Molecular Microbiology (1995) **15**(4), 593–600

MicroReview

Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination

Ruth M. Hall* and Christina M. Collis
CSIRO Division of Biomolecular Engineering, Sydney Laboratory, PO Box 184, North Ryde 2113, New South Wales, Australia.

Summary

An integron is a genetic unit that includes the determinants of the components of a site-specific recombination system capable of capturing and mobilizing genes that are contained in mobile elements called gene cassettes. An integron also provides a promoter for expression of the cassette genes, and integrons thus act both as natural cloning systems and as expression vectors. The essential components of an integron are an *int* gene encoding a site-specific recombinase belonging to the integrase family, an adjacent site, *attI*, that is recognized by the integrase and is the receptor site for the cassettes, and a promoter suitably oriented for expression of the cassette-encoded genes. The cassettes are mobile elements that include a gene (most commonly an antibiotic-resistance gene) and an integrase-specific recombination site that is a member of a family of sites known as 59-base elements. Cassettes can exist either free in a circularized form or integrated at the *attI* site, and only when integrated is a cassette formally part of an integron. A single site-specific recombination event involving the integron-associated *attI* site and a cassette-associated 59-base element leads to insertion of a free circular cassette into a recipient integron. Multiple cassette insertions can occur, and integrons containing several cassettes have been found in the wild. The integrase also catalyses excisive recombination events that can lead to loss of cassettes from an integron and generate free circular cassettes. Due to their ability to acquire new genes, integrons have a clear role in the evolution of the genomes of the plasmids and transposons that contain them. However, a more general role in evolution is also likely. Events involving recombination between a specific 59-base-element site and a non-specific secondary site have recently been shown to occur. Such events should lead either to the insertion of cassettes at non-specific sites or to the formation of stable cointegrates between different plasmid molecules, and a cassette situated outside the integron context has recently been identified.

Received 4 July, 1994; revised 17 October, 1994; accepted 19 October, 1994. *For correspondence. Tel. (2) 8864962; Fax (2) 8864805.

Introduction

The term integron was originally coined to describe the group of apparently mobile elements which contain one or more antibiotic-resistance genes located at a specific site, and also contain the determinants of the site-specific recombination system responsible for insertion of the resistance genes (Stokes and Hall, 1989). The definitive features of integrons were the site-specific recombination system and the consequent ability to capture many different resistance genes. However for the most extensively studied family of integrons, two levels of mobility are clearly involved. First, it was possible to infer that the resistance genes were mobile and had been acquired by site-specific recombination (Cameron *et al.*, 1986; Hall and Vockler, 1987; Ouellette *et al.*, 1987; Sundström *et al.*, 1988; Hall *et al.*, 1991; Sundström *et al.*, 1991). Second, the structure into which the resistance genes were inserted also appeared to be mobile as it was found in several independent locations in different plasmids and transposons (Hall and Vockler, 1987; Stokes and Hall, 1989). As initially our knowledge of integrons was derived mainly from comparisons of the sequences surrounding the different antibiotic-resistance genes found in naturally occurring integrons (see Fig. 1), they were inevitably described from an essentially structural standpoint as consisting of two conserved regions flanking a variable region which contains one or more resistance genes (Stokes and Hall, 1989). However, it is now clear that the integrons found in clinical isolates have several distinct structures (Hall *et al.*, 1994; Rådström *et al.*, 1994; H. J. Brown, R. M. Hall and H. W. Stokes, unpublished) and that a more general definition is needed. A description of the details of the structure and mobility of integrons lies outside the scope of this review, which focuses on the mobile gene cassettes and the cassette

