The EMBO Journal (2010) 29, 2623–2634 | © 2010 European Molecular Biology Organization | All Rights Reserved 0261-4189/10
www.embojournal.org

THE EMBO JOURNAL

Cellular pathways controlling integron cassette site folding

Céline Loot[1,2,3], David Bikard[1,2,3], Anna Rachlin[1,2,4] and Didier Mazel[1,2,*]

[1]Institut Pasteur, Unité Plasticité du Génome Bactérien, Paris, France and [2]CNRS, URA2171–Génétique des génomes, Paris, France

By mobilizing small DNA units, integrons have a major function in the dissemination of antibiotic resistance among bacteria. The acquisition of gene cassettes occurs by recombination between the *attI* and *attC* sites catalysed by the IntI1 integron integrase. These recombination reactions use an unconventional mechanism involving a folded single-stranded *attC* site. We show that cellular bacterial processes delivering ssDNA, such as conjugation and replication, favour proper folding of the *attC* site. By developing a very sensitive *in vivo* assay, we also provide evidence that *attC* sites can recombine as cruciform structures by extrusion from double-stranded DNA. Moreover, we show an influence of DNA superhelicity on *attC* site extrusion *in vitro* and *in vivo*. We show that the proper folding of the *attC* site depends on both the propensity to form non-recombinogenic structures and the length of their variable terminal structures. These results draw the network of cell processes that regulate integron recombination.

The EMBO Journal (2010) **29,** 2623–2634. doi:10.1038/emboj.2010.151; Published online 13 July 2010
Subject Categories: genome stability & dynamics; microbiology & pathogens
Keywords: cruciform; palindrom; recombination; single-stranded DNA; superhelicity

Introduction

Integrons are genetic elements commonly found in bacteria from diverse phyla and environments (Mazel, 2006). They are defined as gene capture systems, which incorporate exogenous open reading frames and convert them to functional genes by ensuring their correct expression (Hall and Collis, 1995). The integron platform consists of three major elements: an *intI* gene coding a site-specific recombination enzyme belonging to the tyrosine-recombinase family (Azaro and Landy, 2002) called an integrase, a primary recombination site (*attI*), and an outwards oriented promoter (P_c), which directs transcription of captured gene cassettes (Hall and Collis, 1995). Gene cassettes generally contain a single gene and an imperfect inverted repeat at the 3′ end called an *attC* site. Recombination events between the *attI* and *attC* sites, leading to the insertion of gene cassettes in the platform, are the most common and efficient reactions performed by integron-associated integrases (Collis *et al*, 2001).

The length of natural *attC* sites varies from 57 to 141 bp. They include two regions of inverted homology, R″–L″ and L′–R′, that are separated by a central region that is highly variable in length and sequence (Stokes *et al*, 1997) (Figure 1A). Contrasting with their sequence heterogeneity, the *attC* sites display a strikingly conserved palindromic organization (Hall *et al*, 1991; Stokes *et al*, 1997; Rowe-Magnus *et al*, 2003) that can form secondary structures through the self-pairing of DNA strands (Figure 1B). On folding, single-stranded (ss) *attC* sites present an almost canonical core site consisting of R and L boxes separated by an unpaired central segment, two or three extrahelical bases (EHB) and a variable terminal structure (VTS) (Bouvier *et al*, 2009). The VTSs vary in length among the various *attC* sites from three predicted unpaired nucleotides such as for $attC_{aadA7}$ to a complex branched secondary structure in the larger sites such as the VCRs (*Vibrio cholerae attC* sites; Figure 1B).

It has been shown using a DNA-binding assay that IntI1 binds strongly and specifically to the bottom strand (bs) of ss *attC* DNA (Francia *et al*, 1999). *In vivo*, it has also been shown that only the bs of the *attC* site is used as a substrate during the integration of gene cassettes (Bouvier *et al*, 2005), thereby creating a pseudo-Holliday Junction (pHJ) between the ss *attC* and the double-stranded (ds) *attI* site. It is currently believed that this pHJ is resolved through host processes, but these have yet to be identified.

Many questions about the specific nature of these unique gene capture systems still remain unanswered. Specifically, we do not yet know how and when the *attC* sites fold inside the cell, or the identity of the factors influencing these processes.

It has long been clear that structures more complex than the canonical double helix B-form DNA are biologically important. These include unpaired or mismatched bases, triplex DNA, Z-form DNA, hairpin loops, cruciforms and Holliday junctions (Bacolla and Wells, 2004). DNA secondary structures have been detected in both prokaryotes and eukaryotes and can be assembled essentially by two pathways: from an ssDNA by generating a hairpin structure or from dsDNA by extrusion of cruciform structures (Figure 1C).

The first pathway requires ssDNA, which can be produced during conjugation, natural transformation, viral infection, replication, transcription and DNA repair. Here, we examined the effect of ss availability mediated by two cellular processes considered as the principal sources of ssDNA in bacterial cells, conjugation and replication (Figure 1C). During conjugation, the transferred DNA is in essence single-stranded, whereas during replication, the lagging strand contains a

*Corresponding author. Département Génomes et Génétique, Institut Pasteur, Unité Plasticité du Génome Bactérien, 25 rue du Dr Roux, Paris F-75015, France. Tel.: +33 1 40 61 32 84; Fax: +33 1 45 68 88 34; E-mail: mazel@pasteur.fr
[3]These authors contributed equally to this work
[4]Present address: McLean Hospital, Behavioral Genetics Lab, MRC 215, Mill St Belmont, MA 02478, USA

Received: 25 November 2009; accepted: 11 June 2010; published online: 13 July 2010

