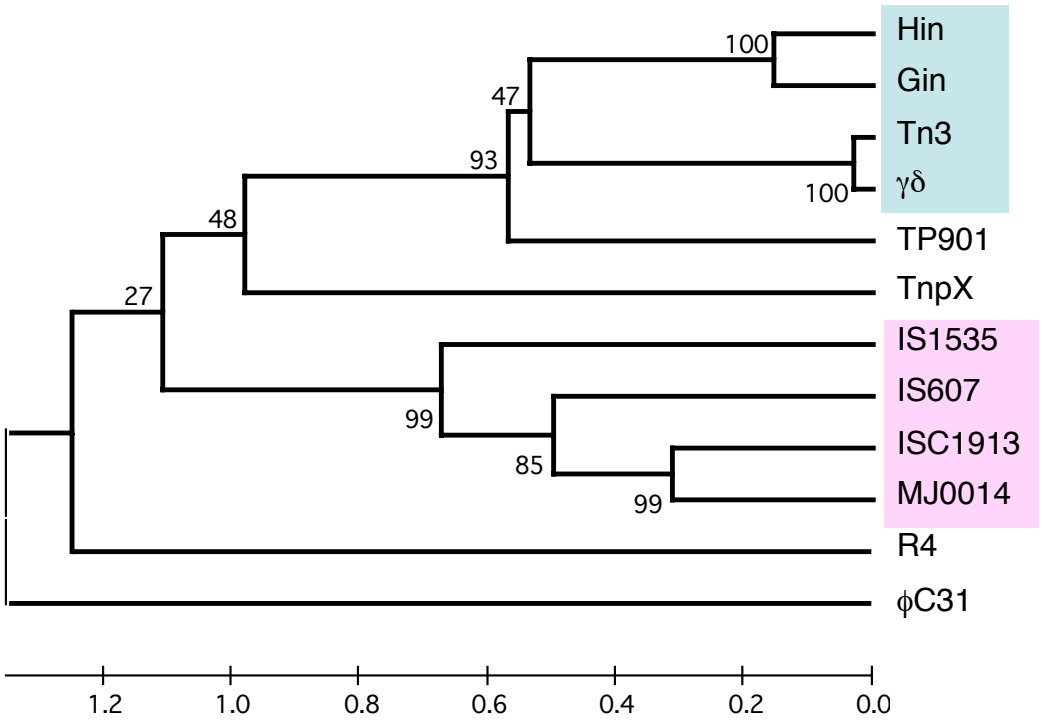

Hin
Gin
Tn3
γδ
TP901
TnpX
IS1535
IS607
ISC1913
MJ0014
R4
ϕC31
100
47
93
100
48
27
99
85
99
1.2
1.0
0.8
0.6
0.4
0.2
0.0

**Figure 1.4.** Serine recombinase reaction mechanism. Circles represent catalytic recombinase subunits; DNA strands are on the outsides of the complex. $S_{OH}$ shows the serine nucleophile not bound to DNA; S shows when bound. The serine nucleophile of the recombinase subunits cleave all four DNA strands and form covalent linkages with the 5' ends, leaving 3' hydroxyls. The complex rotates prior to strand exchange. The 3' hydroxyls attack the phosophoserine linkages to ligate the DNA (68).

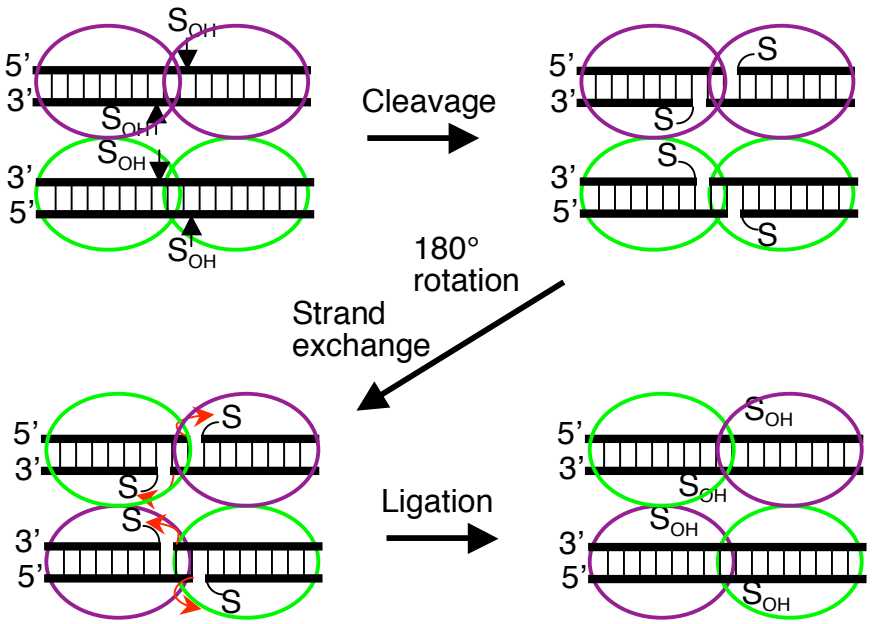

S
OH
5'
3'
Cleavage
S
180°
rotation
Strand
exchange
Ligation

**Figure 1.5.** *res* DNA recombination site for the Tn3 and γδ resolvases. The *res* site is 114 bp in total length with three core-type binding sites (subsites I-III, red boxes). The purple arrows show the 12 bp left (L) and right (R) half-site inverted repeats for recombinase subunit binding in each subsite. The blue triangles indicate the cleavage sites within subsite I. The numbers show the length in base pairs of spacer regions within and between subsites (168).

**Tn3 and γδ *res* site**

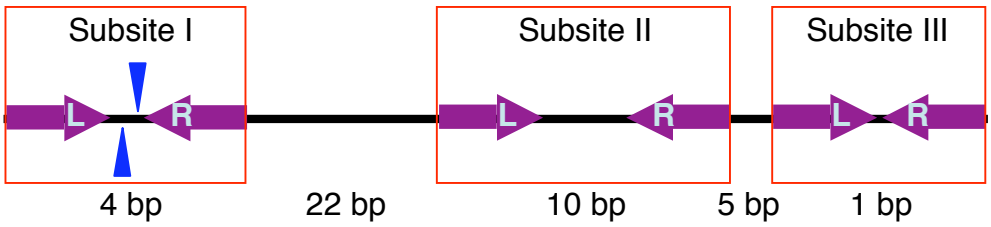


**Figure 1.6**. Protein domains of specific tyrosine recombinases. The core-type DNA-binding domains are blue, and the catalytic domains are green, and the length in residues of each domain are given. Positions of the RKHRHY conserved residues are indicated. **A.** Cre resolvase (72, 200). **B.** XerD (9, 187). **C.** Lambda integrase (1, 69). The arm-binding domain is shown in orange. **D.** VchIntIA integron integrase (114). The unique integron integrase region within the catalytic domain is shown in purple (positions 192-210).

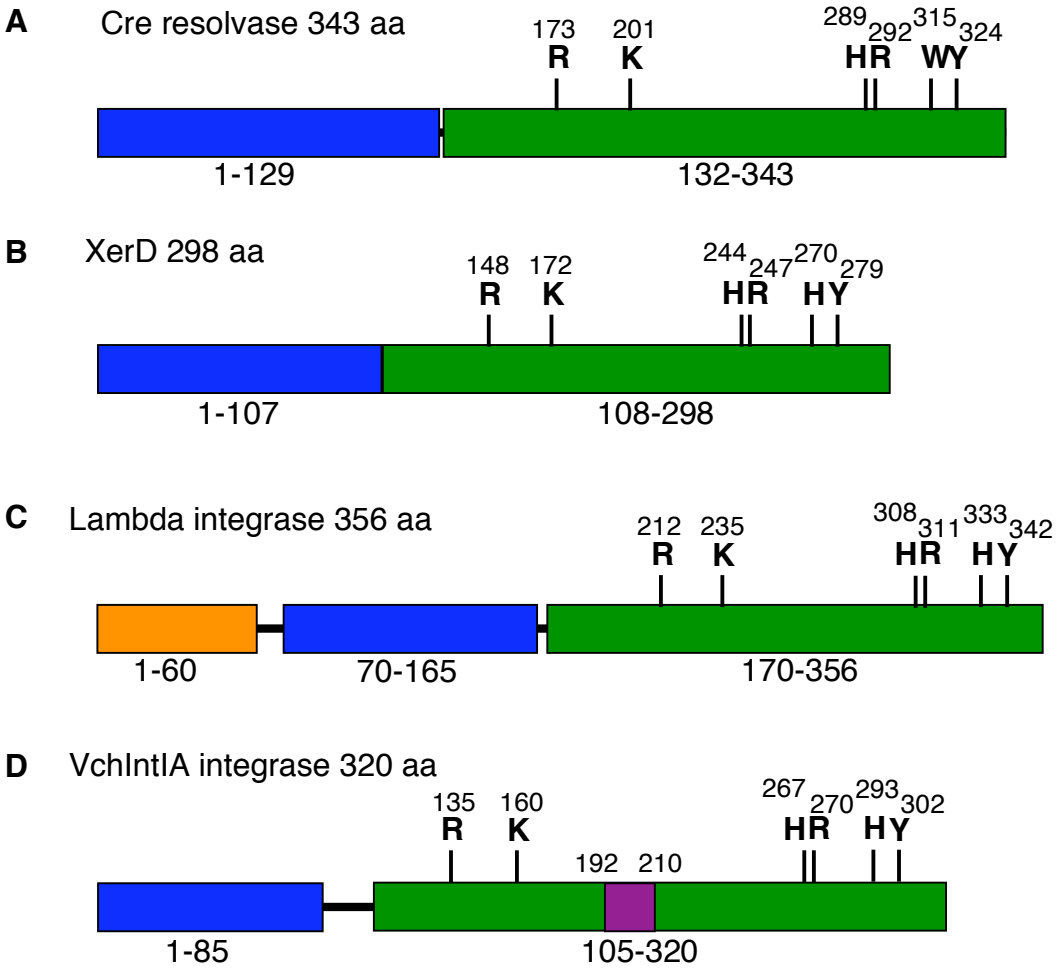

A
Cre resolvase 343 aa
173 R
201 K
289 HR 292
315 WY 324
1-129
132-343
B
XerD 298 aa
148 R
172 K
244 HR 247
270 HY 279
1-107
108-298
C
Lambda integrase 356 aa
212 R
235 K
308 HR 311
333 HY 342
1-60
70-165
170-356
D
VchIntIA integrase 320 aa
135 R
160 K
267 HR 270
293 HY 302
192
210
1-85
105-320

**Figure 1.7.** Tyrosine recombinase reaction mechanism. The mechanism shown is based on the strand-swapping isomerization model and Cre-*loxP* interactions. Ovals represent recombinase subunits; the active subunits are shown in orange and inactive subunits are gray. The active recombinase subunits cleave one strand of DNA, forming phosphotyrosine covalent linkages with the 3' ends, leaving 5' hydroxyls. The 5' hydroxyl then attacks the 3' phosphotyrosine linkage of the opposite DNA molecule, ligating to form a Holliday junction intermediate. The inactive recombinase subunits are activated by isomerization of the complex, and cleavage and ligation of the second pair of DNA strands resolves the Holliday junction to complete the reaction. The black and green lines are the continuous strands in the left side of the figure and the crossing strands in the right side of the figure, and the blue and red lines are the crossing strands on the left and continuous strands on the right. The strand exchange steps shown where cleaved DNA not bound to recombinase anneals to the complementary strand of the opposite DNA molecule is based specifically on the Cre-*loxP* reaction (200).

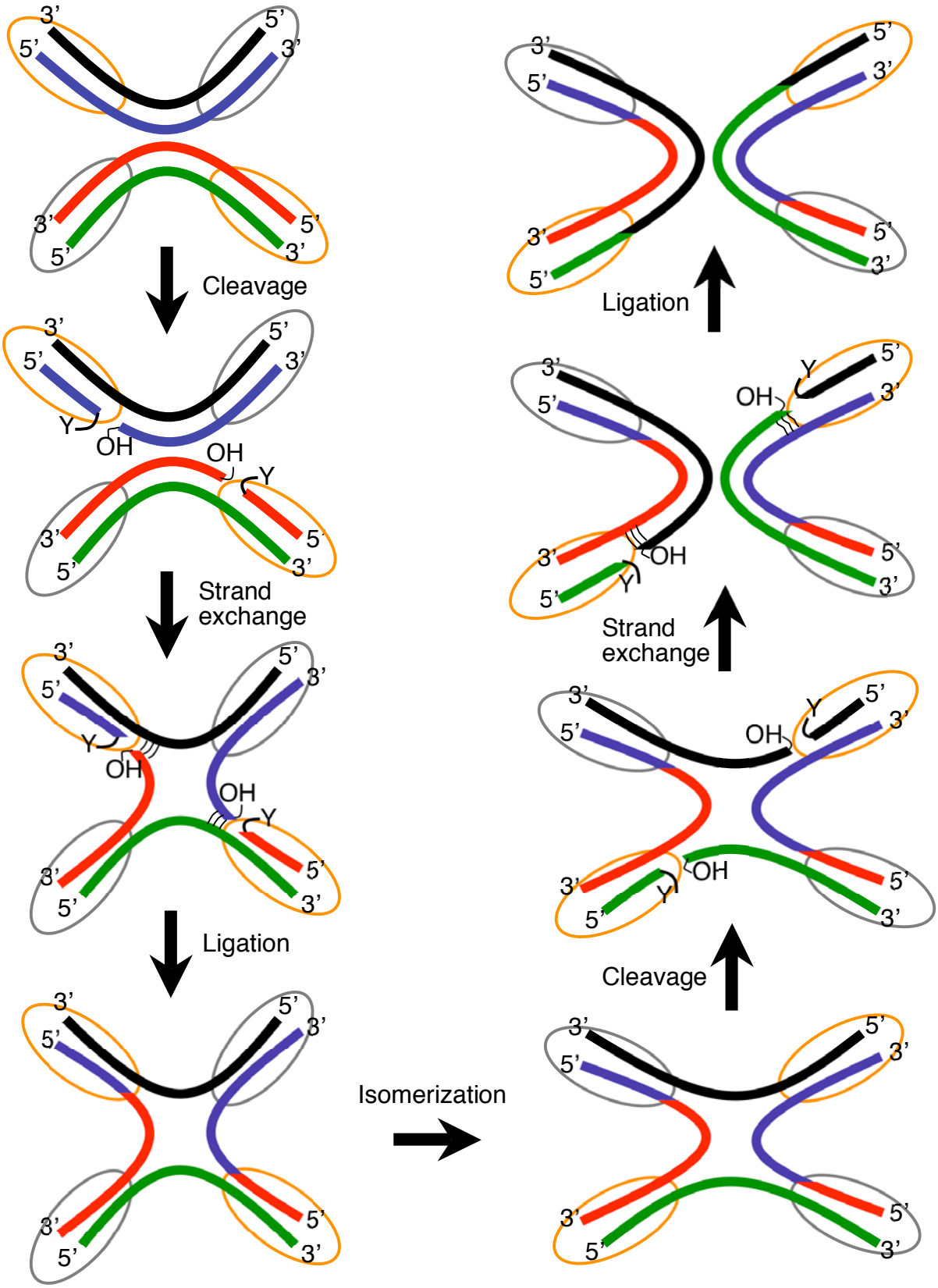

3'
5'
Cleavage
Y
OH
Strand exchange
Ligation
Isomerization
Cleavage
Strand exchange
Ligation

**Figure 1.8.** DNA recombination sites for the tyrosine recombinases Cre, XerCD, and lambda integrase. **A.** The 34 bp *loxP* core site sequence for the Cre resolvase (200). One Cre subunit binds to each 14 bp inverted repeat half site (purple arrows), also known as a recombinase binding element (RBE). Cleavage occurs on either side of the crossover region shown in bold-type. The closed triangle indicates the top strand cleavage site, and the open triangle shows bottom strand cleavage. **B.** Cartoons and core site sequences of XerCD DNA recombination sites (9, 63). Cartoons show accessory binding regions relative to the core sites (orange and green). PepA binding regions are shown in black, and ArcA and ArgR binding regions are shown in aqua and blue, respectively. For the core site sequences, orange arrows show XerC recognition sequences, and green arrows show XerD recognition sequences. **C.** Cartoons and *attB* and *attP* core site sequences of lambda integrase DNA recombination sites (9, 68, 101). Cartoons show core sites in purple. Regions for lambda Int arm-type binding are gray, IHF binding regions are brown, and regions for Xis and Fis binding are red and black, respectively.

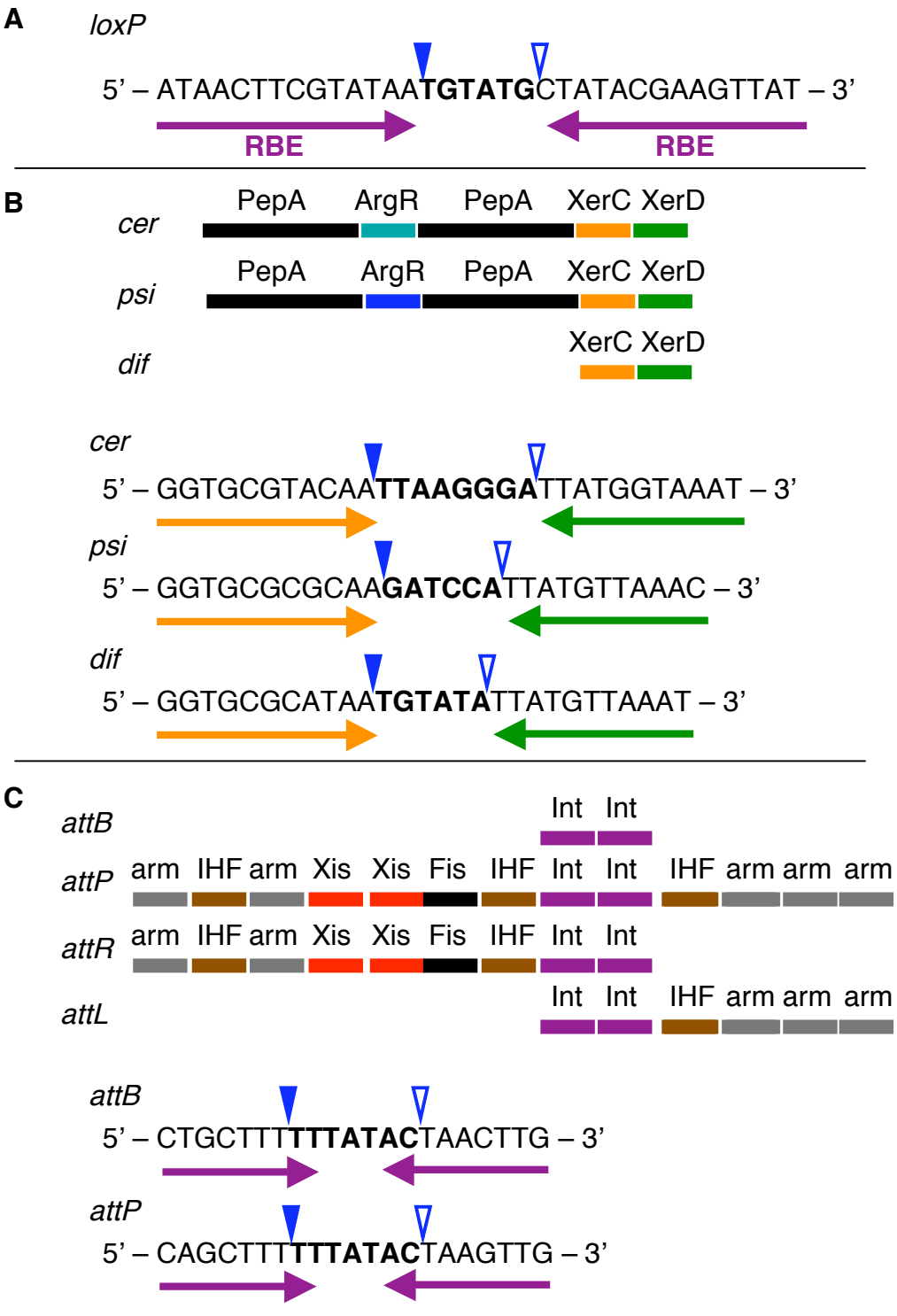

A
loxP
5' – ATAACTTCGTATAATGTATGCTATACGAAGTTAT – 3'
RBE
RBE
B
cer
PepA ArgR PepA XerC XerD
psi
PepA ArgR PepA XerC XerD
dif
XerC XerD
cer
5' – GGTGCGTACAATTAAGGGATTATGGTAAAT – 3'
psi
5' – GGTGCGCGCAAGATCCATTATGTTAAAC – 3'
dif
5' – GGTGCGCATAATGTATATTATGTTAAAT – 3'
C
attB
Int Int
attP
arm IHF arm Xis Xis Fis IHF Int Int IHF arm arm arm
attR
arm IHF arm Xis Xis Fis IHF Int Int
attL
Int Int IHF arm arm arm
attB
5' – CTGCTTTTTTATACTAACTTG – 3'
attP
5' – CAGCTTTTTTATACTAAGTTG – 3'

**Figure 1.9.** Examples of integrons from different classes. Integrases are shown in aqua, and *attI* sites are red ovals. Gene cassettes are green with yellow ovals showing *attC* sites. Arrows indicate the direction of transcription. **A.** Cartoon of a general integron. Approximate positions of the integrase ($P_{int}$) and cassette ($P_c$) promoters are shown. **B.** Class 1 integrons Tn*402*, In0 from pVS1 in *Pseudomonas aeruginosa*, and In2 from Tn*21* (25,143). 3' conserved segments (CS) are shown in black. *tni* transpose modules are shown in white, and insertion sequences (IS) are brown and blue. The gray box within the *tni* module indicates a transposase *res* site, and the vertical black bars (IRi, IRt) show inverted repeats for transpostion. **C.** Class 2 integrons from Tn7 (83) and *Providencia stuartii* (12). The *intI2* stop codon is indicated by a black dot. **D.** The class 3 integron from *Serratia marcescens* (39). The dashed black line indicates an unsequenced region. **E.** The *Vibrio cholerae* superintegron (118). **F.** The *Shewanella oneidensis* chromosomal integron (52).

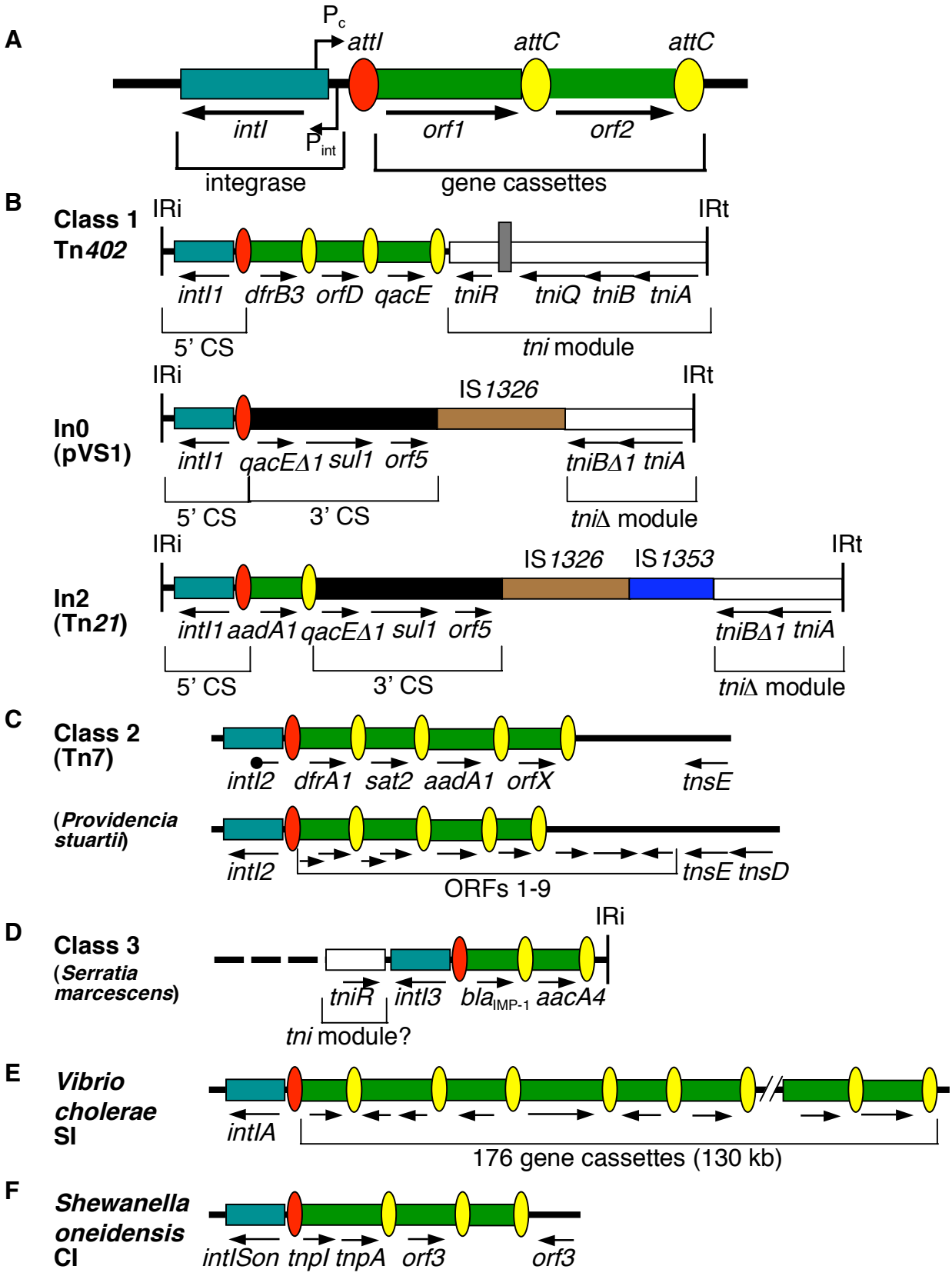

A
P c
attI
attC
attC
intI
P int
orf1
orf2
integrase
gene cassettes
B
Class 1
Tn402
IRi
IRt
intI1
dfrB3
orfD
qacE
tniR
tniQ
tniB
tniA
5' CS
tni module
In0
(pVS1)
IS1326
qacEΔ1
sul1
orf5
tniBΔ1
3' CS
tniΔ module
In2
(Tn21)
IS1353
aadA1
C
Class 2
(Tn7)
intI2
dfrA1
sat2
aadA1
orfX
tnsE
(Providencia stuartii)
tnsD
ORFs 1-9
D
Class 3
(Serratia marcescens)
tniR
intI3
bla IMP-1
aacA4
tni module?
E
Vibrio cholerae SI
intIA
176 gene cassettes (130 kb)
F
Shewanella oneidensis CI
intISon
tnpI
tnpA
orf3

**Figure 1.10.** The excision/integration mechanism of cassette capture. The gray ovals represent *attI*, and the striped ovals show *attC*. The donor integron encodes two gene cassettes, *aacA4* and *aadA2*, and the recipient integron contains an *attI* site and no mobile cassettes. The *aadA2* of the donor integron is excised by the integrase into a covalently closed circle. Then, the integrase inserts the cassette into the *attI* site of the recipient integron. The result is a donor integron lacking the *aadA2* cassette, and the recipient gaining the *aadA2* cassette.

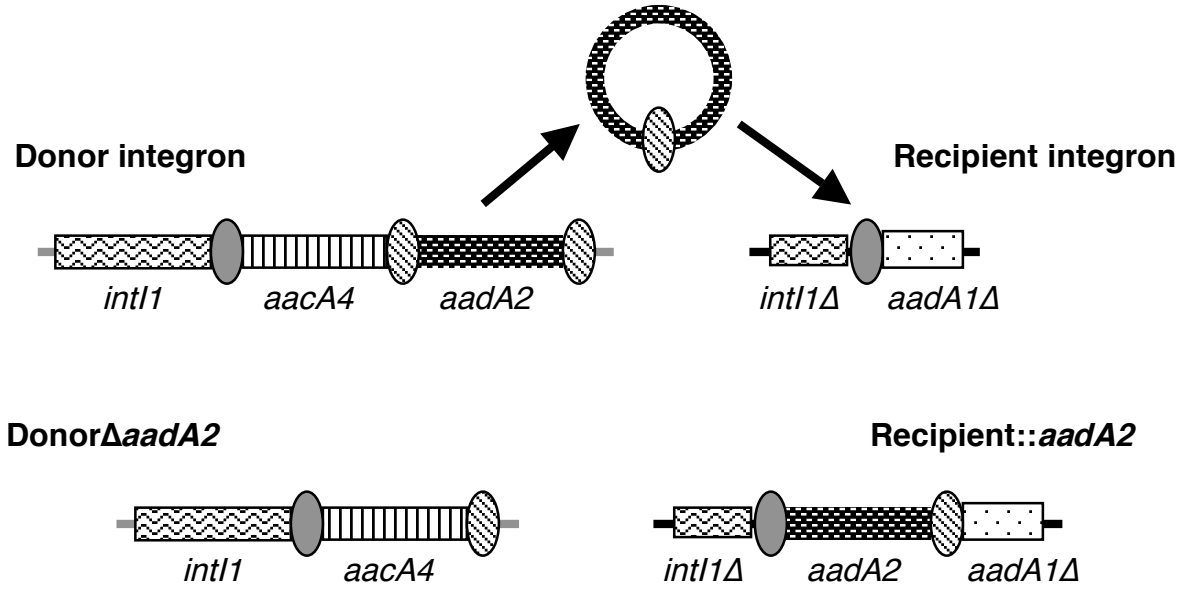

Donor integron
Recipient integron
intI1
aacA4
aadA2
intI1Δ
aadA1Δ
DonorΔaadA2
Recipient::aadA2
intI1
aacA4
intI1Δ
aadA2
aadA1Δ

**Figure 1.11.** The cointegrate formation and resolution mechanism of cassette capture. **A.** Integrase-mediated recombination occurs between one site on pDonor and one site on pRecipient to form a cointegrate. **B.** The cointegrate resulting from *attI* x *attI* recombination is shown. **C.** A second recombination event then resolves the cointegrate into two separate plasmids. The example shown is from *aacA4 attC* x *attI* resolution.