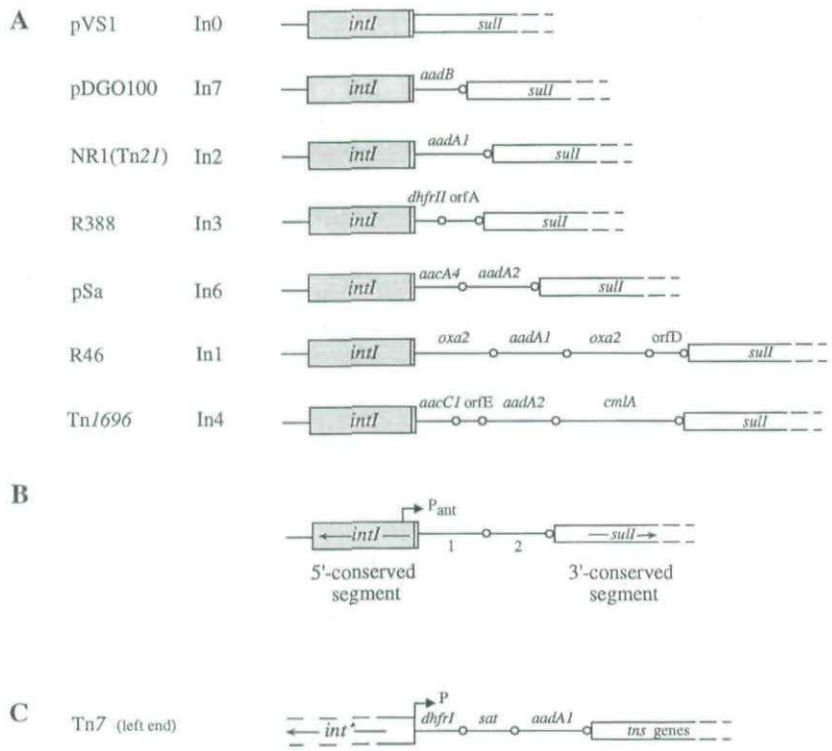

Fig. 1. Structural features of integrons.
A. Structures of some naturally occurring integrons. Genes associated with each inserted cassette are indicated. The 5'-CS is represented by a stippled box and the extent of *attI* indicated by vertical lines at the inner end. Gene cassettes are shown as narrow lines, with a circle to represent the 59-base-element recombination site. The 3'-CS is shown as an open box. The *intI* gene encodes the integrase and the *sulI* gene confers resistance to sulphonamides.
B. General structure of *sulI*-associated integrons. The position of P_{ant}, the common promoter responsible for transcription of the inserted cassettes, is indicated.
C. Structure of the left end of Tn*7*. *int** represents the ORF whose product shows similarity to Int *of sulI*-associated integrons but which is interrupted by a stop codon. The position of the Tn*7* transposition genes (*tns*) is indicated. The promoter P is presumed to be responsible for transcription of the genes in inserted cassettes.

integration and expression systems which are the definitive features of integrons.

Over the past few years, site-specific recombination catalysed by the integron-encoded integrase (IntI) has been demonstrated experimentally, and the main components of the site-specific recombination system have now been identified (Martinez and de la Cruz, 1988; 1990; Hall *et al.*, 1991; Recchia *et al.*, 1994). The gene cassettes which contain the resistance genes have been defined structurally (Hall *et al.*, 1991) and have been shown to be mobile (Collis and Hall, 1992a). Cassettes can exist in a free circular form (Collis and Hall, 1992b), and insertion of circularized cassettes into a recipient integron structure has also been demonstrated (Collis *et al.*, 1993). Thus it is now possible to define both the mobile gene cassettes and the integrons that capture them on a strictly functional basis. In this review we introduce a definition of integrons as any genetic unit that contains the determinants of the site-specific recombination and expression systems. The gene cassettes are members of a family of mobile elements that are unlike other known mobile elements in that they do not include all functions required for their mobility, and cassettes are only formally part of an integron when they are integrated at the integron-receptor site.

The aims of this review are to provide a simple overview of the integron/cassette system, to summarize the evidence for the mobilization of cassettes by the integron-determined site-specific recombination system, and to highlight some of the novel features of this system that may be important in the evolution of bacterial and plasmid genomes.

Integrons are natural cloning and expression vectors

Features of integrons found in the wild

It is now well established that many of the antibiotic-resistance genes found in clinical isolates of Gram-negative organisms (particularly *Enterobacteriaceae* and *Pseudomonads*) are situated in the same immediate sequence environment (see Cameron *et al.*, 1986; Hall and Vockler, 1987; Ouellette *et al.*, 1987; Sundström *et al.*, 1988; Stokes and Hall, 1989; Hall *et al.*, 1991 and references therein). The genes identified to date determine resistance to a range of antibiotics (aminoglycosides, trimethoprim, chloramphenicol and penicillins and cephalosporins), although the preponderance of antibiotic resistance genes in the pool of genes identified to date presumably reflects the fact that the genes studied were all derived from antibiotic-resistant organisms. For each antibiotic family several distinct genes have been found. The encoded enzymes confer resistance by distinct mechanisms, antibiotic modification (aminoglycosides, chloramphenicol and β-lactams), metabolic by-pass (trimethoprim) and efflux of the antibiotic (chloramphenicol). In addition, several open reading frames (ORFs) with no known function have been identified (Hall *et al.*, 1991). Not only do these genes all occur in the same location

but many arrays containing more than one gene have been found in the wild (Fig. 1A). The most common structure found associated with the antibiotic-resistance genes is that illustrated in Fig. 1A, in which a *sul*I gene (conferring resistance to sulphonamides) is located adjacent to the variable region. In this group of integrons the inserted genes are all in the same orientation with respect to the flanking sequences, which are identified by their relationship to the insert genes as the 5′- and 3′- conserved segments (5′-CS and 3′-CS). However, integrons which do not contain inserted genes have been found (Bissonnette and Roy, 1992; Collis and Hall, 1992a), and in this 'cassette-free' integron configuration the 5′-CS and 3′-CS are contiguous (Fig. 1A).