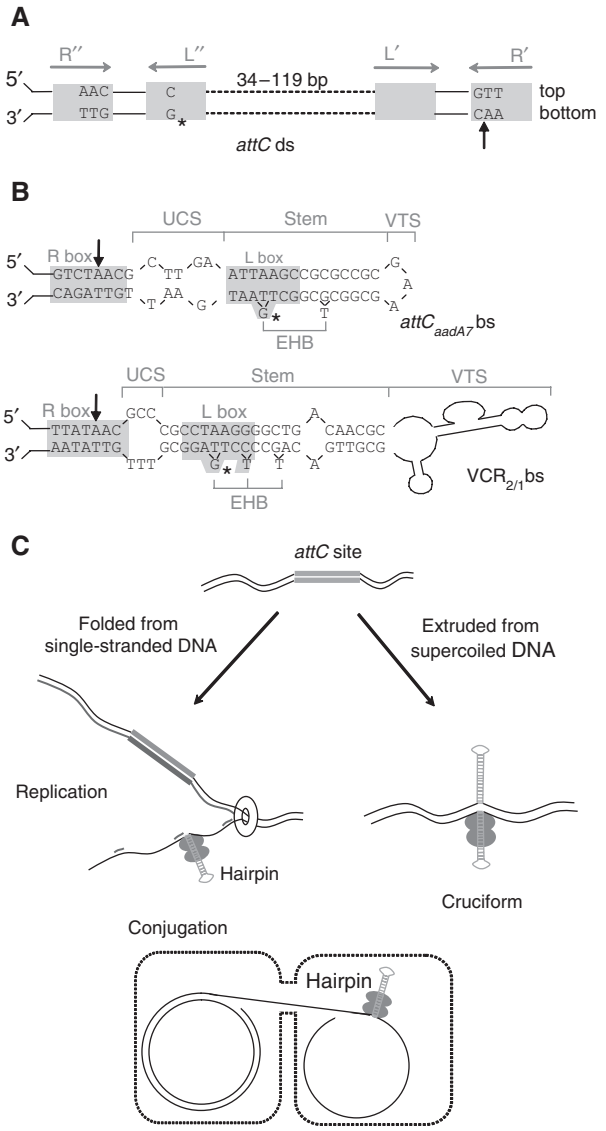


Figure 1 *attC* recombination sites and model for the *attC* folding. (**A**) Schematic representation of a double-stranded (ds) *attC* site. Inverted repeats R″, L″, L′ and R′ are indicated by grey boxes. The dotted lines represent the variable central part. The conserved nucleotides are indicated. Asterisks (*) show the conserved G nucleotides, which generate extrahelical bases (EHB) in the folded *attC* site bottom strand (bs). The black arrow shows the cleavage point. (**B**) Secondary structures of the $attC_{aadA7}$ and $VCR_{2/1}$ sites bottom strands (bs). Structures were determined by the UNAFOLD online interface at the Institut Pasteur. The four structural features of *attC* sites, namely, the unpaired central segment (UCS), the EHB, the stem and the variable terminal structure (VTS) are indicated. Black arrows show the cleavage points. Primary sequences of the *attC* sites are shown (except for the VTS of the $VCR_{2/1}$ site). (**C**) Cellular processes possibly allowing the *attC* site folding. The different possible cellular pathways allowing proper folding of the *attC* site bottom strand are shown: from single-stranded DNA during replication and conjugation, and cruciform extrusion from supercoiled double-stranded DNA. IntI monomers are represented as grey ovals.

region of ssDNA reflecting the length of an Okazaki fragment (1–2 kb). These may, therefore, provide an opportunity for the formation of DNA secondary structure. Our results show that the availability of ssDNA in the cell during conjugation and/or during replication (lagging strand) could favour the folding, and thus recombination, of *attC* sites of various lengths. However, the results could not be fully explained by these ssDNA production pathways, suggesting that other processes are likely involved. We thus considered the possibility that *attC* sites could be extruded from double-strand DNA into a cruciform structure (Figure 1C). Indeed, inverted repeats (perfect or imperfect) have the potential to form branched structures called cruciforms, in which inter-strand base pairing within the symmetric region is replaced by intra-strand base pairing (Courey, 1999). The formation of a cruciform by an inverted repeat involves a great deal of structural disruption as it requires a complete reorganization in base pairing. Furthermore, superhelicity is expected to directly influence cruciform extrusion. We tested this possibility as a second *attC*-folding pathway, and found that the *attC* sites could recombine from DNA essentially under ds form, after extrusion of the cruciform structure in a superhelicity-dependent manner. We confirmed these results by *in vitro* cruciform detection. We also showed that two parameters are implicated in the proper folding of the bs of the *attC* site: the length of the VTS and the propensity of the *attC* site to form a non-recombinogenic structure. Our results suggest that the contribution of these different processes varies as a function of the length and the sequence of the *attC* sites.

These results show that ssDNA structures, be they generated from replication, conjugation or extruded from dsDNA, can be recruited for specific processes such a site-specific recombination. The interplay between these cellular processes governs folding of *attC* sites, and certainly allows regulation of integron recombination by the host.

Results

Influence of single-strand DNA availability during conjugation on attC folding

We have earlier shown using a conjugation-based assay that the *attC* sites, contrarily to the *attI* sites, recombine as a folded structure generated from the bs of the *attC* site (Bouvier *et al*, 2005) (Figure 1B).

Natural *attC* sites sizes vary from 57 to 141 nt. To determine whether these size limits result from the constraints linked to the folding of ssDNA, we made a series of 21 *attC* site derivatives (Figure 2; Supplementary Figure S1 and Supplementary Table S1 and S2). Starting from the $VCR_{2/1}$ (the 123 bp signature *attC* site of the *V. cholerae* superintegron) (Mazel *et al*, 1998), we increased the length of the stem and/or VTS (up to sites of 180 nt). Inversely, we serially deleted the central part of the $VCR_{2/1}$ site to make shorter derivatives (down to 56 nt). Three wild-type *attC* sites were also tested: $attC_{aadA7}$, $attC_{ereA2}$ and $attC_{oxa2}$. We first tested them in the suicide-conjugation assay we earlier developed (Bouvier *et al*, 2005). This assay uses conjugation, which proceeds exclusively through ssDNA transfer, to deliver the *attC* site in ss form to a recipient cell expressing the IntI1 integrase and carrying the *attI1* recombination site. The *attC* site provided by conjugation is carried on pSW plasmid that cannot replicate in the recipient (Demarre *et al*, 2005). This system permits the delivery of DNA in ss form to provide a substrate for recombination. We verified that all the constructed pSW::*attC* plasmids were transferred at similar rates ($\sim 3 \times 10^{-1}$) using a recipient strain able to sustain pSW replication.

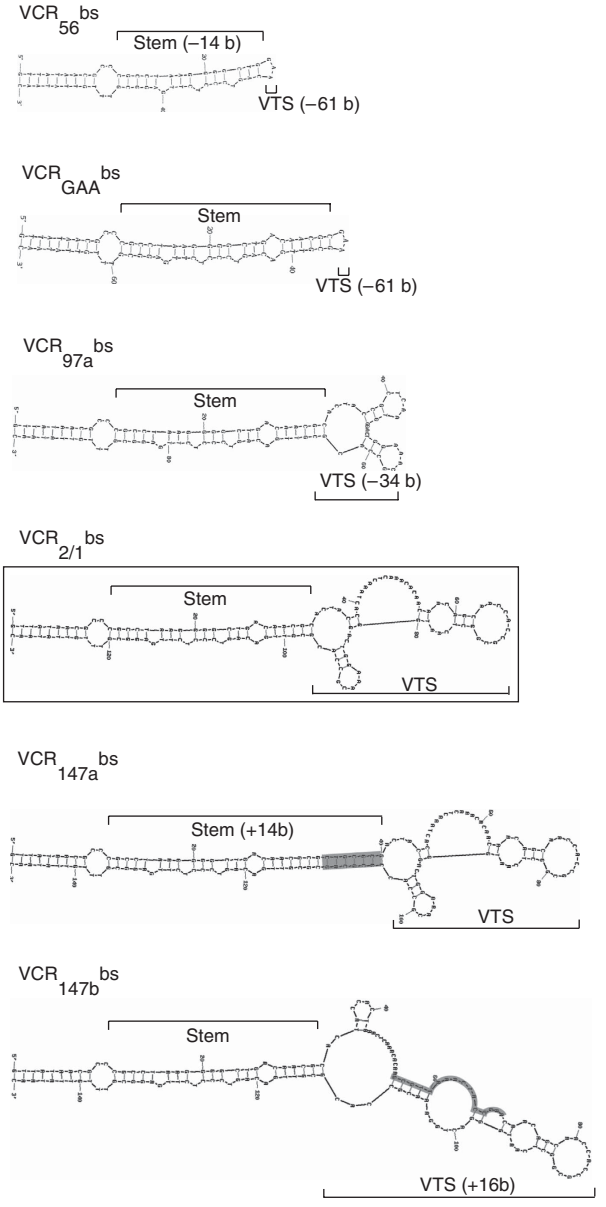


Figure 2 Predicted secondary structures of several constructed $VCR_{2/1}$ site derivatives. Secondary structures were determined using the UNAFOLD online interface of the Institut Pasteur. The *attC* sites are classified according to their size (smallest to largest). The natural $VCR_{2/1}$ site is boxed. Modifications made from the natural $VCR_{2/1}$ site in the stem and/or in the VTS are indicated.