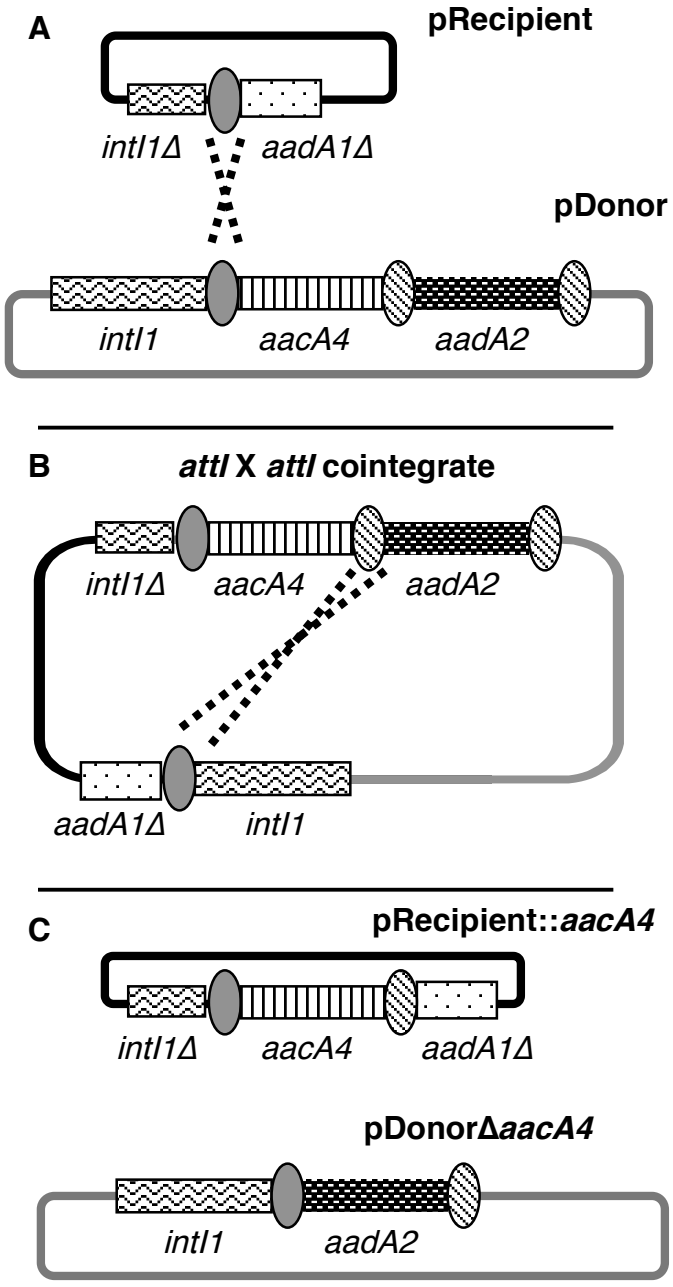
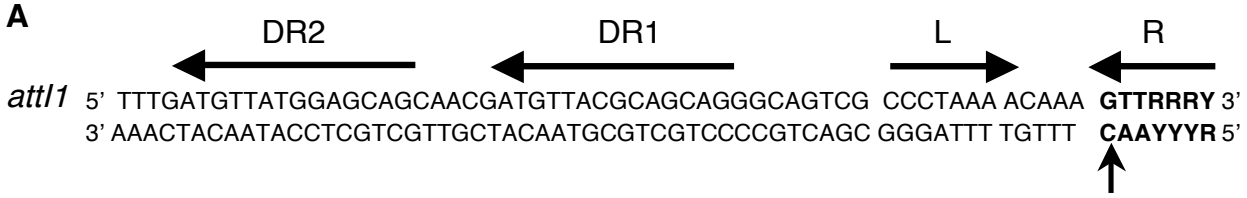

A
pRecipient
intI1Δ
aadA1Δ
pDonor
intI1
aacA4
aadA2
B
attI X attI cointegrate
intI1Δ
aacA4
aadA2
aadA1Δ
intI1
C
pRecipient::aacA4
intI1Δ
aacA4
aadA1Δ
pDonorΔaacA4
intI1
aadA2

**Figure 1.12.** *attI* recombination sites. The "TTRRRY" sequence is part of the gene cassette found adjacent to *attI*. **A.** The *attI1* sequence (40). The direct repeats DR1 (strong IntI1 binding site) and DR2 (weak IntI1 binding site) are shown by arrows, as are the L and R imperfect inverted repeats of the core site. The right hand half site where crossover occurs is in bold-type and the vertical arrow shows the site of cleavage. **B.** Sequences of five classes of *attI* sites (40, 52, 156).

**A**




**B**

*attI1* TTTGAT GTT–ATGGAG CAGCAACGAT GTTACGCAGC AGGGCAGTCG CCCTAAA ACAAA **GTTRRRY**

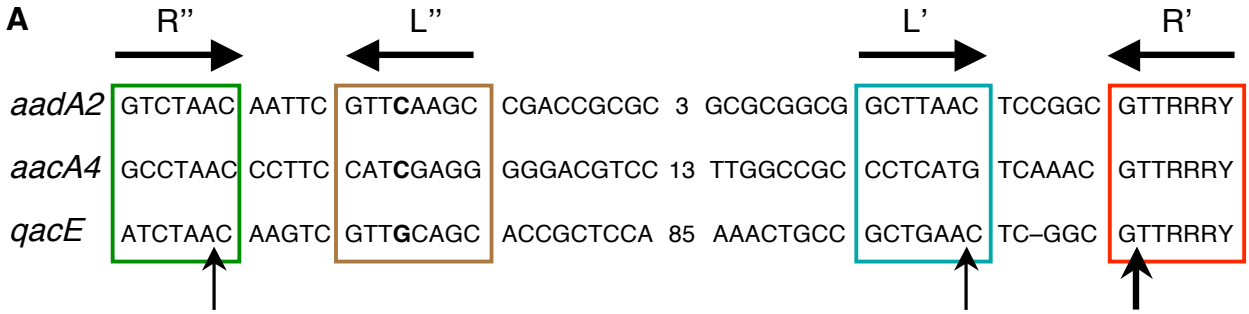
*attI3* ACTTTG TTTAACGACC ACGGTT–GTG GGTATCCGGT –GTTTGGTCA GATAAAC CACAA **GTTRRRY**

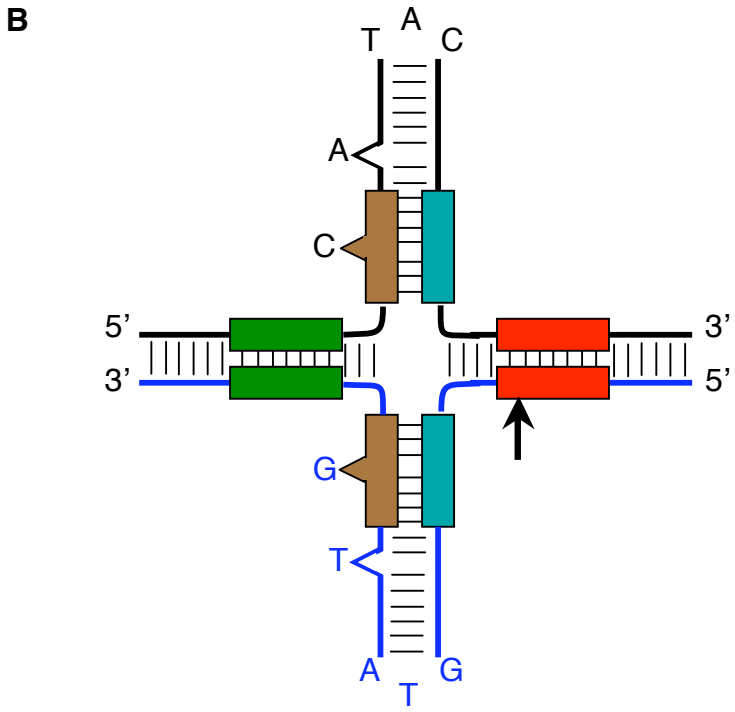
*attN*ch CTATTA ATTAGATAGC GGTAGCCTAC CAGCTGGAAA GGTAAGAAGC TGTCTAG AAAGC **GTTRRRY**

*attI2* ATATTA ATTAACGGTA AGCATCAGCG GGTGACAAAA CGAGCATGCT TACTAAT AAAAT **GTTRRRY**

*attI*Son TAACTG TGCGCGTTTT CACTGGTCTA ATGAGGTAAA CGCGCATGCG CACGAAT AAAAT **GTTRRRY**

**Figure 1.13.** *attC* recombination sites. **A.** *attC* sequences for three gene cassettes (90, 185). The inverted repeats R'' (green), L'' (brown), L' (aqua), and R' (red) are boxed. The sequence shown is the top strand from 5' to 3', and the numbers in the middle of each sequence indicate the number of bases in the middle of the sequence that are not shown. The extra base in L'' is shown in bold. The thick vertical arrow shows the position where cleavage occurs on the bottom strand, and the thin vertical arrows show rare alternative cleavage sites (90, 185). **B.** Possible cruciform structure of the *aadA2 attC* site. The top strand is black, and the bottom strand is blue. The rectangles show the inverted repeats (same color scheme as A). Paired bases are shown by thin black lines, and unpaired nucleotides are indicated by bulges (90). **C**. Possible structure formed by the bottom strand of the *aadA2 attC* site (23, 185).


A
R''
L''
L'
R'
aadA2 GTCTAAC AATTC GTTCAAGC CGACCGCGC 3 GCGCGGCG GCTTAAC TCCGGC GTTRRRY
aacA4 GCCTAAC CCTTC CATCGAGG GGGACGTCC 13 TTGGCCGC CCTCATG TCAAAC GTTRRRY
qacE ATCTAAC AAGTC GTTGCAGC ACCGCTCCA 85 AAACTGCC GCTGAAC TC–GGC GTTRRRY


B
T A C
A
C
5'
3'
3'
5'
G
T
A T G

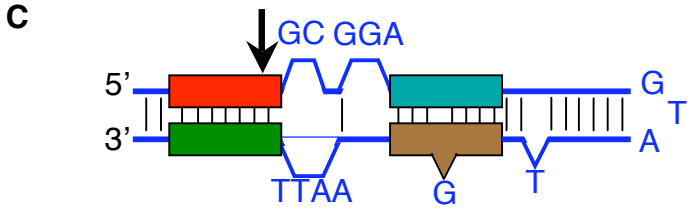
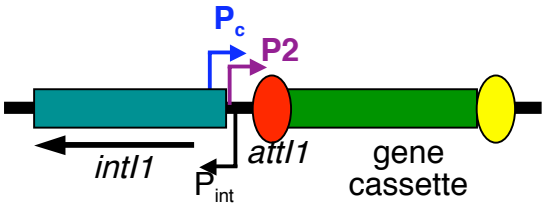

C
GC GGA
5'
3'
G T
A
TTAA
G
T

**Figure 1.14.** $P_c$ promoters of class 1 integrons. **A.** Cartoon of a class 1 integron showing the relative positions of the cassette promoters, $P_c$ and P2. **B.** Cassette promoter sequences of three integrons. In2 has the weak $P_c$ promoter sequence and a functional P2 with a 17 bp spacer region (the GGG insert), In58 has the strong $P_c$ sequence and a nonfunctional P2 with a 14 bp spacer, and In111 has the hybrid $P_c$ with the weak -35 sequence and strong -10 sequence (111, 149, 202). The number indicates 86 base pairs left out of the alignment. Nucleotides differences are shown in bold-type. The -35 and -10 regions of the Pc promoter are boxed in blue, and P2 regions are boxed in purple.

**A**




**B**

| | Pc -35 | | Pc -10 | | P2 -35 | | P2 -10 |
|---|---|---|---|---|---|---|---|
| In2 | T**G**GACA | TAAGCCTGTTCGGTTCG | TAA**G**CT | GT 86 GC | TTGTTA | TGACTGTTTTTTTG **GGG** | TACAGT |
| In58 | T**T**GACA | TAAGCCTGTTCGGTTCG | TAA**A**CT | GT 86 GC | TTGTTA | TGACTGTTTTTTTG – – – | TACAGT |
| In111 | T**G**GACA | TAAGCCTGTTCGGTTCG | TAA**A**CT | GT 86 GC | TTGTTA | TGACTGTTTTTTTG – – – | TACAGT |

## CHAPTER 2

### IntI1 can mediate cointegrate formation at low concentrations *in vivo*

### Overview

The integron integrase IntI1 mediates the transfer of antibiotic resistance gene cassettes between plasmid-carried integrons by two mechanisms, integration and excision or the formation and resolution of cointegrates. Although cointegrates have been isolated from cells expressing integrase under the natural promoter $P_{int}$ (7, 22, 23) and cassette capture has never been experimentally observed without excess integrase, excision and integration has been reported as the major mechanism used to transfer gene cassettes (5, 7, 15). We constructed an integron capture vector pICV8 that acquires and donates antibiotic resistance gene cassettes in strains that overexpress IntI1. PCR of recombinant junctions revealed that pICV8 readily forms cointegrates *in vivo* in strains expressing integrase under its natural promoter on a low copy number plasmid, though cassette capture products are not detected without excess integrase.

### Introduction

Bacterial resistance to antibiotics has become a major challenge in healthcare. Acquisition of antibiotic resistance genes by horizontal gene transfer has been implicated as the major cause of dissemination of these resistances throughout bacterial

populations (10, 26). One genetic system that transfers multigene antibiotic resistances is the integron. In fact, integrons are the most prevalent genetic system through which multiple resistances are spread at one time in the Enterobacteriaceae (26). Integrons carry gene cassette arrays encoding multiple antibiotic resistances and are associated with multiresistant Enterobacteriaceae isolated from hospitals and the environment (12, 19, 20, 27, 28, 31, 32, 36, 40, 42). Nearly 100 different antibiotic resistance gene cassettes have been described (14, 25). PCR of environmental DNA has identified integrases and antibiotic resistance gene cassettes associated with integrons in diverse environments such as in soil and in the fecal matter of livestock and companion animals (12, 28, 31, 38). Previously thought to be carried almost exclusively by Gram-negative Enterobacteriaceae, more recent studies have identified integron-type integrases in both low-GC and high-GC Gram-positive bacteria and show that Gram-positive bacteria may represent a larger reservoir for integrons than Gram-negative bacteria in certain ecosystems (4, 28, 30). These studies demonstrate that integrons are more widespread than previously imagined and may be responsible for transmitting antibiotic resistances between diverse bacterial species.

Integrons are naturally occurring genetic elements found on plasmids and on the chromosomes of bacteria. They contain a site-specific recombination system mediated by an integrase, a site-specific recombinase of the tyrosine recombinase family, where gene cassettes are excised or inserted at the *attI* integrase recognition site located in between the integrase gene and the gene cassettes (Fig. 2.1). The gene cassettes are short DNA elements that consist of a recombination site (*attC*) and an open reading

frame, often encoding antibiotic resistance (13, 15, 22, 37). Integrons are classified by their integrase sequence; several classes have been identified but only 5 have been well characterized, including the multiresistant class 1, 2, and 3 integrons and the *Vibrio cholerae* chromosomal integron (5, 10, 17, 31, 35). The gene cassettes are transcribed by a common promoter located just upstream of the *attI* site (Fig, 2.1). Integrase-mediated recombination effects gene cassette movement between integrons by two different mechanisms: excision and integration (Fig. 2.2) or formation and resolution of cointegrates (Fig. 2.3).

In excision and integration, gene cassettes are excised into covalently closed circles and then inserted into the *attI* site of a different integron in the same cell (5, 6, 15). This process is called "cassette capture" (Fig. 2.2). In the second mechanism, formation and resolution of cointegrates, the integrase mediates RecA-independent recombination between the donor and recipient plasmids at the *attI* and/or *attC* sites to form a cointegrate plasmid (Fig. 2.3). The cointegrate is resolved by integrase, generating two separate plasmids either of which can contain the cassette(s) adjacent to the insertion site; this a conservative process, and no replication is involved (7, 16, 22, 34). Complete cassette capture has only been experimentally observed when integrase is overexpressed (5-8, 16, 22, 23, 34), but cointegrates have been isolated from cells expressing IntI1 under its natural promoter $P_{int}$ (7, 22, 24).

I constructed a vector carrying a class 1 integron to serve as a tool for measuring cassette movement between integrons. This integron capture vector, pICV, is low copy,

mobilizable, wide-host range and carries zeocin resistance for selection and an *attI* site, and I constructed versions with and without an active *intI1* gene.

## Methods

**Bacterial strains and plasmids.** *E. coli* strains used were CB454 [*gal lacZ lacY thi str rpsl- recA56*] and SK2267A [*gal thi T1 recA1 hsm+ hsrR4 endA sbcB15*]. Table 2.1 lists the plasmids used and their phenotypes. Plasmids were transformed into CB454 and SK2267A by electroporation using a Gene Pulser (BioRad, Inc.) (2). Transformants were selected on Luria-Bertani (LB) plates containing antibiotics. Concentrations of antibiotics used were: streptomycin 250 μg/ml for chromosomal resistance and 25 μg/ml for plasmid resistance, zeocin 25 μg/ml, kanamycin 50 μg/ml, trimethoprim 1000 μg/ml, ampicillin 50 μg/ml. Antibiotics were obtained from Sigma Inc. (St. Louis, MO) except zeocin (Invitrogen Corporation, Carlsbad, CA).

**Plasmid construction.** The construction of the pICV vectors was based on the near-zero background cloning method (41). The Tn*21* (21) integrase gene, *attI* site, and part of the *aadA1* cassette were amplified by PCR and cloned into pJRD215 (11) (Fig. 2.4 line 1) to make pICV3 (11,837 bp) (Fig. 2.4 line 2). Then, the zeocin resistance gene from pEM7/Zeo (Invitrogen) was amplified by PCR and cloned into pICV3 to make pICV4 (10,385 bp) (Fig. 2.4 line 3). To minimize the risk of homologous recombination between the donor plasmids and pICV during eventual use of these vectors in wild bacterial strains, 686 bp of the integrase gene flanked by *BsaI* and *BsmBI* sites were

removed by restriction digest, the ends were blunted using DNA Polymerase I large fragment, Klenow (NEB), and ligated to generate pICV6 (9699 bp) (Fig. 2.4 line 4). The kanamycin resistance gene and multicloning region from pJRD215 were removed to make pICV7 (8567 bp) and pICV8 (7584 bp), respectively (Fig. 2.4 lines 5-6). Isolation of plasmid DNA was performed using the QIAquick Spin Miniprep Kit (QIAGEN) using conditions for low copy number plasmids or the QIAGEN Plasmid Midi Kit. Inserted regions of pICV3 and pICV4 and all of pICV8 were sequenced and were found to be as intended.

**Detection of cassette acquisition by PCR.** Briefly, the donor plasmid (R388 or pRMH14) was electroporated first into CB454 (or SK2267A). Then, the pICV recipient (pICV4 or pICV8) was electroporated into CB454 (or SK2267A) containing the donor, and lastly, in some cases, pSU2056 was electroporated into CB454 (or SK2267A) containing the donor and recipient plasmids. In strains containing the pICV recipient plasmid and a cassette donor plasmid with *intI1*, integrase-mediated cointegrate formation and cassette capture were detected by PCR of the recombinant junctions. As indicated, in some experiments additional integrase was supplied by the IntI1-overexpressing plasmid, pSU2056. Table 2.2 shows PCR primer sequences, and Fig. 2.5 shows the primer locations on pICV8::*aadA2* and pICV4::*dfrB2*. PCR primers were synthesized by Sigma-Genosys. pRMH14::pICV8 cointegrate junctions were detected with primers 8E-U1 and int-L1. pICV8::*aadA2* 5' junctions were detected with primers aadA2-U1 and 4E-L2, and pICV8::*aadA2* 3' junctions were detected with primers 8E-U1

and aadA2-L2. pICV4::*dfrB2* 5' junctions were detected with primers dfr-U2 and 4E-L2, and pICV4::*dfrB2* 3' junctions were detected with primers 4E-U1 and dfr-L1. Control strains contained only a pICV plasmid or only the cassette donor plasmid, with or without pSU2056. PCR was done on whole cells taken periodically from LB broth cultures grown up to 40-45 hours at 37°C, 250-275 rpm. Cells from 1.5 ml of culture were pelleted by centrifugation, suspended in 1.0 ml nuclease free water (NFW), and diluted 100-fold in NFW. A 25 μl aliquot of the dilute cell suspension, lysed at 95°C for 5 minutes, was used in a 75 μl PCR which contained 1X Magnesium-free Buffer, 1.5 mM $MgCl_2$, 0.25 mM each dNTP (Applied Biosystems, Foster City, CA), 20 pmol of each primer, and 2.5 units of Taq DNA Polymerase. All reagents except dNTPs were from Promega, Inc., Madison, WI. 30 cycle PCRs were performed in PTC-100 and PTC-200 thermocyclers (MJ Research, Inc., Waltham, MA). 30 μl of PCR products were visualized by agarose gel electrophoresis with detection by a Molecular Dynamics Fluoroimager after staining with ethidium bromide.

**Detection of cassette acquisition by plasmid isolation.** Using the QIAprep Spin Miniprep Kit (QIAGEN), plasmid DNA was extracted from 25 ml 40-45 hour LB broth cultures of a strain containing the cassette donor, the pICV recipient, and pSU2056. The plasmid extract was digested with restriction enzymes to eliminate the donor plasmid and pSU2056 (*BsrGI, PvuII, KpnI* for pICV8::*aadA2* isolation, *AhdI* for pICV4::*dfrB2* isolation, *FseI* and *AhdI* for pICV3::*aadA2* isolation), and electroporated into SK2267A. Transformants were selected on LB plates with zeocin to select for the

pICV recipient plasmid and with streptomycin for *aadA2* cassette selection or trimethoprim for *dfrB2* cassette selection.

**Detection of recombination in broth cultures.** Aerobic LB broth cultures of strains containing the pRMH14 cassette donor plasmid and pICV recipient plasmid, with or without pSU2056, were monitored from one hour to 45 hours of growth. Culture samples (1.5 ml) were taken at 5, 7, and 45 hours of growth for PCR detection of pRMH14::pICV8 cointegrate junctions and pICV8::*aadA2* 5' and 3' junctions as described above. The PCR product was visualized by agarose gel electrophoresis, and its mass was estimated by densitometry using mass standards on the same gel. Aliquots of same undiluted broth culture suspensions used were lysed by sonication, and the protein concentration was determined by the Bradford assay. Protein concentration was used to estimate the number of cells per mililiter based on the assumption that 16% of wet cell mass is protein and the total wet mass of one cell equals $9.5 \times 10^{-13}$ g (29).

## Results

**Acquisition of cassettes by pICV.** Although plasmids capable of acquiring antibiotic resistance gene cassettes from integrons have been constructed previously (6, 7, 9, 22), none have been designed for examining cassette acquisition from natural environments. The pICV plasmids are based on a wide-host-range IncQ replicon in order to allow their movement throughout bacterial populations. They also carry a

relatively rare antibiotic resistance marker, zeocin, for eventual detection of cells carrying them in a mixed cultures or wild populations. The optimized vector, pICV8, contains an *attI* site for cassette capture, a truncated version of the integrase gene including the divergently oriented common cassette promoter ($P_c$) and a fragment of the *aadA1* gene cassette. Thus, pICV8 only contains a single integrase recombination site, *attI*. Integrase is supplied by the low copy number cassette donor plasmid (variously R388 or pRMH14) and, in some experiments, pSU2056, which overexpresses integrase but has neither cassettes nor recombination sites.

When a pICV vector captures a cassette by excision/integration or by cointegrate formation and resolution, the two unique junctions formed can be amplified by specific PCR primer pairs (Fig. 2.6). For each junction amplicand, one primer anneals to the pICV vector sequence and the other primer anneals to the specific gene cassette; these primers will not produce an amplicand with the original pICV vector alone or with the cassette donor plasmid alone. Thus, amplification of both junctions (called the 5' junction and 3' junction) formed between pICV8 and the *aadA2* gene cassette from donor pRMH14 indicates *aadA2* cassette capture (Fig. 2.7 A). The 3' junction is only observed when excess integrase is supplied by pSU2056 (Fig. 2.7 B lanes 6-7). For the 5' junction (Fig. 2.7 C lanes 10-16), the smaller band reports on the capture of the *aadA2* cassette alone, and the larger band arises from the capture of two cassettes from pRMH14, *aacA4* and *aadA2*. The capture products arising from the 5' junction are observed in *E. coli* cells containing pICV8 and pRMH14 (Fig. 2.7 lanes 13-14), but unresolved cointegrates can also give the same 5' junction PCR products as a simple

cassette capture event (Fig. 2.8). Amplifying the unique cointegrate junction by PCR can show cointegrate formation between pICV8 and pRMH14 (Fig. 2.8). The detection of both the large and small 5' junction amplicands (Fig. 2.7 C lanes 13-14) and the absence of the 3' junction amplicand (Fig. 2.7 B lanes 4-5) indicates the presence of *attI* x *attI* and *aacA4 attC* x *attI* cointegrates (Fig. 2.8) in the strain expressing integrase under the natural promoter. Since the 3' junction is only observed when excess integrase is supplied by pSU2056 (Fig. 2.7 B lanes 6-7), neither *aadA2* cassette capture nor formation of the *aadA2 attC* x *attI* pRMH14::pICV8 cointegrate occur at a detectable level without excess integrase.

PCR detection of recombinant junctions was also performed on strains in which both donor and recipient integrons express *intI1* under the natural $P_{int}$ promoter. R388 was used as a donor of the *dfrB2* gene cassette with pICV4 as the recipient. The 5' junction PCR product of pICV4::*dfrB2* can arise from single or double cassette capture by pICV8 (*dfrB2* and *orfA*) (Fig. 2.9 B lanes 6-9), or from the R388::pICV4 *attI* x *attI* cointegrate. The pICV4::*dfrB2* 3' junction product shows simple cassette capture of *dfrB2* as well as capture of both *dfrB2* and *orfA* (Fig. 2.9 C lanes 18-19). As with capture of *aadA2* by pICV8 from pRMH14, the pICV4::*dfrB2* 3' junction product is observed only when excess integrase is supplied by pSU2056 (Fig. 2.9 C lanes 18-19).

**pICV8::*aadA2* can act as a cassette donor.** Since pICV8 can acquire gene cassettes, we next asked whether it could donate an acquired cassette to another vector. Plasmid pICV3, which has a full-length integrase gene, was used as the cassette recipient.

Junction PCR was performed on lysates of SK2267A(pICV8::*aadA2*, pICV3) and SK2267A(pICV8::*aadA2*, pICV3, pSU2056). Cassette capture by pICV3 from pICV8::*aadA2* was only observed in the strain containing the donor and recipient plasmids and pSU2056 (Fig. 2.10 B lane 7). The plasmid pICV3::*aadA2* was isolated by electroporation of plasmid DNA purified from SK2267A(pICV8::*aadA2*, pICV3, pSU2056), and showed the expected 3' junction (Fig. 2.10, lanes 8-9) indicating insertion of the cassette.

**Integrase-mediated recombination products accumulate from late log to late stationary phase.** Since we had detected integron-mediated recombination products in late stationary phase broth cultures, we wanted to observe whether these products were equally present during log phase. In aerobic broth cultures, junction amplicands in the low integrase strain CB454(pRMH14, pICV8) increase in intensity to 45 hours (Fig. 2.11 A, lanes 2-7). This trend also occurs in the high integrase strain CB454(pRMH14, pICV8, pSU2056) (Fig. 2.11 A, lanes 8-13), which gives more intense cointegrate junction amplicands than the low integrase strain at any time point. Densitometric quantification of the amplicand band in each lane normalized to the number of cells in the PCR by determining the protein content (see Methods) shows a slight decrease in the junction amplicands from 5 to 7 hours and an increase up to 45 hours for both low and high integrase strains (Fig. 2.11 B). The 5' junction amplicands per number of cells in the PCR reaction show a similar trend. The low integrase strain shows a slight increase from 5 to 7 hours and a greater increase at 45 hours, and the high integrase

strain shows a decrease from 5 to 7 hours and an increase up to 45 hours (Fig. 2.12 A). Thus, there are more cointegrate junctions and 5' junctions per cell in late stationary phase than in late exponential phase. The 3' junction also decreases from 5 to 7 hours and increases to 45 hours in the high integrase strain (Fig. 2.12 B), but no 3' junction amplicand is seen in the low integrase strain.

**Discussion**

The constructed vector pICV can acquire antibiotic resistance gene cassettes from other plasmid-carried integrons and form cointegrates, as well as donate an acquired cassette in the presence of high integrase. pICV readily forms cointegrates when integrase is expressed under the natural promoter $P_{int}$, but overexpressed integrase is needed to detect complete cassette acquisition. This suggests that cassette capture is dependent on the intracellular concentration of integrase. Even with both donor and recipient plasmids expressing integrase from natural promoters, as seen in recombination between pICV4 and R388, there is not enough integrase to detect cassette capture without pSU2056 providing excess integrase. This is consistent with earlier experiments where cointegrates were isolated from strains expressing integrase under $P_{int}$ but cassette capture products were only detected when integrase was overexpressed (7, 22, 23).