A similar situation to that described above also applies to the Tn*7* family of transposons (Fig. 1C). Tn*7* contains three antibiotic resistance genes (*dhfrI–sat–aadA1*) close to its left end (see Hall *et al.*, 1991; Sundström *et al.*, 1991), and in two closely related transposons (Tietze *et al.*, 1987) the *dhfrI* gene has either been lost or replaced by a new DNA segment of unknown function (see Hall *et al.*, 1991 for details). Two of the genes found in Tn*7* (*dhfrI and aadA1*) have also been found in *sulI*-associated integrons (Sundström and Sköld, 1990) indicating that exchange of genes between these locations is possible.

Expression of cassette genes

The orientation specificity of the integrated resistance genes allows them to be transcribed from a common promoter. In *sulI*-associated integrons, a promoter, P_{ant}, that is located in the 5′-CS and is responsible for transcription of the antibiotic-resistance genes has been identified (Fig. 1B). This promoter is one of the locations where minor variations in the sequence of 5′-CS are found (Stokes and Hall, 1989), and these alterations have recently been shown to affect promoter strength over at least a 20-fold range (Collis and Hall, 1995; Lévesque *et al.*, 1994). A second common promoter, P2, that has arisen by the insertion of three G residues to increase the spacing between potential −10 and −35 sequences to 17, is present in a few naturally occurring integrons and is also active (Schmidt *et al.*, 1988). To date P2 has been found only in conjunction with the weakest variant of the P_{ant} promoter, and in this configuration P2 is the major promoter (Collis and Hall, 1995, Lévesque *et al.*, 1994).

When more than one cassette is present, the cassette-encoded genes are co-transcribed from P_{ant}. However, expression of distal genes is reduced by the presence of the upstream cassette(s), and this effect appears to be due to premature termination of transcripts, particularly at or near the end of the cassettes (Collis and Hall, 1995).

Integron-determined components of the site-specific recombination system

The site-specific recombinase encoded by the 5′-CS of *sulI*-associated integrons was initially identified from the similarity of the predicted product of ORF3 (*intI*) in the 5′-CS (Hall and Vockler, 1987) to other integrases (Ouellette and Roy, 1987). That this putative integrase in fact catalyses site-specific recombination events was clearly demonstrated by Martinez and de la Cruz (1988; 1990), and confirmed by subsequent studies (Hall *et al.*, 1991; Collis and Hall, 1992a,b; Collis *et al.*, 1993; Recchia *et al.*, 1994). Given the weight of evidence, a claim that the *intI* gene is not required for site-specific recombination (Nücken *et al.*, 1991) is very unlikely to be correct, and an alternative explanation for the observations of these authors needs to be sought.

The integron-associated recombination site that acts as a receptor site for the insertion of cassettes has only recently been examined in detail (Recchia *et al.*, 1994), and designated *attI*. This site was previously assumed to be largely confined to the conserved sequences immediately 5′ to the inserted cassettes, and evidence that it participates in recombination events was first reported by Martinez and de la Cruz (1990). (The sites designated RHS-1 and RHS-3 by these authors correspond to *attI sites, but only RHS-3 was found to be active.*) As there is essentially no resemblance between the sequences near the inner boundary of the 5′-CS and the recognizable features of the cassette-associated 59-base-element sites (see Hall *et al.*, 1991; Collis and Hall, 1992b; and below for details of the features of 59-base elements), the extent of the *attI* site has until recently remained a matter for conjecture. A 14-base sequence (designated hs2) was proposed to be the extent of this site (Schmidt *et al.*, 1989), and supporting experimental evidence was subsequently claimed (Nücken *et al.*, 1991). However, the extent of sequences required for *attI* site activity has recently been examined using deletion analysis (Recchia *et al.*, 1994). A region of at least 40 bases was essential for full site activity and the maximum length required was 70 bases. The bulk of the *attI* site lies within the 5′-CS, but a further 6 bases at the insertion point are formally part of the adjacent 3′-CS or first cassette. These 6 bases comprise 6 of the 7 bases that were originally identified as the only sequence feature common to both boundaries of the integron insert region and designated the core site (Stokes and Hall, 1989). This core site consists of the conserved triplet GTT and four further bases conforming to the consensus AGGC or RRRY (R = purine, Y = pyrimidine).