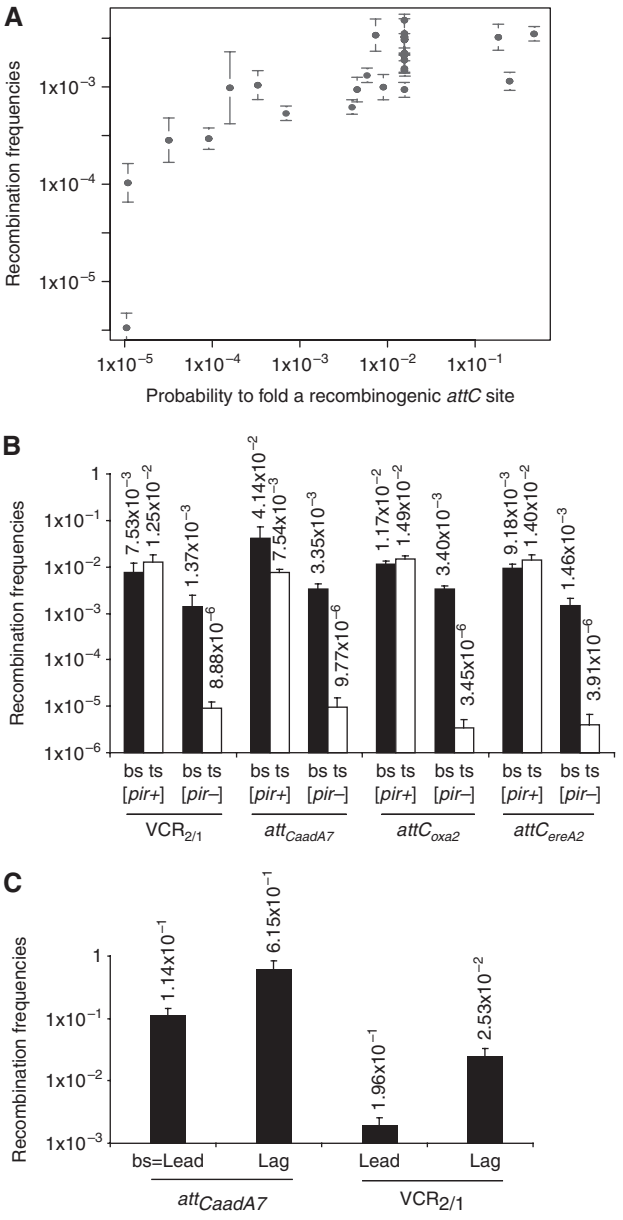


Figure 3 Influence of processes delivering single-stranded DNA on *attC* recombination. (**A**) Recombination frequencies of the different *attC* sites in the 'suicide-conjugation' assay as a function of their probability to fold a recombinogenic *attC* site (see Additional materials: suicide-conjugation assay). The UNAFOLD software was used (see Additional materials). (**B**) Recombination frequencies of four natural *attC* sites after conjugation in either a replication permissive *pir*+ or non-permissive *pir*− recipient (see Additional materials: suicide-conjugation assay). Black and white columns correspond, respectively, to the recombination frequencies established when the bottom strand (bs) or the top strand (ts) is injected by conjugation. (**C**) Recombination frequencies of the $VCR_{2/1}$ and $attC_{aadA7}$ when the recombinogenic bs is carried on the lagging (lag) or leading (lead) strand of the replicating molecule (see Materials and methods: recombination assay with unidirectional-replicative substrate). Error bars show s.d.

The results are shown in Figure 3A. Recombination frequencies (see Supplementary Table S3) are plotted as a function of the probability of the *attC* site to fold into a recombinogenic site. We define a recombinogenic site as forming the R box, as well as having the G16 EHB of the L box (see Additional materials). The UNAfold software was used to estimate the probability to form active sites (Zuker, 2003). We observed that the recombination frequencies tended to drop for sites with very low probabilities to fold properly. Nevertheless, most of the *attC* sites displayed similar recombination frequencies when delivered by conjugation, regardless of their size or VTS length (Figure 2; Supplementary Figure S1 and Supplementary Table S3). This was in agreement with earlier observations (Bouvier *et al*, 2009).

These results suggested that if an upper limit to the length of *attC* sites exists, it is not constrained by the availability of ssDNA during conjugation.

In a second set of experiments, we compared the recombination frequencies of the *attC* sites when the bs or the top

strand (ts) was injected into recipient cells that either could or could not sustain their replication. Four natural *attC* sites were tested: $VCR_{2/1}$, $attC_{aadA7}$, $attC_{oxa2}$ and $attC_{ereA2}$. They were all cloned into the pSW::*attC*-B and pSW::*attC*-T plasmids, which permit the delivery of the bs and the ts, respectively, through conjugation (Bouvier *et al*, 2005). As earlier observed, when the ts of the *attC* site was injected into a *pir*− recipient strain, we obtained a much lower recombination frequency (Figure 3B). This confirmed that the *attC* sites are essentially ss in this assay. If they were not, one would expect similar recombination frequencies for the ts and the bs (Bouvier *et al*, 2005). We then performed the same experiment using a *pir*+ recipient cell that permitted the replication of the pSW::*attC* substrate once transferred. In these conditions, we observed for both pSW::*attC*-B and pSW::*attC*-T plasmids an equivalent high efficiency of recombination (Figure 3B). This suggested that replication can induce recombination at a similar or higher frequency than conjugation, bringing the recombination frequencies of the two transferred strands in this assay to the same level.

We conclude that recombination can happen both during the delivery of ssDNA (the bs) by conjugation and from other processes involving a replicating molecule. Recombination may indeed occur during replication-mediated DNA melting and/or during the extrusion of the *attC* site in a cruciform structure from dsDNA.

Influence of single-strand DNA availability during replication on attC folding

Replication is a process that transiently produces ssDNA. This in turn could regulate folding of the *attC* site (Figure 1C). On the basis of the differences in the dynamics between the lagging and leading strands during DNA replication (Wolfson and Dressler, 1972), proper bs *attC* folding is probably favoured by its localization on the lagging strand in which large regions of ssDNA are available (the length of the Okazaki fragments: 1–2 kb; Trinh and Sinden, 1991). If there are differences in recombination frequency based on which strand the *attC* is on, this would support the theory that the bs can be folded independently of the ts as a hairpin structure.

