Microbiology Spectrum
American Society for Microbiology Press

# The Integron: Adaptation On Demand

JOSÉ ANTONIO ESCUDERO,[1,2,*] CÉLINE LOOT,[1,2,*] ALEKSANDRA NIVINA,[1,2,3] and DIDIER MAZEL[1,2]

[1]Institut Pasteur, Unité Plasticité du Génome Bactérien, 75724 Paris, France; [2]CNRS UMR3525, 75724 Paris, France; [3]Paris Descartes, Sorbonne Paris Cité, Paris, France

**ABSTRACT** The integron is a powerful system which, by capturing, stockpiling, and rearranging new functions carried by gene encoding cassettes, confers upon bacteria a rapid adaptation capability in changing environments. Chromosomally located integrons (CI) have been identified in a large number of environmental Gram-negative bacteria. Integron evolutionary history suggests that these sedentary CIs acquired mobility among bacterial species through their association with transposable elements and conjugative plasmids. As a result of massive antibiotic use, these so-called mobile integrons are now widespread in clinically relevant bacteria and are considered to be the principal agent in the emergence and rise of antibiotic multiresistance in Gram-negative bacteria. Cassette rearrangements are catalyzed by the integron integrase, a site-specific tyrosine recombinase. Central to these reactions is the single-stranded DNA nature of one of the recombination partners, the *attC* site. This makes the integron a unique recombination system. This review describes the current knowledge on this atypical recombination mechanism, its implications in the reactions involving the different types of sites, *attC* and *attI*, and focuses on the tight regulation exerted by the host on integron activity through the control of *attC* site folding. Furthermore, cassette and integrase expression are also highly controlled by host regulatory networks and the bacterial stress (SOS) response. These intimate connections to the host make the integron a genetically stable and efficient system, granting the bacteria a low cost, highly adaptive evolution potential "on demand".

## INTRODUCTION

Integrons are genetic platforms that allow bacteria to evolve rapidly through the acquisition, stockpiling, excision, and reordering of open reading frames found in mobile elements named cassettes. The evolutionary potency that integrons provide for bacteria is based on the variety of functions encoded in the cassettes, as well as on the intricate coupling of integron activity to bacterial stress (1).

The structure of any integron includes a stable platform and a variable cassette array. The platform contains: (i) the gene encoding the integrase (IntI), a type of tyrosine recombinase (Y-recombinase) that has evolved structural features to specifically perform the integration and excision of cassettes; (ii) a recombination site for integration of cassettes, the *attI* site, that is found upstream of the *intI* gene (except in integrons from the *Treponema* genus, where it is located downstream); and (iii) a promoter, $P_C$, within the *intI* gene or between *intI* and the *attI* site, that is oriented towards the integration point and drives the expression of cassettes (2–4) (Fig. 1A). Cassettes are circular nonreplicative elements (5) that generally contain a promoterless gene and a second recombination site: the *attC* site. These genes are rendered functional through their integration into the platform and expression from the $P_C$ promoter. Successive integration events result in the assembly of an array of cassettes, of which only the subset that is closest to the

**Received:** 27 March 2014, **Accepted:** 28 March 2014, **Published:** 12 March 2015

**Editors:** Phoebe Rice, University of Chicago, Chicago, IL, and Nancy Craig, Johns Hopkins University, Baltimore, MD

**Citation:** Escudero JA, Loot C, Nivina A, Mazel D. 2014. The integron: adaptation on demand. *Microbiol Spectrum* 3(2):MDNA3-0019-2014. doi:10.1128/microbiolspec.MDNA3-0019-2014.

**Correspondence:** Didier Mazel, mazel@pasteur.fr

* These authors contributed equally to this work

© 2014 American Society for Microbiology. All rights reserved.



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

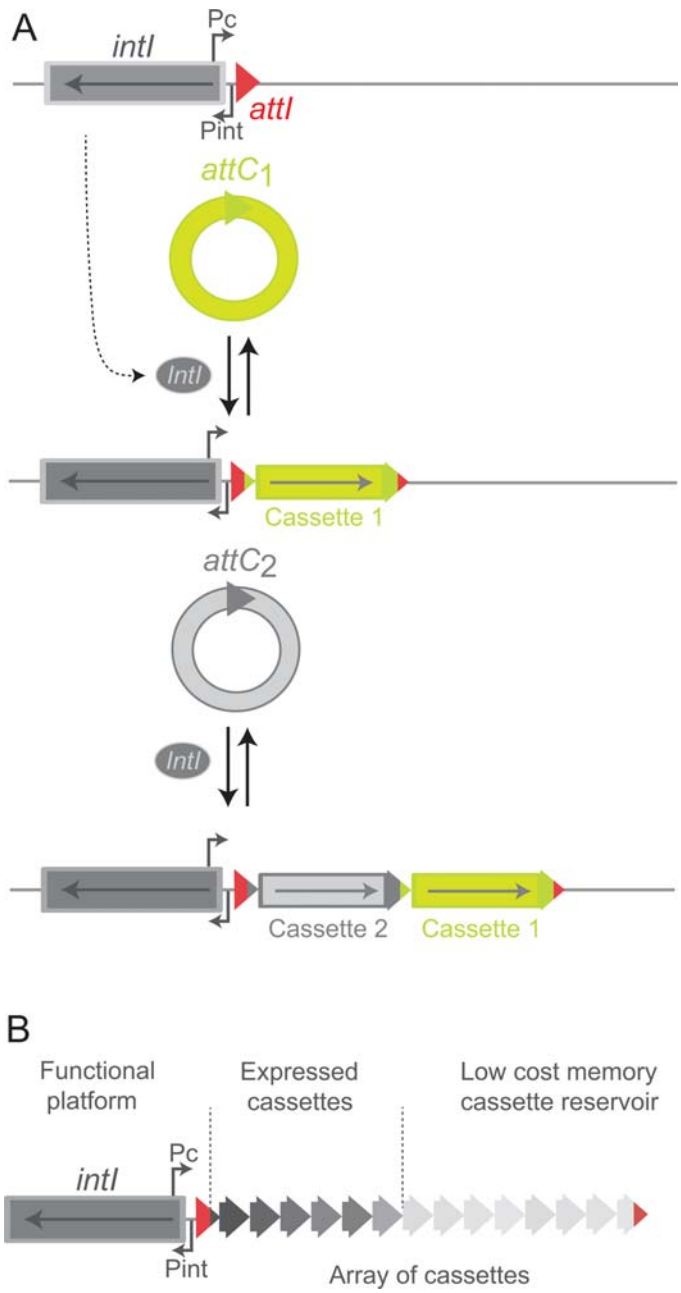



**FIGURE 1** Organization of integrons. (A) Insertion and excision of cassettes: the functional platform, composed of the integrase-encoding *intI1* gene, the cassette ($P_C$) and integrase promoters (Pint), and the primary *attI* recombination site (red triangle), is shown. Cassette insertion (*attC*×*attI*) and excision (*attC*×*attC*) catalyzed by the IntI integrase are represented. Hybrid *attI* and *attC* sites are indicated. Arrows inside the cassettes indicate the direction of the open reading frame. (B) Expression of cassettes: cassettes of the array are represented by small arrows. Their expression level is reflected by the color intensity of each arrow. Only the first cassettes of the array are expressed, and the subsequent ones can be seen as a low-cost cassette reservoir. doi:10.1128/microbiolspec.MDNA3-0019-2014.f1

integration site is expressed (Fig. 1B). Cassettes can be randomly excised and further integrated into the first position of the array where their expression is maximal. Hence, the cassette content of an integron is variable, reflecting a history of adaptive events, and represents a low-cost memory of valuable functions for the cell (Fig. 1B).

## Origin of integrons

Integrons are ancient structures that have shaped the evolution of bacteria for hundreds of millions of years (6). They are present in the chromosomes of approximately 17% of the bacterial species for which the genome sequence is available (7), spanning several bacterial phyla. These integrons are commonly referred to as chromosomal integrons (CIs), to distinguish them from plasmid-borne integrons (see below). Phylogenetic studies show that the branching of integrases is coherent with the organismal phylogeny, proving the ancestry and stability of integrons (6, 8, 9). Three large groups can be distinguished in the phylogeny of integrases: (i) the soil–freshwater proteobacteria, (ii) the marine γ-proteobacteria, and (iii) the inverted integrase group (7, 10). The first two groups form ecologically relevant taxons (11), suggesting, together with some phylogenetic incongruences in the branching of integrases, that horizontal transfer of CIs occurs over long evolutionary periods, among bacteria sharing the same environment.

Chromosomal integrons can be very large, embodying a significant percentage of a species' genome. Such is the case of the so-called superintegrons (SI) found in the *Vibrio* genus that can stockpile more than 200 cassettes. The best-studied SI is that of *Vibrio cholerae* (12), which harbors around 175 cassettes, representing 3% of the genome. Although some cassettes of the SI have been characterized (13, 14), mainly through their role in pathogenicity, the function of the majority is yet to be elucidated (a lack of knowledge that holds true for almost all CI cassettes). For instance, of the genes found in *Vibrio* species cassettes, the majority have no known function (10), and the remainder are homologous to proteins with a wide variety of functions, underscoring the genomic plasticity that integrons confer to their hosts. Another interesting feature of many CIs is that the array of *attC* sites shows a high sequence identity, suggesting a relationship between the sequence of the recombination site and the host (6, 15). Although this sequence conservation is not true for all CIs (10), its existence is relevant to our understanding of the hitherto unexplored phenomenon of *de novo* generation of



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

cassettes, as discussed below (see Genesis of integron cassettes).

### Mobilization of integrons: from the environment to the hospitals

It is now broadly accepted that sedentary CIs have found their way into clinically relevant bacteria through their association with transposable elements and conjugative plasmids (7, 10, 11). This has allowed for the efficient dispersal of integrons among members of the same and different species, and so their classification as mobile integrons (MIs). It is not surprising that such an evolvable equipment as the integron would be useful for rapid bacterial adaptation in the changing environment shaped by humans. Indeed, MIs played a major role in the early rise of multidrug resistance among clinically relevant bacteria in the 1960s (16), triggering the research that led to the discovery of integrons in the late 1980s (1, 17). Five different MIs (classes 1 to 5) have been described to date, classified according to the sequence of their integrase. Integron classes 1 and 3 (18) are found associated with Tn*402* (19, 20), whereas class 2 integrons are almost exclusively linked to Tn*7* derivatives (21). These three classes are the ones that are historically involved in the spread of multiresistance, and are definitively the most important from a medical point of view. Classes 4 and 5 were identified for their role in the rise of trimethoprim resistance in *Vibrio* spp. and were respectively located in a subset of the SXT elements in *V. cholerae* (22) and on the pRSV1 plasmid of *Aliivibrio salmonicida* (GenBank AK277063) (23).

The cargo of MIs is small, with arrays of up to eight cassettes (24). The spread of MIs grants them access to a large pool of cassettes from a variety of genetic backgrounds, as is evidenced by their GC content, codon usage and *attC* sequences. Functions of MI cassettes are strikingly homogeneous, since they encode, almost exclusively, antibiotic resistance genes against almost all families of clinically relevant antibiotics and antiseptics of the quaternary ammonium-compound family (11, 25, 26). In the clinical setting, where antibiotic pressure is high, MIs confer such an adaptive advantage that their presence is nowadays commonplace among Gram-negative pathogens. Class 1 integrons are the most widespread and clinically relevant, since they are detected in 22% to 59% of Gram-negative clinical isolates (27) and have occasionally been identified in some Gram-positive bacteria (28–31). It is also the only class for which evidence of activity *in persona* is available (32). Hence, they represent the major experimental model of the integron.