In Tn*7* the product of an ORF located 5′ to the antibiotic-resistance genes was found to show significant similarity to the product of the *intI* gene of *sulI*-associated integrons

(Hall and Vockler, 1987; A. Pelletier and P. H. Roy, unpublished). The Tn*7* ORF (*int** in Fig. 1) is shorter than the *intI* gene, but the similarity extends beyond the stop codon, the presence of which presumably renders the putative Tn*7* integrase inactive. The Tn7 *att* site is presumably located between the *int* gene and the first inserted resistance gene, but the activity and extent of this site have not been examined.

Definition of integrons

From the evidence now available it is possible to define integrons as elements that contain the genetic determinants of the components of a site-specific recombination system that recognizes and captures mobile gene cassettes. An integron then includes the *int* gene and the adjacent *attI* recombination site. Integrons thus do not necessarily include a gene cassette, but when integrated, a cassette is logically part of the integron. As expression of the cassette genes relies on the promoter in the integron, P_{ant} must also be viewed as a definitive feature of integrons. Any genetic unit that contains these features is thus an integron, whether it is a transposon such a Tn*7* or Tn*402*, or a defective transposon derivative as appears to be the case for the majority of *sulI*-associated integrons. As this definition is simpler and more generally applicable than the structural definition used to date, we hope that its use will prevail in the future.

Gene cassettes: a family of mobile elements

Gene cassettes are discrete genetic units

The cassettes which contain the antibiotic-resistance genes were defined as discrete units by comparison of the sequences flanking a limited number of genes which were either found alone or in different positions in naturally occurring cassette arrays (Hall *et al.*, 1991). The conserved triplet GTT, which is part of a 7-base core site, is found at both ends of a cassette in a linear array, but only one GTT is contained within the cassette. As this causes a slight ambiguity in identifying the cassette boundaries, for simplicity it was assumed that the GTT at the 5′ end of the gene is cassette-associated (Hall *et al.*, 1991). However, the position of the recombination crossover has now been localized to between the G and T residues of this triplet (D. O'Gorman, R. M. Hall, and H. W. Stokes, unpublished) and the cassette-derived sequences in a linear array would in fact have this precise boundary.

Cassettes normally include only one complete gene (or ORF) followed by a recombination site belonging to a family of sites known as 59-base elements (Hall *et al.*, 1991). Each gene is associated with a particular 59-base element, and though the 59-base elements associated with different genes can differ substantially in both sequence and length, they are identifiable from the features they share (see below). In general the distance between the cassette boundary and the first in-frame initiation codon is quite short (as few as nine bases) and the cassettes do not include a promoter preceding the gene. Only one exception is known; the *cmlA* gene is preceded by both a promoter and translational attenuation signals (Stokes and Hall, 1991). The absence of promoters internal to the cassettes makes expression of the cassette *genes dependent on an upstream promoter which is* located in the 5′-conserved segment as described above (see Fig. 1B). In most cases little or no untranslated sequence is found between the termination codon and the 59-base element, and in several cases the reading frame extends into the 59-base element (see Hall *et al.*, 1991 for details). Thus a cassette includes essentially only a gene and a specific recombination site.

Gene cassettes are mobile

Each cassette is not only a discrete unit but is also mobile. Direct evidence for cassette mobility came first from studies of the loss of resistance genes from cloned integron fragments containing an array of two or more cassettes (Collis and Hall, 1992a and references therein; A. Gravel and P. H. Roy, unpublished). Loss of cassettes was shown to be dependent on the presence of the integrase, and when the sequences of new boundaries were determined precise excision of a complete cassette was found to have occurred (Collis and Hall, 1992a). Cassettes have been shown to exist free as covalently-closed circular molecules (Collis and Hall, 1992b), and these molecules are generated by IntI-mediated direct excision of an integrated cassette from an integron. Though free circular cassettes are not able to replicate, they are likely to be important participants in the process which leads initially to the association of a cassette with an integron, and subsequently to the spread of resistance genes from one integron to another as well as to non-specific sites as described below.