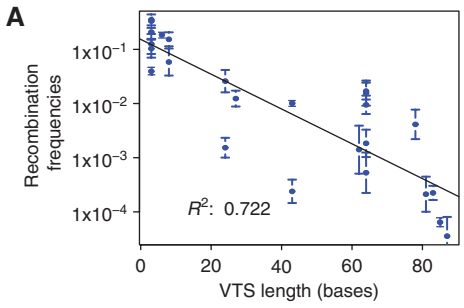
To test this hypothesis, we inserted an *attC* site (either $attC_{aadA7}$ or $VCR_{2/1}$) in both orientations into a unidirectional-replicating pTSC plasmid (Phillips, 1999) (Supplementary Table S1), so that the bs of the *attC* site is either on the leading or on the lagging strand. For the purpose of the experiment, the origin of replication is thermosensitive (oripSC101ts). These pTSC::*attC* plasmids were introduced into strain UB5201 with two others plasmids, one containing the *attI* site (pSU38Δ::*attI1*) and the other carrying the *intI1* gene (pBAD::*intI1*). Assays were performed at 30°C in the presence of arabinose ensuring the expression of the integrase gene. Recombination events between *attI* and *attC* sites were then selected on plates at 42°C with the pTSC::*attC* plasmid resistance marker (Cm). At 42°C, the temperature-sensitive pTSC::*attC* plasmids are unable to replicate and, therefore, cells containing these plasmids do not grow on Cm, unless there was a recombination event producing a cointegrate between the pTSC::*attC* and pSU38Δ::*attI1* plasmids. The results are shown in Figure 3C. We obtained for pSW::$attC_{aadA7}$ and pSW::$VCR_{2/1}$ a 4.2- and 12.1-fold increase in recombination, respectively, when the bs of these *attC* sites corresponds to the lagging strand. As a control, the same experiment was performed on a bidirectional-replicating plasmid. As expected, we did not observe any significant differences between the two orientations for the *attC* site (data not shown).

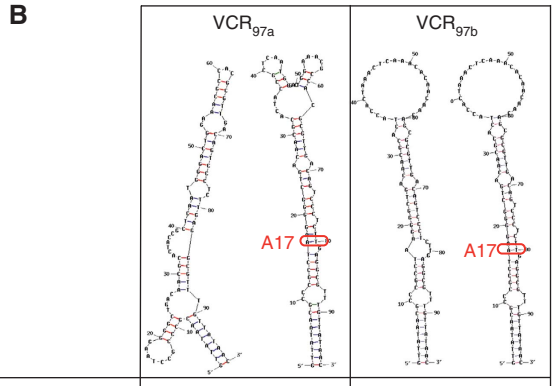
Remarkably, there was still a high rate of recombination when bs of $attC_{aadA7}$ and $VCR_{2/1}$ were carried on the leading strand (1.14×10^{-1} and 1.96×10^{-3}, respectively; Figure 3C). Very little ssDNA is produced on the leading strand, making hairpins unlikely to fold. This led us to suspect another pathway.

Finally, these results show that ss production in both conjugation and replication influences and favours folding of the *attC* site, but other mechanisms are implicated. Specifically, we then studied the ability of the *attC* sites to extrude from dsDNA as cruciform structures able to recombine (Figure 1C).

Influence of double-strand and single-strand DNA availability on attC folding

We used all earlier tested *attC* sites derivatives (Figure 2; Supplementary Figure S1) in the replicative condition assay. Contrarily to the conjugation assay, *attC* sites are provided on a replicative plasmid and are thus mostly ds over the cell cycle, being only transiently ss during replication. In these conditions, the various sites display markedly different recombination frequencies (Supplementary Table S3). We observe a correlation between the VTS length and the recombination frequency (Figure 4A); *attC* sites containing a minimal VTS (3 nt) recombined at a frequency ranging from 1.01×10^{-1} (VCR_{GC}) to 3.53×10^{-1} ($attC_{aadA7}$), whereas *attC* sites containing a larger VTS recombined at a lower frequency, from 4.37×10^{-5} (VCR_{147c}) to 1.01×10^{-2} (VCR_{116b}). A negative effect of the VTS length on recombination is in agreement with the hypothesis of cruciform formation. Indeed, to go from a ds state to a cruciform, an *attC* site needs to melt at least the length of its VTS (see Discussion). The energy to melt this region can thus be considered as the energy of activation of cruciform formation, and thus is directly correlated to the probability of forming a cruciform after the Arrhenius equation (Supplementary Figure S2). Therefore, the larger the VTS is, the lower the probability of folding into a cruciform, which is what we observed. Furthermore, we found that the propensity of the *attC* site to form non-recombinogenic secondary structures directly impacts the recombination frequency. This is illustrated in Figure 4B, which represents the most favourable structure for the two 'a' and 'b' versions of VCR_{97} (see ΔGc and ΔG). The 'a' and 'b' versions only differ by a few bases substitutions in the VTS, but these point mutations have an impact on the formation of improperly folded structures. Indeed, the most favourable structure formed by VCR_{97a} differs greatly from the recombinogenic one and is most likely not an active substrate for recombination. On the other hand, the most stable structure for VCR_{97b} is very similar to that of the recombinogenic site. It might even be recognized by the integrase, which could help reach the active conformation. Not surprisingly, VCR_{97a} has a 20-fold lower recombination frequency than VCR_{97b}. The same effect of non-recombinogenic structures can also explain the 40-fold difference between the VCR_{116a} and VCR_{116b} (Supplementary Figure S1 and Table S3).





ΔGc (kcal mol^{-1})	27.05	31.41	28.07	30.72
ΔG (kcal mol^{-1})	−22.22	−17.86	−19.54	−16.89
Probability to fold a recombinogenic *attC* site	7.31×10^{-3}		1.56×10^{-2}	
Recombination frequencies	1.66×10^{-3}		2.81×10^{-2}	

Figure 4 *attC* recombination, in 'replicative' recombination conditions. (**A**) Recombination frequencies of the different *attC* sites in the 'replicative' assay as a function of their VTS length (see Additional materials: recombination assay with a replicative double-stranded substrate). Error bars show s.d. (**B**) Secondary structures of VCR97a and VCR97b as predicted by UNAFOLD. The free energy of single-strand *attC* site folding (ΔG) and of cruciform extrusion (ΔGc), the probability to fold a recombinogenic *attC* site and its measured recombination frequencies are indicated for each site. The A17, whose probability to bind the corresponding T is used as *a proxi* for the probability of the *attC* site to fold in a recombinogenic structure, is highlighted. For the calculation of the probability to fold a recombinogenic *attC* site and the free energy of cruciform formation, see Additional materials.