## A UNIQUE RECOMBINATION SYSTEM

### The *attI* site

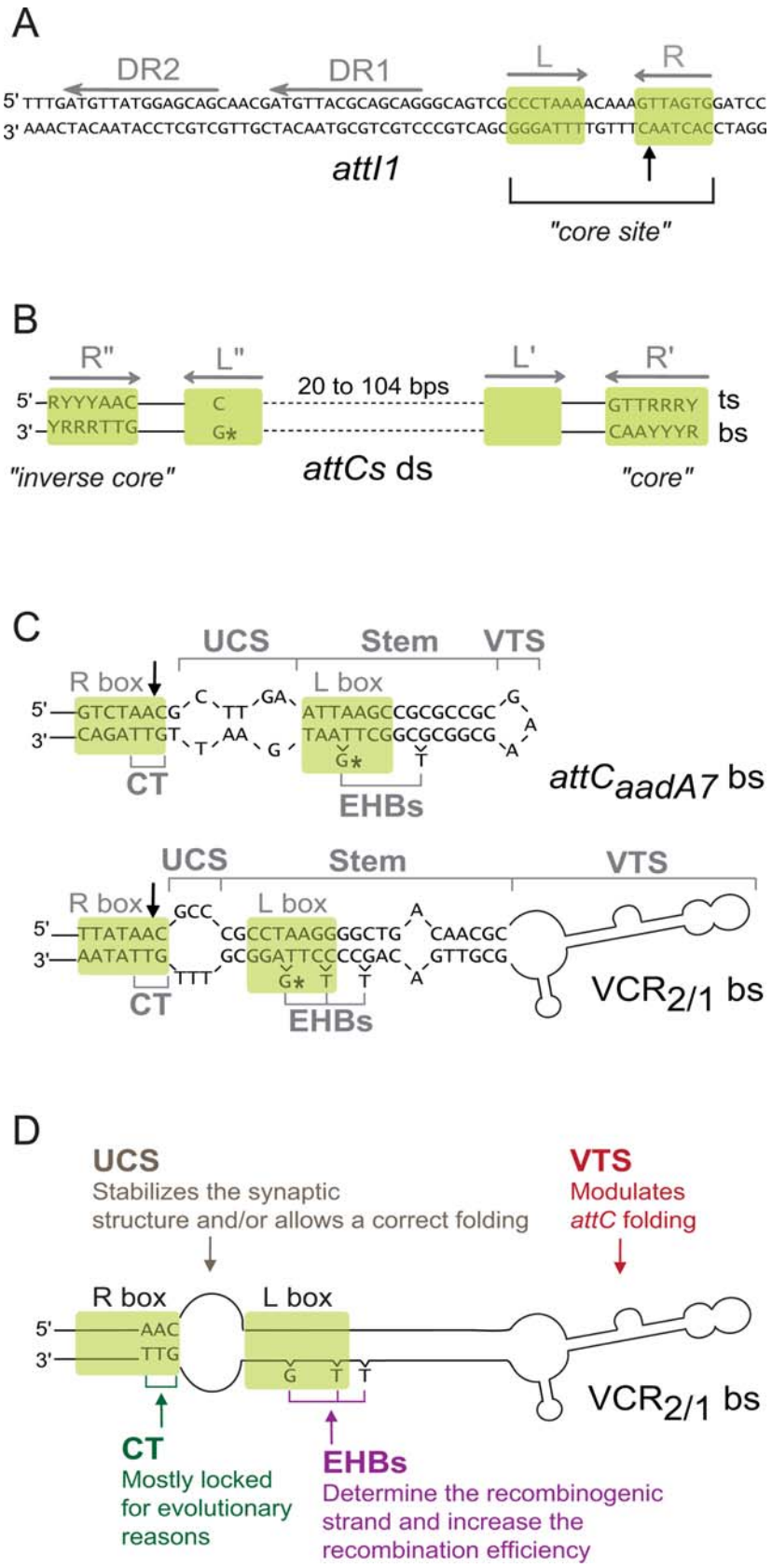
The *attI* site of the integron, in which cassettes are integrated, is minimally composed of two integrase binding sites termed L and R, forming the *core* site. The recombination point is in the 5′-GTT-3′ triplet located in the R box, with the cleavage taking place between the A and the C on the complementary strand, the bottom strand (bs) (Fig. 2A). The *attI* sites are hardly recognizable, because L binding domains are always degenerate with respect to R and the central regions differ greatly between *attI* sites. *In vitro* experiments have proven the specific binding of integrases to their cognate *attI* sites. In the case of IntI1/*attI1*, four monomers bind to the site, two in the R and L boxes of the core site, as expected, and two within imperfect direct repeats (DR) dubbed DR1 and DR2, located upstream of the core site (33) (Fig. 2A). These accessory sequences enhance, but are dispensable for, recombination in *attI1*. It has been suggested that they could serve as a trap to keep the integrase monomers in the vicinity of the core site (33), but the fact that they have a higher impact in the reaction with an *attC* site compared with that with another *attI* site (34), suggests a role as a topological filter. These accessory sequences do not seem to be a common feature of *attI* sites. In a recent study on the diversity of these sites among environmental samples from Suez and Tokyo Bay only three putative *attI* sites had accessory sequences, summing to a total of only five out of more than 40 sites (taking into account *attI2*, *attI3* and the sites in ref. (35)). Nevertheless, these numbers are merely speculative, as experimental data on the binding of the cognate integrases to these sites is lacking, and some large direct repeats have been overlooked in this study, probably for their lack of a GTT triplet, a feature of unknown importance in these sequences (36). *attI* sites from different integrons diverge significantly, paralleling the pattern observed for integrases (6). Cross recombination assays between non-cognate *attI*/IntI partners have proven that integrases recognize preferentially their related site, suggesting the coevolution of both partners. Nevertheless, in some cases cross-talk between two systems can occur. IntI1 from class 1 MIs can, for instance, integrate cassettes into *attI2* and *attI3* sites, albeit 100 times less efficiently than into *attI1*, but not at all into the *attI* site of the *V. cholerae* SI (37).

### The *attC* site

*attC* sites are an integral part of integron cassettes, terminating each of them, and are necessary for their mobility. These sites differ significantly from canonical



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20






Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

Y-recombinase sites. As we will see through the chapter, many of the features that make the integron a unique recombination system are made possible through the peculiarities of *attC* sites.

Two regions of inverted homology, R"-L" and L'-R', are found in all *attC* sites, separated by a central region that is highly variable in size and sequence, leading to *attC* sites of broadly variable lengths, from 57 to 141 bp (38) (Fig. 2B). *attC* sites have almost no sequence conservation, but instead they display a strikingly conserved palindromic organization that can form secondary structures through DNA intra-strand pairing (Fig. 2C) (39). Upon folding, *attC* sites show a hairpin structure resembling a canonical core site consisting of R and L boxes. This specific single-stranded (ss) structure is the substrate recognized (40, 41) and recombined by the integrase (42). Of both strands in any *attC* site, the bs is approximately $10^3$ times more recombinogenic than the top strand (ts) (42).

A comparison of *attC* sites shows that sequence conservation is restricted to two inverted triplets, 5′-**AAC**-3′ and 5′-**GTT**-3′ located in the R" and R' boxes, respectively (Fig. 2B, see ts). Nevertheless, these conserved sequences can be extended to inverted repeat sequences of 7 bp designated as *inverse core* and *core* (Fig. 2B). The *inverse core* consists of an RYYY**AAC** (R: purine; Y: pyrimidine) sequence and is located at the 5′ end of the ds *attC* site and the complementary *core* consisting of **GTT**RRRY is located at the 3′ end (Fig. 2B). Once folded, a conserved CAA triplet is reconstituted within the R box and, consistent with the *attI* sites, the recombination point is located between the C and A nucleotides.

In *attC* sites, the genetic information required for proper recombination is not entirely contained in the primary sequence, but rather determined by specific features of their secondary structures (43, 44). The folding permits the inclusion of a new layer of information and regulation in the site. Three structural features that are common to all known ss *attC* sites emerge from the folding of the bs of *attC* sites (defined in ref. (43)). First, the presence of a set of single bases located on the R"-L" arm of the symmetrical *attC* sequence that have no complementary nucleotides on the R'-L' arm and are present as extrahelical bases (EHBs) in the structured site (Fig. 2C). Depending on the *attC* sites, two or three EHBs can be found (43). These bases have three major roles in the recombination reaction: (i) they determine the recombinogenic strand in the site (the bottom strand), (ii) they serve as stabilizers of the synaptic complex, establishing contacts with protein monomers across the synapse, and (iii) they avoid a second cut in the *attC* site by pulling apart the tyrosine residue of the integrase monomer bound to the L"-L' box (see *attI*×*attC* reaction below) (43, 44). The second structural feature arises from the annealing of the R"-L" and L'-R' arm sequences, which contain spacer regions that are not complementary, leading to the formation of the unpaired central spacer (UCS) between the R"-R' and L"-L' boxes. The structure of the UCS is essential to achieve high-level recombination of the bottom strand, suggesting a dual role for this structure in active site exclusion and in hindering the reverse reaction after the first-strand exchange. The last structural feature is defined as the variable terminal structure (VTS) and corresponds to the sequence located at the end of the stem (Fig. 2C) (43). VTSs vary in length among the various *attC* sites, going from three predicted unpaired nucleotides as in $attC_{aadA7}$, to a complex branched secondary structure in the larger sites such as the *Vibrio cholerae* repeats (VCRs; the *attC* sites of the SI (12)).

**FIGURE 2** Integron recombination sites. The putative IntI1 binding domains are marked with green boxes. The black arrows show the cleavage points. (A) Sequence of the double-stranded *attI1* site: inverted repeats (R and L) and direct repeats (DR1 and DR2) are indicated with gray arrows. (B) Schematic representation of double-stranded (ds) *attC* sites; inverted repeats (R", L" L', and R') are indicated with gray arrows. The dotted lines represent the variable central part. The conserved nucleotides are indicated. Asterisks (*) show the conserved G nucleotides, which generate extrahelical bases (EHB) in the folded *attC* site bottom strand (bs). The top strand (ts) and bottom strand (bs) are marked. (C) Proposed secondary structures of the $attC_{aadA7}$ and $VCR_{2/1}$ bottom strands: structures were determined by the UNAFOLD online interface at the Institut Pasteur. Structural features of *attC* sites, namely, the Unpaired Central Spacer (UCS), the ExtraHelical Bases (EHBs), the stem and the Variable Terminal Structure (VTS) are indicated. Asterisks (*) show the conserved G extrahelical base. The conserved triplet (CT) is indicated. Primary sequences of the *attC* sites are shown (except for the VTS of the $VCR_{2/1}$ site). (D) Schematic representation of structural features of the $VCR_{2/1}$ site and their roles: the structural features and their roles are indicated. doi:10.1128/microbiolspec.MDNA3-0019-2014.f2



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

VTSs have a regulatory role through the modulation of *attC* folding (45), because the length and sequence of VTSs have an impact on the tendency of intrastrand pairing of the *attC* site and hence the propensity to form the recombinogenic structure. This influence is critical in cruciform formation (the extrusion of a hairpin on each strand of a symmetric and paired dsDNA molecule), where *attC* sites containing very large VTSs have a highly unfavorable energy to extrude and form a cruciform even in highly supercoiled DNA. It is noteworthy that during conjugation, where *attC* sites are delivered as ssDNA and folding is favored, the VTS size does not impact the recombination efficiency of the *attC* site (45).

This atypical sequence-independent recognition of ss *attC* sites readily explains how cassettes containing different *attC* sites can be efficiently recombined by the same integrase as well as how the proper orientation with respect to the $P_C$ promoter is ensured through the recombination of the bs, since the ts recombination would put promoterless genes in antisense orientation relative to $P_C$.

### The integrase

Integrases belong to the family of Y-recombinases (46), but form a specific subfamily within it, the closest relatives being the chromosome dimer resolution proteins, Xer (47). Integrases possess structural features typical of Y-recombinases, such as patches I to III, boxes I and II and the active site residues RKHRHY (48). Interestingly, compared with the rest of the members of the family, an extra domain of about 19 amino acids is present between patches II and III, containing an α-helix dubbed I2, that is essential for the activity of the protein (49). The crystal structure of the synaptic complex formed by the SI integrase of *V. cholerae* (VchIntIA) with two $VCR_{bs}$-derivative substrates, has shown that this extra domain has a specific role in the assembly of the complex, in particular in the folding of the hydrophobic pockets, that allows it to accommodate and stabilize the extrahelical bases of the *attC* sites bs (44). The catalytic domain of Y-recombinases is very conserved, and is shared with type IB topoisomerases and telomere resolvases. In all cases, one monomer can only cleave one strand, hence necessitating four monomers to perform the complete recombination reaction. The study of integrases has involved more than 50 mutants of IntI1 and VchIntIA recombinases that have been obtained through directed mutagenesis or directed evolution experiments (49–52). The study of the recombination efficiency and DNA binding affinity of the mutants has confirmed the role of catalytic residues (44), as well as the integrase-specific αI2 helix, that serves to accommodate EHBs. It is noteworthy that *attI* sites do not have EHBs, highlighting the fact that integrases recognize substrates that are structurally different. Little is known about the structural basis of this dual site-recognition, but some results suggest that it is dependent on the strength and flexibility of the interactions among integrase monomers in the synaptic complex, and that this parameter is finely balanced, because a more efficient recognition of the ds *attI* site brings a decrease in the recognition of *attC* sites as a trade-off (50).

## RECOMBINATION REACTIONS

Y-recombinases process the recombination between two DNA molecules in a set of archetypal steps. The process starts with the assembly of the synaptic complex, formed by the two DNA substrates and four monomers of the protein. It proceeds through the exchange of a set of strands from each partner molecule, forming a transient Holliday junction (HJ), and ends with the resolution of the junction through a second strand exchange (48). In this section we will see that recombination in the integron is an exception to this model. This is mainly due, as we mentioned before, to the peculiarities of *attC* sites, that impose changes in the process and the mechanism of the recombination reactions. Several of these changes are common to the recombination of the CTX phage of *V. cholerae*, because the ss genome of the phage adopts an *attC*-like structure. Nevertheless, this is not a unified recombination system, since the phage hijacks the host recombinases XerCD (53).

In the next few paragraphs we will try to give a comprehensive view of the integron by providing a detailed explanation of the recombination process followed by the biological meaning of each reaction. As an overview, three different recombination reactions are possible between the *attC* and *attI* sites: (i) the *attI*×*attC* reaction, that integrates cassettes into the integron (Fig. 3), (ii) the *attC*×*attC* reaction leading to the excision of cassettes from the array (Fig. 4); and (iii) the *attI*×*attI* reaction (Fig. 5), a rather cryptic one, that can have biological consequences in the case of multicopy MIs. The combination of excision and integration reactions leads to the shuffling of integron cassettes in the array.