If cassette acquistion is dependent on high integrase, how does cassette capture occur in nature where there is no pSU2056 to provide excess integrase? The ready occurrence of cointegrate formation with low integrase indicates that cointegrate

formation is a preferred recombination mechanism. This could mean a more important role for cointegrates in natural ecosystems. However, since high integrase is needed to observe complete cassette acquisition, cointegrate resolution does not occur at a detectable level without excess integrase. Intracellular recombinase concentration affects recombination reactions for other tyrosine recombinases, as well. The recombinase Cre resolves dimers of the P1 prophage to ensure stable inheritance, but Cre causes P1 instability when overexpressed *in vivo*, likely by mediating multimer formation rather than resolution (33). The lambda phage Int uses integrase concentation, relative to concentration of excisionase (Xis), to regulate the recombination choice of integration or excision (18). Thus, it would not be unusual for integron integrases to alter reaction preference based on integrase concentration. Clearly integron integrase regulation mechanisms require further study.

Integrase-mediated recombination products accumulate during stationary phase in broth cultures. Particularly in the low integrase strain, junction PCR products per cell increase from log phase to late stationary phase. Unfortunately, using protein concentration levels in cell lysates is not the most accurate method to determine cell number since the amount of protein per total wet cell mass may not be 16% throughout the growth cycle (29), and the total wet mass of a cell may also change according to growth cycle since cells change size and shape in stationary phase (1). The number of cells could be measured more accurately by PCR of total DNA isolated from a culture rather than cell lysates, using a quantitative PCR method of detecting a gene in single copy on the chromosome (see Chapter 3). In addition, the amount of template in a PCR

could affect the results and using total DNA isolated from cultures at various timepoints in the growth cycle would allow measurement of the amount of template in a PCR prior to the reaction instead of being measured afterwards; thus, PCRs of different timepoints would contain the same amount of template. The accumulation of cointegrate products shown by the cell lysate PCR used here may occur in the low integrase strain because of inadequate cointegrate resolution since the lack of 3' junction products indicates that complete cassette capture products are not detectable, but cointegrates are also abundant in the high integrase strain where cointegrate resolution presumably occurs. The high amounts of recombinant junctions per cell observed at 5 hours in the high integrase strain may be from recombination products produced in the stationary phase plate cultures used to inoculate the broth cultures, or they may indicate that recombination increases at lag or early log phase when integrase is overexpressed. These possibilities could be resolved by examining recombination products in newly constructed strains rather than subcultures (see Chapter 3), so there would be no carryover recombination products during lag and early log phase. Integrase-mediated recombination has never before been examined throughout a culture cycle. The accumulation of recombination products during stationary phase could indicate that integrase-mediated recombination reactions are differentially regulated by growth phase.

## References


1. **Akerlund, T., K. Nordstrom, and R. Bernander.** 1995. Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. J Bacteriol **177:**6791-7.
2. **Ausubel, F. A., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1988. Current Protocols in Molecular Biology. John Wiley & Sons, New York, N. Y.
3. **Avila, P., and F. de la Cruz.** 1988. Physical and genetic map of the IncW plasmid R388. Plasmid **20:**155-7.
4. **Clark, N. C., O. Olsvik, J. M. Swenson, C. A. Spiegel, and F. C. Tenover.** 1999. Detection of a streptomycin/spectinomycin adenylyltransferase gene (*aadA*) in *Enterococcus faecalis*. Antimicrob. Agents Chemother. **43:**157-60.
5. **Collis, C. M., G. Grammaticopoulos, J. Briton, H. W. Stokes, and R. M. Hall.** 1993. Site-specific insertion of gene cassettes into integrons. Mol. Microbiol. **9:**41-52.
6. **Collis, C. M., and R. M. Hall.** 1992. Gene cassettes from the insert region of integrons are excised as covalently closed circles. Mol. Microbiol. **6:**2875-85.
7. **Collis, C. M., and R. M. Hall.** 1992. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. J. Bacteriol. **174:**1574-85.

8. **Collis, C. M., M. J. Kim, S. R. Partridge, H. W. Stokes, and R. M. Hall.** 2002. Characterization of the class 3 integron and the site-specific recombination system it determines. J. Bacteriol. **184:**3017-26.

9. **Collis, C. M., M. J. Kim, H. W. Stokes, and R. M. Hall.** 2002. Integron-encoded IntI integrases preferentially recognize the adjacent cognate *attI* site in recombination with a 59-be site. Mol. Microbiol. **46:**1415-27.

10. **Davies, J.** 1994. Inactivation of antibiotics and the dissemination of resistance genes. Science **264:**375-82.

11. **Davison, J., M. Heusterspreute, N. Chevalier, V. Ha-Thi, and F. Brunel.** 1987. Vectors with restriction site banks. V. pJRD215, a wide-host-range cosmid vector with multiple cloning sites. Gene **51:**275-80.

12. **Goldstein, C., M. D. Lee, S. Sanchez, C. Hudson, B. Phillips, B. Register, M. Grady, C. Liebert, A. O. Summers, D. G. White, and J. J. Maurer.** 2001. Incidence of class 1 and 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. Antimicrob. Agents Chemother. **45:**723-6.

13. **Hall, R. M., D. E. Brookes, and H. W. Stokes.** 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Mol. Microbiol. **5:**1941-59.

14. **Hall, R. M., C. M. Collis, M. J. Kim, S. R. Partridge, G. D. Recchia, and H. W. Stokes.** 1999. Mobile gene cassettes and integrons in evolution. Ann. N. Y. Acad. Sci. **870:**68-80.

15. **Hall, R. M., and H. W. Stokes.** 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica **90:**115-32.

16. **Hansson, K., O. Skold, and L. Sundstrom.** 1997. Non-palindromic *attI* sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. **26:**441-53.

17. **Hansson, K., L. Sundstrom, A. Pelletier, and P. H. Roy.** 2002. IntI2 integron integrase in Tn7. J. Bacteriol. **184:**1712-1721.

18. **Leffers, G. G., Jr., and S. Gottesman.** 1998. Lambda Xis degradation *in vivo* by Lon and FtsH. J. Bacteriol. **180:**1573-1577.

19. **Leverstein-van Hall, M. A., M. B. HE, T. D. AR, A. Paauw, A. C. Fluit, and J. Verhoef.** 2003. Multidrug resistance among Enterobacteriaceae is strongly associated with the presence of integrons and is independent of species or isolate origin. J. Infect. Dis. **187:**251-9.

20. **Lévesque, C., L. Piche, C. Larose, and P. H. Roy.** 1995. PCR mapping of integrons reveals several novel combinations of resistance genes. Antimicrob. Agents Chemother. **39:**185-91.

21. **Liebert, C. A., R. M. Hall, and A. O. Summers.** 1999. Transposon Tn*21*, flagship of the floating genome. Microbiol. Mol. Biol. Rev. **63:**507-522.

22. **Martinez, E., and F. de la Cruz.** 1990. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. **9:**1275-81.

23. **Martinez, E., and F. de la Cruz.** 1988. Transposon Tn*21* encodes a RecA-independent site-specific integration system. Mol. Gen. Genet. **211:**320-5.

24. **Martinez, J. L., and F. Baquero.** 2002. Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance. Clin. Microbiol. Rev. **15:**647-79.

25. **Mazel, D.** 2006. Integrons: agents of bacterial evolution. Nat Rev Micro **4:**608-620.

26. **Mazel, D., and J. Davies.** 1999. Antibiotic resistance in microbes. Cell Mol. Life Sci. **56:**742-54.

27. **Mazel, D., B. Dychinco, V. A. Webb, and J. Davies.** 2000. Antibiotic resistance in the ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob. Agents Chemother. **44:**1568-74.

28. **Nandi, S., J. J. Maurer, C. Hofacre, and A. O. Summers.** 2004. Gram-positive bacteria are a major reservoir of Class 1 antibiotic resistance integrons in poultry litter. Proc. Natl. Acad. Sci. USA **101:**7118-7122.

29. **Neidhardt, F. C., and H. E. Umbarger.** 1996. Chemical Composition of *Escherichia coli*, p. 13-16. *In* F. C. Neidhardt, R. C. III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarge (ed.), *Escherichia coli* and *Salmonella* Cellular and Molecular Biology, 2nd ed, vol. 1. ASM Press, Washington DC.

30. **Nesvera, J., J. Hochmannova, and M. Patek.** 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. FEMS Microbiol. Lett. **169:**391-5.

31. **Nield, B. S., A. J. Holmes, M. R. Gillings, G. D. Recchia, B. C. Mabbutt, K. M. H. Nevalainen, and H. W. Stokes.** 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. **195:**59-65.

32. **Partridge, S. R., C. M. Collis, and R. M. Hall.** 2002. Class 1 integron containing a new gene cassette, *aadA10*, associated with Tn*1404* from R151. Antimicrob. Agents Chemother. **46:**2400-8.

33. **Paul, S., and D. Summers.** 2004. ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. Plasmid **52:**63-68.

34. **Recchia, G. D., H. W. Stokes, and R. M. Hall.** 1994. Characterisation of specific and secondary recombination sites recognised by the integron DNA integrase. Nucleic Acids Res. **22:**2071-8.

35. **Rowe-Magnus, D. A., A.-M. Guerout, and D. Mazel.** 2002. Bacterial resistance evolution by recruitment of super-integron gene cassettes. Molecular Microbiology **43:**1657-1669.

36. **Sandvang, D.** 1999. Novel streptomycin and spectinomycin resistance gene as a gene cassette within a class 1 integron isolated from *Escherichia coli*. Antimicrob. Agents Chemother. **43:**3036-8.

37. **Stokes, H. W., and R. M. Hall.** 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol. Microbiol. **3:**1669-83.

38. **Stokes, H. W., A. J. Holmes, B. S. Nield, M. P. Holley, K. M. Nevalainen, B. C. Mabbutt, and M. R. Gillings.** 2001. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. Appl. Environ. Microbiol. **67:**5240-6.

39. **Stokes, H. W., C. Tomaras, Y. Parsons, and R. M. Hall.** 1993. The partial 3'-conserved segment duplications in the integrons In6 from pSa and In7 from pDGO100 have a common origin. Plasmid **30:**39-50.

40. **Tosini, F., P. Visca, I. Luzzi, A. M. Dionisi, C. Pezzella, A. Petrucca, and A. Carattoli.** 1998. Class 1 integron-borne multiple-antibiotic resistance carried by IncFI and IncL/M plasmids in *Salmonella enterica* serotype Typhimurium. Antimicrob. Agents Chemother. **42:**3053-8.

41. **Zeng, Q., M. K. Eidsness, and A. O. Summers.** 1997. Near-zero background cloning of PCR products. Biotechniques **23:**412-4, 416, 418.

42. **Zuhlsdorf, M. T., and B. Wiedemann.** 1992. Tn*21*-specific structures in gram-negative bacteria from clinical isolates. Antimicrob. Agents Chemother. **36:**1915-21.

**Table 2.1.** Phenotypes of plasmids

| **Plasmid name** | **Role** | **Phenotype** | **Reference** |
|---|---|---|---|
| pJRD215 | Cloning | KmR, SmR | (11) |
| pICV3 | pRecipient | KmR, SmR, IntI1$^{+}$ | this work |
| pICV4 | pRecipient | KmR, ZeoR, IntI1$^{+}$ | this work |
| pICV6 | Cloning | KmR, ZeoR | this work |
| pICV7 | Cloning | ZeoR | this work |
| pICV8 | pRecipient and pDonor | ZeoR | this work |
| pRMH14 | pDonor | KmR, SmR, SuR, IntI1$^{+}$ | (39) |
| R388 | pDonor | TpR, SuR, IntI1$^{+}$ | (3) |
| pSU2056 | Integrase expression | ApR, IntI1$^{+}$ | (22) |

**Table 2.2.** PCR primer sequences

| Primer name | Sequence | Target |
|---|---|---|
| 4E-U1 | 5' GGAGGCCATCAAACCA CGTCAAATA 3' | pICV3, pICV4 backbone |
| dfr-L1 | 5' CCTGAAGGCTATGCGG TCGAGTC 3' | *dfrB2* |
| dfr-U2 | 5' GTGGCACTCAGGGGAA GCGCAAACT 3' | *dfrB2* |
| 4E-L2 | 5' GCCTATGCCTACAGCA TCCAGGGTGAC 3' | pICV3, pICV4, pICV8 backbone |
| 8E-U1 | 5' CCTCGTTAAAGGACAA GGACCTGAG 3' | pICV3, pICV4, pICV8 backbone |
| aadA2-L2 | 5' GCGAGCTGCAATTTGG AGAATGG 3' | *aadA2* |
| aadA2-U1 | 5' GCCGGTTATTGCGCTG TACCAAATG 3' | *aadA2* |
| int-L1 | 5' CGCGCTGAAAGGTCTG GTCATAC 3' | *intI1* |

**Figure 2.1.** Typical integron structure. $P_c$ represents the cassette promoter and $P_{int}$ the integrase promoter. The gray oval shows *attI*, and the striped ovals show *attCs*. Each gene cassette contains a gene, often encoding antibiotic resistance, and an *attC* recombination site. Arrows indicate direction of transcription.

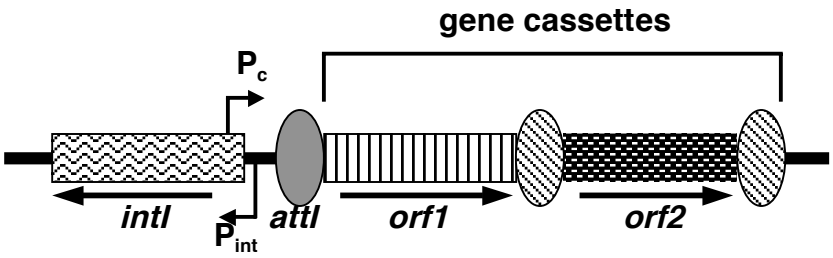

gene cassettes
P c
intI
P int
attI
orf1
orf2

**Figure 2.2.** The excision/integration mechanism of cassette capture. The gray ovals represent *attI*, and the striped ovals show *attC*. pDonor encodes two gene cassettes, *aacA4* and *aadA2*, and pRecipient contains an *attI* site and no mobile cassettes. The *aadA2* of pDonor is excised by the integrase into a covalently closed circle. Then, the integrase inserts the cassette into the *attI* site of pRecipient. The result is a pDonor lacking the *aadA2* cassette, and pRecipient gaining the *aadA2* cassette.

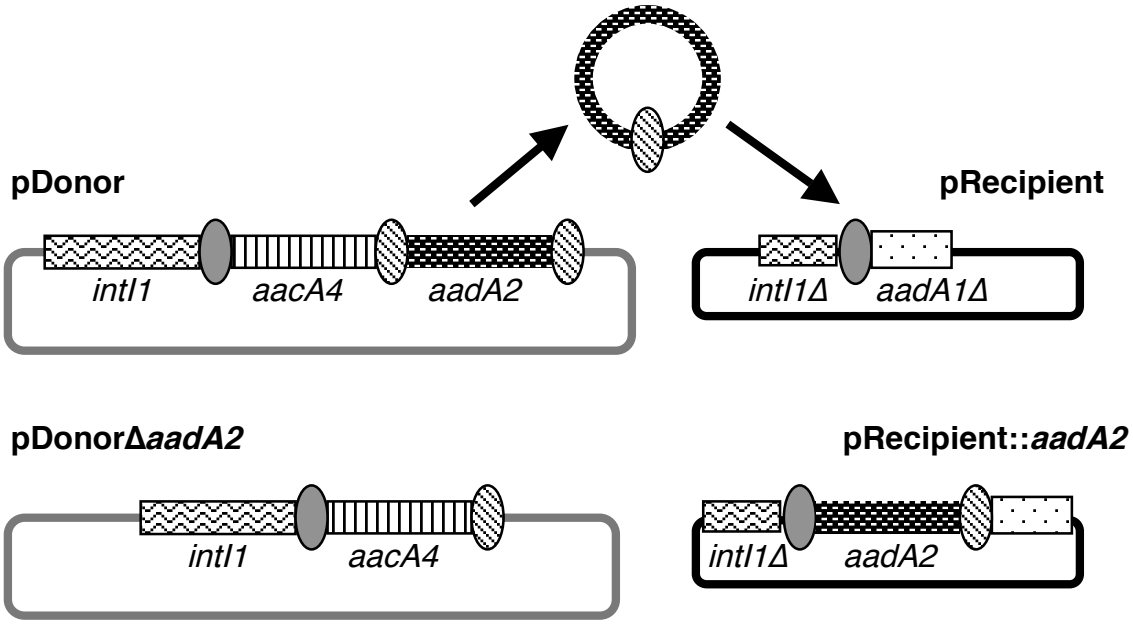

pDonor
pRecipient
intI1
aacA4
aadA2
intI1Δ
aadA1Δ
pDonorΔaadA2
pRecipient::aadA2
intI1
aacA4
intI1Δ
aadA2

**Figure 2.3.** The cointegrate formation and resolution mechanism of cassette capture. **A.** Integrase-mediated recombination occurs between one site on pDonor and one site on pRecipient to form a cointegrate. **B.** The cointegrate resulting from *attI* x *attI* recombination is shown. **C.** A second recombination event then resolves the cointegrate into two separate plasmids. The example shown is resolution of the cointegrate in B by *aacA4 attC* x *attI* recombination.

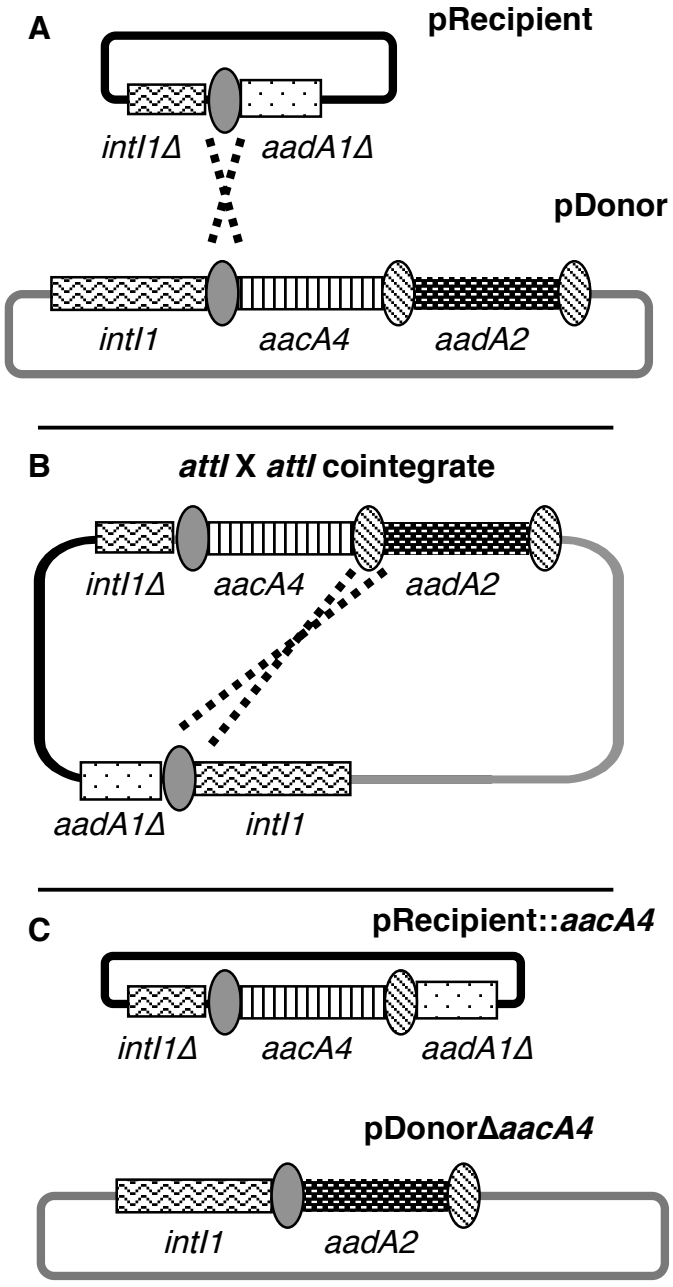
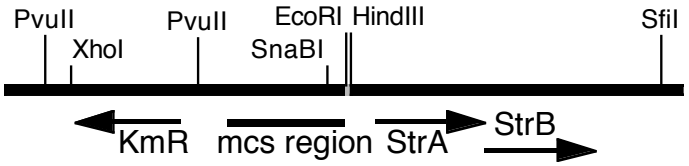
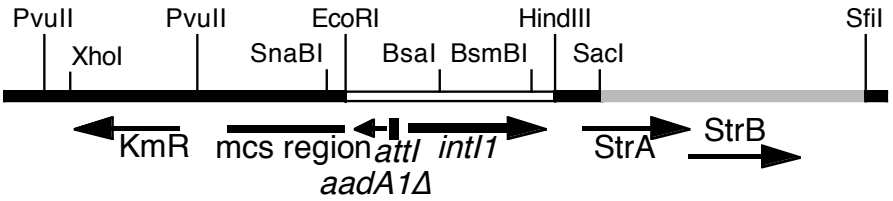

A
pRecipient
intI1Δ
aadA1Δ
pDonor
intI1
aacA4
aadA2
B
attI X attI cointegrate
intI1Δ
aacA4
aadA2
aadA1Δ
intI1
C
pRecipient::aacA4
intI1Δ
aacA4
aadA1Δ
pDonorΔaacA4
intI1
aadA2

**Figure 2.4.** Construction of pICV plasmid vectors. Bases 5222 through the end of each vector sequence are shown, with position 1 (not shown) set at the beginning of pJRD215 *oriV* (11). At each stage, the white areas are insertions, and the gray areas will be deleted in the subsequent stage. The pICV3 *attI-intI1* insert is from Tn*21* (21), and the pICV4 ZeoR insert is from pEM7/Zeo (Invitrogen). Arrows indicate direction of transcription (see Table 2.1 for phenotypes). mcs region=multiple cloning site. Note that in this figure the plasmids are drawn so that the integron is oriented opposite to other figures, except Fig. 2.5.

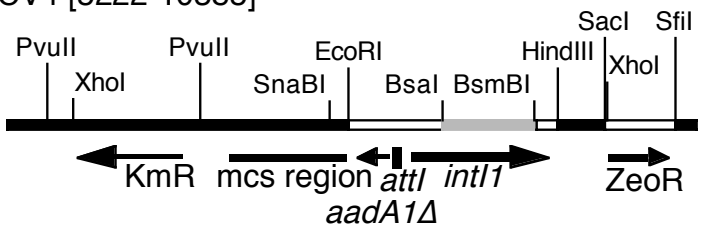
1. pJRD215 [5222-10299]



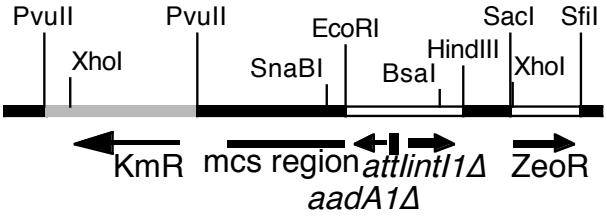

2. pICV3 [5222-11837]



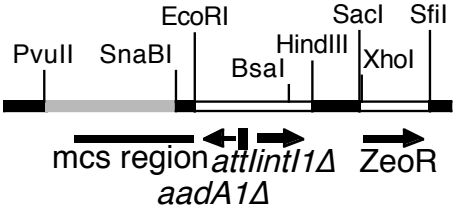

3. pICV4 [5222-10385]



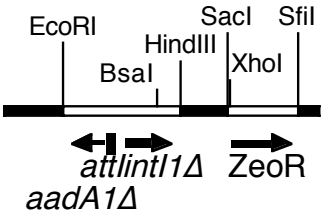

4. pICV6 [5222-9699]




5. pICV7 [5222-8567]




6. pICV8 [5222-7584]



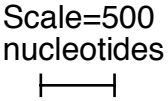
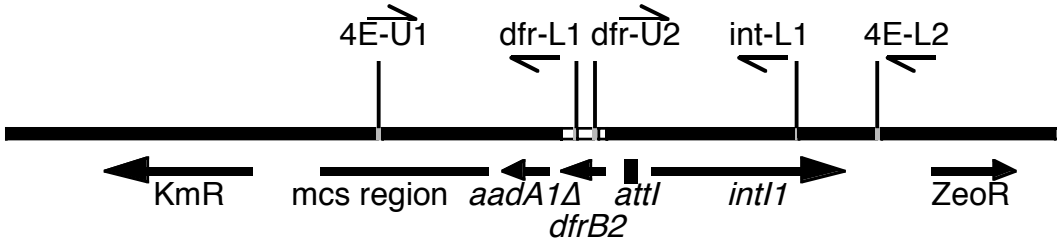



**Figure 2.5.** Location of PCR primers on pICV4::*dfrB2* and pICV8::*aadA2.* Bases 5222 through the end of each vector sequence are shown. Black indicates pICV sequence, and the acquired cassette sequence is shown in white. Arrows indicate direction of transcription. Primer positions are noted by vertical bars and directions noted by half arrows. Note that in this figure the plasmids are drawn so that the integron is oriented opposite to other figures, except Fig. 2.4.

pICV4::*dfrB2* [5222-10769]




pICV8::*aadA2* [5222-8440]

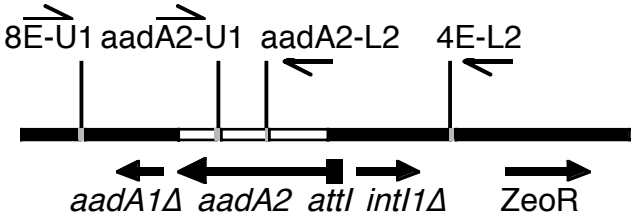
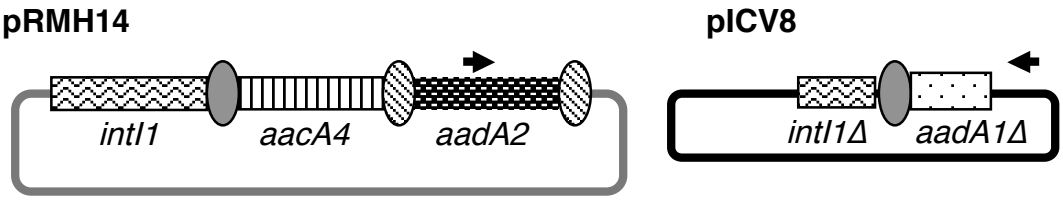


**Figure 2.6.** PCR detection of pICV8 cassette capture. One primer anneals to the gene cassette while the other primer anneals to a nonmobile region of pICV8. Thus, no PCR product will form without cassette capture or cointegrate formation. See text for details.

**No PCR product possible**




**PCR product possible**

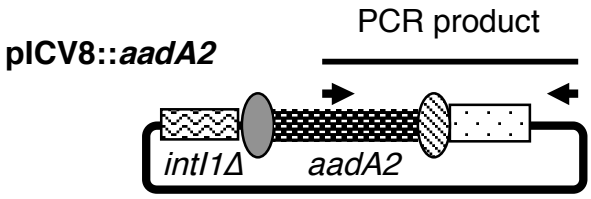


**Figure 2.7.** PCR amplification of pICV8::*aadA2* 5' and 3' junctions distinguishes single and double insertion events. **A.** Schematic of PCR products. **B.** 3' junction PCR. The 3' junction amplicand is 1000 bp for a single or double insertion event. Lanes 4-5 and 6-7 show duplicate PCRs. **C.** 5' junction PCR. The 5' junction amplicand is 1276 bp for insertion of *aadA2* and is 1915 bp for insertion of both *aacA4* and *aadA2*. Lanes 13-14 and 15-16 show duplicate PCRs. CB=CB454, pI8=pICV8, pRMH=pRMH14, and pSU=pSU2056.