It thus seems that two main parameters are instrumental for the recombination frequency of *attC* sites in these conditions. First, the longer the VTS is, the lower the recombination frequency. Second, the accumulation of non-recombinogenic (improperly folded) sites dilutes the number of recombinogenic (properly folded) sites available for recombination. Note that these two parameters are not independent, as the longer is the VTS, the higher is the chance to form stable non-recombinogenic structures. An analysis of covariance performed with these two regressors (the length of the VTS and the probability of folding properly) explains 82.5% (R^2) of the experiment variance with a *P*-value of 4.72×10^{-9}, confirming the implication of these two parameters.

This, combined with the fact that high recombination frequencies can be obtained in replicative conditions when the bs of the *attC* site is on the leading strand of replication, strengthens the hypothesis that recombination can occur with *attC* sites that are extruded as cruciform structures.

Detection of attC cruciform structures

To test whether the *attC* sites could be directly extruded as cruciform structures from a dsDNA molecule, we performed complementary *in vitro* and *in vivo* analysis.

In vitro mapping of S1 nuclease-sensitive sites. Inverted repeat sequences in natural plasmids and phages have been shown to be centrally hyper-sensitive to cleavage by single-strand selective nucleases (Lilley, 1980; Panayotatos and Wells, 1981). Indeed, when folded into hairpins, they exhibit not only ssDNA in their loop, but also potential bulges of their stem. S1 nuclease, which cleaves ssDNA, was earlier used to probe the formation of cruciform structures *in vitro* (Noirot *et al*, 1990). We carried out the same kind of experiments to detect *in vitro* formation of *attC* cruciform structures (Figure 5A). The method consists of treating supercoiled plasmid DNA containing the *attC* sites with S1 nuclease. This is followed by a restriction digest with an enzyme having a single site in the molecule, which produces pairs of fragments arising from molecules linearized by S1 nuclease. Restriction enzymes used and expected band sizes (corresponding to the cruciform detection) are mentioned in Figure 5B. This assay showed that cruciform formation was occurring at detectable rates from the $attC_{aadA7}$ site, but not from the $VCR_{2/1}$ site (Figure 5C). These results coincided with the fact that $attC_{aadA7}$ has a very short VTS (3 nt) relative to that of $VCR_{2/1}$ (61 nt), allowing cruciform extrusion at a much higher frequency. To confirm this observation, we performed the same experiment, but using a mutant derivative of $attC_{aadA7}$, the $attC_{aadA7Mut3}$ site and a mutant derivative of VCR, the VCR_{GAA} site (Figure 5A and C). The mutations in the $attC_{aadA7Mut3}$ disrupt the base pairing of the upper part of the stem and have been shown to decrease the recombination by >100-fold (Bouvier *et al*, 2005). As expected, we failed to detect bands indicating cruciform extrusion of this $attC_{aadA7}$ site derivative. On the contrary, VCR_{GAA}, which contains a small VTS, ensured cruciform extrusion at a higher level than the natural VCR site. These results coincide with the high recombination frequency of $attC_{aadA7}$ (3.53×10^{-1}) and VCR_{GAA} (2.16×10^{-1}) in 'replicative' conditions (Supplementary Table S3).

To strengthen our hypothesis of $attC_{aadA7}$ site cruciform extrusion, we determined the precise boundaries of the S1 nuclease-generated fragments from a representative sample by sequencing. Figure 5D shows the distribution of the cleavage sites within the $attC_{aadA7}$ site and its neighbouring sequences. Among 91 sequenced clones, 73% (66/91) of the S1-cleavage sites clearly localize into $attC_{aadA7}$ at/near expected structural features, drawing a high-resolution structure map of the cruciform. The 25 remaining clones revealed S1-cleavage sites on both sides of the $attC_{aadA7}$ site up to a distance of 40 bases and could be explained by non-*attC*-specific extrusions of nearby inverted repeats (Supplementary Figure S4).

Recombination from non-replicative ds plasmid. In the 'replicative' assay described above, the *attC* site is carried by a dsDNA plasmid able to replicate and, therefore, to transiently produce ssDNA. To precisely study the ability of the *attC* site to recombine as a cruciform structure, we developed an assay ensuring the delivery of the *attC* site exclusively from dsDNA. To this end, Pir-dependent plasmids containing the *attC* sites were introduced by transformation in *pir*− deficient strains.