### *attI*×*attC* recombination

#### Mechanistic view

The *attI*×*attC* reaction is the most efficient of the reactions catalyzed by the integrase. From a structural



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

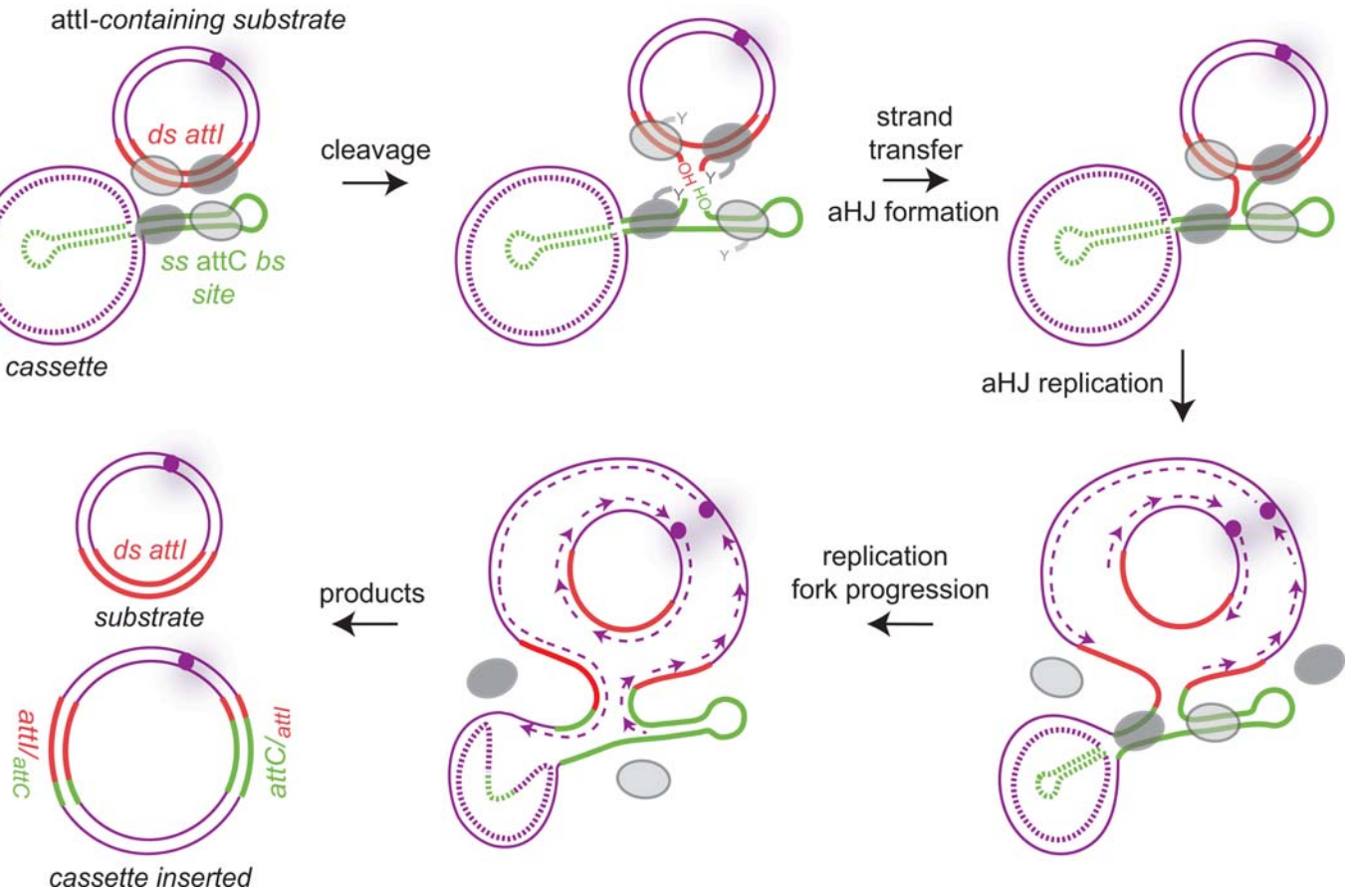



**FIGURE 3** Replicative resolution of integron cassette insertion. Recombination between a double-stranded *attI* site (bold red lines) and a single-stranded bottom *attC* site (bold green lines) terminating a cassette is shown. The top strand of the *attC* site is represented as a dotted line because we do not exactly know the nature of the cassettes (ss or ds). The synaptic complex comprises two DNA duplexes bound by four integrase protomers. The two activated protomers are represented by dark gray ovals. One strand from each duplex is cleaved and transferred to form an atypical Holliday junction (aHJ). Classical resolution gives rise to covalently closed abortive molecules. The non-abortive resolution implies a replication step. The origin of replication is represented by a purple circle and the newly synthesized leading and lagging strands by dashed purple lines. Both products are represented: the initial substrate resulting from the top strand replication, and the molecule containing the inserted cassette resulting from the bottom strand replication. Hybrid *attC* and *attI* sites are indicated. doi:10.1128/microbiolspec.MDNA3-0019-2014.f3

point of view it is an atypical one because it involves double-stranded (*attI*) and single-stranded (*attC*) substrates. This poses a mechanistic problem, since the first strand exchange generates an asymmetric and therefore atypical Holliday junction (aHJ) that cannot be resolved through the classical second strand exchange, because cutting the *attC* site twice would lead to linear and thus abortive products (Fig. 3). Hence, a second strand exchange in the bs *attC* must somehow be avoided and the aHJ must be resolved differently. The crystal structure of the synaptic complex of VchIntIA/VCRs bs revealed that the extra-helical "T" in *attC* sites acts to pull the catalytic tyrosine of one integrase monomer away from the phosphate link, avoiding the nucleophilic attack on the L box of the *attC* site (44). Resolution of the junction is then carried out by a replicative process without the involvement of a second strand exchange of any kind (Fig. 3) (54). Therefore, cassette insertion is a semiconservative process and implies the release of the initial *attI*-containing substrate.

At the nucleotide level, the DNA cleavage is localized in both substrates between the C and AA (i) on the bs



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

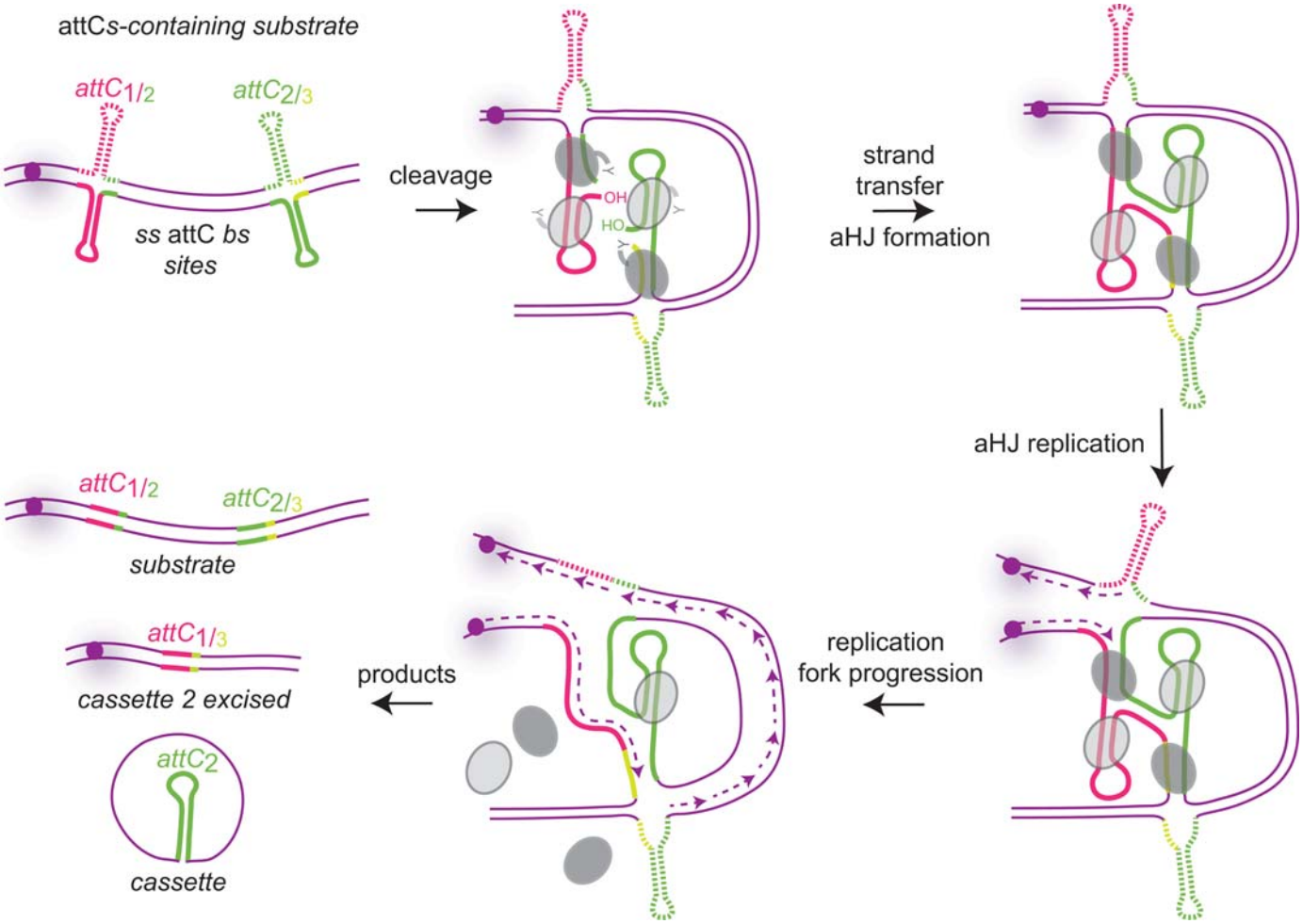



**FIGURE 4** Replicative resolution model of integron cassette excision. Recombination between two single-stranded bottom *attC* sites (bold green and pink lines) is shown. Top strands of *attC* sites are represented as dotted lines. The synaptic complex comprises two DNA duplexes bound by four integrase protomers. The two activated protomers are represented by dark gray ovals. One strand from each duplex is cleaved and transferred to form an atypical Holliday junction (aHJ). The proposed aHJ resolution model implying a replication step is based on the *attC*×*attI* recombination. The origin of replication is represented by a purple circle and the newly synthesized leading and lagging strands by dashed purple lines. Products are represented: on one hand, the initial substrate resulting from the top strand replication, and on the other, the excised cassette (cassette) and the molecule devoid of the excised cassette (cassette 2 excised) both resulting from the bottom strand replication. Hybrid *attC* sites are indicated. doi:10.1128/microbiolspec.MDNA3-0019-2014.f4

within the R box for the ds *attI* site, and (ii) on the bs within the R box for the ss *attC* site. The reaction results in the formation of a "chimeric" *attI* site in which the last six base-pairs of the R box belong to the last six base-pairs of the *attC* site's *core* of the newly inserted cassette. For the cassettes in the array, their *attC* sites will be ended by the six base-pairs of the *attC* site's *core* of the cassette previously inserted (Fig. 1A). Note that the *attC* site of the last cassette in the array will always be terminated by the last six base-pairs of the *attI* site (Fig. 1A, B).

The imperfect complementarity between *cores* and *inverse cores* contained in "chimeric" *attC* sites does not seem to impede either the propensity of *attC* sites to efficiently fold, or their stability once folded, especially since the integrase is capable of capturing *attC* sites at the beginning of the folding process, stabilizing them and facilitating their complete extrusion (55).



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

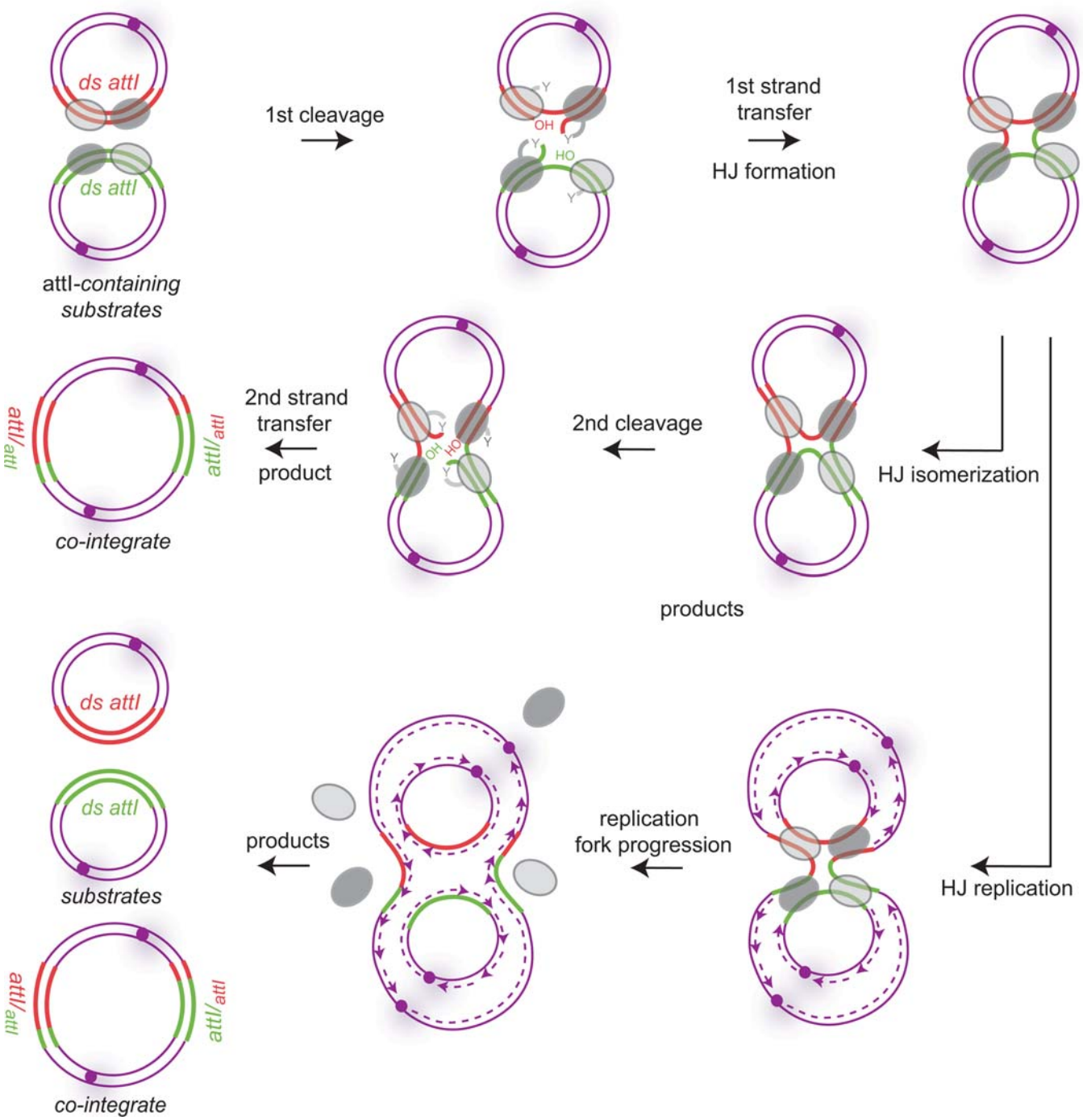



**FIGURE 5** Two resolution pathways proposed for *attI*×*attI* recombination. Recombination between two double-stranded *attI* sites (bold green and pink lines) is shown. The first proposed pathway is similar to the classical site-specific recombination catalyzed by Y-recombinases. The synaptic complex comprises two DNA duplexes bound by four recombinase protomers. The first two activated protomers are represented by dark gray ovals. One strand from each duplex is cleaved and transferred to form a HJ. Isomerization of this junction alternates the catalytic activity between the two pairs of protomers (dark and light-gray ovals) ensuring the second strand exchange and recombination product formation (co-integrate). The second pathway proposes a resolution of the HJ by replication. The origin of replication is represented by a purple circle and the newly synthesized leading and lagging strands by dashed purple lines. Products are represented: two initial substrates resulting from the top strand replication and co-integrate resulting from the bottom strand replication. Hybrid *attI* sites are indicated. doi:10.1128/microbiolspec.MDNA3-0019-2014.f5