59-base elements, the cassette-associated recombination sites

The cassette-associated recombination sites were first identified as imperfect inverted repeats that were found at the 3′ end of inserted genes in integrons and were related to a consensus of 59 bases (Cameron *et al.*, 1986), hence the name 59-base elements. However it is now recognized that the 59-base-element family also includes longer imperfect repeats (up to 110 bp) that are related to the consensus at their outer ends (Hall *et al.*,

1991; Collis and Hall, 1992b). While several 59-base elements are closely related to a 60-base consensus sequence (see Fig. 3 in Hall *et al.*, 1991), no two of the 59-base elements sequenced to date are identical. The most conserved features of members of the 59-base-element family are an inverse core site (RYYYAAC) *closest to the 3′ end of the gene and a core site* (GTTRRRY) at the other end. About 20 bases at each end are related to a consensus sequence (Collis and Hall, 1992b), and between these sequences lies an inverted-repeat region of variable length and sequence. Several distinct 59-base elements have been shown to function as specific recombination sites recognized by the integron integrase (Martinez and de la Cruz, 1990; Hall *et al.*, 1991; Collis and Hall, 1992a). (The sites designated RHS-2, RHS-4 and RHS-5 by Martinez and de la Cruz (1990) are 59-base elements.) IntI-mediated recombination can take place between either two 59-base-element sites or between *attI* and a 59-base element, and the recombination crossover occurs within the seven-base core site at the 3′ end of the 59-base element (or *attI*). In the linear cassette arrays found in integrons, the 59-base element is a thus a composite site with the last six bases of the core site supplied by the adjacent 3′-CS or cassette, leading to variability in the sequence of these bases. However, in free circular cassettes a 59-base element with a unique sequence is associated with each cassette gene, and the core site and inverse core site of these 59-base elements are generally perfect inverted repeats (Collis and Hall, 1992b).

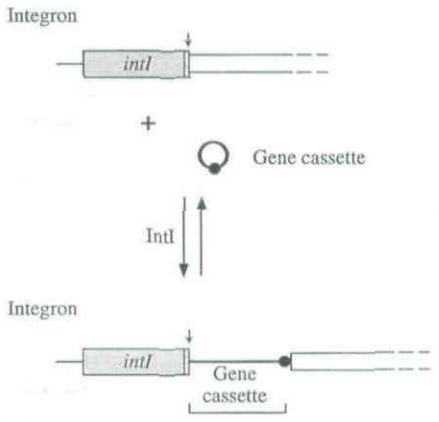

Fig. 2. Insertion of circular gene cassettes into integrons. The recipient structure is shown as an integron containing no inserted cassettes, with the 5′-CS represented by a stippled box and the 3′-CS by an open box. The *attI* site is shown bounded by two vertical lines and the point of insertion is indicated by an arrow. The *intI* gene encodes the integrase (IntI) which catalyses the insertion of a circular gene cassette as shown. The resulting integron contains a single integrated cassette and two recombination sites, *attI* and the 59-base element (closed circle) at the 3′ end of the inserted cassette.

Site-specific insertion of gene cassettes into integrons

In 1991, it was proposed that the primary event leading to the association of a cassette with an integron was likely to be a single IntI-mediated recombination event between the integron-associated receptor site (*attI*) and the 59-base element in a circular gene cassette (Hall *et al.*, 1991), as shown in Fig. 2. The demonstration that the reverse reaction, i.e. the excision of a cassette to generate a free circularized molecule can occur supported this notion (Collis and Hall, 1992b), and direct insertion of cassettes has since been demonstrated (Collis *et al.*, 1993). However, once cassettes have become associated with an integron, alternative routes for the movement of cassettes from one integron to another are possible, and it is likely that in the wild these routes are important for the spread of cassettes.

Direct insertion of circular cassettes

Circular cassettes generated *in vitro* have been used to transform cells containing a plasmid that includes an integron fragment containing the *attI* site and a second plasmid which expresses the integrase (Collis *et al.*, 1993). Precise IntI-dependent insertion of the cassette was observed, though the frequency of insertion was extremely low. When the recipient integron fragment contained a complete cassette and thus a choice of insertion sites (*attI* and a 59-base element) was available, the incoming cassette was inserted exclusively at *attI*. Similar observations have also been made by others (K. Hansson, O. Sköld, P. Rådström, G. Swedberg, J. Flensberg and L. Sundström, unpublished). When *in vitro*-generated circular molecules composed of two cassettes were used to transform cells, insertion of each cassette was observed but the two cassettes were not inserted together (Collis *et al.*, 1993). This observation presumably reflects a tendency for the excisive intramolecular event leading to resolution of the transforming molecules into two discrete circular cassettes to occur more readily than the integrative intermolecular event leading to insertion of cassettes into the recipient integron.