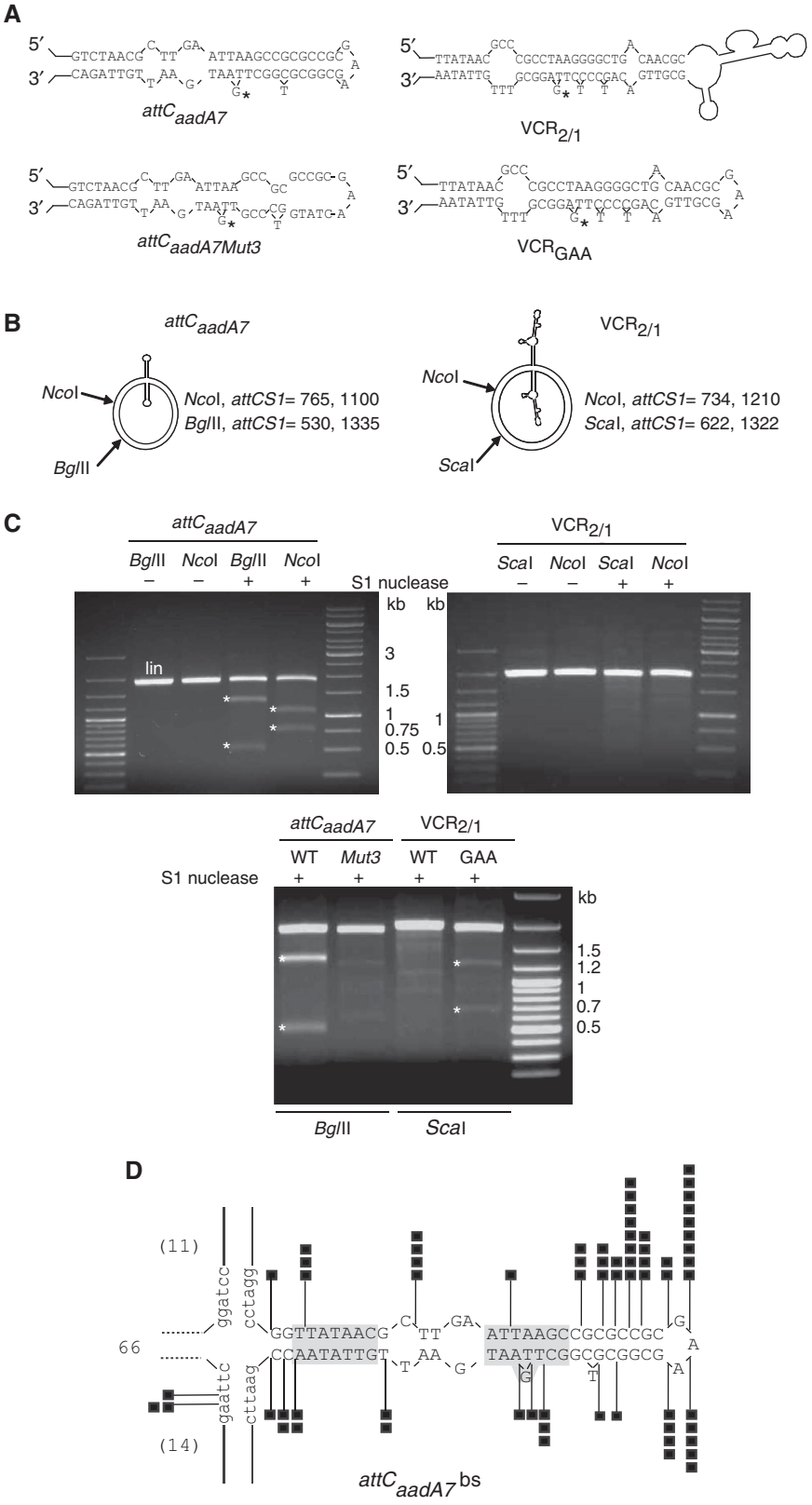


Figure 5 *In vitro* extrusion of the *attC* sites secondary structures from double-stranded DNA. (**A**) Schemes of the folded natural and mutants *attC* sites used in the S1 nuclease assay. (**B**) Schemes of the plasmids with extruded $attC_{aadA7}$ and $VCR_{2/1}$ sites and the expected S1 fragments sizes (in bp) after cleavage by the different restriction enzymes. (**C**) *In vitro* mapping of S1 nuclease-sensitive sites. S1 nuclease-treated (+) or -untreated (−) plasmids were subjected to restriction digest (enzymes are indicated above each lane) and submitted to electrophoresis. The bands (white asterisks) smaller than the linear monomer (lin) result from S1 nuclease action and thus of cruciform extrusion. Kb, marker DNA. (**D**) Precise mapping of S1-cleavage position on the $attC_{aadA7}$-folded site. Each square corresponds to one S1-cleavage position in the sequenced representative sample.

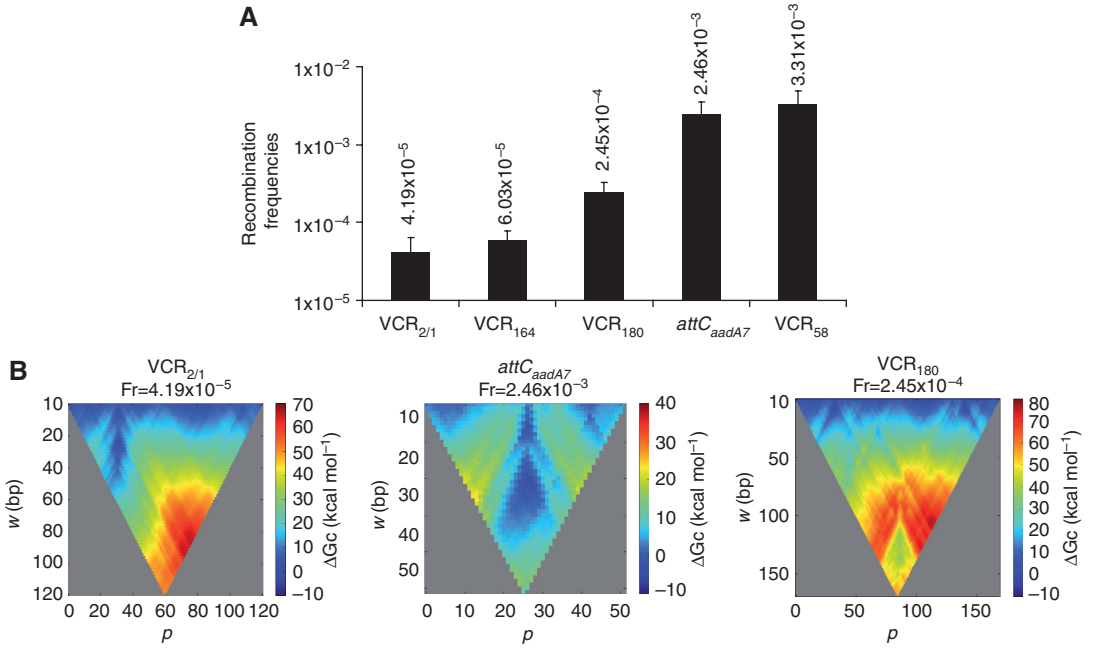


Figure 6 *In vivo* extrusion of the *attC* sites secondary structures from double-stranded DNA. (**A**) Recombination frequencies of different *attC* sites after transformation of *attC*-containing plasmids in non-replicative permissive recipient cells (see Materials and methods: recombination assay with a non-replicative substrate). Error bars show s.d. (**B**) Energy landscape of cruciform formation. The frequency of recombination (Fr) for each *attC* site is indicated. Each colour dot represents the free energy of cruciform formation (ΔGc in kcal/mol) for a portion of the *attC* site of length *w* (bp) at the position *p* (from favourable in blue to unfavourable in red). The tip of the triangle is the free energy of the whole *attC* site. As *attC* sites are symmetrical, favourable energies that are not along the centre of the triangle represent favourable non-recombinogenic structures. $attC_{aadA7}$ folds in a favourable recombinogenic site (blue colour from the middle of the base of the triangle). $VCR_{2/1}$ folds in a favourable non-recombinogenic structure (blue colour shifted from the middle). VCR_{180} presents neither favourable recombinogenic nor favourable non-recombinogenic structures.

These plasmids can only be maintained on recombination with the *attI* site carried by a Pir-independent replicon. To establish a recombination frequency, we transform in parallel a *pir*+ permissive strain (UB5201-Pi) with the same samples of pSW-*attC* plasmids preparation. Note that the transformation efficiency of both UB5201 and UB5201-Pi strains are determined beforehand and used to adjust the recombination frequency (see Materials and methods). We found that all tested *attC* sites could lead to detectable cointegrate formation through recombination with the *attI* site (Figure 6A). These results strongly suggested that *attC* sites substrate could be formed by cruciform extrusion from dsDNA. Nevertheless, we cannot exclude the presence of nicked/damaged molecules in our plasmid preparation. Those molecules could allow *attC* site folding from the ssDNA generated during their repair and explain the obtained recombination events. However, as we failed to detect these nicked molecules by electrophoretic gel analysis, we concluded that, if they exist, these molecules represent a very minor part of the supercoiled plasmid preparation. If this minor part accounts for the majority of the recombination events, we should observe a much higher recombination frequency when all molecules are damaged. To test this hypothesis, we transformed the same cellular setup with identical quantities of either nicked or supercoiled molecules containing the $attC_{aadA7}$ site (see Additional materials; Supplementary Table S1 and S2). We modified the plasmids to introduce a single Nb-Bts1 endonuclease site either at 33 or 296 nt away from the *attC* site, and in the two orientations at each locations. This endonuclease only cuts one strand, producing plasmids carrying a nick on either the bs or ts, depending on the orientation of the endonuclease site. The nicked molecules showed recombination frequencies similar to the supercoiled plasmid (Supplementary Figure S3). As none of these nicked molecules presented a higher frequency of recombination than the supercoiled ones, these results confirmed that the presence of a minor part of damaged/nicked molecules could not account for the frequency of recombination obtained in this assay. It is yet to be determined how these nicked molecules recombine. The most likely explanation is that nicked plasmids are rapidly ligated (Heitman *et al*, 1989), allowing the introduction of supercoils and recombination through *attC* site cruciform extrusion. On the other hand, we observed variations in the efficiency of recombination depending on the *attC* sites. It can certainly be explained by the formation of non-recombinogenic structures. This is illustrated by comparison of the recombination frequencies and the energy landscapes of cruciform formation (ΔGc) from $VCR_{2/1}$ and $attC_{aadA7}$ (Figure 6B). An improperly folded structure for the natural $VCR_{2/1}$ site is clearly visible (blue colour shifted from the middle of the base of the triangle) and correlates with its relatively low recombination frequency. In contrast, the most efficient site, $attC_{aadA7}$, presents a very favourable energy landscape to fold as a recombinogenic site (blue colour from the middle). The VCR_{180} site presents neither, favourable energy landscape to fold a recombinogenic site, nor improperly folded structure, and recombines to an intermediate frequency of recombination. Here again, both the length of the VTS and the propensity to fold into non-recombinogenic