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

### Biological meaning: cassette insertion

Insertion of a circular cassette preferentially occurs at the *attI* site of the integron, rather than at an arbitrary *attC* site within the array (56, 57). This feature is essential for the adaptive role of integrons: it ensures the expression of newly acquired promoterless cassettes from the promoter $P_C$ within the integron, allowing the immediate testing of their adaptive effect. Successive insertions of integron cassettes at the *attI* site lead to the assembly of a cassette array encoding a set of functions that have been useful for adaptation in previous conditions. Thus, it represents a *memory* of adaptive functions that have been valuable for the cell in the past and which can be recalled on demand.

Considering that the *attI*×*attC* reaction is the most efficient one, one could imagine that integrons have a tendency to empty their content by the sequential loss of cassettes in the first position, especially considering that intramolecular reactions are more efficient than intermolecular ones. Here, the different structural nature of the sites allows for an additional layer of regulation of the system. Indeed, it is extremely unlikely that two adjacent *attI* and *attC* sites will be found simultaneously in their reactive form (ds and ss, respectively). Therefore, the equilibrium in the *attI*×*attC* reaction is shifted towards intermolecular rather than intramolecular reactions; i.e., towards integration rather than excision.

Integron cassette insertion can also occur at *attC* sites. As an example, the *catB* cassette of the *V. cholerae* SI moved under experimental conditions to the second position of the array (58). Nevertheless, this type of event is probably disfavored relative to the insertion of integron cassettes at the *attI* site, especially because it requires the *attC* sites carried in the array to be ss at the same time as the incoming cassette. This likely decreases the efficiency of such intermolecular events.

## *attC*×*attC* recombination

### Mechanistic view

Recombination events between two *attC* sites located in the same cassette array lead to the excision of covalently closed integron cassettes (5, 59). Two ss folded *attC* sites must be recruited at the same time by the integrase to generate, after the first strand exchange, the HJ. Once again, the generated HJ is atypical because a second strand exchange of the bottom *attC* site strand would not lead to cassette excision but only to the exchange of the bs of *attC* sites between cassettes. Therefore, we propose, as for the *attI*×*attC* reaction, a resolution of the aHJ by a replicative, and hence semiconservative, mechanism. In this model, the replication of the recombined bs would release a covalently closed single-stranded cassette and the original DNA molecule lacking the excised cassette, while the replication of the nonrecombined top strand would reconstitute the initial substrate (Fig. 4).

If we consider the reaction at the nucleotide level, recombination between two adjacent "chimeric" *attC* sites releases a cassette containing an *attC* site with perfect complementarity between the *core* and *inverse core* (Fig. 4). Therefore, folded *attC* sites contained in circular excised cassettes have a longer stem and probably a more stable secondary structure. This feature could increase the rate of reinsertion of the excised cassettes (see Cassette shuffling below).

### Biological meaning: integron cassette excision

Integron cassette excision requires the simultaneous folding of two ds *attC* sites into their recombinogenic ss form. This event is dependent on the tendency of the site to extrude from ds DNA (see the role of the VTS, in the *attC* section), but can be favored if the fragment containing the *attC* sites is found in ss form. Once the bottom strands of both *attC* sites have recombined, the cassette is presumably separated from the ts and released from the integron via replication (Fig. 4).

### Cassette shuffling

After their excision from the array of an integron, cassettes can be reintegrated into the *attI* site. This relocation is favored, as the *attI*×*attC* reaction is the most efficient of the reactions catalyzed by integrases. Coupling the two reactions allows bacteria to render functional genes with adaptive functions that were kept silent as part of the *low-cost memory* of integrons (see the Expression section in Integron cassettes).

Cassettes excised under an ss form could theoretically be rapidly degraded by host factors and, even if the ss cassette is converted into ds form by DNA synthesis, the resulting molecule is devoid of a replication origin and so is lost after cell division. Therefore, if not rapidly reinserted, an excised cassette is not maintained in the cell. We cannot exclude that both cassette excision and insertion processes are coincidental or almost simultaneous events.

The proposed atypical mechanism of integron cassette excision can even account for cassette duplications if the cassette, after excision, is reintegrated at the *attI* site of the conserved integron (deriving from the replication of the ts). Notably, large integrons such as *V. cholerae* SI, contain duplicated cassettes. Moreover, cassette duplication may lead to generation of cassette diversity, e.g. *aadA1* and *aadA2*, which share 89.3% identity and likely arose via duplication (60).



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

### *attI*×*attI* recombination

#### Mechanistic view

The recombination between two *attI* sites has been described (57, 61), but is $10^3$ times less efficient than the *attI*×*attC* reaction. It is, however, considered to follow the same pathway with only one strand exchange and resolution through replication. The cleavage point remains in the CAA of the bs, within the R box of both *attI* sites. Nevertheless, the *attI*×*attI* reaction has structural differences with important implications for the reaction. In this case the reaction is symmetrical, involving two ds partners, for which resolution through a second strand exchange is theoretically possible, since it is neither abortive nor impeded by EHBs. We have recently studied this reaction in depth to elucidate the recombination pathway followed by the integrase when processing two symmetrical *attI1* sites (Fig. 5, HJ replication). We have observed that, in contrast to what happens in the two other reactions, both bottom and top strands of the *attI1* site are reactive. Indeed they can be transferred independently if the transfer is followed by the replicative resolution of the HJ. Interestingly, it is also possible to resolve the *attI1*×*attI1* HJ through the classical second strand exchange (Fig. 5, HJ isomerization) (Escudero et al., in preparation), as for canonical Y-recombinases. The cleavage on the ts occurs within the AAC triplet overlapping the spacer region and the L box (Fig. 2) (Escudero et al., in preparation). We have also studied the influence of each base in the L box on the recombination process and found a higher tolerance for mutation in the central CT base pairs. These results are in accordance with the work from Gravel et al. (33) in which they observed a weak contact between the integrase and the T base in the L box.

#### Biological meaning: inter-integron content rearrangement

The processing of the reaction between two *attI* sites has unveiled an unexpected flexibility of integron integrases, proving that they can recombine structurally distant substrates and switch recombination pathways accordingly. Our data challenge the rather rigid view of site-specific recombination mediated by Y-recombinases. From a biological perspective, it could be argued that this reaction has little evolutionary meaning in CIs, where there is only one copy of the platform and the *attI*×*attI* reaction would most likely produce an undesirable dimer between replicating chromosomes. Interestingly, in the clinical environment, where integrons are plasmid-borne and have become prevalent to the point of redundancy (62, 63), *attI* sites can be found in multiple copies within the cell. In this setting, the reaction between *attI* sites can have biological consequences through the rearrangement of cassette content between different integrons, and is a reaction of unknown relative importance. It is noteworthy that in the case of resolution through second strand exchange, the reaction is not semiconservative.

### Recombination at secondary sites

The *attI* and *attC* sites can recombine outside the integron at non-specific secondary sites containing GNT sequences. These events have occasionally been observed to occur naturally with *attC* sites, and have been reproduced experimentally with *attI* sites (61). The complete *aadB* integron cassette has been found inserted at a secondary site in the IncQ plasmid RSF1010 just downstream of a known promoter (64). Subsequent excision of these cassettes is unlikely to happen, leading to the stable acquisition of the gene, unless the secondary site contains the canonical GTTRRRY sequence, thereby maintaining the integrity of the newly formed *attC* site (65). In these aberrant integration events, the expression of integron cassettes is conditional upon the presence of a promoter at the insertion point (64).

Regarding *attI* sites, recombination at secondary sites was experimentally observed at very low frequency (61). This reaction disrupts the integrity of the structure, separating the cassettes from the $P_C$, possibly explaining why these reactions have not been observed in nature.

## INTEGRON CASSETTES

### Structure

Integron cassettes constitute the variable and mobile part of the integron. They usually contain a single open reading frame (ORF) devoid of a promoter, and are terminated by the *attC* recombination site recognized by the integrase. Carrying a single ORF, the size of integron cassettes is relatively small, generally lying between 500 and 1000 bp. Integron cassettes can be found in a linear ds form as part of an array within an integron, or, when excised, as a free, nonreplicative circular element that can move between integrons.

### The cassette array

The ensemble of cassettes found in CIs, their *cargo*, can vary widely from being empty to containing up to 217 cassettes (in the *Vibrio vulnificus* SI), constituting a prodigious reservoir of readily interchangeable genes. MIs typically bear a small cargo of fewer than six cassettes, with the longest array reported being of eight (24).



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

*attC* sites found in the integron cassette array of CIs generally show a high degree of identity (>80% for the VCRs in *V. cholerae*) suggesting a link between the sequence of the site and the bacterial species containing the integron—see also Cassette genesis (6, 7, 11). In contrast, cassettes in MIs show diversity in the length and sequence of their *attC* sites, as well as an inconsistent codon usage in the ORFs encoded. Hence, it is plausible that MIs have access to the vast pool of cassettes found in CIs and that they gather cassettes from different genomic backgrounds. Integron cassettes contained in MIs could therefore be seen as representatives of a specific CI in the environment (6).

## Expression

As genes in integron cassettes are generally promoterless, their expression is ensured only when inserted into an integron at the *attI* site, by the proximity of the external $P_C$ promoter carried within the nonmobile functional platform of the integron (Fig. 1). The $P_C$ promoter is located either in the *intI* gene or in the *attI* site. *attC* sites, through their imperfect symmetry, ensure correct orientation of cassette-borne genes relative to the $P_C$ promoter when inserted into *attI* sites. Most studies on cassette expression have been performed using the class I integron system, in which it has been shown that the $P_C$ promoter is located within the *intI1* coding sequence. Occasionally, it may be combined with a second promoter, termed $P_C2$, and located in *attI1*. $P_C$ variants of different strengths have been identified for both $P_C$ and $P_C2$ promoters: thirteen for $P_C$ and three for $P_C2$ (2, 4). There are five main $P_C$–$P_C2$ combinations defining five levels of promoter strength. The diversity of strength of these promoters could mediate a differential expression for an identical array of cassettes. The distance between the $P_C$ promoter and integron cassettes affects their expression level. Indeed, expression levels are maximal for the first gene in the array, and gradually decrease for those following, a phenomenon dependent on the distance to the promoter and on the nature of the inserted cassettes (4). For a long time, it was thought that this feature was due to the ability of the folded *attC* sites to act as Rho-independent transcription terminators, impeding the transcription of the full cassette array (4). Finally, contrary to expectations, the transcription of integron cassettes was found not to be affected by the folded *attC* sites. Indeed, mutations of *attC* sites revealed that destabilization of the *attC* secondary structures in the transcript could enhance the expression of the 3′ gene at the translational level but not affect its expression at the transcriptional level. In particular, the presence of a translated ORF was shown to increase translation of the 3′-located gene (66). These results might reflect the capacity of the folded *attC* sites to impede ribosome progression (7, 66), and therefore explain, in part, the expression gradient observed in integron cassette arrays. Altogether, the dependency on $P_C$ for expression of cassettes allows the integron to provide the cell with an adaptive array of functions at low fitness cost. With a large portion of the cassette cargo found too far from the $P_C$ promoter to be expressed, these cassettes are carried at the minimum cost possible, the cost of replication. Although silent, these cassettes remain available for the cell if needed (see cassette reshuffling).

The expression of integron cassettes is also governed by the presence of a binding motif initiating the assembly of ribosomes (67). Some cassette-borne genes are preceded by their own ribosome-binding site (RBS), while others are devoid of this motif. For the latter, translation can be initiated at an upstream RBS site. In class I integrons, a small ORF (*orf11*) preceded by a functional RBS has been found in the *attI1* site. This RBS is present in all transcripts generated from the $P_C$ promoters, and accounts for a significant part of the expression of cassettes devoid of an RBS (68).

Note also that, although rare, integron cassettes sometimes harbor their own promoter, e.g. the *cmlA1* chloramphenicol resistance gene of the In4 class I integron (69, 70), the *ere*(A) erythromycin resistance gene of the pIP1100 plasmid-borne class II integron (71), the *qnrVC1* quinolone resistance genes found in the class I integron (72) and the toxin–antitoxin (TA) gene pairs found in the *V. cholerae* SI—see below for more details (15, 73, 74). The expression of these integron cassettes is assured, regardless of their position in the array.