Alternative routes for loss, gain and rearrangement of gene cassettes

The integrative and excisive modes of IntI-dependent recombination can also lead to the cointegration of two plasmid molecules (either identical or non-identical) present in the same cell (Martinez and de la Cruz, 1988; 1990; Hall *et al.*, 1991; Recchia *et al.*, 1994), and the subsequent resolution of these cointegrates (Hall *et al.*,

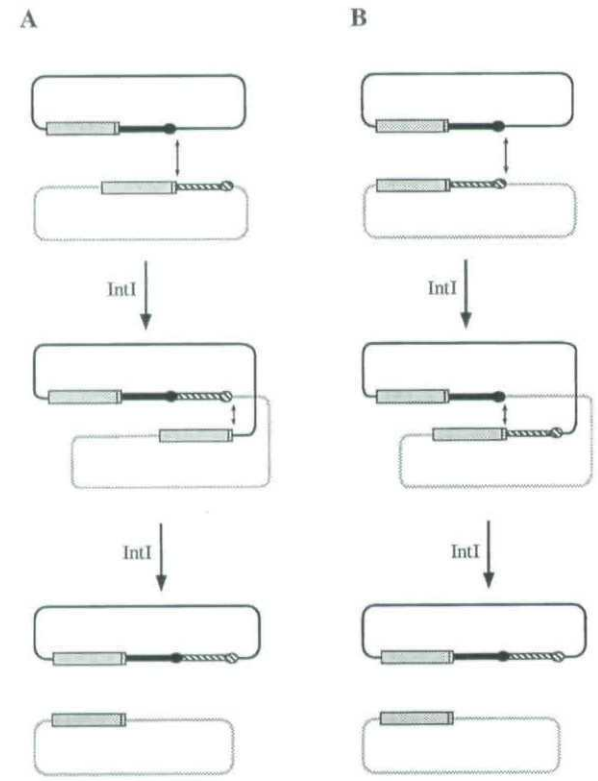

Fig. 3. Reciprocal acquisition and loss of gene cassettes by IntI-mediated cointegration and resolution. The two participating plasmids are represented by black and stippled lines, respectively. The 5′-CS is represented by a stippled box with a vertical line to indicate the extent of *attI*. Cassettes are shown as black and hatched lines with circles to represent the 59-base elements. IntI-catalysed recombination between *attI* (A) or the 59-base element (B) in one plasmid molecule and the 59-base element in the other, as indicated by the small arrow, results in the formation of a cointegrate of the two plasmids. IntI-catalysed resolution of the cointegrate leads again to two separate plasmids. One of these plasmids has lost its cassette, while the other has gained a cassette and now has an array of two integrated cassettes.

1991; Recchia *et al.*, 1994). A cointegration event involving one pair of sites followed by a resolution event involving a different pair of sites can also lead to the reciprocal acquisition and loss of genes from the cassette arrays involved. This pathway is illustrated in Fig. 3. When the participating plasmid molecules are identical, new cassette orders can be generated and duplication of a cassette in the array can also occur (Collis and Hall, 1992a). However when two different plasmids containing integrons with different integrated cassettes are involved, the transfer of cassettes from one plasmid to the other can result in transfer of a particular resistance gene to a new location. This process is likely to be an important factor in the spread of genes during an outbreak of antibiotic resistance in the clinical situation, particularly if the new plasmid host is transfer proficient, has a broader host range or is in some other way fitter than the original plasmid host.

Because of the conservation in the sequences flanking cassettes in an integron context, the role of homologous recombination in exchanging cassettes between integrons in different replicons must also be considered. Indeed, one case in which such an event is likely to have occurred has been documented (Nies *et al.*, 1985; Stokes and Hall, 1992). Combinations of homologous and site-specific recombination events are also clearly possible (see Stokes and Hall, 1992 for detailed discussion). Though a great deal of attention has been paid to attempting to trace the evolutionary relationships between integrons with different cassette configurations (see e.g. Bissonnette and Roy, 1992) the large number of events that can occur and contribute to the dissemination of cassette-encoded antibiotic-resistance genes makes it seem unlikely that relationships between integrons found in the wild can be traced to make a single or even several plausible evolutionary trees.