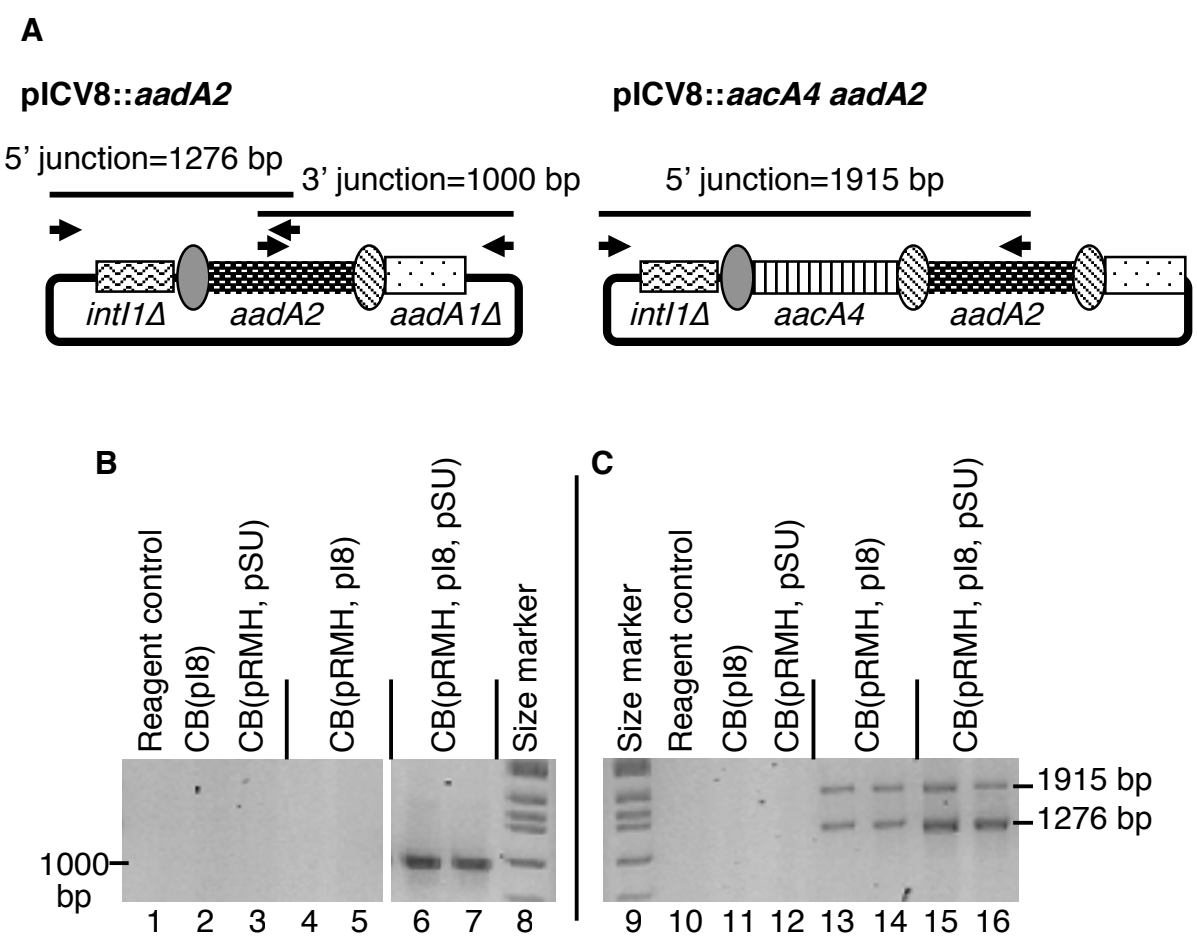

A
pICV8::aadA2
5' junction=1276 bp
3' junction=1000 bp
intI1Δ
aadA2
aadA1Δ
pICV8::aacA4 aadA2
5' junction=1915 bp
intI1Δ
aacA4
aadA2
B
Reagent control
CB(pI8)
CB(pRMH, pSU)
CB(pRMH, pI8)
CB(pRMH, pI8, pSU)
Size marker
1000 bp
1 2 3 4 5 6 7 8
C
Size marker
Reagent control
CB(pI8)
CB(pRMH, pSU)
CB(pRMH, pI8)
CB(pRMH, pI8, pSU)
1915 bp
1276 bp
9 10 11 12 13 14 15 16

**Figure 2.8.** Schematic of possible junctions arising in cointegrate formation. The 5' and 3' junction primers are black arrows, and the cointegrate junction primers are gray arrows. **A.** Cointegrate formed between the two *attI* sites. **B.** Cointegrate formed by recombination between the *aacA4 attC* and the pICV8 *attI* sites. **B.** Cointegrate formed by recombination between the *aadA2 attC* and the pICV8 *attI* sites.

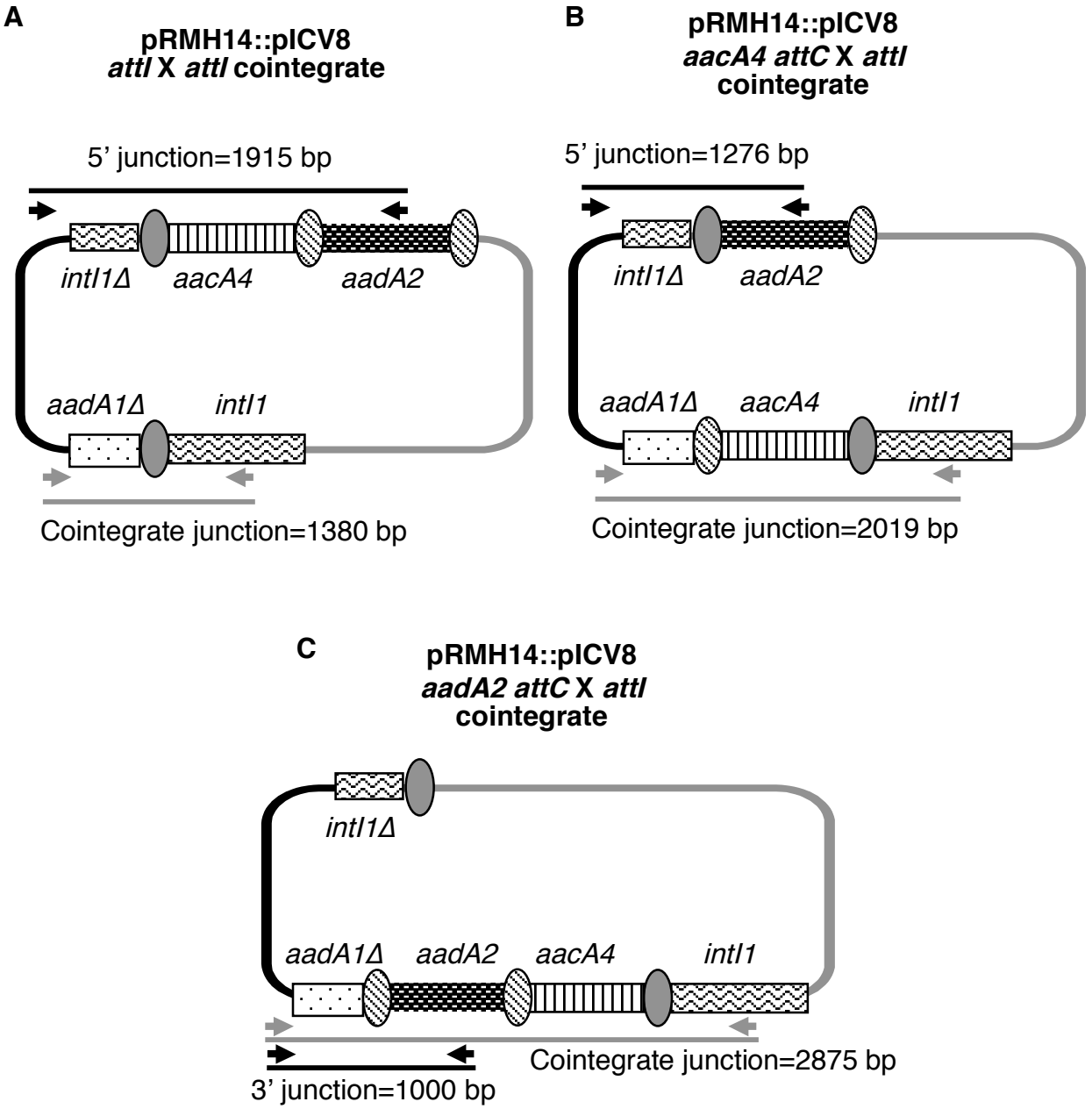

A
pRMH14::pICV8
attI X attI cointegrate
5' junction=1915 bp
intI1Δ
aacA4
aadA2
aadA1Δ
intI1
Cointegrate junction=1380 bp
B
pRMH14::pICV8
aacA4 attC X attI
cointegrate
5' junction=1276 bp
intI1Δ
aadA2
aadA1Δ
aacA4
intI1
Cointegrate junction=2019 bp
C
pRMH14::pICV8
aadA2 attC X attI
cointegrate
intI1Δ
aadA1Δ
aadA2
aacA4
intI1
Cointegrate junction=2875 bp
3' junction=1000 bp

**Figure 2.9.** The pICV4::*dfrB2* 3' junction is only detectable with high intracellular integrase. **A.** Schematic of PCR products. **B.** 5' junction PCR. The 5' junction amplicand of a single insertion of *dfrB2* or an insertion of both *orfA* and *dfrB2* is 1523 bp. **C.** 3' junction PCR. The 3' junction is 1060 bp for a single insertion of *dfrB2* and 1561 bp for insertion of both *orfA* and *dfrB2*. CB=CB454, pI4=pICV4, R3=R388, and pSU=pSU2056. The two lanes shown for each strain are duplicate PCRs.

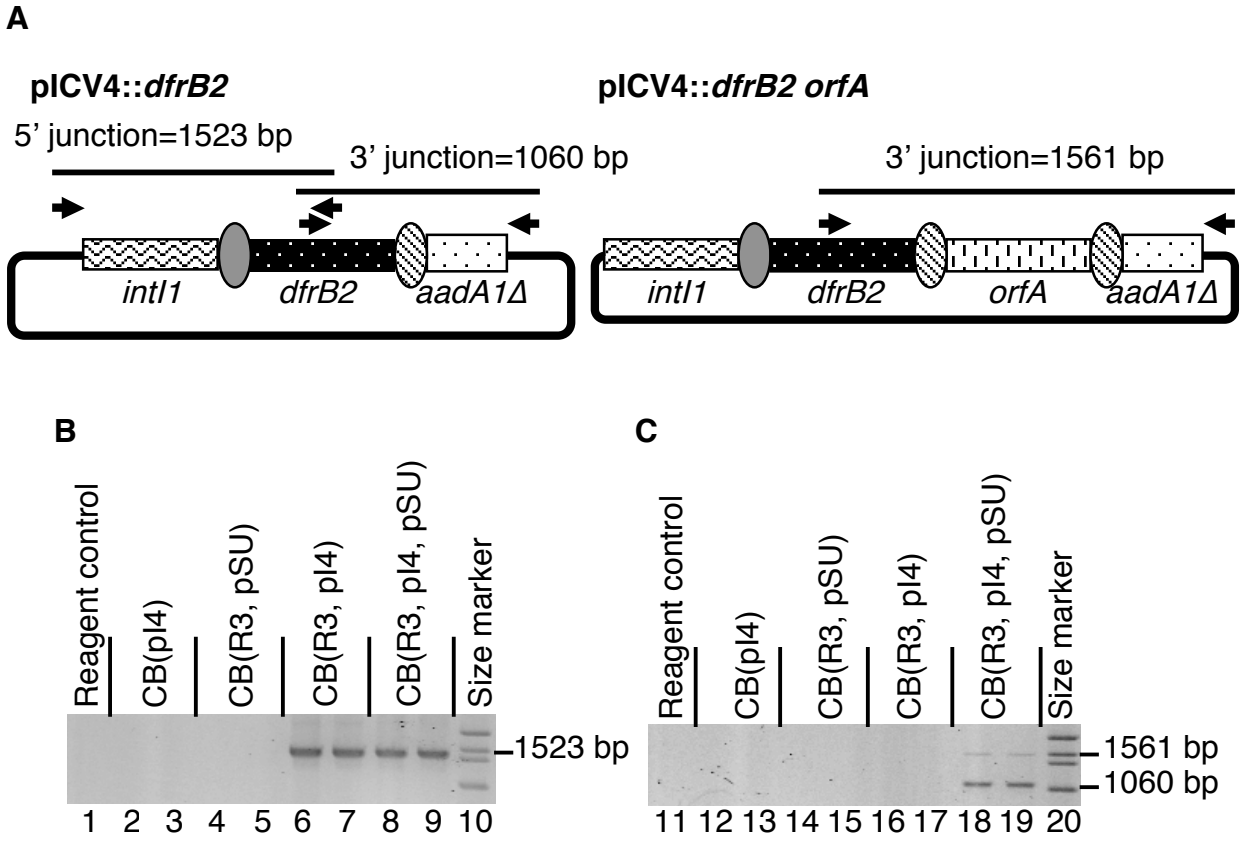

A
pICV4::dfrB2
5' junction=1523 bp
3' junction=1060 bp
intI1
dfrB2
aadA1Δ
pICV4::dfrB2 orfA
3' junction=1561 bp
intI1
dfrB2
orfA
aadA1Δ
B
Reagent control
CB(pI4)
CB(R3, pSU)
CB(R3, pI4)
CB(R3, pI4, pSU)
Size marker
1523 bp
1 2 3 4 5 6 7 8 9 10
C
Reagent control
CB(pI4)
CB(R3, pSU)
CB(R3, pI4)
CB(R3, pI4, pSU)
Size marker
1561 bp
1060 bp
11 12 13 14 15 16 17 18 19 20

**Figure 2.10.** pICV8::*aadA2* can act as a cassette donor. **A.** Schematic of the 3' junction formed by insertion of *aadA2* into pICV3. **B.** 3' junction amplicand of isolated pICV3::*aadA2*. Lanes 3-4, 5-6, 8-9 show duplicate PCRs. CB=CB454, pI3=pICV3, pI8::*A2*=pICV8, SK=SK2267A, pSU=pSU2056, and pI3::*A2*=pICV3::*aadA2*.

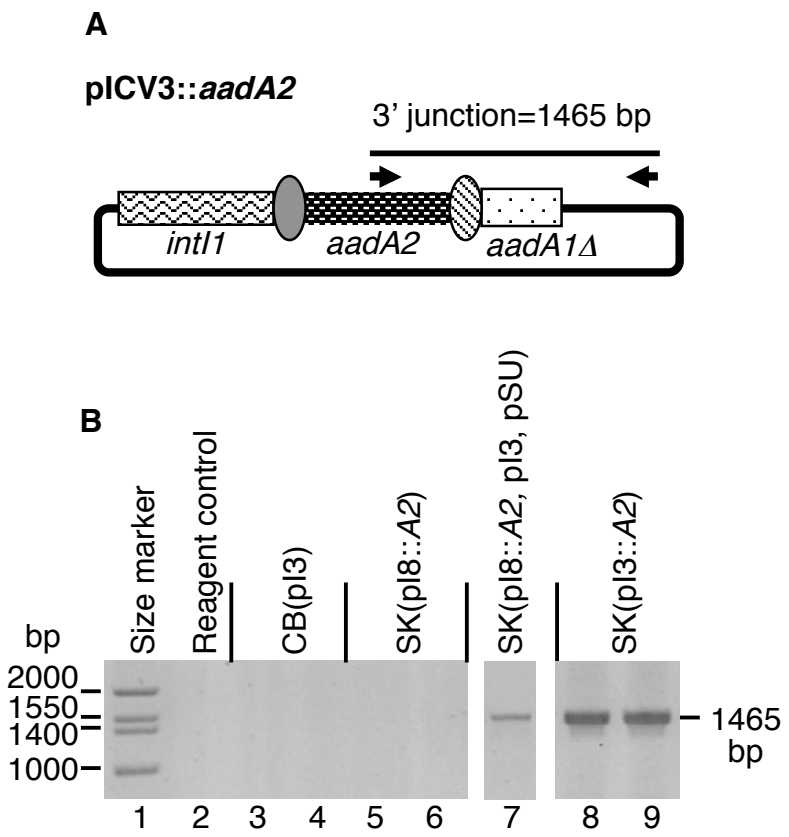
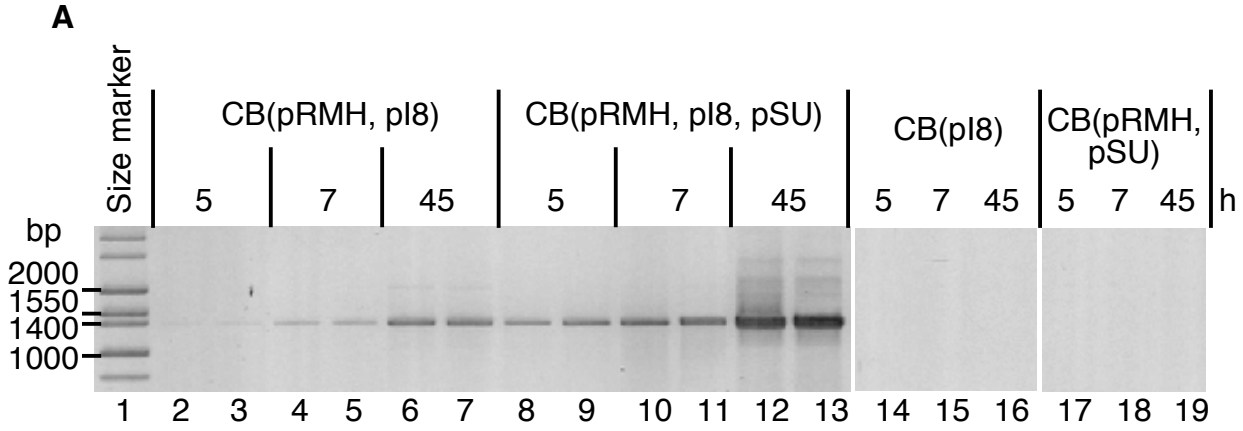

A
pICV3::aadA2
3' junction=1465 bp
intI1
aadA2
aadA1Δ
B
Size marker
Reagent control
CB(pI3)
SK(pI8::A2)
SK(pI8::A2, pI3, pSU)
SK(pI3::A2)
bp
2000
1550
1400
1000
1465 bp
1 2 3 4 5 6 7 8 9

**Figure 2.11.** Cointegrates accumulate from late log to late stationary phase. **A.** pRMH14::pICV8 cointegrate junction amplicand detected by PCR of broth cultures at 5, 7, and 45 hours of growth. Each lane contains 30 ml PCR product. The 1380 bp product represents the *attI* x *attI* cointegrate (Fig. 2.8). Masses of size markers: 2000 bp=150 ng, 1550 bp=100 ng, 1400 bp=100 ng, 1000 bp=120 ng. Lanes 2-3, 4-5, 6-7, 8-9, 10-11, 12-13 show duplicate PCRs. **B.** Mass of the cointegrate junction amplicand determined by densitometry per number of cells in the PCR. CB=CB454, pRMH=pRMH14, pI8=pICV8, pSU=pSU2056.


A
Size marker
CB(pRMH, pI8)
CB(pRMH, pI8, pSU)
CB(pI8)
CB(pRMH, pSU)
5
7
45
5
7
45
5
7
45
5
7
45
h
bp
2000
1550
1400
1000
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

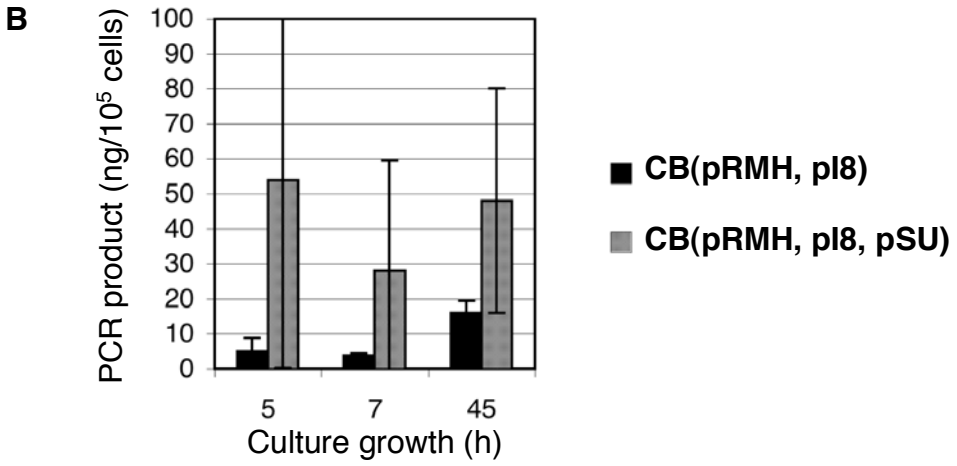
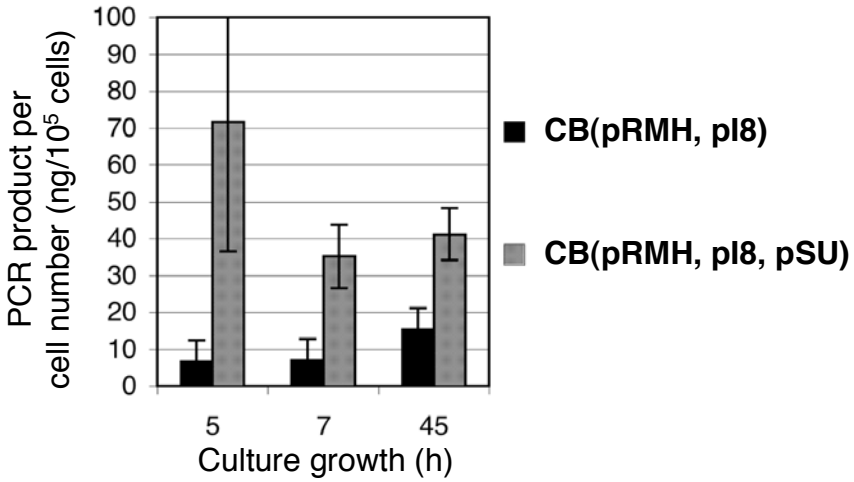
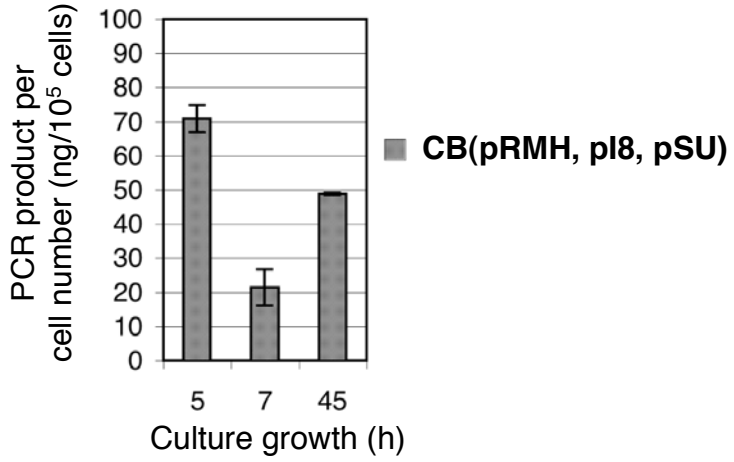

B
PCR product (ng/10^5 cells)
100
90
80
70
60
50
40
30
20
10
0
5
7
45
Culture growth (h)
CB(pRMH, pI8)
CB(pRMH, pI8, pSU)

**Figure 2.12.** 5’ and 3’ junctions accumulate from late log to late stationary phase. **A.** Mass of the 5’ junction pICV8::*aadA2* amplicand per number of cells in the PCR. **B.** Mass of the 3’ junction pICV8::*aadA2* amplicand per number of cells in the PCR.

**A**


PCR product per cell number (ng/10^5 cells)
100
90
80
70
60
50
40
30
20
10
0
5
7
45
Culture growth (h)
CB(pRMH, pI8)
CB(pRMH, pI8, pSU)

**B**


PCR product per cell number (ng/10^5 cells)
100
90
80
70
60
50
40
30
20
10
0
5
7
45
Culture growth (h)
CB(pRMH, pI8, pSU)

# CHAPTER 3

## *In vivo* integrase abundance affects the type and amount of IntI1-mediated recombination products differentially during the growth cycle

### Overview

The integron integrase IntI1 mediates the recombination of antibiotic resistant gene cassettes between integrons, facilitating the dissemination of these genes. Historically, integrase activity has been measured after conjugation of its products to a recipient cell. Here we report the first measurements of the time-dependence of formation of integron intracellular products. Integrase-mediated recombination products were quantified over time using dilution PCR in low and high integrase strains. Only cointegrates were detectable in strains expressing integrase under a natural promoter, and the simple acquisition of gene cassettes could only be seen in strains overexpressing integrase. Recombination products in both the low and high integrase strains increased in late log to early stationary phase. Unexpectedly, more *attI* x *attI* than *attC* x *attI* recombination products were detected in log phase for both strains, but the high integrase strain showed increased *attC* recombination in stationary phase, consistent with earlier observations on integrase crossover preferences (14, 32). Thus, the integrase's intracellular abundance affects the amount and type of recombination events, especially toward the end of log phase.

## Introduction

Horizontal gene transfer is a major mechanism through which bacterial pathogens gain resistance to multiple antibiotics (15, 34). One genetic system that spreads antibiotic multiresistance between bacterial species is the integron. Integrons carrying as many as nine antibiotic resistance gene cassettes have been detected in clinical and agricultural isolates and in bacterial species ranging from Gram-negatives to low- and high-GC Gram-positives (8, 18, 28, 30, 35-38, 40, 44, 47, 50, 57). Class 1 integrons are the most prevalent genetic system through which multiple resistances are spread in the Enterobacteriaceae (34). Although the majority of antibiotic resistances found in class 1 integrons are to older antibiotics (7, 17, 27, 29, 45, 54), demonstrating the persistence of these genes, *qnr* genes conferring resistance to quinolones have also been found within integrons (25, 51), showing that integrons are also spreading resistance to newer antimicrobials (49). The integron-mediated dissemination of antibiotic resistances cannot be slowed without further understanding of how these genes move throughout these diverse ecosystems.

Integrons are found on the plasmids and chromosomes of many different bacteria and contain a site-specific recombination system in which an integrase, a site-specific recombinase of the tyrosine family (46), inserts or removes gene cassettes at an *attI* recombination site, located between the integrase and the cassettes (see Chapter 2, Fig. 2.1). The gene cassettes are short DNA sequences (600-1500 bp) consisting of an open reading frame, often encoding antibiotic resistance, and an *attC* recombination site (20, 21, 32, 46). Gene cassettes are arranged in a tandem array and rarely contain

their own promoter sequence, but, instead, are transcribed by a common promoter ($P_c$) located on the opposite side of *attI*. The integrase gene itself is divergently transcribed from the gene cassette array (see Chapter 2, Fig. 2.1) (9, 46). Integron transfer of gene cassettes, called "cassette capture," occurs by two mechanisms: direct excision and integration of only the cassette or the formation and resolution of cointegrates between the donor and recipient replicons.

In the excision and integration mechanism, gene cassettes are excised from one integron (donor) into covalently closed circles and then inserted into the *attI* site of a different integron (recipient) in the same cell (see Chapter 2, Fig. 2.2) (11, 21). In the other mechanism, formation and resolution of cointegrates, the integrase mediates recombination between the donor and recipient plasmids at the *attI* and/or *attC* sites to form a cointegrate plasmid (see Chapter 2, Fig. 2.3). Unlike cointegrate formation during replicative transposition (22), this is a conservative process with no replication involved. The cointegrate is resolved by the integrase to create two separate plasmids, either of which can contain the cassette(s) adjacent to the insertion site (12, 23, 32, 42). The final cassette capture products resulting from the two different mechanisms are the same and indistinguishable by sequence (12).

Integron recombination has long been thought to be a highly regulated process (23). Related tyrosine recombinases like XerCD and lambda integrase have multiple processes for controlling the outcome of recombination including transcriptional regulation of the recombinase (26, 56), binding of specific accessory proteins (1, 19, 43), and specificity for supercoiled DNA template (4, 19). In addition, under high

integrase conditions, integron gene cassettes insert at secondary sites just 3-bp long, usually “GTT” or “GAT”, which occur frequently in any DNA (14, 23, 42). So, regulation of integrase-mediated recombination is necessary since gene cassette insertion at secondary sites in the chromosome would likely have a deleterious effect (14, 23). Nonetheless, there is much not yet understood about integron regulation. Prior *in vivo* integron studies have assayed a distant endpoint product (i.e. recombination products in plasmids conjugated or transformed into a recipient strain) and required integrase overexpression in the donor strain to see direct transfer of a cassette from one integron to another (10, 12). In cells with normal levels of integrase expressed by the wild-type integrase promoter on a typical low copy plasmid, only cointegrates were detected (12, 32, 33) (and see Chapter 2). The use of conjugation or transformation in the assays could have constrained the kinds of products detected and provides no information on their steady-state distributions in the donor strain. Such methods also do not allow assessments of real time changes in the type and abundance of products during the culture cycle. Kinetic data for integron reactions is of interest because recombinase concentrations affect *in vivo* reactions catalyzed by other tyrosine site-specific recombinases, such as Cre resolvase (41) and the lambda phage integrase (26), by reducing reaction specificity or by favoring one type of reaction over another. Moreover, since cointegrate formation is the only integrase-mediated recombination reaction observed *in vivo* without excess integrase (12, 32, 33), it is itself worthy of examination. Given that cointegrates are observed with or without extra integrase, cointegrate formation may dominate in natural systems. Thus, we wanted to examine the relative

frequencies at which the products of both processes occur in strains lacking or containing an integrase-overexpressing plasmid without additional complexities arising from transfer to a recipient cell.

Here we examine integrase-mediated recombination using an integron capture vector, pICV8, capable of receiving and donating antibiotic resistance gene cassettes (see Chapter 2). pICV8 is a wide-host range, low copy number plasmid containing only an *attI* recombination site. During initial qualitative work we noted that novel junctions were most abundant in late stationary phase (see Chapter 2). To better define this phenomenon we used dilution PCR of recombinant junctions to quantify distinct integrase-mediated recombination products *in vivo* throughout the culture cycle using pICV8 and an integron cassette donor, pRMH14 (Fig. 3.1 A), with and without excess integrase supplied by pSU2056.

## Methods

**Bacterial strains and plasmids.** The *E. coli* strain used was CB454 [*gal lacZ lacY thi str rpsl- recA56*]. Table 3.1 lists the plasmids used and their phenotypes (6, 16, 32, 48, 52). Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN). All plasmids were transformed by electroporation using a Gene Pulser (BioRad, Inc.) set for 25 μFD, 200 ohms, and 2.5 kV (2). Transformants were selected on Luria-Bertani (LB) plates or broth containing antibiotics. Concentrations of antibiotics used were: zeocin 25 μg/ml (LB pH 7.5), kanamycin 50 μg/ml, ampicillin 50 μg/ml, streptomycin 250 μg/ml. Antibiotics were obtained from Sigma Inc. (St. Louis, MO) except zeocin (Invitrogen

Corporation, Carlsbad, CA). pRMH14::pICVm *attI*x*attI* cointegrates were isolated using a mating-out procedure (33). SK2267A(pRMH14, pICVm, pSU2056) was mated with CB454, and CB454(pRMH14::pICVm) was selected on LB plates with streptomycin, kanamycin, and zeocin. Cointegrate plasmid DNA was isolated, and the type of crossover that had occurred in each case was determined by PCR and restriction digestion.