## Diversity and functions

Integron cassettes seem ubiquitous—they have been recovered from every environment investigated, including soil, riverine sediment, seawater, biofilms, plant surfaces and even eukaryotes' symbionts (36, 75–81).

As mentioned in the Introduction, distinct sets of functions are encoded by cassettes found in CIs and in MIs. On the one hand, MIs carry, almost exclusively, antimicrobial resistance genes. For instance, class I integrons have been associated with more than 130 integron cassettes comprising resistance determinants against almost all antibiotic families, including β-lactams, all aminoglycosides, trimethoprim, chloramphenicol, streptothricin, fosfomycin, macrolides, rifampin, quinolones and antiseptics of the quaternary ammonium-compound family (25, 44). On the other hand, CIs mainly contain



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

cassettes of unknown function. Combined analyses of metagenomic and CI cassettes from *Vibrio* species revealed that up to 65% of cassette-encoded proteins had no known homologs and that 13% had homologs of unknown function (10). The remainder showed a wide range of non-specific functions in metabolism, cellular processes, and information storage. Moreover, cassettes in *Vibrio rotiferianus* DAT722 have been found to be involved in host surface polysaccharide modifications suggesting that integron cassettes may be important in processes such as bacteriophage resistance, adhesion/biofilm formation, protection from grazers and bacterial aggregation (82). Other data also reveal functions mediating interactions with the external environment (i.e., the presence of a signal peptide region, or signatures of multiple transmembrane domains (15, 77)).

### Toxin–antitoxin cassettes

Among the cassettes of CIs, a distinct type is notable, encoding members of TA families (15, 83, 84). TAs are addiction systems encoding a stable toxin and its labile neutralizing antitoxin. The perpetuity of the system in the genome is assured by the difference in half-lives of toxin and antitoxin. If the TA system is lost, the antitoxin degrades first and the cell suffers toxin-mediated postsegregational killing. TAs have been classified, depending on the nature and mode of action of the antitoxin, into three types: types I and III in which antitoxins are small RNAs that impede the translation of the toxin; and type II systems, in which the antitoxin is a protein that inhibits the toxin through protein complex formation (85). Seventeen cassettes carrying type II TAs have been found in the SI of *V. cholerae* N16961 ((2), Iqbal et al., in preparation). A peculiarity of nine of them is that they are integrated in the opposite orientation relative to the array, and transcription is ensured by their own promoter. Therefore, their expression is independent of the promoter(s) contained in the integron platform. Our laboratory has studied the TA array in *V. cholerae*, and found a remarkable specificity of every toxin for its cognate antitoxin, with no cross talk between toxins and antitoxins of different systems, even between those belonging to the same family (Iqbal et al., in preparation).

From these observations it is tempting to speculate that these TA systems play a role in the stabilization of SIs, preventing the loss of silent cassettes and allowing for the formation of large arrays (15, 74). However, TAs have also been shown to mediate phage resistance (86), and a dual role for these elements cannot be ruled out.

### Conventional annotation

Several web databases have been developed to provide easy access to integron integrase and cassette DNA sequences. Among them, RAC is specialized in the annotations of integron cassettes mainly encoding antibiotic resistance (87), Integrall database has listed more than 8,500 cassettes ((88), http://integrall.bio.ua.pt/), and ACID has collected and stored in its first version 5,622 integron cassettes (89). Annotation systems allowing cassette identification have also been developed, such as XXR ((http://mobyle.pasteur.fr) (15)).

## A SYSTEM INTIMATELY CONNECTED TO CELL PHYSIOLOGY

Integrons allow bacteria to adapt rapidly by granting access to an almost infinite array of functions encoded in cassettes. But the elegance of this genetic platform is more easily appreciated in the context of the complex and subtle coupling of its activity to the physiology and needs of its host. In this section we will try to give an overview of the integron–host connection, of which Fig. 6 provides a snapshot scheme that should help the reader throughout the following paragraphs.

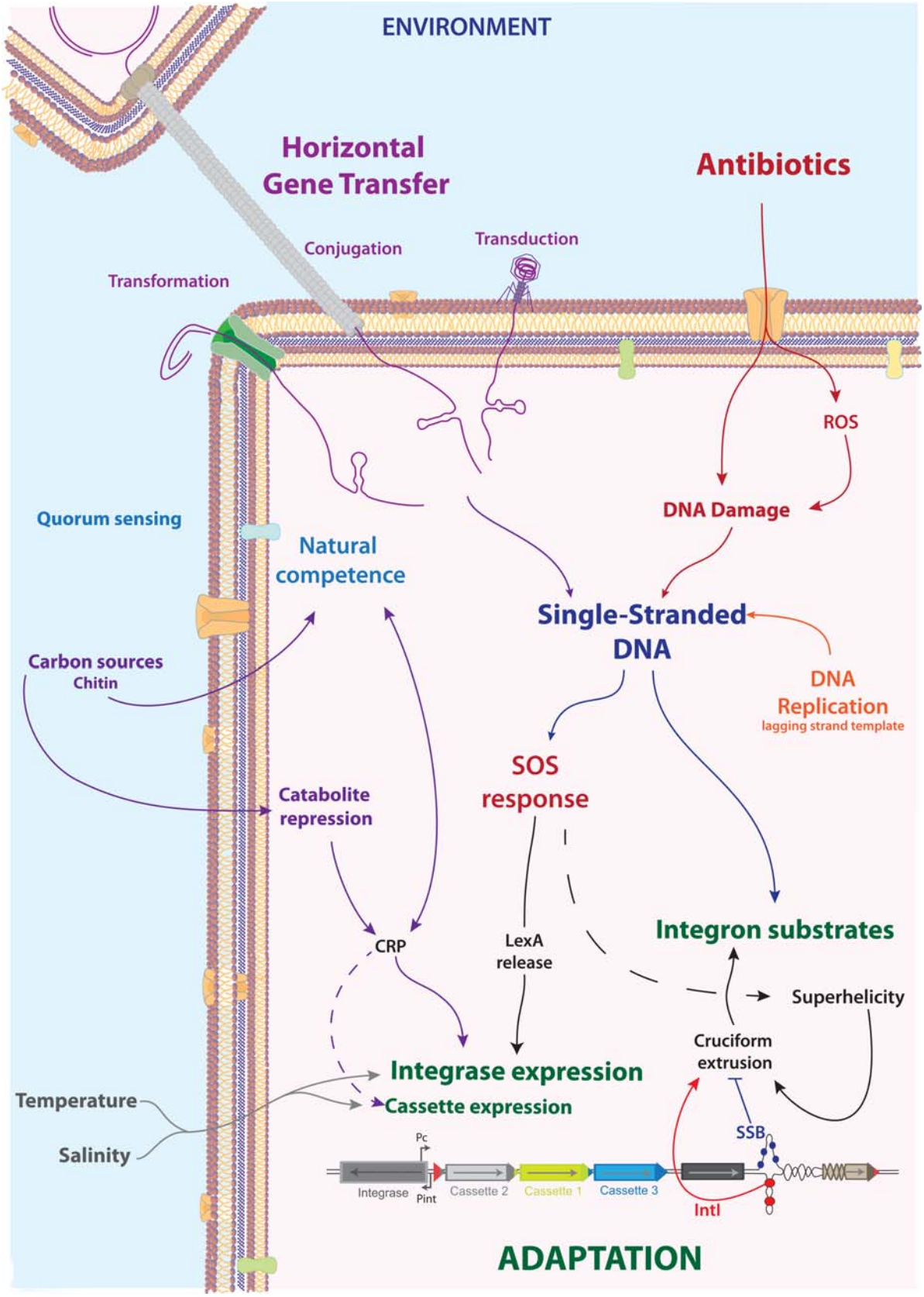
### Expression of the integron elements

#### The integrase

Like all elements promoting genetic variation in a cell, integrons must be well controlled (domesticated) to avoid the deleterious effects of an overwhelming recombination activity. The regulation of the integrase has remained elusive until recently, but it had been noticed that integrons and their cassette content are stable in laboratory conditions, while SIs are the most variable loci in the genome of natural isolates of *V. cholerae* (90–93). This paradox suggested a controlled regulation of integrase expression. Careful analysis of the promoter region of integrases from CIs and MIs revealed the presence of a LexA-binding motif overlapping the –10 box of the *intI* promoter. LexA is the master repressor of the SOS response, a widespread regulatory network aimed at addressing DNA damage by repairing or bypassing lesions (94). LexA represses the SOS regulon by allosteric interference with the RNA polymerase, through its binding to specific DNA motifs (LexA boxes) in the promoter regions of the regulated genes. DNA damage and repair ultimately produce an increase of ss DNA in the cell that is recognized by RecA and promotes its polymerization. RecA nucleofilaments induce autocatalysis of LexA, releasing the repression of the SOS genes and triggering the SOS response. The functionality of the LexA box has



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20






Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

been assessed *in vitro* for the class 1 integron and the *V. cholerae* CI. The SOS response increases 4.5-fold the expression of the class 1 integrase, and 37-fold that of the *V. cholerae* integrase (95). The LexA box has been identified in the sequences of a large proportion of integrases in the databases, suggesting that the SOS control of integrase expression is a conserved feature in integrons (96).

The control of integrase expression by the SOS response allows for a subtle coupling of integron activity with bacterial physiology. On one hand, by linking the activity of the integron to an alarm signal, recombination events are limited to the times when bacteria need to evolve and adapt. It is at these precise moments that cassette acquisition or reshuffling can have a dramatic impact on cell survival. Repressing the expression of the integrase when the bacterium is well adapted is also beneficial, or even necessary, because it prevents the reshuffling of cassettes when the configuration of the array is optimal and avoids random recombination events at secondary sites that could be deleterious (97, 98). The selective pressure leading MIs to exclusively capture and spread antimicrobial genes is clearly understood through this link to the SOS response. Indeed MIs carry resistance genes against most classes of antibiotics. Some antibiotic molecules, such as quinolones and trimethoprim, target DNA-related systems (DNA topoisomerases and nucleotide synthesis, respectively). The prevalence of cassettes containing resistance genes against such antibiotics can be explained through the link between the DNA damage produced by the antibiotic and the SOS-induced integron. A matter that had remained elusive was the presence of cassettes containing genes that confer resistance against antibiotics that do not damage DNA and do not induce the SOS response in *Escherichia coli*. This intriguing question led us to the discovery that aminoglycosides and β-lactams, antibiotics targeting the ribosome and the cell wall respectively, do induce the SOS response (even at subinhibitory concentrations) in species other than *E. coli*, such as *V. cholerae*, *Klebsiella pneumoniae* and *Photorhabdus luminescens* (99–101). The underlying mechanism for this seems to be the generation of reactive oxygen species intermediates that oxidize nucleotides and cause DNA damage, ultimately triggering the SOS response.

On the other hand, there is another aspect of the coupling to the SOS response that underscores the subtle adaptation of the integron to the host's needs: ssDNA is simultaneously the triggering signal for integron activity (through the SOS response) and its substrate. As we will see in depth in this chapter, one of the main sources of ssDNA and therefore drivers of integron activation is the entrance of new DNA during conjugation and natural transformation (58, 102), enabling screening of the incoming DNA for new cassettes to capture.

In the case of the *V. cholerae* integrase coding gene (*intIA*), a second layer of control, through the cAMP receptor protein (CRP), has been identified. CRP is the master regulator of the carbon catabolite repression response, adapting cellular metabolism to the type of carbon sources available in the environment. A CRP binding box is present in the promoter region of *intIA* (between $P_{int}$ and $P_C$), controlling its expression independently of the SOS response and connecting the integron to environmental conditions. It is noteworthy that CRP is also directly and indirectly linked to the uptake of ssDNA, via TfoX and HapR (regulators of natural competence), respectively (102). Finally, low-level induction of the expression of *intIA* has also been observed at high temperature (42°C) (103).

The mechanisms of integrase expression regulation reveal that the integron is intimately connected to bacterial physiology and the environment, allowing it to be active when evolving is either mandatory or possible.

## The cassettes

As seen previously, cassettes are expressed from the $P_C$ promoter in the integron. The extensive work of Ploy's laboratory on the expression of cassettes in the class I integron has led to a more complex view of the system, with several strength variants of the $P_C$ promoter and an indirect influence of this promoter on the expression of the integrase (2). Indeed, by facing each other, $P_{int}$ and $P_C$ are subjected to transcriptional interference. Strong $P_C$ variants will have a negative effect on $P_{int}$ transcription when the latter is de-repressed during SOS induction. It has been suggested that this can serve as a repression system for the integrase in bacterial species lacking LexA, such as *Acinetobacter baumannii* (Couvé-Deacon, personal communication), where the

**FIGURE 6** Intimate connection between the integron and cell physiology. A snapshot representation of the links between integrons' activity and bacterial physiology is shown. The main triggering signal for integrase expression is the bacterial SOS response. A detailed description of these connections is depicted in the section entitled: *A system intimately connected to cell physiology.* doi:10.1128/microbiolspec.MDNA3-0019-2014.f6



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

unregulated expression of the integrase is harmful (98). Also, when present, the $P_C2$ promoter alters the LexA box, disrupting the SOS regulation of the integrase, but does not interfere transcriptionally with $P_{int}$, probably because they are close enough for their transcription starts (+1) not to face each other (104). The transcriptional interference between $P_C$ and $P_{int}$ establishes a trade-off between the expression of cassettes and of the integrase. The higher prevalence of weak $P_C$ variants among clinical and nonclinical *E. coli* isolates, and bacteria from wastewater environments, suggests that a flexible cassette array is more important than enhanced expression (105, 106). Since $P_C$ is encoded within the class 1 integrase, $P_C$ variants also have an impact on the sequence of *intI1*. Interestingly, regardless of the transcriptional interference between promoters, the amino acid substitutions associated with the presence of strong promoters entail a decrease in the excision (but not in the integration) activity of the integrase (2).