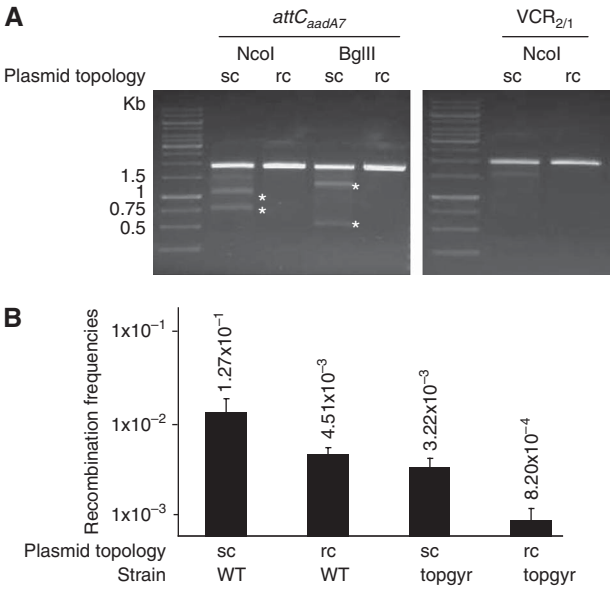


Figure 7 Influence of superhelicity on *attC* folding. (**A**) Mapping of S1 nuclease-sensitive sites of topoisomers products. Plasmid topological status is indicated by: sc, supercoiled and rc, relaxed by topoisomerase I treatment. The estimated band sizes for the extrusion of $attC_{aadA7}$ and $VCR_{2/1}$ in the pSW plasmid are given in Figure 5B. Kb, marker DNA. (**B**) Effect of topoisomerase I activity on the recombination frequency of the $attC_{aadA7}$ site in 'non-replicative' recombination assay. The assay was performed in two genetic backgrounds, wild type (WT) and *topA10 gyrB266* (topgyr, toposisomerase I and DNA gyrase-deficient strain) and using supercoiled and relaxed *attC*-containing plasmids.

structures seem to explain the recombination frequencies of the *attC* sites.

These results confirmed that substrate *attC* sites could be formed by cruciform extrusion from dsDNA and, therefore, lead us to test the effect of superhelicity on the recombination of *attC* sites under dsDNA form.

Influence of superhelicity on integron recombination

Cruciform extrusion requires the opening of the DNA double helix to allow intra-strand base pairing. The free energy held by the supercoiled molecule in the form of torsional underwinding is required for stabilization of cruciforms (Lilley *et al*, 1985). We first tested the effect of supercoiling on *attC* cruciform extrusion *in vitro*. For these, we used topoisomerase I, which catalyses the relaxation of negatively supercoiled DNA by introducing single-strand breaks that are subsequently religated. Relaxed DNA cannot stabilize cruciform structures. As expected, we observed an inhibitory effect of topoisomerase I treatment on $attC_{aadA7}$ cruciform extrusion *in vitro* (Figure 7A).

Second, to investigate the influence of supercoiling on the *attC* site folding *in vivo*, we used the earlier sample of negatively supercoiled and relaxed plasmids to transform a JTT1 WT strain containing the *attI* site and expressing the integrase. As the JTT1 strain is a *pir*− deficient strain, the Pir-dependent plasmids containing the *attC* sites cannot be selectively maintained without recombination with the *attI* site carried by the Pir-independent replicon. To establish a recombination frequency, we transformed in parallel a *pir*+ permissive strain (UB5201-Pi) with the same negatively supercoiled and relaxed plasmids samples. Here again, the transformation efficiency of the both JTT1 and UB5201-Pi strains were determined beforehand and used to adjust the final ratio and normalize the results (see Materials and methods).

Negatively coiled plasmids appeared to recombine 2.8 times better than plasmids that are relaxed (Figure 7B). However, in this experiment, the topoisomerases and gyrases of the WT strain could act on the transformed plasmids by changing their supercoiling state before they might be recombined. Therefore, we repeated the experiment in the SD7 strain, a *topA10 gyrB266* derivative of JTT1, which, because of its mutations in topoisomerase I and DNA gyrase, has a lower intracellular level of negative supercoil density (Napierala *et al*, 2005). Calculation of the average superhelical density ($-\sigma_{av}$) of the pUC19 reference plasmid isolated from JTT1 and SD7 strains were 0.057 and 0.049, respectively (Napierala *et al*, 2005). In SD7, supercoiled plasmids recombined 3.9 times more efficiently than the relaxed ones (Figure 7B). Thus, superhelicity directly affects integron recombination, supporting the cruciform extrusion pathway for *attC* site recombination.

Discussion

DNA secondary structures

Extensive studies of DNA secondary structure during the past decades have shown that DNA is a dynamic molecule whose structure depends on the underlying nucleotide sequence and is influenced by the environment and the overall DNA topology. Several non-B-DNA structures have been described (Z-DNA, triplex DNA, unpaired DNA bases and hairpin or cruciform structures), which can be formed under physiological conditions. Circumstantial evidence suggests that secondary structures may serve functions in processes such as transcription regulation (Hartvig and Christiansen, 1996; Dai and Rothman-Denes, 1999), conjugation (Guasch *et al*, 2003), initiation of replication (Khan, 2005) or replication slippage (d'Alencon *et al*, 1994; Bierne *et al*, 1997). More recently, it has been shown that secondary structures are implicated in the recombination of genetic elements. Indeed, Xer recombinases can promote the direct integration of the (+)ssDNA genome of CTX into the ds *dif* site of *V. cholerae*. ssDNA substrates for integration can fold into a stem-loop structure, creating a small region of duplex DNA that is the target of site-specific recombinases (Val *et al*, 2005). Secondary structures are also implicated in the transposition of IS*608* of *Helicobacter pylori*, in which the TnpA transposase recognizes and cleaves only the top strand of the IS*608* ends that have folded into hairpin structures (Ton-Hoang *et al*, 2005; Guynet *et al*, 2008).