**Template detection limits of PCR primers.** PCRs contained 1X Magnesium-free Buffer, 1.5 mM $MgCl_2$, 0.25 mM each dNTP (Applied Biosystems, Foster City, CA), 20 pmol of each primer, and 2.5 units of Taq DNA Polymerase. All reagents except dNTPs were from Promega, Inc., Madison, WI. PCRs were performed in PTC-100 and PTC-200 thermocyclers (MJ Research, Inc., Waltham, MA). 30 μl of PCR products were visualized by agarose gel electrophoresis after staining with ethidium bromide and detection by a Molecular Dynamics Fluoroimager 575. PCR primers were synthesized by Sigma-Genosys, and PCR conditions were optimized by gradient PCR for all primer pairs. Large PCR amplicands were required to detect recombinant junctions in order to avoid homologous regions between donor and recipient plasmids. Table 3.2 shows each primer pair used for 50 μl reactions in 40-cycle PCR, the size of the amplified fragment quantified, and the minimum number of copies detectable as determined by dilution PCR of pure chromosomal or standard plasmid DNA with the relevant structure. Six separate PCRs were performed in duplicate reactions using a DNA dilution series with a dilution factor 2, for a total of 12 reactions to determine detection limits. The

detection limit was defined as the least amount of DNA in a 50 μl PCR yielding a visible band when 30 μl was run on an agarose gel. The mean mass of DNA used in each reaction yielding gel bands was calculated from the known concentration of the corresponding stock and converted to the number of copies of each target replicon, based on the molecular weight of each plasmid or chromosome derived from its known length in base pairs (660 Da/bp). One copy of the CB454 chromosome is 5.17 fg DNA; 1 copy of pICV8 is $8.53 \times 10^{-3}$ fg DNA; 1 copy of pICV8::*aadA2* is $9.28 \times 10^{-3}$ fg DNA; and 1 copy of pRMH14::pICVm is $4.35 \times 10^{-2}$ fg DNA.

**Cultivation of strains, viable cell counts, and DNA extraction.** Plasmid DNA was electroporated into 40 μl electrocompetent cells using conditions described above. The amount of plasmid DNA used was: 55 ng of pICV8, 55 ng of pICV8::*aadA2*, and/or 27.5 ng pSU2056. pICV8 was electroporated into CB454 and into CB454(pRMH14), and pICV8::*aadA2* was electroporated into CB454. pICV8 and pSU2056 combined were electroporated together into CB454(pRMH14), and pSU2056 was electroporated into CB454(pICV8). After pulsing, 1ml SOC media (2) was added to each electroporated aliquot and two CB454(pRMH14, pICV8, pSU2056) aliquots were pooled in order to have sufficient volume for later inoculation of a large culture. Electroporated cultures were incubated at 37 °C, shaking at 200 rpm for a 1 hour recovery. We define as time zero the point at which 500 μl (or 1 ml for CB454(pRMH14, pICV8, pSU2056) of the electroporated culture was added to 700 ml of LB broth containing antibiotics selecting for the plasmids in each strain—zeocin for pICV8, ampicillin for pSU2056, and

kanamycin for pRMH14 (Fig. 3.2). An aliquot of culture (100–150 ml) was removed at time zero for DNA extraction. The cultures were grown at 37 °C, shaking at 250 rpm, and aliquots were removed at 12, 14, 16, 20, 24, 36, and 48 hours of growth for DNA extraction (between 2 ml and 175 ml) and viable cell counts (1 or 2 ml) (Fig. 3.2). The 14 hour timepoint was collected for CB454(pRMH14, pICV8), CB454(pICV8), and CB454 (pICV8::aadA2) only in experiments 2 and 3. Aliquots for DNA extraction were centrifuged immediately at 9000 g in a Beckman J2-21 or an Eppendorf 5415C centrifuge for 10 minutes, and pellets were held at 4 °C. For viable cell counts, the $OD_{600}$ of an aliquot portion was measured on a Cary 100 spectrophotometer, and the remaining sterile aliquot was diluted in duplicate series in LB broth at 10-fold dilution steps and plated in triplicate on LB agar with antibiotic selection for the plasmids present (six plates per dilution or three plates of undiluted aliquot plated for each strain). Colony forming units (cfu) per ml were determined after incubation at 37 °C for 24 hours. DNA was isolated from each culture sample pellet by phenol/chlorform extraction and ethanol precipitation (2), followed by dialysis against TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) in a 10,000 MWCO Slide-A-Lyzer® Mini Dialysis Unit (Pierce). DNA was quantified at $OD_{260}$ on a Varian Cary 100 Bio UV/vis spectrophotometer using Cary Win/UV RNA/DNA software. This entire experiment (electroporation, cultivation, time sample collections, and assays) was repeated three times per strain.

**Quantification of recombinant junctions during the growth cycle by dilution PCR.** Dilution PCR with serial steps of 2 for each primer set (Table 3.2) was performed using total DNA isolated from each culture timepoint sample for strains CB454(pRMH14, pICV8) and CB454(pRMH14, pICV8, pSU2056). Three PCRs were performed with duplicate reactions of DNA from each timepoint per primer pair, giving six observations of limiting dilution detection values at each timepoint. The detection value (units=ng) is operationally defined as the least amount of total DNA (ng) in a 50 μl PCR yielding a visible band when 30 μl of the reaction mix was run on an agarose gel and stained with ethidium bromide. The detection values (ng) were recorded for each reaction, and the mean of all was determined at each timepoint for each strain. Less than 500 ng total DNA was used in 5' and cointegrate junction PCRs because using 500 ng or more total DNA yielded hybrid or false products in PCR of negative controls—e.g. CB454(pICV8) DNA combined with pRMH14 pure plasmid DNA and/or PCR of CB454(pRMH14, pSU2056). No false PCR products were seen in 3' junction PCRs using up to 36 μg DNA. Simple PCR (presence/absence of product in 40 cycles) was performed in triplicate on DNA isolated from each timepoint for CB454(pICV8), CB454(pICV8, pSU2056), and CB454(pICV8::*aad2*) for all primer pairs to confirm the presence of chromosome (*bioB*) and pICV8 in all strains (10-50 ng total DNA per reaction), the absence of recombinant junctions (250-500 ng total DNA per reaction for 5' and cointegrate junctions and up to 25 μg for 3' junctions) in the former two strains, the presence of 5' and 3' junctions (10-100 ng total DNA) and the absence of cointegrate junctions (250-500 ng) in the latter strain.

**Calculating recombinant junctions per backbone or per cell.** The number of recombinant junctions (5', 3' or cointegrate), chromosomal *bioB* genes, and pICV8 plasmids per μg of total culture DNA (Tables 3.3 and 3.4) were calculated by dividing the primer detection limit (units=number of copies/standard reaction) (Table 3.2) by the average detection value for each timepoint (units=μg/reaction) (i.e. the least amount of DNA (ng) giving rise to a band on an agarose gel of the standard reaction). The standard reaction is defined as a 50 μl PCR where 30 μl product is run on an agarose gel. For timepoints where no junction amplicand was detected (Table 3.3, gray values), the theoretical possible number of recombinant junctions per μg DNA was calculated based on the maximum amount of DNA in the PCR. The number of recombinant junctions per replicon (either pICV8 or *bioB*) was calculated by dividing the recombinant junctions per μg total DNA by the number of pICV8 copies or the number of *bioB* copies per μg total DNA (Tables 3.5, 3.6). For clarity its reciprocal, the copies of pICV8 or cells (*bioB)* per recombinant junction (an integer rather than a fraction), was plotted versus time (Fig. 3.5) and versus number of generations up to 24 hours (Fig. 3.6). Generations elapsed were calculated for each strain using the equation: $n$=3.3(log N - log $N_0$), where N is the cfu/ml of a culture at a specific timepoint and $N_0$ is the cfu/ml at time zero.

**Cloning and sequencing PCR products.** PCR products for CB454(pRMH14, pICV8, pSU2056) (with all 5 primer pairs) and for CB454(pRMH14, pICV8) (with all primer pairs except 3' junction) were cloned for sequencing. PCR product bands were cut from an agarose gel stained with Sybr Green I (Molecular Probes). The DNA was removed from

the gel slices by electroelution in 1" dialysis tubing that had been boiled in 25 mM EDTA for 10 minutes, rinsed in water, and stored at 4 °C in 30% ethanol. Tubing was cut to 3-4 inches and rinsed in 0.1X TBE immediately before use. Gel slices were placed along one wall inside dialysis tubing with 250 μl 0.1X TBE and all air removed before clamping. Tubing was placed in an electorphoresis apparatus so the gel slice was perpendicular to the path of the current. The dialysis tubing was covered with 0.1X TBE and electrophoresed at 100 V for 2 hours. The polarity was then reversed for two minutes to disassociate DNA from tubing. The electoeluted DNA was removed by pipette. Cloning of electroeluted PCR products into pCR2.1-Topo and electroporation of clones into Top10 cells were performed using the TOPO TA cloning kit (Invitogen). White colonies on LB agar plates containing kanamycin (50 μg/ml) and spread with 80 μl Xgal (20 mg/ml) were selected for plasmid DNA isolation and restriction digest. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (QIAGEN). Plasmids showing inserts by restriction digest were sent to Agencourt, Inc. (Beverly, MA) for seqencing and results were as expected.

**Calculations**

**Determining the number of copies of a gene present by limiting detection PCR.**

*N1-N12 determined experimentally, all other values known*

**Abbreviations:**

*N1-N12*= smallest mass (in fg) of pure DNA in a 50 μl PCR giving a visible band when 30 μl reaction is run on an agarose gel and stained with ethidium bromide (for each

timepoint in each strain 6 PCRs were done in duplicate reactions, giving 12 values in units of fg DNA)

$N_{av}$=the average of *N1-N12* in fg

*b*=size of target molecule (either plasmid DNA or chromosome) in base pairs

660 Da=molecular weight of one base pair of DNA

*M*=molecular weight (g/mol)

$6 \times 10^{23}$=number of molecules (copies) in one mole (Avogadro's number)

$1 \text{ g} = 1 \times 10^{15}$ fg

**Relationships:**

$N_{av}$ fg = (*N1*+*N2*+…+*N12*) / 12

*M* g/mol = *b* x 660 Da/bp

Detection limit of primer pair (# copies) = $N_{av}$ / (*M* x (1/(6 x $10^{23}$) x (1 x $10^{15}$))

**Example:**

pICV8 plasmid quantification using pure pICV8 plasmid DNA (*b*=7584 bp):

$N_{av}$ of pICV8 = (6.25 fg + 3.125 fg + 3.125 fg + 3.125 fg + 6.25 fg + 3.125 fg + 6.25 fg + 3.125 fg + 6.25 fg + 12.5 fg + 12.5 fg + 6.25 fg) / 12

$N_{av}$ of pICV8 = 5.99 fg ± 3.39

*M* = 7584 bp x 660 Da/bp = 5.005 x $10^6$ g/mol pICV8

Detection limit = 5.99 fg / ((5.005 x $10^6$ g/mol) x (1 mol / (6 x $10^{23}$ copies)) x (1 x $10^{15}$ fg/g))

**detection limit** = **718.1 copies pICV8 plasmid per reaction ± 406** (this means that when 30 μl of a 50 μl PCR is run on an agarose gel and yields a visible band at least 718.1 copies of pICV8 plasmid are present)

**Determining the number of recombinant junctions and pICV8 plasmids per μg total cellular DNA for each strain at each timepoint.**

**Abbreviations:**

*N1-N6*= smallest mass of total cellular DNA (in ng or pg) in a 50 μl PCR giving a visible band when 30 μl is run on an agarose gel and stained with ethidium bromide (3 PCRs were done in duplicate reactions, giving 6 detection values in ng or pg DNA)

$N_{av}$(in ng or pg per reaction) = the average of *N1-N6*

**Relationships:**

$N_{av}$(in ng or pg per reaction) = (*N1*+*N2*+…+*N6*) / 6

Given that $N_{av}$ is defined as the average number of copies per μg starting DNA and must have at least the number of copies previously determined by the detection limit of the primer pair:

#copies/μg total DNA ≈ (detection limit in #copies per reaction) / $N_{av}$(μg)

**Example:**

Copies pICV8 plasmid per μg DNA for CB454(pRMH14, pICV8, pSU2056) trial #1, 24 hour timepoint

$N_{av}$ = (7.8 pg + 3.9 pg + 3.9 pg + 1.95 pg + 3.9 pg + 1.95 pg) / 6 = 3.9 pg ± 2.1

3.9 pg total DNA has at least 718.1 copies pICV8

718.1 copies / 3.9 pg ≈ 184 copies/pg x (1 x $10^6$ pg/μg)

**≈ 1.84 x $10^8$ copies of pICV8 /μg total DNA**

3' junctions per μg DNA for CB454(pRMH14, pICV8, pSU2056) trial #1, 24 hour timepoint

$N_{av}$ = (5302 ng + 5302 ng + 5302 ng + 5302 ng + 5302 ng + 5302 ng) / 6

= 5302 ng ± 0

5302 ng total DNA has at least 12342 copies 3' junction

12342 copies / 5302 ng ≈ 2.33 copies/ng (1 x $10^3$ ng/μg)

**≈ 2.33 x $10^3$ 3' junctions /μg total DNA**

**Determining the number of recombinant junctions per pICV8 molecule.**

**Definition:**

Number of recombinant junctions per μg total DNA / number of copies pICV8 plasmid per μg total DNA (at that timepoint for that strain trial)

**Example:**

CB454(pRMH14, pICV8, pSU2056) trial #1, 24 hour timepoint

Number of 3' junctions per μg total DNA / number pICV8 plasmids per μg total DNA = (2.33 x $10^3$) / (1.84 x $10^8$)

**= 1.27 x $10^{-5}$ [or the reciprocal, 7.90 x $10^4$ copies pICV8 / 3'junction—i.e. 3' junction only present in 1/79,000 pICV8 molecules]**

**(same method is used to determine "per chromosome")**

**Determining the number of generations (*n*) of a culture at a particular timepoint.**

**Abbreviations:**

*n*=number of generations

$N_0$=cfu/ml of broth culture at time zero, determined by aerobic plate count of electroporate used to inoculate 700 ml broth culture

N=cfu/ml of broth culture at time t, determined by aerobic plate count

**Relationships:**

$N_0$=[(average # cfu per plate / volume of electroporate plated in ml) x (volume electroporate added to 700 ml broth)] / 700 ml

N=(average # cfu per plate) / [(dilution plated) x (volume plated in ml)]

$n=3.3(\log N - \log N_0)$

**Example:**

CB454(pRMH14, pICV8, pSU2056) trial #1, 24 hour timepoint

Calculate $N_0$:

$N_0$=[(((266+251+210)/3) / 0.1 ml) x (1 ml)] / (700 ml)

=3.46 cfu/ml ± 29.0

Calculate N:

$N=((74+71+58+187+59+30)/6) / [(1 \times 10^{-2}) \times (0.1 \text{ ml})]$

$=7.98 \times 10^4$ cfu/ml ± 54.8

Calculate *n*:

$\boldsymbol{n}=3.3(\log (7.98 \times 10^4) - \log (3.46))$

$=3.3 \times (4.902 - 0.539)$ =**14.4**

## Results

### There are similar numbers of cointegrate and 5' recombinant junctions per μg total cell DNA in low and high integrase strains during log phase

When pICV8 captures a cassette by excision/integration or by cointegrate formation and resolution, the two unique junctions formed can be amplified by specific PCR primer pairs (Fig. 3.1). Amplification of both the 5' and 3' junctions formed between pICV8 and the *aadA2* gene cassette from donor pRMH14 indicates *aadA2* cassette capture (Fig. 3.1 B). The 5' and 3' junctions can also indicate the presence of pRMH14::pICV8 cointegrates (Fig. 3.1 C). Using the cointegrate junction primers, a 1380 bp amplicand detects pRMH14::pICV8 *attI* x *attI* cointegrates. We previously observed that only the 5' junction and cointegrate junction amplicands, and not the 3' junction amplicand, are detected in a strain expressing integrase under its natural promoter, CB454(pRMH14, pICV8) but all three junctions are seen in a strain overexpressing integrase, CB454(pRMH14, pICV8, pSU2056) (see Chapter 2). Thus, in the low integrase strain, *attI* x *attI* and *attC1* x *attI* cointegrates (Fig. 3.1 C) are present, but complete cassette capture products are not observed without excess integrase provided by pSU2056. In order to quantify recombination products by dilution PCR, we first determined the PCR primer detection limits by dilution PCR of pure plasmid (or chromosomal) DNA for the three junction primer pairs, as well as for primer pairs respectively detecting pICV8 plasmids and chromosomes (*bioB*) (Table 3.2). Though the cointegrate junction and 3' junction share a primer that anneals to the pICV8 plasmid backbone (Fig. 3.1), the 3' junction requires 44 times more DNA for detection

(Table 3.2). The low detectability of the 3' junction may be caused by the secondary structure formed by the *aadA2 attC* site which is not present in the cointegrate junction (Fig. 3.1 B, C).

We then proceeded to quantify recombination products using dilution PCR of total DNA isolated at various times during the growth of a culture. So as to define the onset of integrase-mediated recombination in cells during culture growth, strains were constructed afresh in each experiment. The recipient plasmid pICV8 (with or without pSU2056 for high integrase expression) was (were) electroporated into a *recA*$^-$ *E. coli* strain already containing the donor plasmid pRMH14. Time zero was defined as the timepoint immediately after a one hour recovery period when antibiotic selection was imposed, and culture samples were acquired at several timepoints up to 48 hours of growth at 37ºC with aeration (Fig. 3.2). The viable cell counts for three independent trials each of CB454(pRMH14, pICV8) and CB454(pRMH14, pICV8, pSU2056) (Fig. 3.3) show that the strain containing pSU2056 has a longer lag phase and enters stationary phase with 30 – 200-fold lower cell density than the strain with natural integrase expression, but the growth rates of both strains are similar between 12 and 24 hours and both reached stationary phase between 24 and 36 hours of growth.

As expected, dilution PCR (Fig. 3.4) shows there are more cells (as reported by the copies of the chromosomal gene *bioB*) than there are pICV8 plasmids (as reported by copies of a nonmobile region of this plasmid) at time zero because it is immediately after electroporation and before selection has enriched the transformed subpopulation carrying pICV8 (initially approximately 0.1–1% of cells). Subsequently, the copy

numbers of *bioB* and of pICV8 per μg of total DNA are very similar in each strain in all three trials, especially from 20 to 36 hours of growth. There were ~10-fold more copies of pICV8 than of *bioB* per μg total DNA as expected (16). The viable cell counts (cfu/ml) for CB454(pRMH14, pICV8) and for CB454(pRMH14, pICV8, pSU2056) differ by 100-fold in stationary phase (Figs. 3.3 and 3.4 A-C, D-F) and colonies of the latter were mucoid, especially after 24 hours, although liquid cultures were not viscous (data not shown).

Recombinant junctions first became detectable at 16 or 20 hours in both strains (Table 3.3). The most striking difference between the two strains is that 3' junctions are not observed in the strain expressing integrase at natural levels (Fig 3.4 A-C). In contrast, 5' and cointegrate junctions are abundant in both the low and high integrase strains at 36 and 48 hours, although they are 1000 – 10,000-fold lower in the former. However, *in vivo* at 36 hours in the low integrase strain, there are at least 14.6, 56.9, and 84.6 times (Table 3.3) more cointegrate junctions (*attI* x *attI* cointegrate) (Fig. 3.1 C) than 3' junctions (*attC2* x *attI* cointegrate, pICV8::*aadA2*, and/or pICV8::*aacA4aadA2*) (Figs. 3.1 B) for the three repeated experiments (Fig. 3.4 A, B, C). In addition, there are at least 5-fold more 5' junctions (*attC1* x *attI* cointegrate and/or pICV8::*aadA2*) than 3' junctions (*attC2* x *attI* cointegrate, pICV8::*aadA2*, and/or pICV8::*aacA4aadA2*) at 36 hours in the low integrase strain (Table 3.3). The 5' junctions are similarly abundant in both strains at 20 and 24 hours, but there is more cointegrate junction in the high integrase strain at all timepoints. The apparent absence of 3' junction in the low integrase strain suggests that the 5' junctions observed represent pRMH14::pICV8

*attC1* x *attI* cointegrates rather than *aadA2* cassette capture by pICV8 (Fig. 3.4 A-C). This agrees with our initial qualitative observations that pICV8 did not acquire *aadA2* from pRMH14 and pICV4 did not acquire *dfrB2* from R388 without increased integrase provided by pSU2056 (see Chapter 2).

Moreover, in the lower integrase strain CB454(pRMH14, pICV8), there is as much or more *attI* x *attI* cointegrate junction as there is 5' junction per μg DNA at all timepoints (Fig. 3.4 A-C), suggesting, surprisingly, that in natural levels of integrase there are more *attI* x *attI* cointegrates than *attC1* x *attI* cointegrates (Fig. 3.1 C). This was unexpected since, using a mating-out assay with a strain highly expressing integrase, *attI* x *attI* recombination was found to occur at least 10-fold less often than *attC* x *attI* recombination (14). The high integrase strain (Fig 3.4 D-F) also has more copies of the cointegrate junction than of the 5' junction during log phase (at 16, 20, and 24 hours), but has similar amounts of each junction or more 5' junction in stationary phase (>24 hours). When both 5' and 3' junctions are detectable (in the high integrase strain; Fig. 3.4 D-F), there are similar amounts of 5' and 3' junctions per μg total DNA at 24 hours, but, curiously, the 3' junction becomes more prominent during stationary phase, especially at 48 hours, indicating an increase in pRMH14::pICV8 *attC2* x *attI* cointegrates relative to *attC1* x *attI* cointegrates (Fig. 3.1 C) during stationary phase.

**With excess integrase, the ratios of pICV8 or cells to recombinant junctions decrease greatly from late log phase into stationary phase**

For cells over-expressing integrase (black squares, Fig. 3.5 A-D), on first appearance at 16-20 hours there is approximately one cointegrate per 3,000 plasmids (Fig. 3.5 A, Table 3.5) and per 400-1000 cells (Fig. 3.5 B, Table 3.6). Similarly, early on there is about one 5' junction per 100,000 plasmids (Fig. 3.5 C, Table 3.5) and per 10,000 cells (Fig. 3.5 D, Table 3.6). Both junctions increase relative to the plasmid and the chromosome, and well into stationary phase (48 h) nearly one in 50 plasmids has a cointegrate junction (Fig. 3.5 A, Table 3.5) and one in 30 plasmids a 5' junction (Fig. 3.5 C, Table 3.5). Correspondingly, by stationary phase nearly one in every five cells contains a cointegrate junction (Fig. 3.5 B, Table 3.6) and one in three cells a 5' junction (Fig. 3.5 D, Table 3.6). In cells with natural levels of the integrase (gray diamonds, Fig. 3.5 A-D), the first detected abundances of the cointegrate and 5' junctions are quite like those in the integrase-overexpressing strain; there is one cointegrate junction per ~30,000 plasmids (Fig. 3.5 A, Table 3.5) and per 2,000-10,000 cells (Fig. 3.5 B, Table 3.6) and one 5' junction per 100,000 plasmids (Fig. 3.5 C, Table 3.5) and per 10,000 cells (Fig. 3.5 D, Table 3.6). However, in sharp contrast, in cells naturally expressing integrase, the cointegrate and 5' junctions increase by no more than 5-fold relative to the plasmid and chromsomal replicons when the cells are well into stationary phase (Fig. 3.5 A-D, Tables 3.5, 3.6).

Detection of the 3' junction (Fig. 3.5 E) was only possible in the integrase-overexpressing strain from 24 hours. At this point, the 3' junction is similar in

abundance to the 5' junction at one copy per 50,000 plasmids (Fig. 3.5 E, black dots left axis, Table 3.5) or per 5,000 cells (Fig. 3.5 E, gray triangles right axis, Table 3.6). However, the 3' junctions increase even more sharply than the cointregrate and 5' junctions, and when the cells are well into stationary phase, these initially rare junctions are actually more abundant than the others at one 3' junction per 7 plasmids and two per cell (Fig. 3.5 E, Tables 3.5, 3.6).

The numbers of all three junctions per pICV8 in stationary phase (Table 3.5) are considerably higher than the "cointegration frequencies" (i.e. the number of cointegrate-containing transconjugants per total number of transconjugants) found by the mating-out assay to range from one in 10,000 to one in 100 with high integrase (14, 32), suggesting that conjugation of these recombinant products to a recipient cell does not accurately reflect their steady state concentrations from 36 to 48 hours in the donor cytosol.

**Ratios of pICV8 and of cells to recombinant junctions decrease over generations**

Since the high integrase strain reaches stationary phase at a lower cell density, we wanted to compare junction formation in terms of generations as well as clock time. The viable cell counts for each culture trial (Fig. 3.3) were used to calculate the cumulative generations in each strain up to the onset of stationary phase, and the copies of pICV8 or of *bioB* per recombinant junction were plotted by generations (Fig. 3.6 A-D).

Interestingly, when plotted by generations, the ratio of plasmid and chromosome copies to cointegrate and 5' recombinant junctions in the low integrase strain trends

clearly downward (Fig. 3.6 A-D, gray diamonds). However, in the high integrase strain a similar downward trend in replicons per junction occurs only with the cointegrate junctions (Fig. 3.6 A-B, black squares) and not with the 5' junctions (black squares, Fig. 3.6 C-D). Moreover, unlike cointegrate formation where the low integrase strain has 10-fold fewer events per replicon than the high integrase strain (Fig. 3.6 A-B), for the 5' junction the ratio of plasmids or chromosomes is very similar when both strains are at the same culture age in log phase (Fig. 3.6 C-D). Prior to the onset of stationary phase, the 3' junction was only observable in the high integrase strain at the 24 hour timepoint, so its change per generation cannot be calculated.

Thus, when viewed from the perspective of cell generations (divisions), for the low integrase strain the ratios of replicons to the cointegrate and 5' junctions both decrease at similar rates in each generation (i.e. the junctions per replicon increase). In the high integrase strain, the replicons per cointegrate junction also decrease in each generation, but this is not the case for the 5' junctions. In addition, in midlog, the low integrase strain has ~10-fold more replicons per cointegrate junction than the high integrase strain (Fig. 3.6 A-B), but their replicon-to-5' junction values are indistinguishable (Fig. 3.6 C-D); the replicon-to-5' junction ratio for the low integrase strain is within 2-fold of the replicon-to-cointegrate ratio for that strain—i.e. ca. $10^4$-$10^5$ per junction for the plasmid (Table 3.5) and between 1000 and 10,000 chromosomes per junction (Table 3.6). However, for the high integrase strain, the replicon-to-5' junction ratio is 10- to 50-fold higher—i.e. 5' junctions are more rare than *attI* x *attI* cointegrates.

## Discussion

### Complete cassette capture is only detectable with high intracellular integrase

Within the limits of the assay system employed here, cointegrate formation by pICV8 is readily detectable when natural levels of the integrase are provided by the cassette donor plasmid pRMH14, but cassette acquisition (as reported by the 3' junction amplicand (Fig. 3.1 B)) is only detectable with much higher than natural integrase expression. Unfortunately, to design a 3' primer that would not also pick up regions of the cassette donor plasmid pRMH14 required a relatively large amplicand with a high detection limit (Table 3.2), possibly due to secondary structure formed by the *aadA2* *attC* site. In the strain with natural integrase expression, the cointegrate and 5' junctions remain relatively flat during the culture cycle and are at their most abundant at 36 hours (Fig. 3.4, Table 3.3). The 3' junctions, though not detectable in the low integrase strain, are at least 5-fold less abundant than 5' junctions and at least 14-fold less abundant than cointegrate junctions at 36 hours (Table 3.3). Thus, pRMH14::pICV8 *attI* x *attI* cointegrates (as reported by the cointegrate junction) and *attC1* x *attI* cointegrates (as reported by the 5' junction) are, respectively, at least 14-fold and 5-fold more abundant than the total of all three recombination products that give a 3' junction: pICV8::*aadA2*, pICV8::*aacA4aadA2*, and *attC2* x *attI* cointegrates (Fig. 3.1 B, C). Given the low detectability of the 3' junction, we cannot say that natural levels of integrase do not effect direct cassette capture, but we can say that pRMH14::pICV8 *attI* x *attI* and *attC1* x *attI* cointegrates are more abundant than direct capture products in stationary phase, suggesting that cointegrate formation is not a rare event as has been suggested (12)

and may be more important in natural ecosystems than has been assumed. However, the amounts of pRMH14::pICV8 *attI* x *attI* and *attC1* x *attI* cointegrates per cell (shown respectively by the ratios of cointegrate and 5' junctions per cell) seen in the low integrase strain (only one present in 1000-10,000 cells) are low enough to be difficult to detect using a mating-out assay, consistent with only a few cointegrates being isolated by that method under natural integrase expression conditions (12, 32, 33).