Insertion of cassettes at secondary recombination sites

It has recently been shown that the integron integrase is also able to catalyse recombination between one specific site (a 59-base element) and secondary sites that conform to the consensus GNT (Gt/aT, Ga/tTNa/t) (Francia *et al.*, 1993; Recchia *et al.*, 1994) which has some similarity to the core-site consensus. These events occur at very low frequency, but are nonetheless likely to be important in the spread of cassette-associated genes to many different locations. Because the mobile cassettes do not generally include a promoter, expression of the cassette gene would only be observed if the cassette is inserted in the correct orientation with respect to a pre-existing promoter. We have recently examined a case in which a gene (*aadB*) that is normally cassette associated, and thus integron associated, is found in a plasmid that does not contain an integron (Recchia and Hall, 1995). Sequence analysis has shown that the complete cassette is present, and has been precisely inserted at a secondary site downstream of a promoter in the recipient plasmid RSF1010. Insertion of a cassette at a secondary site leads to a stable situation as the cassette is not flanked by two specific recombination sites and the composite 59-base element formed on insertion may be inactive. These features make removal of the cassette by the integrase extremely unlikely, and IntI-mediated excision of the *aadB* cassette from its secondary-site location could not be detected (Recchia and Hall, 1995).

Events involving a secondary site can also potentially lead to other genomic rearrangements and detailed pathways for some of these are described elsewhere (Recchia and Hall, 1995). When the participating sites are in different genomes, recombination leads to fusion

of the two replicons to form a cointegrate (Francia *et al.*, 1993; Recchia *et al.*, 1994) and resolution of these cointegrates by other routes could lead to the reassortment of the two original replicons. This pathway can be invoked to explain the fact that the 5′-CS of *sulI*-associated integrons is occasionally found without the downstream 3′-CS that contains the *sulI* gene, though whether the structures of such integrons are consistent with this explanation remains to be established. Recombination between a specific site in an integron (*attI* or a cassette 59-base element) and a secondary site in the same genome could lead to deletion (or possibly inversion) of part of the genome, though such events have not yet been detected experimentally.

Origin of gene cassettes

While the origin of the gene cassettes is a fascinating question, essentially no pertinent information is currently available. Any pathway proposed for cassette creation must necessarily account for all of their features, namely that cassettes normally include a single complete gene and, as the orientation of the cassettes is determined by the 59-base element, a correctly oriented 59-base element at the 3′ end of the gene. The most attractive hypothetical routes (Hall *et al.*, 1991) involve reverse transcription of an mRNA molecule, as this would explain the fact that cassettes contain a complete gene with little flanking sequence. The 59-base elements could originate either from transcription terminators contained in the transcript or from some other structure that is added later. In this context it is interesting to note that preliminary evidence that 59-base elements may act as transcription terminators has recently been obtained (Collis and Hall, 1995).

Francia *et al.* (1993) have argued that IntI-mediated recombination events involving secondary sites explain in part how new cassettes are created, as they can lead to the association of new genes with integrons. However, while it is clear that such events can replace the 3′-CS with sequences containing new genes, it is difficult to envisage how these structures would subsequently give rise to new functional cassettes.

Comparison with other site-specific recombination systems

Site-specific recombination is used to effect a variety of distinct reactions. Members of the integrase family of site-specific recombinases are utilized for the integration and excision of small genomes (e.g. phage, integrating plasmids and conjugative transposons) into and from chromosomes, for the resolution of plasmid or chromosomal multimers and for the inversion of a segment of the 2μ circle in yeast (Sadowski, 1986). While it may be tempting to compare the integration and excision of gene cassettes into and from integrons with other integration/excision systems, the basic features of the biology of the integron/cassette system differ from these other systems in at least two fundamental ways. First, the small genomes that are known to be integrated/excised by integrases invariably encode the relevant integrase, whereas the integron integrase is encoded by the recipient (integron), not the mobile units (cassettes). Second, many distinct mobile units are recognized by the integron integrase, allowing the integration of different cassettes at the same site and, as multiple integration events can occur, several different cassettes can be simultaneously resident in an integron. Thus other systems do not provide a useful conceptual framework for understanding the cassette/integron system.

At a more detailed mechanistic level, it should indeed ultimately be possible to compare the functioning of the integron integrase with that of other well-studied integrases (e.g. lambda Int, P1 Cre and 2μ circle F LP). However, to date such comparisons have yielded little or no information. The lambda Int, Cre and FLP enzymes all recognize a core-site region of approximately 30 bases made up of a pair of imperfect inverted repeats (that bind the integrase) separated by an overlap region of 6–8 bases that is identical in all the recombination sites, and recombinational breakage and reunion of the top and bottom strands occurs at opposite ends of the overlap region (Sadowski, 1986; 1993; Stark *et al.*, 1992). However, equivalent features are not distinguishable when the sequences of the many different sites (*attI* and more than 20 distinct 59-base elements) recognized by the integron integrase are compared (see compilation in Collis and Hall, 1992b), and the recombination crossover can be confined to between two adjacent nucleotide pairs (D. O'Gorman, R. M. Hall and H. W. Stokes, unpublished) rather than to an overlap region of 6–8 bases. Whether the latter feature reflects breakage and reunion of both top and bottom strands at the same position, or IntI catalyses breakage and reunion of only one pair of strands (top or bottom) with the reaction being completed by some other mechanism, remains to be established. Further studies, both *in vivo* and *in vitro* will be necessary to resolve these dilemmas.