The analysis of the $P_C$ in the *V. cholerae* SI has recently been carried out in our group (103). As stated before, a CRP binding box is present between $P_{int}$ and $P_C$, from where the cAMP–CRP complex activates not only the integrase, but also the expression of the cassette array. We have tested several conditions mimicking the different environments in which *V. cholerae* thrives, including different temperatures and salinities. We have observed a CRP-dependent increase in the expression levels of the cassette array during entry into stationary phase, as well as a correlation with temperature.

## *attC* site folding

We previously described that *attC* sites are recognized and recombined as ss folded structures, a peculiarity conferring certain advantages to the cell. Indeed, secondary structures represent a way to expand information storage in DNA, in addition to the primary base sequence (107), and more importantly they allow the cell to control integron recombination through a variety of physiological processes that normally regulate the formation of secondary structures. This makes the integron an integrated system. Regulation of *attC* sites is controlled differently depending on whether the site folds as a hairpin from an ss molecule, or as a cruciform from dsDNA.

### Single-strand pathway

Large fragments of ssDNA, from which *attC* sites can easily fold, can be found in the cell during some of its physiological processes or during the entry of exogenous DNA. Indeed, all horizontal gene transfer mechanisms, namely natural transformation, conjugation and transduction, involve the entry of only one strand of DNA into the recipient cell. Conjugative transfer of DNA containing *attC* sites favors the folding of *attC* sites of various lengths, and hence their recombination (45). At the same time, conjugation induces the expression of the integrase through the SOS response (58), a phenomenon also observed during transformation (102). Altogether, this allows the integron to recruit incoming cassettes.

Three essential cellular processes are a source of ssDNA: transcription, DNA repair, and replication. During replication, the lagging strand is synthesized discontinuously through the assembly of Okazaki fragments, a process in which short stretches of up to 2 kb of the template strand are found transiently in ss form, possibly favoring the folding of *attC* sites. Since the leading strand is synthesized continuously, it is possible that the effect of replication on *attC* site folding depends on the orientation of the integron relative to the replication fork. Indeed, published observations suggest that secondary structures are more easily formed when carried on the template of the lagging strand compared with the template of the leading strand (108). Accordingly, we have demonstrated that the *attC* bs is recombined with an *attI* site at a higher rate when carried on the template of the lagging strand. This suggests that the orientation of the cassette array in the chromosome influences the recombination frequency of *attC* sites, and more generally, that the cellular availability of ssDNA directly impacts the efficiency of integron recombination (107). It is noteworthy that CIs carry the *attC* site bs always on the leading strand template, limiting the excision of cassettes (45). The limited length of ssDNA stretches on the lagging strand template should favor cassette excision events between *attC* sites at less than 2kb distance from each other, i.e. those resulting in the excision of a limited number of cassettes (1 or 2).

Interestingly, ssDNA production can be triggered in response to stress or specific environmental determinants. For example, in the Gram-positive bacterium *Streptococcus pneumoniae*, competence (and therefore single-strand availability through DNA uptake) is induced by an antibiotic stress response (109). In the Gram-negative human pathogen, *Helicobacter pylori*, competence is increased in conditions of low $CO_2$ and high pH (110).

### Double-strand pathway

Another pathway for *attC* site structuring is its extrusion as a cruciform from a dsDNA molecule, a process largely dependent on superhelicity. Formation of cruciforms by annealing of inverted repeats involves deep structural



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

disruption and reorganization of base pairing. Until recently, only extrusions from perfect palindromes have been observed *in vivo*, because imperfections have major negative effects on the overall dynamics of cruciform extrusion (111). We have proven the cruciform extrusion of *attC* sites from dsDNA, including sites with very large VTSs such as VCRs. The positive impact of superhelicity on *attC* folding has been established using topoisomerase I-deficient and gyrase-deficient (the enzymes maintaining supercoiling levels in *E. coli*) strains (112, 113).

Cruciforms are seldom extruded at significant rates under average *in vivo* supercoiling conditions. However, many factors may transiently increase local superhelical density to a level sufficient for cruciform extrusion (107), and hence can impact integron recombination by favoring *attC* site extrusion. Among these factors, we can find biological processes such as transcription and replication (114); growth phase; environmental stimuli, such as antibiotics and growth conditions (115–117); and/or internal stimuli such as the induction of the SOS response that entails higher levels of negative supercoiling (118). Finally, bacterial species can present different levels of superhelicity, resulting in different levels of integron recombination efficiencies according to the host strain (119).

### SSB and integrase duality

Bacteria have to find a subtle balance between the benefit provided by encoding biological functions in secondary structures and the detrimental effects of possessing an excess of them. Indeed, overly long and stable palindromes could not be maintained *in vivo,* either because they are inviable (i.e. intrinsically toxic to the cell) or because they are genetically unstable (i.e. partially mutated or deleted) (120). It is assumed that *inviability* is caused by an arrest of the replication fork, as it is unable to process these secondary structures, and *instability* is caused by the presence of proteins such as SbcCD, which destroys these structures. This leads to constraints on the size and perfection of the inverted repeats and/or the need of host factor regulation of secondary structure folding.

Among these host factors, the single-stranded DNA-binding protein (SSB) is able to bind cellular ssDNA without sequence specificity. SSB plays extensive cellular roles in DNA replication, repair, and homologous recombination, where it prevents premature annealing, stabilizes and protects the single-stranded DNA, and removes secondary structures (121) by migrating along ssDNA and melting unstable hairpins while stimulating RecA filament elongation (122). SSB assures the stability and viability of *attC* sites by flattening the hairpin. We also demonstrated that the integrase could capture *attC* sites at the moment of their extrusion, efficiently stabilizing and recombining them. Therefore, when expressed, the integrase is able to counteract the effect of SSB (55, 123, 124).

### Holliday junction resolution

As we have seen, the *attC*×*attI* reaction, which leads to the integration of cassettes, forms an atypical HJ that is resolved through replication, and this is likely to also be the case for the *attC*×*attC* reaction. However, it is unknown whether the replicative resolution is a passive process, involving a replisome that was assembled at the replication origin and implying that the aHJ remains stable until its arrival; or an active one, for which a new replisome is assembled *ad hoc*. The set of proteins involved would be different for both cases: in the first scenario, in *E. coli* it would include DnaA (the replication initiation factor that promotes the unwinding or denaturation of DNA), DnaE (the catalytic subunit of the *pol*III polymerase essential for processive replication) and DnaN (the clamp, presumably essential for all kinds of replication). The second scenario considers that the aHJ could mimic an arrested replication fork, thus involving the local recruitment of replication complexes able to restart halted forks. In bacteria, this essential activity is orchestrated by the PriA DNA helicase, which identifies stalled replication forks via structure-specific DNA binding. We propose that DNA replication start could be initiated by PriA at the aHJ, permitting its resolution (125, 126).

### Accessory host factors

It has been observed that, contrary to the class I integron system, which recombines equally in *E. coli* and *V. cholerae*, the VchIntIA integrase of the *V. cholerae* CI recombines at a 2,600-fold higher rate the VCR sites in its original genomic background than in *E. coli* strains (37). Such results suggest either the involvement of one or more host factors in *V. cholerae* that would be absent or too divergent in *E. coli* strains, or the presence of an inhibitory factor in *E. coli*. The nature of CIs and MIs suggests that the first case is more likely, and that the lack of dependence on host factors is what has made class 1 integrons so successful upon mobilization.

## A BIOTECHNOLOGICAL TOOL

Several applications of the integron as a bioengineering tool have been developed. For instance, it has been suggested to harness the capacity of the integron to in-



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

corporate new cassettes to perform sequence-independent recovery of integron cassettes from genomic libraries (127). The particular utility of this tool lies in its ability to access integron cassettes from unculturable organisms, and the interest of their retrieval is supported by a recent analysis suggesting that the integron cassette metagenome contains a repertoire of genes belonging to new, currently uncharacterized protein families with possible novel functions (128).

An integron-based cloning technique has also been developed, based on the delivery of synthetic cassettes into the genetic backbone of both MIs and CIs via natural transformation (129). This technique can be used with wild-type environmental bacteria, does not require a vector, can yield high frequencies of recombinants in favorable conditions, and recombinants are stable in the absence of selection. The lack of a cloning vector and associated antibiotic resistance genes makes the technique particularly appealing from the biosafety perspective.

The capacity of the integron system to rearrange cassettes has led to its application as an *in vivo* genetic shuffling device (130). In fact, the synthetic integron allows the generation of a large number of genetic combinations and arrangements using site-specific recombination, which is of particular interest for metabolic pathway optimization, where the selection of optimal arrangements of genetic elements can lead to higher production yields. Moreover, the flexibility of *attC* sites suggests the possibility of designing synthetic *attC* sites "à la carte" and embedding them into elements having a distinct function on the sequence level, for instance coding for a protein or a promoter. This avenue of synthetic integron development is being pursued, and might lead to new applications of the system for bioengineering purposes.

## EVOLUTIONARY IMPLICATIONS

Integrons are powerful agents of bacterial evolution, granting access to a vast variety of functions encoded in cassettes. As we have seen, integrons are seamlessly coordinated with the physiology of bacteria, so that evolution becomes a somewhat domesticated function rather than a stochastic inevitability. The power of integrons has changed our view on the adaptive capabilities of bacteria as evidenced by their role in the unforeseen rise of multiresistance during the 1960s.

### On the success of MIs

Antibiotic pressure of anthropogenic origin has been extraordinarily high during the last 70 years, leading to the selection of integron mobilization events. Indeed, integrons were not originally present in the genomes of pathogenic enterobacteria where they are today in some cases redundant. It is through the association of integrons with transposable elements and conjugative plasmids that integrons have entered circulation among clinically relevant strains. Yet, it remains unclear why not all MIs have had the same spread or impact on antimicrobial resistance. The success of horizontal gene transfer events has been suggested to depend on the *penetration*, *promiscuity*, *plasticity* and *persistence* capacities (the four Ps) of any given determinant (131). It is not known whether the success of the different MI classes is due to the platform to which each class is associated, and/or to intrinsic characteristics of the integron. Mobilizing platforms can affect the penetrance and promiscuity of MIs through differential transposition/conjugation rates or host range; and their persistence through differences in the fitness cost they impose on the host. On their side, it is also possible that MI-integrases show differential properties, such as distinct rates of activity or differences in the range of *attC*-sites recognized. Broader substrate recognition has indeed been observed to be the case for IntI1 compared with VchIntIA and IntIPstQ from *Pseudomonas stutzeri* (37, 132). It is tempting to speculate that the success of class 1 integrons could be due, at least partly, to IntI1 recognizing a broader variety of *attC* sites and conferring access to a larger portion of cassette-encoded functions. Regarding the influence of integron characteristics in their persistence, it could be argued that MI are stabilized by high antibiotic pressure alone. Nevertheless, the picture we have of CIs suggests that the stability of integrons is probably assured through both regulation of integrase activity by host regulatory networks (98), and control over *attC* site folding (SSB flattening the folded *attC* sites to avoid their instability and/or their unviability, integrase stabilizing them, supercoiling, horizontal gene transfer and replication favoring their folding).

### Genesis of integron cassettes

We now have a good understanding of the recombination reactions catalyzed by integrases, and an increasing knowledge of the connection of the integron with the host. Still, some major questions remain unanswered, such as the *de novo* creation of cassettes. In cassettes, the general absence of promoters, together with the paucity of pseudogenes or noncoding sequence, has been interpreted to be representative of an RNA origin (133). In this model, the transcriptional terminator of the gene of interest is fused to an *attC* site through homologous recombination and further retrotranscribed into a DNA cassette. This is performed by a Group IIC-*attC* intron,



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

a type of intron with affinity for palindromic sequences including *attC* sites, as well as retrotranscriptase activity (133). Unfortunately, direct evidence supporting this model remains scarce, and these elements have seldom been found within integrons (133, 134). Although the rationale for the RNA origin of cassettes is interesting, some arguments against this theory are also solid. For instance, in this scenario, what mechanism would allow for the creation of cassettes that do encode promoters, such as TA cassettes? Indeed, if the RNA origin of cassettes were to be true, at least one additional mechanism would be needed to produce cassettes from DNA so that promoters are present. Also, and bearing in mind the millions of years of evolution of integrons and the subtleties of their intertwining with the host machinery, it seems unlikely that a function as important as the creation of cassettes relies on an independent entity that is found in a very low percentage of integrons. Instead, if Group IIC-*attC* introns were cassette generators, one would expect them to be found as part of the constant platform of integrons. This should be especially the case in CIs in which *attC* sites show high levels of identity, a feature suggesting a link of some kind between cassette creation and the host. This is the case, as we have already mentioned, of the VCRs within the SI of *V. cholerae*, a species in which the presence of Group IIC-*attC* introns has not been reported. All in all, the creation of cassettes remains a subject of the utmost importance for understanding integrons, for which an undoubted model is not available.

## ACKNOWLEDGMENTS

This work was supported by the Institut Pasteur, the Centre National de la Recherche Scientifique (CNRS-UMR3525), the European Union Seventh Framework Programme (FP7-HEALTH-2011-single-stage), the "Evolution and Transfer of Antibiotic Resistance" (EvoTAR) and the French Government's Investissement d'Avenir program Laboratoire d'Excellence "Integrative Biology of Emerging Infectious Diseases" (ANR-10-LABX-62-IBEID) and the French National Research Agency (ANR-12-BLAN-DynamINT). JAE is supported by the Marie Curie Intra-European Fellowship for Career Development (FP-7-PEOPLE-2011-IEF).