In contrast, overexpression of the integrase does produce detectable 3' junctions, either by integration of the gene cassette(s) or by resolution of the very abundant cointegrates also produced under these conditions. Also, quite strikingly, unlike the natural integrase expression strain, the high integrase strain continues producing increasing amounts of all three junction types well into stationary phase (Fig. 3.4 D-F). Intracellular recombinase concentration affects recombination reactions for other tyrosine recombinases, as well. The recombinase Cre resolves dimers of the P1 prophage to ensure stable inheritance, but high concentrations of Cre destabilize P1, possibly by causing multimers to form rather than be resolved (41). This suggests that high concentrations of recombinase can overwhelm the regulatory mechanisms that normally prevent Cre-mediated intercellular recombination. In addition, for the lambda bacteriophage, the integrase concentation relative to concentration of excisionase (Xis) regulates the recombination choice between integration and excision (26). Thus, high concentrations of integron integrase might affect reaction specificity as well as the frequency and type of recombination (cointegrate formation/resolution or cassette excision/insertion). The increased recombination products seen in the higher integrase

strains likely indicates that high integrase concentrations overwhelm the putative natural regulatory mechanisms that prevent rampant recombination. Clearly regulation of integron integrase expression and enzyme activity require further study.

**Recombination site choices are affected by integrase concentration**

In newly transformed CB454(pRMH14, pICV8) and CB454(pRMH14, pICV8, pSU2056) strains, the growth lag and lower cell density in stationary phase of the high integrase strain indicates that pSU2056 interferes with cell growth (Fig. 3.3), owing either simply to its high copy number or to the integrase recombination activity or both. Strains containing pSU2056 are known to be unstable, possibly due to the insertion of gene cassettes into secondary sites in the chromosome (14).

Recombination products were not detectable in either strain until at least 16 hours (Fig. 3.4, Table 3.3). That recombinant junctions were not detected at early timepoints in the high integrase strain may arise more from its slow growth and consequently sparse DNA yield than from actual absence of recombinant junctions (Table 3.3, gray values).

In the low integrase strain the 1380 bp amplicand reporting the pRMH14::pICV8 *attI* x *attI* cointegrate junction is as abundant or more abundant than the 1276 bp 5' junction amplicand at all timepoints (Figs. 3.1 B, C and 3.4 A-C) despite the latter amplicand having a lower detection limit (Table 3.2). This is unexpected since prior work (14) shows a 10-fold lower frequency for *attI* x *attI* recombination than for *attC* x *attI* recombination. However, those experiments measured recombination by

conjugation from high integrase donors in stationary phase, an additional step which may not accurately reflect the population of recombinant structures in the donor strain. However, it is possible that integrase concentration influences the recombination crossover site choices resulting in a preference for *attI* x *attI* recombination over *attC* x *attI* in the low integrase strain.

The high integrase strain also has more cointegrate junctions than 5' or 3' junctions during log phase, but similar or more 5' junctions than cointegrate junctions during stationary phase (Fig. 3.4 D-F), suggesting that *attI* x *attI* recombination is also preferred over *attC* x *attI* recombination with overexpressed integrase in log phase. The replicon-to-5' junction ratio is 10- to 50-fold higher than the replicon-to-cointegrate junction in the high integrase strain in log phase (Fig. 3.6 A-D, Tables 3.5, 3.6); thus, there are more *attI* x *attI* cointegrates per cell (and per plasmid) than *attC1* x *attI* cointegrates or pICV::*aadA2* products (Fig. 3.1 B, C). This difference in ratios between the 5' and cointegrate junctions may be explained by an *attI* x *attI* recombination crossover preference during log phase. *attI* x *attI* recombination would increase the amount of cointegrate junctions, and, if *attC1* x *attI* cointegrates were biased to resolve by *attI* x *attI* recombination, it would decrease the amounts of 5' and 3' junctions, as well (Fig. 3.7).

In the high integrase strain (Fig. 3.4 D-F), we cannot distinguish whether cassette capture products or cointegrates generated the 5' and 3' junction amplicands because observing both junctions indicates that the 5' junctions are a mix of pICV8::*aadA2* cassette capture products and pRMH14::pICV8 *attC1* x *attI* cointegrates and the 3'

junctions are a mix of pICV8::*aadA2*, pICV8::*aacA4aadA2*, and pRMH14::pICV8 *attC2* x *attI* cointegrates (Figs. 3.1 B, C). It is especially notable, however, that in late stationary phase (48 h) the 3' junctions occur at more than one per 10 plasmids and more than one per chromosome (Fig. 3.5 E, Tables 3.5, 3.6), more abundant than either the 5' (Fig. 3.5 C-D) or cointegrate junctions (Fig. 3.5 A-B). In the high integrase strain, the 5' and 3' junctions are equally abundant at 24 hours, but there are more 3' junctions by 48 hours (Fig. 3.4 D-F), indicating that *attC2* x *attI* cointegrates and cassette capture at the *attC2* site both increase during stationary phase and that crossover position choices are influenced by the culture cycle in high integrase conditions.

**Integrase concentration most affects the amount and type of recombination products per cell at the end of log phase**

In the low integrase strain, the ratios of 5' and cointegrate junctions per pICV8 plasmid (or per chromosome) peak during late log phase and plateau or decrease after 36 hours (Fig. 6 A-D). This suggests some regulation of integrase-mediated recombination during different growth phases, with *attI* x *attI* and *attC1* x *attI* cointegrate (Fig. 3.1 C) recombination products becoming most abundant during late log and early stationary phase. The low detectability of the 3' junction, which indicates cassette capture products resulting from cointegrate resolution and/or cassette insertion, limited conclusions about whether these processes also vary during the growth cycle under natural integrase expression.

Similarly, for the high integrase strain, the most rapid junction increase occurs at the end of late log phase (between 24 and 36 hours), particularly for 5' and 3' junctions (Fig. 3.4 D-F and Fig. 3.5 C-E) and then begins to plateau after 36 hours. Therefore, even in this artificial expression system the end of log phase marks a sharp increase in all integrase-mediated recombination products.

Growth phase-dependent effects on recombination could result from differential regulation of integrase activity (transcriptional, translational, or allosteric) at different stages of culture growth. Some transposons increase movement in response to changes in nutrient availability and/or temperature in order to induce adaptive changes that aid under these stressful conditions (5, 24, 53). In bacteria, for example, the Tn*4652* transposase is transcriptionally upregulated during stationary phase by RpoS (24). The upregulation of integron gene cassette expression could also affect recombination. Strength of the cassette promoter affects recombination frequency in class 1 integrons, possibly by RNA polymerase physically blocking recombination (13, 14), or, alternatively, the convergence arrangement of the cassette and integrase promoters could also mean that an increase in cassette expression causes a decrease in integrase expression. Gene cassettes in the *Vibrio cholerae* chromosomal integron are positively regulated by RpoS, which indicates an increase in expression during stationary phase (55), though it is not known whether recombination frequencies are affected by growth cycle in that system.

Growth phase-dependent regulation of integrase host cofactors could also affect recombination rates in different growth phases. Most site-specific recombinases use

host cofactors like FIS and/or IHF in bending DNA to form a recombination synapse (22). FIS is upregulated in log phase but is nearly undetectable in stationary phase (3, 39). Though it is not yet known whether integron integrases use host cofactors, the widespread use of these accessory proteins by other recombinases suggests it is a reasonable possibility.

**Summary**

In summary, the direct quantification of integrase recombination products reveals that in a strain with natural integrase expression, *attI* x *attI* recombination is, surprisingly, preferred over *attC* x *attI* recombination throughout the growth cycle and most especially from late log to early stationary phase. In addition, *attI* x *attI* recombination crossovers are more common than *attC* x *attI* crossovers in log phase in a strain overexpressing integrase, but *attC* recombination increases in stationary phase. Thus, integron-integrase recombination site choices are affected by intracellular integrase concentrations and by growth phase. With either natural levels or overexpressed integrase, all detectable recombination products increase during late log phase, and the most dramatic increase in recombination products occurs from late log to early stationary phase with high integrase, also indicating some growth-phase dependent regulation of recombination. Due to the low detectability of the 3' junction, we cannot conclude that cassette capture does not occur in the low integrase strain, but cointegrates are clearly more abundant than any possible cassette capture products,

suggesting that cointegrate formation is not a rare event and may be more important in natural ecosystems than previously imagined.

**References**


1. **Aussel, L., F.-X. Barre, M. Aroyo, A. Stasiak, A. Z. Stasiak, and D. Sherratt.** 2002. FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. Cell **108:**195-205.
2. **Ausubel, F. A., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1988. Current Protocols in Molecular Biology. John Wiley & Sons, New York, N. Y.
3. **Ball, C. A., R. Osuna, K. C. Ferguson, and R. C. Johnson.** 1992. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. J. Bacteriol. **174:**8043-8056.
4. **Blakely, G., G. May, R. McCulloch, L. K. Arciszewska, M. Burke, S. T. Lovett, and D. J. Sherratt.** 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. Cell **75:**351-361.
5. **Capy, P., G. Gasperi, C. Biemont, and C. Bazin.** 2000. Stress and transposable elements: co-evolution or useful parasites? Heredity **85 ( Pt 2):**101-6.

6. **Chang, A. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:**1141-1156.

7. **Chang, C. Y., L. L. Chang, Y. H. Chang, T. M. Lee, and S. F. Chang.** 2000. Characterisation of drug resistance gene cassettes associated with class 1 integrons in clinical isolates of *Escherichia coli* from Taiwan, ROC. J Med Microbiol **49:**1097-102.

8. **Clark, N. C., O. Olsvik, J. M. Swenson, C. A. Spiegel, and F. C. Tenover.** 1999. Detection of a streptomycin/spectinomycin adenylyltransferase gene (*aadA*) in *Enterococcus faecalis*. Antimicrob. Agents Chemother. **43:**157-60.

9. **Collis, C., and R. Hall.** 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. Antimicrob. Agents Chemother. **39:**155-162.

10. **Collis, C. M., G. Grammaticopoulos, J. Briton, H. W. Stokes, and R. M. Hall.** 1993. Site-specific insertion of gene cassettes into integrons. Mol. Microbiol. **9:**41-52.

11. **Collis, C. M., and R. M. Hall.** 1992. Gene cassettes from the insert region of integrons are excised as covalently closed circles. Mol. Microbiol. **6:**2875-85.

12. **Collis, C. M., and R. M. Hall.** 1992. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. J. Bacteriol. **174:**1574-85.

13. **Collis, C. M., M. J. Kim, H. W. Stokes, and R. M. Hall.** 2002. Integron-encoded IntI integrases preferentially recognize the adjacent cognate *attI* site in recombination with a 59-be site. Mol. Microbiol. **46:**1415-27.

14. **Collis, C. M., G. D. Recchia, M. J. Kim, H. W. Stokes, and R. M. Hall.** 2001. Efficiency of recombination reactions catalyzed by class 1 integron integrase IntI1. J. Bacteriol. **183:**2535-42.

15. **Davies, J.** 1994. Inactivation of antibiotics and the dissemination of resistance genes. Science **264:**375-82.

16. **Davison, J., M. Heusterspreute, N. Chevalier, V. Ha-Thi, and F. Brunel.** 1987. Vectors with restriction site banks. V. pJRD215, a wide-host-range cosmid vector with multiple cloning sites. Gene **51:**275-80.

17. **Fluit, A. C., and F. J. Schmitz.** 2004. Resistance integrons and super-integrons. Clinical Microbiology and Infection **10:**272-288.

18. **Goldstein, C., M. D. Lee, S. Sanchez, C. Hudson, B. Phillips, B. Register, M. Grady, C. Liebert, A. O. Summers, D. G. White, and J. J. Maurer.** 2001. Incidence of class 1 and 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. Antimicrob. Agents Chemother. **45:**723-6.

19. **Gourlay, S. C., and S. D. Colloms.** 2004. Control of Cre recombination by regulatory elements from Xer recombination systems. Molecular Microbiology **52:**53-65.

20. **Hall, R. M., D. E. Brookes, and H. W. Stokes.** 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Mol. Microbiol. **5:**1941-59.

21. **Hall, R. M., and H. W. Stokes.** 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica **90:**115-32.

22. **Hallet, B., and D. J. Sherratt.** 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. FEMS Microbiology Reviews **21:**157-178.

23. **Hansson, K., O. Skold, and L. Sundstrom.** 1997. Non-palindromic *attI* sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. **26:**441-53.

24. **Ilves, H., R. Horak, and M. Kivisaar.** 2001. Involvement of $\sigma^S$ in starvation-induced transposition of *Pseudomonas putida* transposon Tn*4652*. J. Bacteriol. **183:**5445-8.

25. **Jacoby, G. A., K. E. Walsh, D. M. Mills, V. J. Walker, H. Oh, A. Robicsek, and D. C. Hooper.** 2006. *qnrB*, another plasmid-mediated gene for quinolone resistance. Antimicrob. Agents Chemother. **50:**1178-1182.

26. **Leffers, G. G., Jr., and S. Gottesman.** 1998. Lambda Xis degradation *in vivo* by Lon and FtsH. J. Bacteriol. **180:**1573-1577.

27. **Leverstein-van Hall, M. A., A. T. Box, H. E. Blok, A. Paauw, A. C. Fluit, and J. Verhoef.** 2002. Evidence of extensive interspecies transfer of integron-mediated

antimicrobial resistance genes among multidrug-resistant Enterobacteriaceae in a clinical setting. J Infect Dis **186:**49-56.

28. **Leverstein-van Hall, M. A., M. B. HE, T. D. AR, A. Paauw, A. C. Fluit, and J. Verhoef.** 2003. Multidrug resistance among Enterobacteriaceae is strongly associated with the presence of integrons and is independent of species or isolate origin. J. Infect. Dis. **187:**251-9.

29. **Leverstein-Van Hall, M. A., A. Paauw, A. T. Box, H. E. Blok, J. Verhoef, and A. C. Fluit.** 2002. Presence of integron-associated resistance in the community is widespread and contributes to multidrug resistance in the hospital. J Clin Microbiol **40:**3038-40.

30. **Levesque, C., L. Piche, C. Larose, and P. H. Roy.** 1995. PCR mapping of integrons reveals several novel combinations of resistance genes. Antimicrob. Agents Chemother. **39:**185-91.

31. **Liebert, C. A., R. M. Hall, and A. O. Summers.** 1999. Transposon Tn*21*, flagship of the floating genome. Microbiol. Mol. Biol. Rev. **63:**507-522.

32. **Martinez, E., and F. de la Cruz.** 1990. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. **9:**1275-81.

33. **Martinez, E., and F. de la Cruz.** 1988. Transposon Tn*21* encodes a RecA-independent site-specific integration system. Mol. Gen. Genet. **211:**320-5.

34. **Mazel, D., and J. Davies.** 1999. Antibiotic resistance in microbes. Cell Mol. Life Sci. **56:**742-54.

35. **Mazel, D., B. Dychinco, V. A. Webb, and J. Davies.** 2000. Antibiotic resistance in the ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob. Agents Chemother. **44:**1568-74.

36. **Nandi, S., J. J. Maurer, C. Hofacre, and A. O. Summers.** 2004. Gram-positive bacteria are a major reservoir of Class 1 antibiotic resistance integrons in poultry litter. Proc. Natl. Acad. Sci. USA **101:**7118-7122.

37. **Nesvera, J., J. Hochmannova, and M. Patek.** 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. FEMS Microbiol. Lett. **169:**391-5.

38. **Nield, B. S., A. J. Holmes, M. R. Gillings, G. D. Recchia, B. C. Mabbutt, K. M. H. Nevalainen, and H. W. Stokes.** 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. **195:**59-65.

39. **Osuna, R., D. Lienau, K. T. Hughes, and R. C. Johnson.** 1995. Sequence, regulation, and functions of fis in Salmonella typhimurium. J. Bacteriol. **177:**2021-2032.

40. **Partridge, S. R., C. M. Collis, and R. M. Hall.** 2002. Class 1 integron containing a new gene cassette, *aadA10*, associated with Tn*1404* from R151. Antimicrob. Agents Chemother. **46:**2400-8.

41. **Paul, S., and D. Summers.** 2004. ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. Plasmid **52:**63-68.

42. **Recchia, G. D., H. W. Stokes, and R. M. Hall.** 1994. Characterisation of specific and secondary recombination sites recognised by the integron DNA integrase. Nucleic Acids Res. **22:**2071-8.

43. **Robertson, C. A., and H. A. Nash.** 1988. Bending of the bacteriophage lambda attachment site by *Escherichia coli* integration host factor. J. Biol. Chem. **263:**3554-3557.

44. **Sandvang, D.** 1999. Novel streptomycin and spectinomycin resistance gene as a gene cassette within a class 1 integron isolated from *Escherichia coli*. Antimicrob. Agents Chemother. **43:**3036-8.

45. **Sandvang, D., and F. M. Aarestrup.** 2000. Characterization of aminoglycoside resistance genes and class 1 integrons in porcine and bovine gentamicin-resistant *Escherichia coli*. Microb Drug Resist **6:**19-27.

46. **Stokes, H. W., and R. M. Hall.** 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol. Microbiol. **3:**1669-83.

47. **Stokes, H. W., A. J. Holmes, B. S. Nield, M. P. Holley, K. M. Nevalainen, B. C. Mabbutt, and M. R. Gillings.** 2001. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. Appl. Environ. Microbiol. **67:**5240-6.

48. **Stokes, H. W., C. Tomaras, Y. Parsons, and R. M. Hall.** 1993. The partial 3'-conserved segment duplications in the integrons In6 from pSa and In7 from pDGO100 have a common origin. Plasmid **30:**39-50.

49. **Summers, A. O.** 2006. Genetic linkage and horizontal gene transfer, the roots of the antibiotic *multi*-resistance problem. Animal Biotechnology **17:**125 - 135.

50. **Tosini, F., P. Visca, I. Luzzi, A. M. Dionisi, C. Pezzella, A. Petrucca, and A. Carattoli.** 1998. Class 1 integron-borne multiple-antibiotic resistance carried by IncFI and IncL/M plasmids in *Salmonella enterica* serotype Typhimurium. Antimicrob. Agents Chemother. **42:**3053-8.

51. **Tran, J. H., and G. A. Jacoby.** 2002. Mechanism of plasmid-mediated quinolone resistance. Proceedings of the National Academy of Sciences **99:**5638-5642.

52. **Watanabe, T., C. Furuse, and S. Sakaizumi.** 1968. Transduction of various R Factors by phage P1 in *Escherichia coli* and by phage P22 in *Salmonella typhimurium*. J. Bacteriol. **96:**1791-1795.

53. **Wessler, S. R.** 1996. Turned on by stress. Plant retrotransposons. Curr Biol **6:**959-61.

54. **White, P. A., C. J. McIver, and W. D. Rawlinson.** 2001. Integrons and gene cassettes in the Enterobacteriaceae. Antimicrob Agents Chemother **45:**2658-61.

55. **Yildiz, F. H., X. S. Liu, A. Heydorn, and G. K. Schoolnik.** 2004. Molecular analysis of rugosity in a *Vibrio cholerae* O1 El Tor phase variant. Molecular Microbiology **53:**497-515.

56. **Yin, S., W. Bushman, and A. Landy.** 1985. Interaction of the lambda site-specific recombination protein Xis with attachment site DNA. PNAS **82:**1040-1044.

57. **Zuhlsdorf, M. T., and B. Wiedemann.** 1992. Tn*21*-specific structures in gram-negative bacteria from clinical isolates. Antimicrob. Agents Chemother. **36:**1915-21.

**Table 3.1.** Phenotypes of plasmids

| Plasmid | Description | Phenotype | Reference |
|---|---|---|---|
| pRMH14 (~33,600 bp) | Derived from pSa (52), carries integron with *aacA4* and *aadA2*, IncW | KmR, SmR, SuR, $Tra^+$, $IntI1^+$ | (48) |
| pICV8 (7584 bp) | pJRD215 (16) with partial Tn*21* (31) integron (*intI1*Δ, *attI*, *aadA1*Δ) and *zeoR* from pEM7/Zeo (Invitrogen) | ZeoR, $Mob^+$ | this work (see Chapter 2) |
| pICV8::*aadA2* (8440 bp) | pICV8 with *aadA2* cassette from pRMH14 | ZeoR, SmR | this work (see Chapter 2) |
| pICVm (5917 bp) | pACYC184 (6) with integron and *zeoR* region from pICV8 (*NarI* to *BamHI)* | ZeoR, TcR, $Mob^-$ | this work |
| pRMH14::pICVm *attI*x*attI* (~39,517 bp) | *attI* x *attI* cointegrate between pRMH14 and pICVm | ZeoR, TcR, KmR, SmR, SuR, $Tra^+$, $IntI1^+$ | this work |
| pSU2056 (3837 bp) | pUC9 expressing IntI1 from Tn*21* under $P_{lac}$ | ApR, $IntI1^+$ | (32) |

**Table 3.2.** PCR primer sequences

| **Target (product size)** | **Primer names and sequences** | **Detection limit**[†] (#copies± SD%) |
|---|---|---|
| chromosome (511 bp) | bioB-U1: 5' TTACGGCAATATCATCACCAC 3'<br>bioB-L1: 5' CTGTTCAAGACGTTGCTGTTG 3' | 1662 ± 41.4% |
| pICV8 backbone (507 bp) | ICV8-U1: 5' CAATGCGCTCATCGTCATCCTC 3'<br>ICV8-L1: 5' GACCACACCGGCGAAGTCGTC 3' | 718.1 ± 56.5% |
| cointegrate junction (1380 bp) | 8E-U1: 5' CCTCGTTAAAGGACAAGGACCTGAG 3'<br>int-L1: 5' CGCGCTGAAAGGTCTGGTCATAC 3' | 280.5 ± 56.0% |
| 5' junction (1276 bp) | aadA2-U1: 5' GCCGGTTATTGCGCTGTACCAAATG 3'<br>4E-L2: 5' GCCTATGCCTACAGCATCCAGGGTGAC 3' | 100.9 ± 85.3% |
| 3' junction (1000 bp) | 8E-U1: 5' CCTCGTTAAAGGACAAGGACCTGAG 3'<br>aadA2-L2: 5' GCGAGCTGCAATTTGGAGAATGG 3' | 12342 ± 21.2% |

[†] Detection limits were determined by averaging six separate dilution PCR trials each done in duplicate using extracted pure chromosomal or plasmid DNA and are based on the number of copies per 50 μl reaction that gave a visible band when 30 μl was run on a gel (see Methods). CB454 chromosomal DNA was used for *bioB* detection, and pICV8 was used for the pICV8 backbone replicon detection. pRMH14::pICVm *attI* x *attI* cointegrate plasmid DNA was used to detect the 1380 bp cointegrate junction (see Fig. 3.1 C), and pICV8::*aadA2* was used for detection of both 5' and 3' junctions (see Fig. 3.1 B-C). Only the smallest amplicand (size shown above) was quantified for each primer pair (see Fig. 3.1).

**Table 3.3.** Number of recombinant junctions per μg total DNA[a] (data plotted in Fig. 3.4)

| **Cointegrate junctions / μg total DNA** (± % error) | | | | | | |
|---|---|---|---|---|---|---|
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 561) | (< 561) | (< 561) | (< 561) | (< 561) | (< 561) |
| **12 h** | (< 565) | (< 583) | (< 569) | (< 999) | (< 1.04E+03) | (< 1.03E+03) |
| **14 h** | ND[b] | (< 561) | (< 608) | (< 2.36E+03) | (< 2.68E+03) | (< 1.73E+03) |
| **16 h** | 2.99E+03 (±66.9%) | (< 561) | 2.69E+03 (±64.0%) | (< 1.39E+03) | (< 1.52E+03) | 1.96E+04 (±90.0%) |
| **20 h** | 6.73E+03 (±68.1%) | 7.18E+03 (±74.4%) | 5.39E+03 (±64.0%) | 2.70E+04 (±68.1%) | 2.40E+04 (±66.9%) | 3.60E+04 (±78.3%) |
| **24 h** | 3.60E+04 (±78.3%) | 2.16E+04 (±64.0%) | 3.37E+03 (±79.5%) | 1.02E+05 (±73.0%) | 1.57E+05 (±107.8%) | 2.87E+05 (±78.1%) |
| **36 h** | 8.63E+04 (±64.0%) | 7.85E+04 (±60.3%) | 1.27E+04 (±73.1%) | 4.49E+06 (±78.3%) | 4.68E+06 (±101.6%) | 6.73E+05 (±64.0%) |
| **48 h** | 1.96E+04 (±84.9%) | 1.96E+04 (±60.3%) | 1.54E+04 (±81.1%) | 8.29E+06 (±88.1%) | 5.08E+07 (±113.2%) | 2.54E+07 (±73.1%) |
| **5' junctions / μg total DNA** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 202) | (< 202) | (< 202) | (< 202) | (< 202) | (< 202) |
| **12 h** | (< 203) | (< 210) | (< 205) | (< 359) | (< 376) | (< 369) |
| **14 h** | ND[b] | (< 202) | (< 219) | (< 851) | (< 965) | (< 624) |
| **16 h** | 605 (±93.7%) | (< 202) | 969 (±90.8%) | (< 501) | (< 545) | (< 622) |
| **20 h** | 2.04E+03 (±95.1%) | 2.15E+03 (±92.8%) | 2.77E+03 (±92.2%) | (< 202) | 3.88E+03 (±90.8%) | 1.08E+03 (±123.6%) |
| **24 h** | 1.11E+04 (±92.2%) | 9.70E+03 (±93.7%) | 2.42E+03 (±93.7%) | 1.94E+03 (±90.8%) | 4.56E+03 (±97.4%) | 3.88E+03 (±90.8%) |
| **36 h** | 8.17E+03 (±95.1%) | 1.11E+04 (±92.2%) | 4.85E+03 (±102.3%) | 4.31E+06 (±92.8%) | 5.17E+06 (±98.4%) | 2.77E+06 (±103.5%) |
| **48 h** | 7.76E+03 (±90.8%) | 4.56E+03 (±97.4%) | 7.76E+03 (±90.8%) | 2.59E+07 (±85.3%) | 5.64+07 (±106.4%) | 3.45E+07 (±92.8%) |
| **3' junctions / μg total DNA** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 9.16E+03) | (< 7.34E+03) | (< 9.59E+03) | (< 5.62E+03) | (< 6.20E+03) | (< 1.28E+04) |
| **12 h** | (< 2.48E+04) | (< 2.57E+04) | (< 2.50E+04) | (< 4.39E+04) | (< 4.59E+04) | (< 4.51E+04) |
| **14 h** | ND[b] | (< 2.10E+04) | (< 2.68E+04) | (< 1.04E+05) | (< 1.18E+05) | (< 7.63E+04) |
| **16 h** | (< 1.47E+04) | (< 2.16E+04) | (< 1.47E+04) | (< 6.13E+04) | (< 6.67E+04) | (< 7.61E+04) |
| **20 h** | (< 2.12E+03) | (< 8.75E+03) | (< 6.20E+03) | (< 2.44E+04) | (< 4.11E+04) | (< 4.58E+03) |
| **24 h** | (< 504) | (< 1.45E+03) | (< 1.29E+03) | 2.33E+03 (±21.2%) | 5.44E+03 (±21.2%) | 1.14E+04 (±44.2%) |
| **36 h** | (< 1.02E+03) | (< 1.38E+03) | (< 870) | 8.42E+06 (±42.0%) | 1.51E+07 (±37.5%) | 5.84E+06 (±50.0%) |
| **48 h** | (< 1.24E+03) | (< 1.71E+03) | (< 1.15E+03) | 1.58E+08 (±53.4%) | 1.48E+08 (±44.2%) | 1.08E+08 (±30.7%) |

[a] Gray values indicate when no junction amplicand was detected and show the theoretical possible number of junctions per μg DNA below the limit of detection (see Methods)

[b] ND=not done

**Table 3.4.** Copies of pICV8 or *bioB* per μg total DNA (data plotted in Fig. 3.4)

| **Copies pICV8 / μg total DNA** (± % error) | | | | | | |
|---|---|---|---|---|---|---|
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | 6.16E+04 (±81.4%) | 1.08E+05 (±68.5%) | 4.41E+05 (±91.8%) | 4.31E+05 (±68.5%) | 3.83E+05 (±67.3%) | 1.91E+05 (±67.3%) |
| **12 h** | 3.45E+06 (±92.0%) | 1.81E+06 (±98.7%) | 6.89E+06 (±64.4%) | 1.72E+06 (±64.4%) | 3.45E+06 (±64.4%) | 3.83E+06 (±67.3%) |
| **14 h** | ND[a] | 3.45E+06 (±64.4%) | 3.45E+07 (±68.6%) | 5.30E+06 (±72.5%) | 8.62E+06 (±68.5%) | 3.07E+07 (±67.3%) |
| **16 h** | 5.74E+07 (±74.8%) | 3.13E+07 (±85.2%) | 1.23E+08 (±67.3%) | 1.06E+07 (±72.5%) | 1.53E+07 (±67.3%) | 5.52E+07 (±64.4%) |
| **20 h** | 1.01E+08 (±73.4%) | 4.31E+07 (±68.5%) | 1.58E+08 (±66.5%) | 5.52E+07 (±64.4%) | 6.14E+07 (±67.3%) | 1.10E+08 (±64.4%) |
| **24 h** | 2.15E+08 (±68.5%) | 7.83E+07 (±60.7%) | 2.21E+08 (±64.4%) | 1.84E+08 (±78.7%) | 1.23E+08 (±67.3%) | 4.41E+08 (±91.8%) |
| **36 h** | 3.13E+08 (±85.2%) | 2.65E+08 (±72.5%) | 3.15E+08 (±81.3%) | 1.77E+08 (±93.9%) | 4.90E+08 (±67.1%) | 4.90E+08 (±67.1%) |
| **48 h** | 2.30E+08 (±56.5%) | 9.57E+07 (±67.3%) | 2.76E+08 (±68.5%) | 1.10E+09 (±68.5%) | 8.01E+08 (±85.0%) | 9.77E+08 (±85.0%) |
| **Copies *bioB* / μg total DNA** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | 9.39E+06 (±62.6%) | 6.94E+06 (±75.1%) | 1.45E+07 (±47.0%) | 1.23E+07 (±79.6%) | 1.99E+07 (±56.7%) | 9.97E+06 (±56.7%) |
| **12 h** | 7.98E+05 (±89.9%) | 9.97E+05 (±70.0%) | 1.77E+06 (±55.2%) | 6.65E+06 (±41.4%) | 9.97E+06 (±56.7%) | 3.32E+06 (±68.7%) |
| **14 h** | ND[a] | 2.22E+06 (±98.6%) | 3.55E+06 (±55.2%) | 1.60E+06 (±51.7%) | 1.16E+07 (±76.1%) | 7.98E+06 (±56.8%) |
| **16 h** | 9.97E+06 (±56.7%) | 1.05E+07 (±59.0%) | 3.55E+07 (±55.2%) | 1.99E+06 (±56.7%) | 7.09E+06 (±55.3%) | 1.83E+07 (±54.2%) |
| **20 h** | 1.11E+07 (±55.2%) | 1.33E+07 (±64.1%) | 7.60E+06 (±54.2%) | 1.56E+07 (±47.5%) | 8.86E+06 (±55.2%) | 2.90E+07 (±47.0%) |
| **24 h** | 4.99E+07 (±56.7%) | 1.77E+07 (±55.2%) | 1.36E+07 (±95.8%) | 1.99E+07 (±70.0%) | 2.66E+07 (±41.4%) | 4.26E+07 (±64.2%) |
| **36 h** | 6.14E+07 (±61.4%) | 2.85E+07 (±71.7%) | 3.99E+07 (±70.0%) | 2.66E+07 (±41.4%) | 1.99E+07 (±70.0%) | 1.77E+07 (±55.2%) |
| **48 h** | 9.97E+07 (±51.7%) | 2.90E+07 (±76.0%) | 7.09E+07 (±55.3%) | 1.16E+08 (±47.0%) | 1.28E+08 (±51.7%) | 3.55E+07 (±55.2%) |