Acknowledgements

We thank Nancy Craig for pointing out the benefits of a functional definition of integrons. We also thank Gavin Recchia, Hatch Stokes, Trevor Lockett and Paul Roy for critical reading of the manuscript.

References

Bissonnette, L., and Roy, P.H. (1992) *J Bacteriol* **174:** 1248–1257.

Cameron, F.H., Groot Obbink, D.J., Ackerman, V.P., and Hall, R.M. (1986) *Nucl Acids Res* **14:** 8625–8635.
Collis, C.M., and Hall, R.M. (1992a) *J Bacteriol* **174:** 1574–1585.
Collis, C.M., and Hall, R.M. (1992b) *Mol Microbiol* **6:** 2875–2885.
Collis, C.M., and Hall, R.M. (1995) *Antimicrob Ag Chemother* **39:** 155–162.
Collis, C.M., Grammaticopoulos, G., Briton, J., Stokes, H.W., and Hall, R.M. (1993) *Mol Microbiol* **9:** 41–52.
Francia, M.V., de la Cruz, F., and García Lobo, J.M. (1993) *Mol Microbiol* **10:** 823–828.
Hall, R.M., Brookes, D.E., and Stokes, H.W. (1991) *Mol Microbiol* **5:** 1941–1959.
Hall, R.M., Brown, H.J., Brookes, D.E., and Stokes, H.W. (1994) *J Bacteriol* **176:** 6286–6294.
Hall, R.M., and Vockler, C. (1987) *Nucl Acids Res* **15:** 7491–7501.
Lévesque, C., Brassard, S., Lapointe, J., and Roy, P.H. (1994) *Gene* **142:** 49–54.
Martinez, E., and de la Cruz, F. (1988) *Mol Gen Genet* **211:** 320–325.
Martinez, E., and de la Cruz, F. (1990) *EMBO J* **9:** 1275–1281.
Nies, B.A., Meyer, J.F., Kratz, J., and Wiedemann, B. (1985) *Plasmid* **13:** 163–172.
Nücken, E.J., Henschke, R.B., and Schmidt, F.R.J. (1991) *Mol Gen Genet* **229:** 137–146.
Ouellette, M., and Roy, P.H. (1987) *Nucl Acids Res* **15:** 10055.
Ouellette, M., Bissonnette, L., and Roy, P.H. (1987) *Proc Natl Acad Sci USA* **84:** 7378–7382.
Rådström, P., Sköld, O., Swedberg, G., Flensburg, J., Roy, P.H., and Sundström, L. (1994) *J Bacteriol* **176:** 3257–3268.
Recchia, G.D., and Hall, R.M. (1995) *Mol Microbiol* **15:** 179–187.
Recchia, G.D., Stokes, H.W., and Hall, R.M. (1994) *Nucl Acids Res* **22:** 2071–2078.
Schmidt, F.R.J., Nücken, E.J., and Henschke, R.B. (1988) *Mol Microbiol* **2:** 709–717.
Schmidt, F.R.J., Nücken, E.J., and Henschke, R.B. (1989) *Mol Microbiol* **3:** 1545–1555.
Sadowski, P. (1986) *J Bacteriol* **65:** 341–347.
Sadowski, P. (1993) *FASEB J* **7:** 760–767.
Stark, M.W., Boocock, M.R., and Sherratt, D.J. (1992) *Trends Genet* **8:** 432–438.
Stokes, H.W., and Hall, R.M. (1989) *Mol Microbiol* **3:** 1669–1683.
Stokes, H.W., and Hall, R.M. (1991) *Plasmid* **26:** 10–19.
Stokes, H.W., and Hall, R.M. (1992) *Plasmid* **28:** 225–234.
Sundström, L., and Sköld, O. (1990) *Antimicrob Ag Chemother* **34:** 642–650.
Sundström, L., Rådström, P., Swedberg, G., and Sköld, O. (1988) *Mol Gen Genet* **213:** 191–201.
Sundström, L., Roy, P.H., and Sköld, O. (1991) *J Bacteriol* **173:** 3025–3028.
Tietze, E., Brevet, J., and Tschäpe, H. (1987) *Plasmid* **18:** 246–249.