The authors acknowledge Dr Jason Bland and Henry Kemble for critical reading of the review.

## REFERENCES

**1. Stokes HW, Hall RM.** 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions: integrons. *Mol Microbiol* **3:**1669–1683.

**2. Jové T, Da Re S, Denis F, Mazel D, Ploy MC.** 2010. Inverse correlation between promoter strength and excision activity in class 1 integrons. *PLoS Genet* **6:**e1000793.

**3. Levesque C, Brassard S, Lapointe J, Roy PH.** 1994. Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. *Gene* **142:**49–54.

**4. Collis CM, Hall RM.** 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob Agents Chemother* **39:**155–162.

**5. Collis CM, Hall RM.** 1992. Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Mol Microbiol* **6:**2875–2885.

**6. Rowe-Magnus DA, Guerout AM, Ploncard P, Dychinco B, Davies J, Mazel D.** 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. *Proc Natl Acad Sci U S A* **98:**652–657.

**7. Cambray G, Guerout AM, Mazel D.** 2010. Integrons. *Annu Rev Genet* **44:**141–166.

**8. Nemergut DR, Robeson MS, Kysela RF, Martin AP, Schmidt SK, Knight R.** 2008. Insights and inferences about integron evolution from genomic data. *BMC genomics* **9:**261.

**9. Cambray G, Mazel D.** 2008. Synonymous genes explore different evolutionary landscapes. *PLoS Genet* **4(11):**e1000256.

**10. Boucher Y, Labbate M, Koenig JE, Stokes HW.** 2007. Integrons: mobilizable platforms that promote genetic diversity in bacteria. *Trends Microbiol* **15:**301–309.

**11. Mazel D.** 2006. Integrons: agents of bacterial evolution. *Nature RevMicrobiol* **4:**608–620.

**12. Mazel D, Dychinco B, Webb VA, Davies J.** 1998. A distinctive class of integron in the *Vibrio cholerae* genome. *Science* **280:**605–608.

**13. Ogawa A, Takeda T.** 1993. The gene encoding the heat-stable enterotoxin of *Vibrio cholerae* is flanked by 123-base pair direct repeats. *Microbiol Immunol* **37:**607–616.

**14. Barker A, Clark CA, Manning PA.** 1994. Identification of VCR, a repeated sequence associated with a locus encoding a hemagglutinin in *Vibrio cholerae* O1. *J Bacteriol* **176:**5450–5458.

**15. Rowe-Magnus DA, Guerout AM, Biskri L, Bouige P, Mazel D.** 2003. Comparative analysis of superintegrons: engineering extensive genetic diversity in the Vibrionaceae. *Genome Res* **13:**428–442.

**16. Mitsuhashi S, Harada K, Hashimoto H, Egawa R.** 1961. On the drug-resistance of enteric bacteria. 4. Drug-resistance of Shigella prevalent in Japan. *Jpn J Exp Med* **31:**47–52.

**17. Martinez E, de la Cruz F.** 1988. Transposon Tn21 encodes a RecA-independent site-specific integration system. *Mol Gen Genet* **211:**320–325.

**18. Arakawa Y, Murakami M, Suzuki K, Ito H, Wacharotayankun R, Ohsuka S, Kato N, Ohta M.** 1995. A novel integron-like element carrying the metallo-beta-lactamase gene blaIMP. *Antimicrob Agents Chemother* **39:**1612–1615.

**19. Collis CM, Kim MJ, Partridge SR, Stokes HW, Hall RM.** 2002. Characterization of the Class 3 integron and the site-specific recombination system it determines. *J Bacteriol* **184:**3017–3026.

**20. Xu H, Davies J, Miao V.** 2007. Molecular characterization of class 3 integrons from *Delftia* spp. *J Bacteriol* **189:**6276–683.

**21. Ramírez MS, Piñeiro S, Centrón D.** 2010. Novel insights about class 2 integrons from experimental and genomic epidemiology. *Antimicrob Agents Chemother* **54:**699–706.

**22. Hochhut B, Lotfi Y, Mazel D, Faruque SM, Woodgate R, Waldor MK.** 2001, Molecular analysis of antibiotic resistance gene clusters in *Vibrio cholerae* O139 and O1 SXT constins. *Antimicrob Agents Chemother* **45:**2991–3000.

**23. Sørum H, Roberts MC, Crosa JH.** 1992. Identification and cloning of a tetracycline resistance gene from the fish pathogen *Vibrio salmonicida*. *Antimicrob Agents Chemother*. **36:**611–615.

**24. Naas T, Mikami Y, Imai T, Poirel L, Nordmann P.** 2001. Characterization of In53, a class 1 plasmid- and composite transposon-located integron of *Escherichia coli* which carries an unusual array of gene cassettes. *J Bacteriol* **183:**235–249.



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

**25. Partridge SR, Tsafnat G, Coiera E, Iredell JR.** 2009. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol Rev* **33:**757–784.

**26. Fluit AC, Schmitz FJ.** 2004. Resistance integrons and super-integrons. *Clin Microbiol Infect* **10:**272–288.

**27. Labbate M, Case RJ, Stokes HW.** 2009. The integron/gene cassette system: an active player in bacterial adaptation. *Methods Mol Biol* **532:**103–125.

**28. Martin C, Timm J, Rauzier J, Gomez-Lus R, Davies J, Gicquel B.** 1990. Transposition of an antibiotic resistance element in mycobacteria. *Nature* **345:**739–743.

**29. Nandi S, Maurer JI, Hofacre C, Summers AO.** 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.

**30. Nesvera J, Hochmannová J, Pátek M.** 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. *FEMS Microbiol Lett* **169:**391–395.

**31. Shi L, Zheng M, Xiao Z, Asakura M, Su J, Li L, Yamasaki S.** 2006. Unnoticed spread of class 1 integrons in gram-positive clinical strains isolated in Guangzhou, China. *Microbiol Immunol* **50:**463–467.

**32. Hocquet D, Llanes C, Thouverez M, Kulasekara HD, Bertrand X, Plésiat P, Mazel D, Miller SI.** 2012. Evidence for induction of integron-based antibiotic resistance by the SOS response in a clinical setting. *PLoS Pathog* **8:**e1002778.

**33. Gravel A, Fournier B, Roy PH.** 1998. DNA complexes obtained with the integron integrase IntI1 at the attI1 site. *Nucleic Acids Res* **26:**4347–4355.

**34. Partridge SR, Recchia GD, Scaramuzzi C, Collis CM, Stokes HW, Hall RM.** 2000. Definition of the attI1 site of class 1 integrons. *Microbiology* **146:**2855–2864.

**35. Nield BS, Holmes AJ, Gillings MR, Recchia GD, Mabbutt BC, Nevalainen KM, Stokes HW.** 2001. Recovery of new integron classes from environmental DNA. *FEMS Microbiol Lett* **195**: 59–65.

**36. Elsaied H, Stokes HW, Kitamura K, Kurusu Y, Kamagata Y, Maruyama A.** 2011. Marine integrons containing novel integrase genes, attachment sites, attI, and associated gene cassettes in polluted sediments from Suez and Tokyo Bays. *ISME J* **5:**1162–1177.

**37. Biskri L, Bouvier M, Guérout AM, Boisnard S, Mazel D.** 2005. Comparative study of class 1 integron and *Vibrio cholerae* superintegron integrase activities. *J Bacteriol* **187:**1740–1750.

**38. Stokes HW, O'Gorman DB, Recchia GD, Parsekhian M, Hall RM.** 1997. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Mol Microbiol* **26:**731–745.

**39. Hall RM, Brookes DE, Stokes HW.** 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:**1941–1959.

**40. Francia MV, Zabala JC, de la Cruz F, García Lobo JM.** 1999. The IntI1 integron integrase preferentially binds single-stranded DNA of the attC site. *J Bacteriol* **181:**6844–6849.

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

**42. Bouvier M, Demarre G, Mazel D.** 2005. Integron cassette insertion: a recombination process involving a folded single strand substrate. *EMBO J* **24:**4356–4367.

**43. Bouvier M, Ducos-Galand M, Loot C, Bikard D, Mazel D.** 2009. Structural features of single-stranded integron cassette attC sites and their role in strand selection. *PLoS Genetics* **5:**e1000632.

**44. MacDonald D, Demarre G, Bouvier M, Mazel D, Gopaul DN.** 2006. Structural basis for broad DNA-specificity in integron recombination. *Nature* **440:**1157–1162.

**45. Loot C, Bikard D, Rachlin A, Mazel D.** 2010. Cellular pathways controlling integron cassette site folding. *EMBO J* **29:**2623–2634.

**46. Nunes-Düby SE, Kwon HJ, Tirumalai RS, Ellenberger T, Landy A.** 1998. Similarities and differences among 105 members of the Int family of site-specific recombinases. *Nucleic Acids Res* **26:**391–406.

**47. Boyd EF, Almagro-Moreno S, Parent MA.** 2009. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends Microbiol* **17:**47–53.

**48. Grindley N, Whiteson K, Rice P.** 2006. Mechanisms of site-specific recombination. *Annu Rev Biochem* **75:**567–605.

**49. Messier N, Roy PH.** 2001. Integron integrases possess a unique additional domain necessary for activity. *J Bacteriol* 183:6699–6706.

**50. Demarre G, Frumerie C, Gopaul DN, Mazel D.** 2007. Identification of key structural determinants of the IntI1 integron integrase that influence attC×attI1 recombination efficiency. *Nucleic Acids Res* **35:**6475–6489.

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

**52. Johansson C, Boukharta L, Eriksson J, Aqvist J, Sundström L.** 2009. Mutagenesis and homology modeling of the Tn21 integron integrase IntI1. *Biochemistry* **48:**1743–1753.

**53. Val ME, Bouvier M, Campos J, Sherratt D, Cornet F, Mazel D, Barre FX.** 2005. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol Cell* **19:**559–566.

**54. Loot C, Ducos-Galand M, Escudero JA, Bouvier M, Mazel D.** 2012. Replicative resolution of integron cassette insertion. *Nucleic Acids Res* **40:**8361–8370.

**55. Loot C, Parissi V, Escudero JA, Amarir-Bouhram J, Bikard D, Mazel D.** 2014. The integron integrase efficiently prevents the melting effect of *Escherichia coli* single-stranded DNA-binding protein on folded attC sites. *J Bacteriol* **196:**762–771.

**56. Collis CM, Grammaticopoulos G, Briton J, Stokes HW, Hall RM.** 1993. Site-specific insertion of gene cassettes into integrons. *Mol Microbiol* **9:**41–52.

**57. Collis CM, Recchia GD, Kim MJ, Stokes HW, Hall RM.** 2001. Efficiency of recombination reactions catalyzed by class 1 integron integrase IntI1. *J Bacteriol* **183:**2535–2542.

**58. Baharoglu Z, Bikard D, Mazel D.** 2010. Conjugative DNA transfer induces the bacterial SOS response and promotes antibiotic resistance development through integron activation. *PLoS Genet* **6:**e1001165.

**59. Collis CM, Hall RM.** 1992. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. *J Bacteriol* **174:**1574–1585.

**60. Gestal AM, Stokes HW, Partridge SR, Hall RM.** 2005. Recombination between the dfrA12-orfF-aadA2 cassette array and an aadA1 gene cassette creates a hybrid cassette, aadA8b. *Antimicrob Agents Chemother* **49:**4771–4774.

**61. Hansson K, Sköld O, Sundström L.** 1997. Non-palindromic attl sites of integrons are capable of site-specific recombination with one another and with secondary targets. *Mol Microbiol* **26:**441–453.

**62. Roy Chowdhury P, Merlino J, Labbate M, Cheong EY, Gottlieb T, Stokes HW.** 2009. Tn6060, a transposon from a genomic island in a Pseudomonas aeruginosa clinical isolate that includes two class 1 integrons. *Antimicrob Agents Chemother* **53:**5294–5296.

**63. González-Zorn B, Catalan A, Escudero JA, Domínguez L, Teshager T, Porrero C, Moreno MA.** 2005. Genetic basis for dissemination of armA. *J Antimicrob Chemother* **56:**583–585.

**64. Recchia GD, Hall RM.** 1995. Plasmid evolution by acquisition of mobile gene cassettes: plasmid pIE723 contains the aadB gene cassette precisely inserted at a secondary site in the incQ plasmid RSF1010. *Mol Microbiol* **15:**179–187.

**65. Segal H, Francia MV, Lobo JM, Elisha G.** 1999. Reconstruction of an active integron recombination site after integration of a gene cassette at a secondary site. *Antimicrob Agents Chemother* **43:**2538–2541.



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

**66. Jacquier H, Zaoui C, Sanson-le Pors MJ, Mazel D, Berçot B.** 2009. Translation regulation of integrons gene cassette expression by the attC sites. *Mol Microbiol* **72:**1475–1486.

**67. Shultzaberger RK, Bucheimer RE, Rudd KE, Schneider TD.** 2001. Anatomy of Escherichia coli ribosome binding sites. *J Mol Biol* **313:**215–228.

**68. Hanau-Berçot B, Podglajen I, Casin I, Collatz E.** 2002. An intrinsic control element for translational initiation in class 1 integrons. *Mol Microbiol* **44:**119–130.

**69. Bissonnette L, Champetier S, Buisson JP, Roy PH.** 1991. Characterization of the nonenzymatic chloramphenicol resistance (*cmlA*) gene of the In4 integron of Tn1696: similarity of the product to transmembrane transport proteins. *J Bacteriol* **173:**4493–4502.