[a] ND=not done

**Table 3.5.** Number of recombinant junctions per copy of pICV8 backbone[a] (data plotted in Fig. 3.5 A, C, E)

| **Cointegrate junctions / pICV8** (± % error) | | | | | | |
|---|---|---|---|---|---|---|
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 9.11E-03) | (< 5.21E-03) | (< 1.27E-03) | (< 1.30E-03) | (< 1.46E-03) | (< 2.93E-03) |
| **12 h** | (< 1.64E-04) | (< 3.21E-04) | (< 8.26E-05) | (< 5.79E-04) | (< 3.03E-04) | (< 2.68E-04) |
| **14 h** | ND[b] | (< 1.63E-04) | (< 1.76E-05) | (< 4.46E-04) | (< 3.11E-04) | (< 5.65E-05) |
| **16 h** | 5.21E-05 (±100.3%) | (< 1.79E-05) | 2.19E-05 (±92.9%) | (< 1.30E-03) | (< 1.46E-03) | 3.55E-04 (±110.7%) |
| **20 h** | 6.64E-05 (±100.1%) | 1.67E-04 (±101.1%) | 3.41E-05 (±92.3%) | 4.89E-04 (±93.7%) | 3.91E-04 (±94.8%) | 3.25E-04 (±101.4%) |
| **24 h** | 1.67E-04 (±104.1%) | 2.75E-04 (±88.2%) | 1.52E-05 (±102.4%) | 5.51E-04 (±107.3%) | 1.28E-03 (±127.1%) | 6.50E-04 (±120.6%) |
| **36 h** | 2.75E-04 (±106.6%) | 2.96E-04 (±94.3%) | 4.02E-05 (±109.3%) | 0.0254 (±122.3%) | 9.55E-03 (±121.8%) | 1.37E-03 (±92.8%) |
| **48 h** | 8.53E-05 (±102.0%) | 2.05E-04 (±90.3%) | 5.57E-05 (±106.1%) | 7.54E-03 (±111.6%) | 0.0634 (±141.5%) | 0.0260 (±99.3%) |
| **5' junctions / pICV8** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 3.28E-03) | (< 1.87E-03) | (< 4.58E-04) | (< 4.68E-04) | (< 5.27E-04) | (< 1.05E-03) |
| **12 h** | (< 5.89E-05) | (< 1.16E-04) | (< 2.97E-05) | (< 2.08E-04) | (< 1.09E-04) | (< 9.63E-05) |
| **14 h** | ND[b] | (< 5.85E-05) | (< 6.34E-06) | (< 1.60E-04) | (< 1.12E-04) | (< 2.03E-05) |
| **16 h** | 1.05E-05 (±119.9%) | (< 6.44E-06) | 7.89E-06 (±113.0%) | (< 4.72E-05) | (< 3.56E-05) | (< 1.13E-05) |
| **20 h** | 2.01E-05 (±120.1%) | 5.00E-05 (±115.3%) | 1.75E-05 (±113.7%) | (< 3.65E-06) | 6.31E-05 (±113.0%) | 9.74E-06 (±139.4%) |
| **24 h** | 5.15E-05 (±114.9%) | 1.24E-04 (±111.6%) | 1.10E-05 (±113.7%) | 1.05E-05 (±120.1%) | 3.71E-05 (±118.4%) | 8.79E-06 (±129.1%) |
| **36 h** | 2.61E-05 (±127.7%) | 4.18E-05 (±117.3%) | 1.54E-05 (±130.7%) | 0.0244 (±132.0%) | 0.0105 (±119.1%) | 5.65E-03 (±123.4%) |
| **48 h** | 3.38E-05 (±106.9%) | 4.76E-05 (±118.4%) | 2.81E-05 (±113.7%) | 0.0235 (±109.4%) | 0.0704 (±136.1%) | 0.0353 (±114.6%) |
| **3' junctions / pICV8** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 0.149) | (< 0.0682) | (< 0.0217) | (< 0.0131) | (< 0.0162) | (< 0.0669) |
| **12 h** | (< 7.21E-03) | (< 0.0141) | (< 3.63E-03) | (< 0.0255) | (< 0.0133) | (< 0.0118) |
| **14 h** | ND[b] | (< 6.09E-03) | (< 7.75E-04) | (< 0.0196) | (< 0.0137) | (< 2.49E-03) |
| **16 h** | (< 2.56E-04) | (< 6.88E-04) | (< 1.20E-04) | (< 5.78E-03) | (< 4.35E-03) | (< 1.38E-03) |
| **20 h** | (< 2.09E-05) | (< 2.03E-04) | (< 3.93E-05) | (< 4.41E-04) | (< 6.70E-04) | (< 4.15E-05) |
| **24 h** | (< 2.34E-06) | (< 1.85E-05) | (< 5.85E-06) | 1.26E-05 (±81.5%) | 4.43E-05 (±70.5%) | 2.58E-05 (±101.9%) |
| **36 h** | (< 3.27E-06) | (< 5.21E-06) | (< 2.76E-06) | 0.0477 (±102.8%) | 0.0308 (±76.9%) | 0.0119 (±83.7%) |
| **48 h** | (< 5.38E-06) | (< 1.78E-05) | (< 4.17E-06) | 0.144 (±86.8%) | 0.185 (±95.8%) | 0.110 (±74.0%) |

[a] Gray values indicate when no junction amplicand was detected and show the theoretical possible number of junctions (see Methods and Table 3.3) per pICV8 backbone

[b] ND=not done

**Table 3.6.** Number of recombinant junctions per copy of *bioB*[a] (data plotted in Fig. 3.5 B, D, E)

| **Cointegrate junctions / *bioB*** (± % error) | | | | | | |
|---|---|---|---|---|---|---|
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 5.98E-05) | (< 8.09E-05) | (< 3.87E-05) | (< 4.57E-05) | (< 2.81E-05) | (< 5.63E-05) |
| **12 h** | (< 7.08E-04) | (< 5.85E-04) | (< 3.21E-04) | (< 1.50E-04) | (< 1.05E-04) | (< 3.08E-04) |
| **14 h** | ND[b] | (< 2.53E-04) | (< 1.71E-04) | (< 1.48E-03) | (< 2.31E-04) | (< 2.17E-04) |
| **16 h** | 3.00E-04 (±87.7%) | (< 5.34E-05) | 7.59E-05 (±84.5%) | (< 6.98E-04) | (< 2.14E-04) | 1.07E-03 (±105.1%) |
| **20 h** | 6.08E-04 (±87.7%) | 5.40E-04 (±98.2%) | 7.09E-04 (±83.9%) | 1.72E-03 (±83.0%) | 2.70E-03 (±86.7%) | 1.24E-03 (±91.4%) |
| **24 h** | 7.21E-04 (±96.7%) | 1.22E-03 (±84.5%) | 2.48E-04 (±124.5%) | 5.09E-03 (±101.2%) | 5.89E-03 (±115.5%) | 6.74E-03 (±101.1%) |
| **36 h** | 1.41E-03 (±88.7%) | 2.75E-03 (±93.7%) | 3.18E-04 (±101.2%) | 0.169 (±88.6%) | 0.235 (±123.4%) | 0.0380 (±84.5%) |
| **48 h** | 1.97E-04 (±99.4%) | 6.76E-04 (±97.0%) | 2.17E-04 (±98.1%) | 0.0713 (±99.8%) | 0.397 (±124.4%) | 0.716 (±91.6%) |
| **5' junctions / *bioB*** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 2.15E-05) | (< 2.91E-05) | (< 1.39E-05) | (< 1.64E-05) | (< 1.01E-05) | (< 2.02E-05) |
| **12 h** | (< 2.55E-04) | (< 2.10E-04) | (< 1.16E-04) | (< 5.40E-05) | (< 3.77E-05) | (< 1.11E-04) |
| **14 h** | ND[b] | (< 9.11E-05) | (< 6.17E-05) | (< 5.33E-04) | (< 8.31E-05) | (< 7.81E-05) |
| **16 h** | 6.07E-05 (±109.5%) | (< 1.92E-05) | 2.73E-05 (±106.2%) | (< 2.51E-04) | (< 7.69E-05) | (< 3.40E-05) |
| **20 h** | 1.84E-04 (±109.9%) | 1.62E-04 (±112.8%) | 3.64E-04 (±107.0%) | (< 1.29E-05) | 4.37E-04 (±106.2%) | 3.71E-05 (±132.2%) |
| **24 h** | 2.22E-04 (±108.2%) | 5.47E-04 (±108.7%) | 1.78E-04 (±134.0%) | 9.71E-05 (±114.6%) | 1.71E-04 (±105.8%) | 9.11E-05 (±111.2%) |
| **36 h** | 1.33E-04 (±113.2%) | 3.89E-04 (±116.8%) | 1.21E-04 (±124.0%) | 0.162 (±101.6%) | 0.259 (±120.8%) | 0.156 (±117.3%) |
| **48 h** | 7.78E-05 (±104.5%) | 1.57E-04 (±123.5%) | 1.09E-04 (±106.3%) | 0.223 (±97.4%) | 0.441 (±118.3%) | 0.973 (±108.0%) |
| **3' junctions / *bioB*** (± % error) | | | | | | |
| **Timepoint** | **CB454(pRMH14)(pICV8)** | | | **CB454(pRMH14)(pICV8)(pSU2056)** | | |
| **0 h** | (< 9.76E-04) | (< 1.06E-03) | (< 6.61E-04) | (< 4.58E-04) | (< 3.11E-04) | (< 1.28E-03) |
| **12 h** | (< 0.0311) | (< 0.0257) | (< 0.0141) | (< 6.61E-03) | (< 4.61E-03) | (< 0.0136) |
| **14 h** | ND[b] | (< 9.48E-03) | (< 7.54E-03) | (< 0.0652) | (< 0.0102) | (< 9.55E-03) |
| **16 h** | (< 1.47E-03) | (< 2.05E-03) | (< 4.16E-04) | (< 0.0307) | (< 9.40E-03) | (< 4.16E-03) |
| **20 h** | (< 1.91E-04) | (< 6.58E-04) | (< 8.17E-04) | (< 1.56E-03) | (< 4.64E-03) | (< 1.58E-04) |
| **24 h** | (< 1.01E-05) | (< 8.19E-05) | (< 9.51E-05) | 1.17E-04 (±73.1%) | 2.05E-04 (±46.5%) | 2.68E-04 (±77.9%) |
| **36 h** | (< 1.67E-05) | (< 4.85E-05) | (< 2.18E-05) | 0.317 (±59.0%) | 0.758 (±79.4%) | 0.329 (±74.5%) |
| **48 h** | (< 1.24E-05) | (< 5.89E-05) | (< 1.62E-05) | 1.36 (±71.1%) | 1.16 (±68.0%) | 3.04 (±63.2%) |

[a] Gray values indicate when no junction amplicand was detected and show the theoretical possible number of junctions (see Methods and Table 3.3) per copy *bioB*

[b] ND=not done

**Figure 3.1.** Schematics of plasmids and PCR junctions arising in pRMH14 and pICV8 integrase-mediated recombination. PCR primers are shown as arrows, and PCR products are lines. 5' junction primers and products are shown in blue, 3' junctions in orange, and cointegrate junctions in purple. Only the PCR products shown in bold-type (the smallest possible products for each primer pair) were quantified. **A.** Cassette recipient plasmid pICV8 and cassette donor plasmid pRMH14. The BamHI to NarI fragment of pICV8 is also present in pICVm, including the zeocin resistance gene and the integron region (see Table 3.1). The primer shown in both orange and purple is used for both 3' and cointegrate junction PCR (see Table 3.2). **B.** pICV8 plasmids and PCR products resulting from *aadA2* or *aacA4aadA2* cassette acquisition. **C.** Cointegrates formed by recombination between the pICV8 *attI* site and pRMH14 *attI* or *attC* sites.

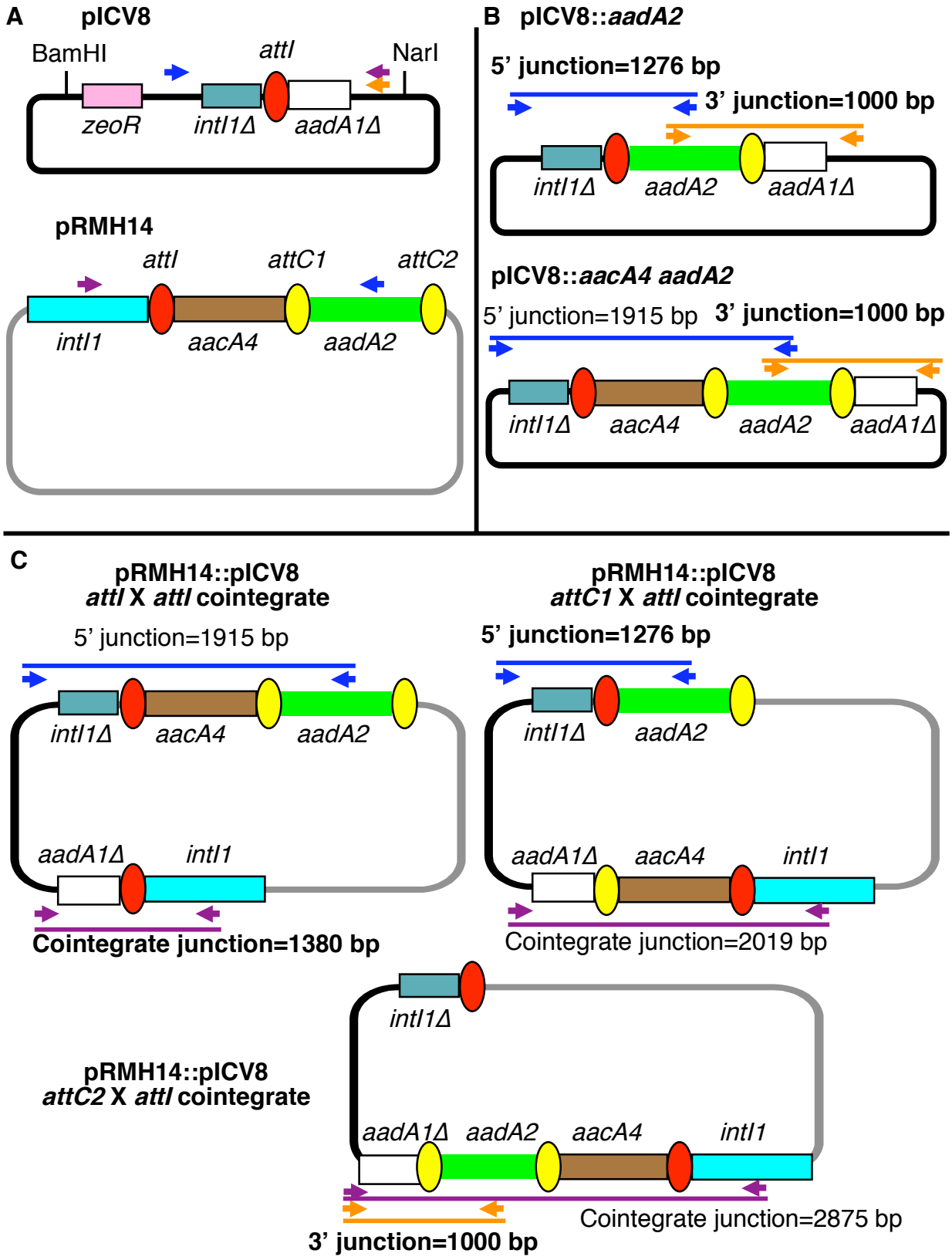
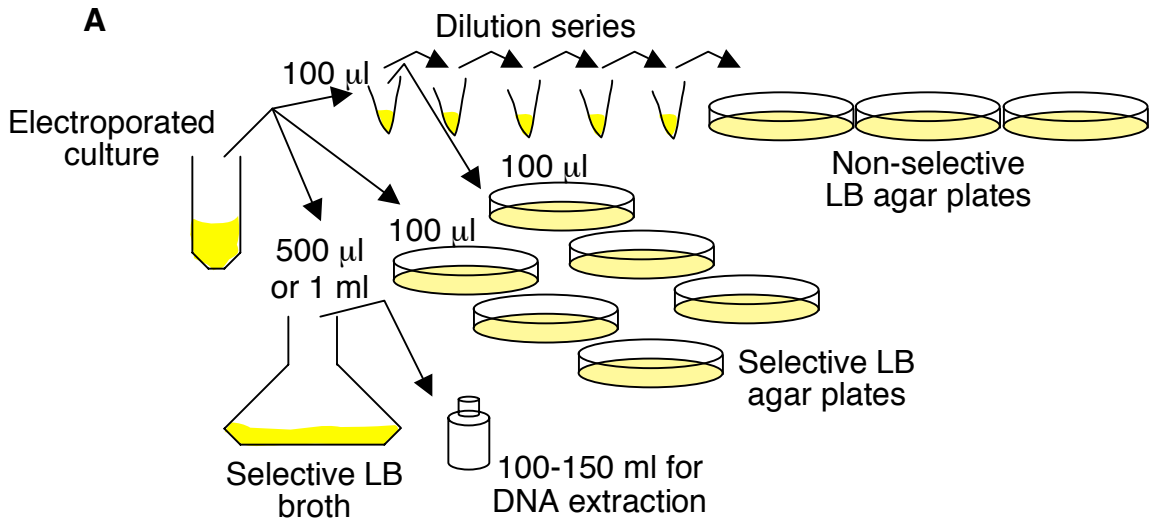

A
pICV8
BamHI
attI
NarI
zeoR
intI1Δ
aadA1Δ
pRMH14
attI
attC1
attC2
intI1
aacA4
aadA2
B
pICV8::aadA2
5' junction=1276 bp
3' junction=1000 bp
intI1Δ
aadA2
aadA1Δ
pICV8::aacA4 aadA2
5' junction=1915 bp
3' junction=1000 bp
intI1Δ
aacA4
aadA2
aadA1Δ
C
pRMH14::pICV8
attI X attI cointegrate
5' junction=1915 bp
intI1Δ
aacA4
aadA2
aadA1Δ
intI1
Cointegrate junction=1380 bp
pRMH14::pICV8
attC1 X attI cointegrate
5' junction=1276 bp
intI1Δ
aadA2
aadA1Δ
aacA4
intI1
Cointegrate junction=2019 bp
pRMH14::pICV8
attC2 X attI cointegrate
intI1Δ
aadA1Δ
aadA2
aacA4
intI1
3' junction=1000 bp
Cointegrate junction=2875 bp

**Figure 3.2.** Inoculation of cultures and timepoint sampling procedures. **A.** Procedure for zero timepoint. 1 ml of electroporated culture was used to inoculate 700 ml LB broth for CB454(pRMH14, pICV8, pSU2056) only, 500 μl inoculum was used for other strains. **B.** Procedure for 12h-48h timepoints. 100 μl was removed for duplicate dilution series, each plated in triplicate. The volume of culture removed for DNA isolation varied according to strain and timepoint.


A
Dilution series
100 µl
Electroporated culture
100 µl
100 µl
500 µl or 1 ml
Non-selective LB agar plates
Selective LB agar plates
Selective LB broth
100-150 ml for DNA extraction

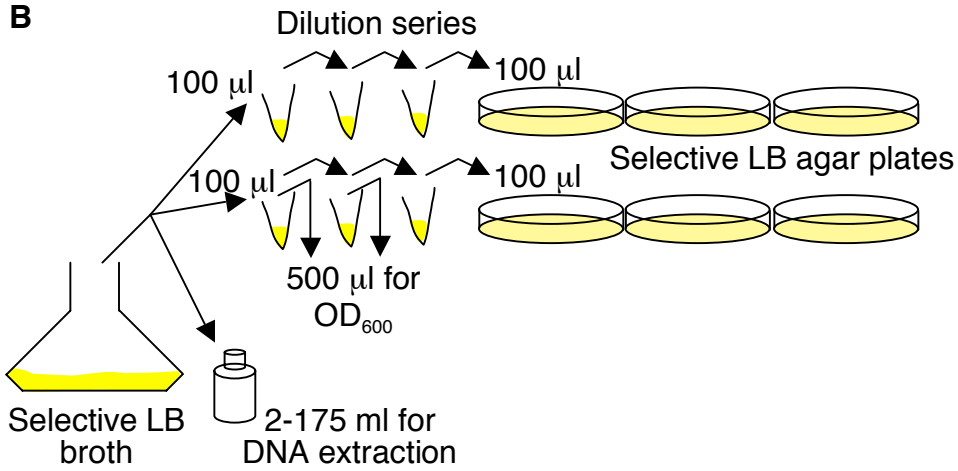

B
Dilution series
100 µl
100 µl
100 µl
100 µl
Selective LB agar plates
500 µl for OD600
Selective LB broth
2-175 ml for DNA extraction

**Figure 3.3**. Broth culture viable cell counts. The cfu/ml of each culture is plotted for every timepoint, from zero to 48 hours. CB(pRI)=CB454(pRMH14, pICV8); CB(pRIS)=CB454(pRMH14, pICV8, pSU2056); #1-3 are three independent experiments for each strain. The error bars show standard deviations for the plate counts (n=3-6 plates) at each timepoint.

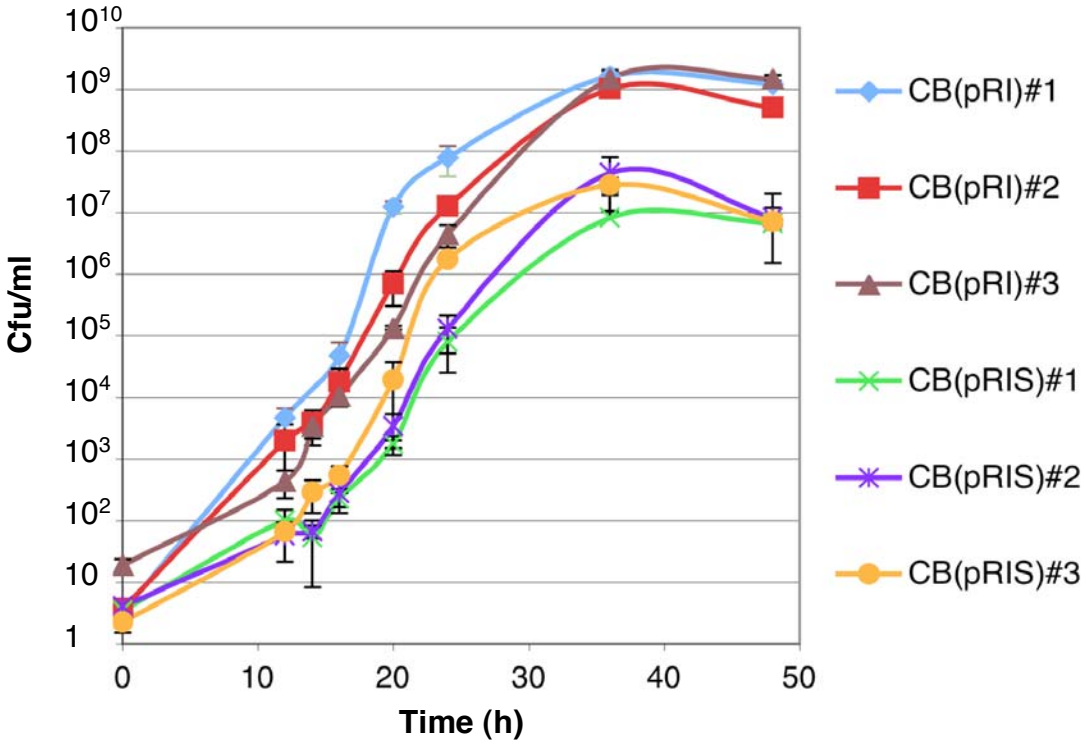

Cfu/ml
10^10
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10
1
0
10
20
30
40
50
Time (h)
CB(pRI)#1
CB(pRI)#2
CB(pRI)#3
CB(pRIS)#1
CB(pRIS)#2
CB(pRIS)#3

**Figure 3.4.** Copies of recombinant junctions and of chromosomal (*bioB*) and pICV8 plasmid marker genes per μg total DNA throughout the growth cycle in three independent experiments. The target copies per μg DNA were determined by dividing the specific primer detection limit (see Methods) by the average limiting dilution detection value (see Methods) determined by dilution PCR at each timepoint. The cfu/ml for each broth culture (Fig. 3.3) is also shown. CB(pRI)=CB454(pRMH14, pICV8); CB(pRIS)= CB454(pRMH14, pICV8, pSU2056); #1-3 are three independent experiments for each strain. **A**-**C** show CB(pRI) experiments 1-3, respectively, and **D**-**F** show CB(pRIS) experiments 1-3, respectively. The standard errors (black bars) are shown in copies per μg DNA at each timepoint.

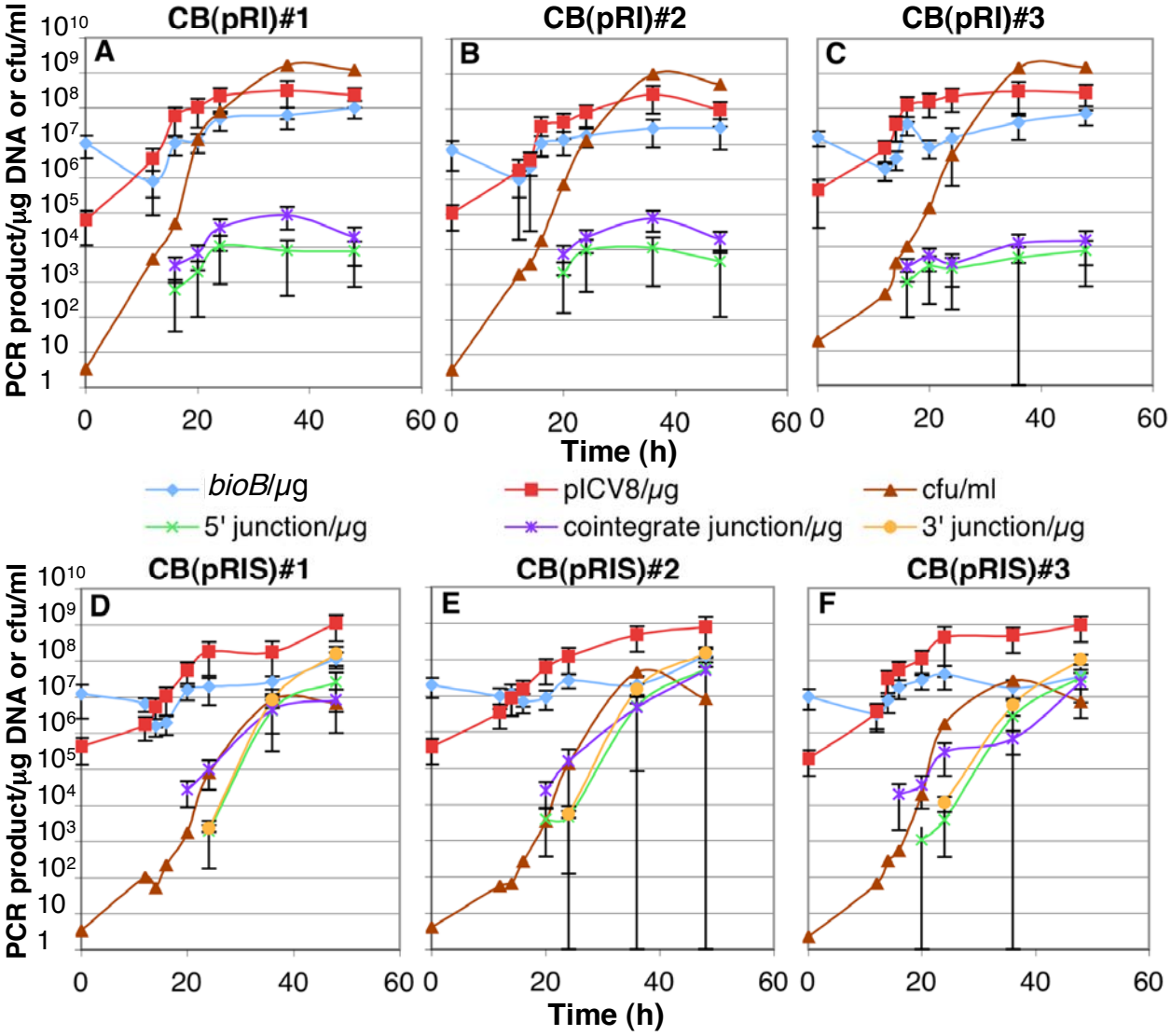

CB(pRI)#1
CB(pRI)#2
CB(pRI)#3
A
B
C
PCR product/µg DNA or cfu/ml
Time (h)
bioB/µg
pICV8/µg
cfu/ml
5' junction/µg
cointegrate junction/µg
3' junction/µg
CB(pRIS)#1
CB(pRIS)#2
CB(pRIS)#3
D
E
F

**Figure 3.5**. Copies of backbone or cells per recombinant junction plotted as a function of the culture age in clock time. Each data point represents the number of copies of pICV8 or of *bioB* per specified recombinant junction on a logarithmic scale at one timepoint (shown in hours) from one of three trials for each strain. The copies of pICV8, cells (*bioB*), and recombinant junctions were determined by the specific primer detection limit (Table 3.2) divided by the average limiting dilution detection value per μg DNA by dilution PCR (see Methods) at every timepoint for each primer pair. In **A**-**D**, gray diamonds show points for the low integrase strain CB(pRI), and black squares show points for the high integrase strain CB(pRIS) on a logarithmic scale. In **E**, black circles show pICV8 copies per 3' junction and gray triangles show *bioB* copies per 3' junction for the high integrase strain CB(pRIS) on a logarithmic scale. Gray lines in **A**-**D** show polynomial fits, and black lines in **A**-**E** and gray lines in **E** show power trendlines. Equations of trendlines are as follows:

| | Polynomial | Power |
|---|---|---|
| **A.** | gray $y=25.21x^2 - 2158.4x + 56795$ | black $y=(1\text{x}10^{9})x^{-4.45}$ |
| **B.** | gray $y=18.21x^2 - 1227.1x + 21014$ | black $y=(1\text{x}10^{10})x^{-5.74}$ |
| **C.** | gray $y=112.95x^2 - 8716.4x + 192608$ | black $y=(2\text{x}10^{18})x^{-10.14}$ |
| **D.** | gray $y=52.29x^2 - 3563.5x + 61707$ | black $y=(6\text{x}10^{18})x^{-11.15}$ |
| **E.** | ---------- | black $y=(1\text{x}10^{22})x^{-12.83}$ |
| | | gray $y=(1\text{x}10^{22})x^{-13.54}$ |

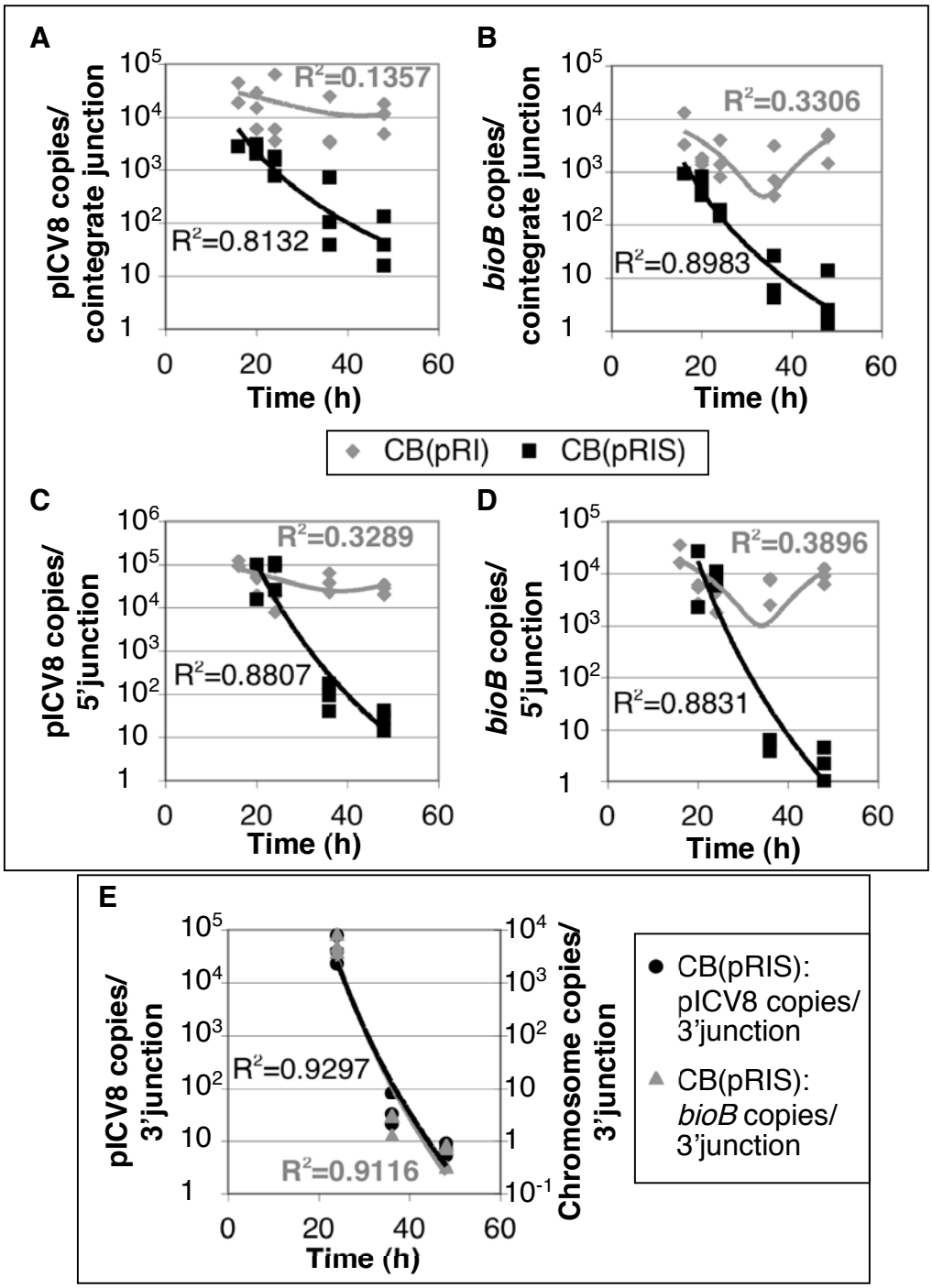

A
pICV8 copies/ cointegrate junction
$10^5$
$10^4$
$10^3$
$10^2$
10
1
$R^2$=0.1357
$R^2$=0.8132
0
20
40
60
Time (h)
B
*bioB* copies/ cointegrate junction
$R^2$=0.3306
$R^2$=0.8983
Time (h)
CB(pRI)
CB(pRIS)
C
pICV8 copies/ 5'junction
$10^6$
$R^2$=0.3289
$R^2$=0.8807
Time (h)
D
*bioB* copies/ 5'junction
$R^2$=0.3896
$R^2$=0.8831
Time (h)
E
pICV8 copies/ 3'junction
$R^2$=0.9297
$R^2$=0.9116
Time (h)
Chromosome copies/ 3'junction
$10^{-1}$
CB(pRIS): pICV8 copies/ 3'junction
CB(pRIS): *bioB* copies/ 3'junction

**Figure 3.6.** Copies of backbone or cells per recombinant junction plotted as a function of the number of generations of each strain. The number of generations was calculated using the viable cell counts of the culture timepoint for each trial (see Methods). The copies of pICV8, chromosome (*bioB*), and recombinant junctions were determined by the specific primer detection limit (Table 3.2) divided by the average limiting dilution detection value per μg DNA by dilution PCR (see Methods) at every timepoint for each primer pair. In **A**-**D**, gray diamonds show points for the low integrase strain CB(pRI), and black squares show points for the high integrase strain CB(pRIS) on a logarithmic scale. Black and gray lines best fit a power model. Fitting equations are as follows:

**A.** gray $y=(5\text{x}10^{6})x^{-2.05}$ black $y=(1\ \text{x}10^{4})x^{0.82}$

**B.** gray $y=(6\text{x}10^{5})x^{-1.99}$ black $y=(4\ \text{x}10^{4})x^{-1.87}$

**C.** gray $y=(1\text{x}10^{7})x^{-2.05}$

**D.** gray $y=(2\text{x}10^{6})x^{-1.99}$

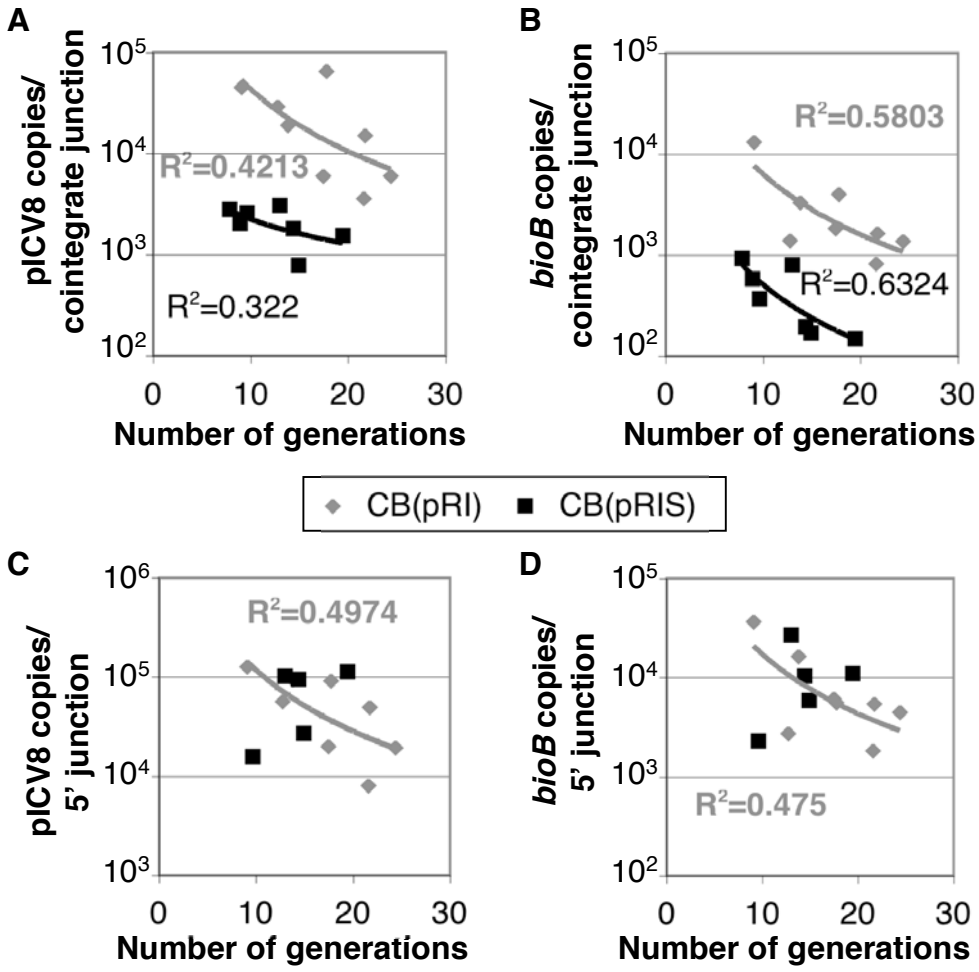
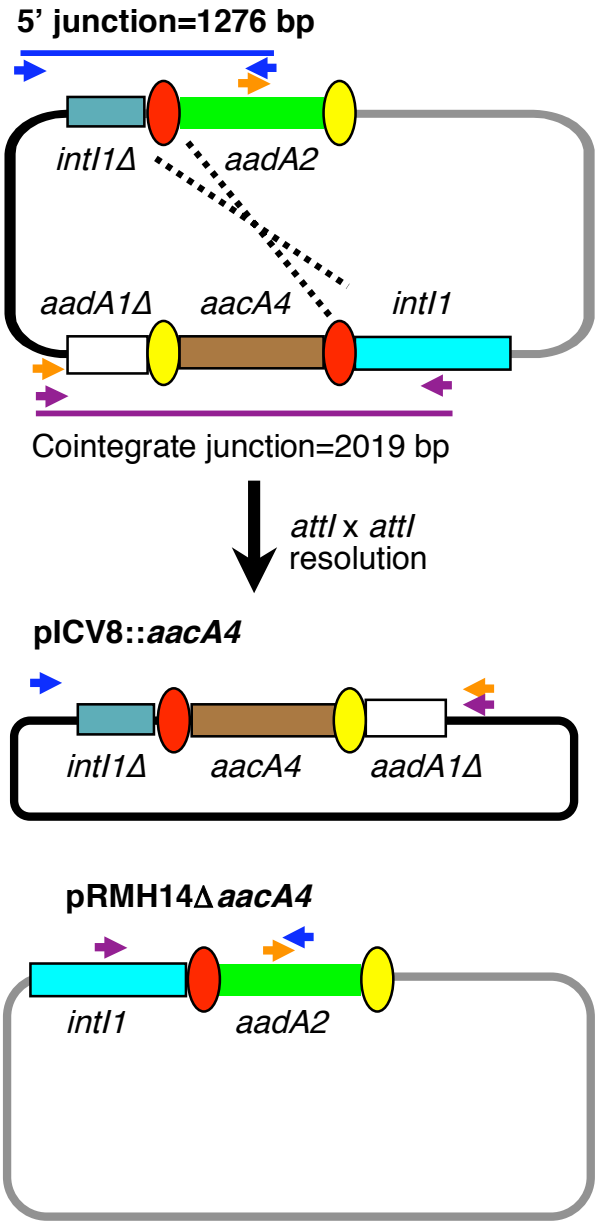

A
pICV8 copies/ cointegrate junction
$10^5$
$10^4$
$10^3$
$10^2$
$R^2$=0.4213
$R^2$=0.322
0
10
20
30
Number of generations
B
bioB copies/ cointegrate junction
$R^2$=0.5803
$R^2$=0.6324
Number of generations
CB(pRI)
CB(pRIS)
C
pICV8 copies/ 5' junction
$10^6$
$R^2$=0.4974
Number of generations
D
bioB copies/ 5' junction
$R^2$=0.475
Number of generations

**Figure 3.7.** Schematics of plasmids arising from *attI* x *attI* resolution of the pRMH14::pICV8 *attC1* x *attI* cointegrate. PCR primers are shown as arrows and PCR products as lines. 5' junction PCR primers and products are blue, cointegrate PCR primers and products are purple, and 3' junction primers are orange. Only the 1276 bp 5' junction product (shown in bold-type) from the *attC1* x *attI* cointegrate is quantified, and no PCR amplicands are detected from the resolution products.

**pRMH14::pICV8**
***attC1* X *attI* cointegrate**



# CHAPTER 4

## Dissertation Summary

Integrons encode a genetic system capable of expressing and mobilizing gene cassettes in natural populations by a site-specific recombinase (11, 20, 32). The site-specific recombinases in integrons (IntI) form a distinct family within the larger tyrosine family of site-specific recombinases (6, 29), and they encode a unique IntI region within the catalytic domain of their protein sequence (22). Integrons are separated into categories based on location, on a chromosome or within a mobile element such as a transposon and/or a plasmid (2, 20, 29). Chromosomal integrons, including superintegrons that contain over 20 gene cassettes, serve as a gene cassette reservoir for the mobile integrons (20, 28, 30). Mobile-element associated integrons then can disseminate antibiotic resistance gene cassettes throughout Gram-negative and Gram-positive bacterial populations in clinical and environmental settings (8, 16, 17, 21, 23, 24, 26, 31, 33, 35), facilitating the acquisition of antibiotic resistances by bacterial pathogens.

The vast majority of mobile integrons are located on plasmids (3, 11). Plasmid carried integrons can capture gene cassettes by two mechanisms: the integration of a closed circular gene cassette that was excised from a different (or the same) integron in the same cell or the formation and resolution of cointegrates (5, 11). The final cassette capture products from the two mechanisms are identical and indistinguishable by

sequence (5). Only plasmid cointegrates have been experimentally isolated from strains expressing integrase under its natural promoter, however, and cassette capture products have only been observed when the integrase is overexpressed (5, 10, 18, 19).

In Chapter 2, I constructed a family of integron capture vectors (pICV) that can acquire and donate antibiotic resistance gene cassettes in cells overexpressing IntI1. Similar to prior work where only cointegrates were isolated under natural integrase conditions by a mating-out assay (5, 18, 19), PCR of novel recombinant junctions formed by integrase-mediated recombination revealed that only cointegrates were detectable without integrase overexpression. Cassette capture products were not detected even when both integron donor and recipient plasmids expressed IntI1 under natural $P_{int}$ promoters, though cointegrates were detected even when only the donor plasmid was expressing IntI1 under $P_{int}$. This suggests that integrase overexpression may overwhelm some putative natural regulatory mechanisms that would normally prevent rampant recombination. Integrase-mediated recombination is thought to be a highly regulated process (13), though the mechanisms of regulation are not yet known, because the integration of gene cassettes at secondary sites in the chromosomes of bacteria would be deleterious (7, 13). In addition, the natural regulation of other tyrosine recombinase-mediated reactions is affected by concentration of recombinase. For example, the resolvase Cre is normally restricted to intramolecular reactions in order to resolve P1 prophage dimers, but high concentrations of Cre *in vivo* destabilize the P1 prophage by creating multimers by intermolecular recombination (27). Also in Chapter 2, PCR of cell lysates from broth cultures sampled at different times in culture growth

showed an accumulation of integrase-mediated recombination products in late stationary phase that may indicate differential regulation of recombination during the growth cycle, which was explored further using quantitative methods.

In Chapter 3, I used dilution PCR to quantify integrase-mediated recombination products from total DNA isolated from broth cultures at several timepoints during culture growth. While previous experiments determining IntI1 recombination frequencies used distant endpoint assays where recombinant plasmids were isolated by conjugation or transformation of extracted DNA (7, 18), the dilution PCR method allows measurement of integrase-mediated recombination products within the cells in which they are produced. Quantification of cointegrate and 5' junction PCR amplicands, respectively showing pRMH14::pICV8 *attI* x *attI* cointegrates and *attC1* x *attI* cointegrates (and/or pICV8::*aadA2* capture products) (Fig. 3.1 B, C), revealed that there is only approximately one cointegrate per 1000-10,000 cells in a strain expressing integrase under a natural promoter. Thus, it is not surprising that only a few cointegrates have been isolated in mating-out experiments (5, 18, 19) without using excess integrase.

More *attI* x *attI* recombination products (shown by the cointegrate junction) than *attC* x *attI* products were seen in the strain expressing integrase under a natural promoter at all timepoints of culture growth, indicating a preference for *attI* x *attI* crossovers in that strain. In addition, more *attI* x *attI* recombination products were seen in the integrase overexpression strain than any other crossover type in log phase, but more *attC* x *attI* recombination products, especially from recombination at *attC2* (as shown by the 3' junction), were seen in stationary phase. This suggests that

recombination crossover preferences may change according to the growth cycle under conditions of integrase overexpression. Observing a preference for *attI* x *attI* crossovers was surprising since previous experiments have shown that *attC* x *attI* crossovers occur 10-fold more frequently than *attI* x *attI* recombinations (7). Those experiments were done on stationary phase cultures overexpressing integrase, however, and dilution PCR shows increased products from *attC* recombination (5' and 3' junctions) relative to *attI* x *attI* products (cointegrate junctions) in stationary phase in the high integrase strain. Another tyrosine recombinase, the bacteriophage lambda integrase, regulates recombination by preferring specific crossovers depending on the relative concentrations of integrase and excisionase in a cell (15). For example, if there is more integrase present, integration reactions are induced by a preference for *attB* x *attP* crossovers, and, if the concentration of Xis is higher than lambda integrase, an *attL* x *attR* crossover is preferred for an excision reaction (14, 15).

In the low integrase strain, dilution PCR revealed that recombination products per cell (and per plasmid) increased during log phase and showed a decrease or plateau after reaching stationary phase, with recombination products most abundant at 36 hours of growth (in stationary phase). This could indicate a growth-cycle dependent regulation of integrase-mediated recombination. Growth cycle effects are more obvious in the strain overexpressing integrase, however, because recombination products increase dramatically from late log (24 hours) to stationary phase (36 hours). All recombination products sharply increasing from late log into stationary phase and *attC* recombination products increasing relative to *attI* x *attI* products (as mentioned above) suggests a

growth phase-dependent regulation mechanism. Many tyrosine recombinases use accessory factors like FIS and/or IHF to bend DNA in a recombination synapse (9, 12) that could could cause a growth cycle-dependent effect on recombination. For example, FIS is abundant in log phase cultures, but is nearly undetectable in stationary phase (1, 25); thus, FIS-dependent recombination reactions would increase in log phase and decrease in stationary phase cultures. No host cofactors are known to be required for integron integrase-mediated reactions, but the majority of site-specific recombinases use some kind of accessory factor in recombination (9, 12), so it is likely that one or more may be identified for integron integrases in the future. A requirement for an unidentified accessory factor may explain why *in vitro* experiments of integron integrase reactions have thus far been unsuccessful (13, 20). In addition, growth cycle effects on recombination could be caused by regulation of gene cassette expression. Decreases in IntI1-mediated recombination frequencies have been seen with increased expression of gene cassettes (6, 7). This reduction in recombination is likely caused by RNA polymerase physically blocking DNA sites or the formation of a recombination synapse, but it could also result from increased cassette expression interfering with integrase expression since the cassette promoter $P_c$ and $P_{int}$ are adjacent and divergently transcribed (4, 6, 7, 32). The gene cassettes of the *Vibrio cholerae* superintegron are upregulated in stationary phase in response to RpoS (34), though the affect of that upregulation on VchIntIA integrase recombination has not yet been investigated.

While the observations that crossover preferences and abundance of products by IntI1-mediated recombination are affected by both integrase concentration and growth

cycle are novel and exciting, much is left to be resolved regarding the regulation of these recombination reactions. In the future, potential accessory factors such as FIS should be investigated for effects on IntI1-mediated recombination. In addition, the effects of RpoS and other stationary phase-associated proteins (or distinctly log phase-associated proteins like FIS) could be determined to further explore growth phase regulation. Also, in future experiments directly following those in Chapter 3, recombination products from late log to early stationary phase (between the 24 and 36 hour timepoints in Chapter 3) could be quantified to further elucidate the specific time in the growth cycle at which the high integrase strain first begins its dramatic increase in recombination and the specific time in the low integrase strain where recombination products are at their maximum and begin to plateau.

**References**


1. **Ball, C. A., R. Osuna, K. C. Ferguson, and R. C. Johnson.** 1992. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. J. Bacteriol. **174:**8043-8056.
2. **Biskri, L., M. Bouvier, A. M. Guerout, S. Boisnard, and D. Mazel.** 2005. Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. J Bacteriol **187:**1740-50.
3. **Boucher, Y., M. Labbate, J. E. Koenig, and H. W. Stokes.** 2007. Integrons: mobilizable platforms that promote genetic diversity in bacteria. Trends Microbiol.

4. **Collis, C., and R. Hall.** 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. Antimicrob. Agents Chemother. **39:**155-162.

5. **Collis, C. M., and R. M. Hall.** 1992. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. J. Bacteriol. **174:**1574-85.

6. **Collis, C. M., M. J. Kim, H. W. Stokes, and R. M. Hall.** 2002. Integron-encoded IntI integrases preferentially recognize the adjacent cognate *attI* site in recombination with a 59-be site. Mol. Microbiol. **46:**1415-27.

7. **Collis, C. M., G. D. Recchia, M. J. Kim, H. W. Stokes, and R. M. Hall.** 2001. Efficiency of recombination reactions catalyzed by class 1 integron integrase IntI1. J. Bacteriol. **183:**2535-42.

8. **Goldstein, C., M. D. Lee, S. Sanchez, C. Hudson, B. Phillips, B. Register, M. Grady, C. Liebert, A. O. Summers, D. G. White, and J. J. Maurer.** 2001. Incidence of class 1 and 2 integrases in clinical and commensal bacteria from livestock, companion animals, and exotics. Antimicrob. Agents Chemother. **45:**723-6.

9. **Grindley, N. D. F., K. L. Whiteson, and P. A. Rice.** 2006. MECHANISMS OF SITE-SPECIFIC RECOMBINATION. Annual Review of Biochemistry **75:**567-605.

10. **Hall, R. M., D. E. Brookes, and H. W. Stokes.** 1991. Site-specific insertion of genes into integrons: role of the 59-base element and determination of the recombination cross-over point. Mol. Microbiol. **5:**1941-59.

11. **Hall, R. M., and H. W. Stokes.** 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica **90:**115-32.

12. **Hallet, B., and D. J. Sherratt.** 1997. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. FEMS Microbiology Reviews **21:**157-178.

13. **Hansson, K., O. Skold, and L. Sundstrom.** 1997. Non-palindromic *attI* sites of integrons are capable of site-specific recombination with one another and with secondary targets. Mol. Microbiol. **26:**441-53.

14. **Landy, A.** 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. Annual Review of Biochemistry **58:**913-941.

15. **Leffers, G. G., Jr., and S. Gottesman.** 1998. Lambda Xis degradation *in vivo* by Lon and FtsH. J. Bacteriol. **180:**1573-1577.

16. **Leverstein-van Hall, M. A., M. B. HE, T. D. AR, A. Paauw, A. C. Fluit, and J. Verhoef.** 2003. Multidrug resistance among Enterobacteriaceae is strongly associated with the presence of integrons and is independent of species or isolate origin. J. Infect. Dis. **187:**251-9.

17. **Lévesque, C., L. Piche, C. Larose, and P. H. Roy.** 1995. PCR mapping of integrons reveals several novel combinations of resistance genes. Antimicrob. Agents Chemother. **39:**185-91.

18. **Martinez, E., and F. de la Cruz.** 1990. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic resistance genes. EMBO J. **9:**1275-81.

19. **Martinez, E., and F. de la Cruz.** 1988. Transposon Tn*21* encodes a RecA-independent site-specific integration system. Mol. Gen. Genet. **211:**320-5.

20. **Mazel, D.** 2006. Integrons: agents of bacterial evolution. Nat Rev Micro **4:**608-620.

21. **Mazel, D., B. Dychinco, V. A. Webb, and J. Davies.** 2000. Antibiotic resistance in the ECOR collection: integrons and identification of a novel *aad* gene. Antimicrob. Agents Chemother. **44:**1568-74.

22. **Messier, N., and P. H. Roy.** 2001. Integron integrases possess a unique additional domain necessary for activity. J. Bacteriol. **183:**6699-6706.

23. **Nandi, S., J. J. Maurer, C. Hofacre, and A. O. Summers.** 2004. Gram-positive bacteria are a major reservoir of Class 1 antibiotic resistance integrons in poultry litter. Proc. Natl. Acad. Sci. USA **101:**7118-7122.

24. **Nield, B. S., A. J. Holmes, M. R. Gillings, G. D. Recchia, B. C. Mabbutt, K. M. H. Nevalainen, and H. W. Stokes.** 2001. Recovery of new integron classes from environmental DNA. FEMS Microbiol. Lett. **195:**59-65.

25. **Osuna, R., D. Lienau, K. T. Hughes, and R. C. Johnson.** 1995. Sequence, regulation, and functions of *fis* in *Salmonella typhimurium*. J. Bacteriol. **177:**2021-2032.

26. **Partridge, S. R., C. M. Collis, and R. M. Hall.** 2002. Class 1 integron containing a new gene cassette, *aadA10*, associated with Tn*1404* from R151. Antimicrob. Agents Chemother. **46:**2400-8.

27. **Paul, S., and D. Summers.** 2004. ArgR and PepA, accessory proteins for XerCD-mediated resolution of ColE1 dimers, are also required for stable maintenance of the P1 prophage. Plasmid **52:**63-68.

28. **Rowe-Magnus, D. A., A.-M. Guerout, and D. Mazel.** 2002. Bacterial resistance evolution by recruitment of super-integron gene cassettes. Molecular Microbiology **43:**1657-1669.

29. **Rowe-Magnus, D. A., A. M. Guerout, P. Ploncard, B. Dychinco, J. Davies, and D. Mazel.** 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. Proc. Natl. Acad. Sci. USA **98:**652-7.

30. **Rowe-Magnus, D. A., and D. Mazel.** 2002. The role of integrons in antibiotic resistance gene capture. Int J Med Microbiol **292:**115-25.

31. **Sandvang, D.** 1999. Novel streptomycin and spectinomycin resistance gene as a gene cassette within a class 1 integron isolated from *Escherichia coli*. Antimicrob. Agents Chemother. **43:**3036-8.

32. **Stokes, H. W., and R. M. Hall.** 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. Mol. Microbiol. **3:**1669-83.

33. **Tosini, F., P. Visca, I. Luzzi, A. M. Dionisi, C. Pezzella, A. Petrucca, and A. Carattoli.** 1998. Class 1 integron-borne multiple-antibiotic resistance carried by IncFI and IncL/M plasmids in *Salmonella enterica* serotype Typhimurium. Antimicrob. Agents Chemother. **42:**3053-8.

34. **Yildiz, F. H., X. S. Liu, A. Heydorn, and G. K. Schoolnik.** 2004. Molecular analysis of rugosity in a *Vibrio cholerae* O1 El Tor phase variant. Molecular Microbiology **53:**497-515.

35. **Zuhlsdorf, M. T., and B. Wiedemann.** 1992. Tn*21*-specific structures in gram-negative bacteria from clinical isolates. Antimicrob. Agents Chemother. **36:**1915-21.