**70. Stokes HW, Hall RM.** 1991. Sequence analysis of the inducible chloramphenicol resistance determinant in the Tn1696 integron suggests regulation by translational attenuation. *Plasmid* **26:**10–19.

**71. Biskri L, Mazel D.** 2003. Erythromycin esterase gene ere(A) is located in a functional gene cassette in an unusual class 2 integron. *Antimicrob Agents Chemother* **47:**3326–3331.

**72. da Fonseca ÉL, Vicente AC.** 2012. Functional characterization of a Cassette-specific promoter in the class 1 integron-associated qnrVC1 gene. *Antimicrob Agents Chemother* **56:**3392–3394.

**73. Szekeres S, Dauti M, Wilde C, Mazel D, Rowe-Magnus DA.** 2007. Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol Microbiol* **63:**1588–1605.

**74. Guérout AM, Iqbal N, Mine N, Ducos-Galand M, Van Melderen L, Mazel D.** 2013. Characterization of the phd-doc and ccd toxin-antitoxin cassettes from *Vibrio* superintegrons. *J Bacteriol* **195:**2270–2283.

**75. Elsaied H, Stokes HW, Nakamura T, Kitamura K, Fuse H, Maruyama A.** 2007. Novel and diverse integron integrase genes and integron-like gene cassettes are prevalent in deep-sea hydrothermal vents. *Environ Microbiol* **9:**2298–2312.

**76. Stokes HW, Holmes AJ, Nield BS, Holley MP, Nevalainen KM, Mabbutt BC, Gillings MR.** 2001. Gene cassette PCR: sequence-independent recovery of entire genes from environmental DNA. *Appl Environ Microbiol* **67:**5240–5246.

**77. Koenig JE, Boucher Y, Charlebois RL, Nesbø C, Zhaxybayeva O, Bapteste E, Spencer M, Joss MJ, Stokes HW, Doolittle WF.** 2008. Integron-associated gene cassettes in Halifax Harbour: assessment of a mobile gene pool in marine sediments. *Environ Microbiol* **10:**1024–1038.

**78. Gillings MR, Holley MP, Stokes HW.** 2009. Evidence for dynamic exchange of qac gene cassettes between class 1 integrons and other integrons in freshwater biofilms. *FEMS Microbiol Lett* **296:**282–288.

**79. Holmes AJ, Gillings MR, Nield BS, Mabbutt BC, Nevalainen KM, Stokes HW.** 2003. The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environ Microbiol* **5:**383–394.

**80. Gillings M, Boucher Y, Labbate M, Holmes A, Krishnan S, Holley M, Stokes HW.** 2008. The evolution of class 1 integrons and the rise of antibiotic resistance. *J Bacteriol* **190:**5095–5100.

**81. Gillings MR, Holley MP, Stokes HW, Holmes AJ.** 2005. Integrons in Xanthomonas: a source of species genome diversity. *Proc Natl Acad Sci USA* **102:**4419–4424.

**82. Rapa RA, Labbate M.** 2013. The function of integron-associated gene cassettes in *Vibrio* species: the tip of the iceberg. *Front Microbiol* **4:**385.

**83. Gerdes K, Christensen SK, Lobner-Olesen A.** 2005. Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* **3:**371–382.

**84. Yamaguchi Y, Park JH, Inouye M.** 2011. Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet* **45:**61–79.

**85. Van Melderen L, Saavedra De Bast M.** 2009. Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genetics* **5:**e1000437.

**86. Sberro H, Leavitt A, Kiro R, Koh E, Peleg Y, Qimron U, Sorek R.** 2013. Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. *Molecular Cell* **50:**136–148.

**87. Tsafnat G, Copty J, Partridge SR.** 2011. RAC: Repository of Antibiotic resistance Cassettes. Database: the journal of biological databases and curation. 2011: bar054.

**88. Moura A, Soares M, Pereira C, Leitão N, Henriques I, Correia A.** 2009. INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics* **25:**1096–1098.

**89. Joss MJ, Koenig JE, Labbate M, Polz MF, Gillings MR, Stokes HW, Doolittle WF, Boucher Y.** 2009. ACID: annotation of cassette and integron data. *BMC Bioinformatics* **10:**118.

**90. Rowe-Magnus DA, Guérout AM, Mazel D.** 1999. Super-integrons. *Res Microbiol* **150:**641–651.

**91. Chowdhury N, Asakura M, Neogi SB, Hinenoya A, Haldar S, Ramamurthy T, Sarkar BL, Faruque SM, Yamasaki S.** 2010. Development of simple and rapid PCR-fingerprinting methods for Vibrio cholerae on the basis of genetic diversity of the superintegron. *J Appl Microbiol* **109:**304–312.

**92. Feng L, Reeves PR, Lan R, Ren Y, Gao C, Zhou Z, Cheng J, Wang W, Wang J, Qian W, Li D, Wang L.** 2008. A recalibrated molecular clock and independent origins for the cholera pandemic clones. *PloS One* **3:**e4053.

**93. Labbate M, Boucher Y, Joss MJ, Michael CA, Gillings MR, Stokes HW.** 2007. Use of chromosomal integron arrays as a phylogenetic typing system for *Vibrio cholerae* pandemic strains. *Microbiology* **153:**1488–1498.

**94. Erill I, Campoy S, Barbé J.** 2007. Aeons of distress: an evolutionary perspective on the bacterial SOS response. *FEMS Microbiol Rev* **31:**637–656.

**95. Guerin E, Cambray G, Sanchez-Alberola N, Campoy S, Erill I, Da Re S, Gonzalez-Zorn B, Barbe J, Ploy MC, Mazel D.** 2009. The SOS response controls integron recombination. *Science* **324:**1034.

**96. Cambray G, Sanchez-Alberola N, Campoy S, Guerin E, Da Re S, González-Zorn B, Ploy MC, Barbe J, Mazel D, Erill I.** 2011. Prevalence of SOS-mediated control of integron integrase expression as an adaptive trait of chromosomal and mobile integrons. *Mob DNA* **2:**6.

**97. Harms K, Starikova I, Johnsen PJ.** 2013. Costly Class-1 integrons and the domestication of the the functional integrase. *Mobile Genetic Elements* **3:**e24774.

**98. Starikova I, Harms K, Haugen P, Lunde T, Primicerio R, Samuelsen Ø, Nielsen KM, Johnsen PJ.** 2012. A trade-off between the fitness cost of functional integrases and long-term stability of integrons. *PLoS pathogens* **8:**e1003043.

**99. Baharoglu Z, Mazel D.** 2011. *Vibrio cholerae* triggers SOS and mutagenesis in response to a wide range of antibiotics: a route towards multiresistance. *Antimicrob Agents Chemother* **55:**2438–2441.

**100. Baharoglu Z, Krin E, Mazel D.** 2013. RpoS plays a central role in the SOS induction by sub-lethal aminoglycoside concentrations in *Vibrio cholerae*. *PLoS Genetics* **9:**e1003421.

**101. Gutierrez A, Laureti L, Crussard S, Abida H, Rodríguez-Rojas A, Blázquez J, Baharoglu Z, Mazel D, Darfeuille F, Vogel J, Matic I.** 2013. β-Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nature Commun* **4:**1610.

**102. Baharoglu Z, Krin E, Mazel D.** 2012. Connecting environment and genome plasticity in the characterization of transformation-induced SOS regulation and carbon catabolite control of the *Vibrio cholerae* integron integrase. *J Bacteriol* **194:**1659–1667.

**103. Krin E, Cambray G, Mazel D.** 2014. The superintegron integrase and the cassette promoters are co-regulated in *Vibrio cholerae*. *PLoS One* **9:** e91194.

**104. Guérin E, Jové T, Tabesse A, Mazel D, Ploy MC.** 2011. High-level gene cassette transcription prevents integrase expression in class 1 integrons. *J Bacteriol* **193:**5675–5682.

**105. Vinué L, Jové T, Torres C, Ploy MC.** 2011. Diversity of class 1 integron gene cassette Pc promoter variants in clinical *Escherichia coli* strains and description of a new P2 promoter variant. *Int J Antimicrob Agents* **38:**526–529.



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20

**106. Moura A, Jové T, Ploy MC, Henriques I, Correia A.** 2012. Diversity of gene cassette promoters in class 1 integrons from wastewater environments. *Appl Environ Microbiol* **78:**5413–5416.

**107. Bikard D, Loot C, Baharoglu Z, Mazel D.** 2010. Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol Mol BiolRev* **74:**570–588.

**108. Trinh TQ, Sinden RR.** 1991. Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*. *Nature* **352:** 544–547.

**109. Prudhomme M, Attaiech L, Sanchez G, Martin B, Claverys JP.** 2006. Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science* **313:**89–92.

**110. Moore ME, Lam A, Bhatnagar S, Solnick JV.** 2014. Environmental determinants of transformation efficiency in *Helicobacter pylori*. *J Bacteriol* **196:**337–344.

**111. Pearson CE, Zorbas H, Price GB, Zannis-Hadjopoulos M.** 1996. Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J Cell Biochem* **63:**1–22.

**112. Lodge JK, Kazic T, Berg DE.** 1989. Formation of supercoiling domains in plasmid pBR322. *J Bacteriol* **171:**2181–2187.

**113. Pruss GJ, Drlica K.** 1986. Topoisomerase I mutants: the gene on pBR322 that encodes resistance to tetracycline affects plasmid DNA supercoiling. *Proc Natl Acad Sci USA* **83:**8952–8956.

**114. Liu LF, Wang JC.** 1987. Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci USA* **84:**7024–7027.

**115. Balke VL, Gralla JD.** 1987. Changes in the linking number of supercoiled DNA accompany growth transitions in *Escherichia coli*. *J Bacteriol* **169:**4499–4506.

**116. Jaworski A, Higgins NP, Wells RD, Zacharias W.** 1991. Topoisomerase mutants and physiological conditions control supercoiling and Z-DNA formation *in vivo*. *J Biol Chem* **266:**2576–2581.

**117. Ferrándiz MJ, Martín-Galiano AJ, Schvartzman JB, de la Campa A.** 2010. The genome of *Streptococcus pneumoniae* is organized in topology-reacting gene clusters. *Nucleic Acids Res* **38:**3570–3581.

**118. Majchrzak M, Bowater RP, Staczek P, Parniewski P.** 2006. SOS repair and DNA supercoiling influence the genetic stability of DNA triplet repeats in *Escherichia coli*. *J Mol Biol* **364:**612–624.

**119. Champion K, Higgins NP.** 2007. Growth rate toxicity phenotypes and homeostatic supercoil control differentiate *Escherichia coli* from *Salmonella enterica* serovar *Typhimurium*. *J Bacteriol* **189:**5839–5849.

**120. Collins J, Volckaert G, Nevers P.** 1982. Precise and nearly-precise excision of the symmetrical inverted repeats of Tn5; common features of recA-independent deletion events in *Escherichia coli*. *Gene* **19:**139–146.

**121. Meyer RR, Laine PS.** 1990. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol Rev* **54:**342–380.

**122. Roy R, Kozlov AG, Lohman TM, Ha T.** 2009. SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* **461:**1092–1097.

**123. Dubois V, Debreyer C, Quentin C, Parissi V.** 2009. In vitro recombination catalyzed by bacterial class 1 integron integrase IntI1 involves cooperative binding and specific oligomeric intermediates. *PloS One* **4:**e5228.

**124. Dubois V, Debreyer C, Litvak S, Quentin C, Parissi V.** 2007. A new in vitro strand transfer assay for monitoring bacterial class 1 integron recombinase IntI1 activity. *PLoS One* **2:**e1315.

**125. Tanaka T, Mizukoshi T, Sasaki K, Kohda D, Masai H.** 2007. *Escherichia coli* PriA protein, two modes of DNA binding and activation of ATP hydrolysis. *J Biol Chem* **282:**19917–19927.

**126. Grompone G, Ehrlich SD, Michel B.** 2003. Replication restart in gyrB Escherichia coli mutants. *Mol Microbiol* **48:**845–854.

**127. Rowe-Magnus DA.** 2009. Integrase-directed recovery of functional genes from genomic libraries. *Nucleic Acids Res* **37:**e118.

**128. Sureshan V, Deshpande CN, Boucher Y, Koenig JE, Midwest Center for Structural G, Stokes HW, Harrop SJ, Curmi PM, Mabbutt BC.** 2013. Integron gene cassettes: a repository of novel protein folds with distinct interaction sites. *PloS One* **8:**e52934.

**129. Gestal AM, Liew EF, Coleman NV.** 2011. Natural transformation with synthetic gene cassettes: new tools for integron research and biotechnology. *Microbiology* **157:**3349–3360.

**130. Bikard D, Julié-Galau S, Cambray G, Mazel D.** 2010. The synthetic integron: an in vivo genetic shuffling device. *Nucleic Acids Res* **38:** e153.

**131. Baquero F, Coque TM, de la Cruz F.** 2011. Ecology and evolution as targets: the need for novel eco-evo drugs and strategies to fight antibiotic resistance. *Antimicrob Agents Chemother* **55:**3649–3660.

**132. Holmes AJ, Holley MP, Mahon A, Nield B, Gillings M, Stokes HW. 2003.** Recombination activity of a distinctive integron-gene cassette system associated with *Pseudomonas stutzeri* populations in soil. *J Bacteriol* **185:**918–928.

**133. Leon G, Roy PH.** 2009. Potential role of group IIC-attC introns in integron cassette formation. *J Bacteriol* **191:**6040–6051.

**134. Quiroga C, Centrón D.** 2009. Using genomic data to determine the diversity and distribution of target site motifs recognized by class C-attC group II introns. *J Mol Evol* **68:**539–549.



Downloaded from www.asmscience.org by
IP: 163.1.28.126
On: Thu, 16 Mar 2017 11:06:20