In the integron recombination process, the *attC* site is recognized by the integrase as an ss-folded substrate (Bouvier *et al*, 2005, 2009; MacDonald *et al*, 2006; Frumerie *et al*, 2010). To adopt this structural state, the DNA double helix needs the opening of inter-strand base pairing. As analysis of the integrase protein sequence failed to identify any helicase domain, we investigated this hypothesis that folding of the *attC* site could either be spontaneous and energy driven in supercoiled DNA, or could be driven by host factors. In this study, we tested the contribution of two pathways that could be involved: hairpin formation linked to

single-strand availability and/or a cruciform extrusion from supercoiled dsDNA.

For these, we used three different *in vivo* approaches, which differ significantly in terms of *attC* copy numbers and replicative status of the DNA. The 'suicide-conjugation assay' delivers by conjugation one copy of ss *attC*-containing plasmid (non-replicative) in the recipient cell. The 'non-replicative assay' delivers by transformation, one or more copies of the *attC*-containing non-replicative plasmid in the recipient cell. In the 'θ-replicative assay', a large number of replicative copies of the *attC*-carrying plasmid is contained in the recipient cell. Note, that the experimental procedures of these approaches are different and that it would be unwise to compare their respective recombination efficiencies.

Single-strand availability

Single-strand DNA production is obviously the most straightforward process allowing the folding of DNA into secondary structures. ssDNA is central for most examples of horizontal gene transfer. During natural transformation of bacteria, one DNA strand is taken up into the cytoplasm, whereas its complementary strand is degraded. During conjugation, ssDNA is unwound from the duplex plasmid and transferred into the recipient bacterium. During infection, filamentous phages such as M13, MV-L51 or ϕX174 are known to inject ssDNA. Inside the cell, there are essentially three processes that create ssDNA, replication, repair and transcription. During transcription, the size of the ssDNA is limited by the maximal size of the transcription bulge (25 nt) (Gamper and Hearst, 1982), and is likely not implicated in *attC* folding, as this size is too small. During DNA repair, the processing of double-strand breaks by the RecBCD complex is a significant source of RecA nucleofilaments, (i.e. the assemblage of RecA monomers on ssDNA) (Spies *et al*, 2005). Replication seems very appropriate to favour DNA secondary structures. During this process, after the melting of dsDNA by the replication machinery, an asymmetric fork is created in which one of the two strands (the lagging strand) provides a large quantity of ssDNA. Published observations suggest that secondary structures are easily made when carried on the lagging strand (Trinh and Sinden, 1991).

We carefully analysed the impact of conjugation and θ-replication processes on the folding of the *attC* site bs and on the integron recombination. We showed that conjugation ensures the folding of *attC* sites containing a larger VTS (e.g. VCR) or a shorter VTS (e.g. $attC_{aadA7}$) with even efficiency. We also showed that when carried on the lagging strand of the replicated DNA, the *attC* bs is recombined at a higher rate, showing that the availability of ssDNA impacts the recombination frequency of *attC* sites. These results show that replication can regulate the efficiency of integron recombination. We, therefore, examined the orientation of *attC* sites on all the chromosomal integrons encountered in sequenced bacterial genomes. We observed that the bs of all *attC* sites were located on the leading strand (Supplementary Figure S5). This specific orientation of *attC* sites could limit cassette rearrangements in chromosomal integrons. Nevertheless, it has been shown that DNA damage can uncouple the replication of the leading and lagging strand forming a partially ds molecule with an ss region of about 1 kb on the leading strand (Pages and Fuchs, 2003; Wang, 2005; Langston and O'Donnell, 2006). So, in this precise situation, which can be associated with stress conditions, we cannot exclude a high frequency of integron recombination even if the *attC* site bs is carried by the leading strand, thus ensuring a rapid adaptation of the stressed cells.

Double-strand extrusion

DNA sequences that possess two-fold symmetry may re-organize their base pairing to form cruciform structures, in which there is local intra-strand hydrogen bonding (Lilley, 1980). Nevertheless, cruciforms are intrinsically less stable than the unbranched duplex DNA from which they are derived. Supercoiling provides free energy that may be used to stabilize unstable structural polymorphs. In particular, numerous *in vivo* methods have indicated that the superhelical density varies between −0.025 and −0.05 (Zheng *et al*, 1991). These values may be too low for cruciform formation. However, many factors (transcription, growth conditions, stress, topoisomerase I…) may transiently increase the local superhelical density to a critical level sufficient for cruciform extrusion (see review; Pearson *et al*, 1996). Indeed, evidence exists for the formation of cruciforms *in vivo* and implicates these non-B-DNA structures in various cell functions (see review; Pearson *et al*, 1996). Here, we studied *attC* folding as a cruciform structure and the implication of supercoiling in integron recombination. We monitored *in vitro* cruciform formation by detection of changes in nuclease sensitivity caused by the formation of these structures. We also presented an *in vivo* study that shows the ability of *attC* sites to extrude from dsDNA as cruciform structures. This assay consists of the transformation of supercoiled pSW::*attC* plasmids into a recipient strain in which they cannot replicate. The *attC* sites are carried by dsDNA and would mostly recombine after cruciform extrusion. In these conditions, we obtained significant recombination frequencies for all the tested *attC* sites. As expected, we observed a correlation between recombination frequency of the *attC* site as a cruciform structure and the length of the VTS. The propensity to form non-recombinogenic structures also seems to influence the formation of the proper cruciform. Nevertheless, it is quite surprising that sites with large VTS recombine at all. Those sites have a very unfavourable energy of cruciform formation even in highly supercoiled DNA, and spontaneous transitions to a cruciform state would be expected to occur with a much smaller probability than the observed recombination frequencies. This suggests that host proteins can favour cruciform formation in a process that is yet to be identified. It is important to note that in the conjugation assay, these two parameters (VTS size and presence of non-recombinogenic structures) do not seem to influence integron recombination. This is probably because of the fact that the *attC* sites are delivered as ssDNA and that, in these conditions, they have ample time to fold and be captured by the integrase even though they might have large VTS and non-recombinogenic structures.

We also studied the influence of superhelicity on *attC* folding. The dynamic balance between the activities of DNA gyrase and DNA topoisomerase I maintains the level of supercoiling in *Escherichia coli* (Pruss and Drlica, 1986; Lodge *et al*, 1989). Therefore, by using topoisomerase I and gyrase-deficient strains, as well as *in vitro* topoisomerase I-treated plasmids, we showed a significant effect of supercoiling on proper *attC* folding and recombination.

Until now, extrusions from only perfect palindromes have been observed *in vivo* (Pearson *et al*, 1996). Extrusion from an imperfect palindrome has only been observed in AT-rich sequences *in vitro* using two-dimensional gel electrophoresis (Benham *et al*, 2002). It has been shown that imperfections have major effects on the overall energetics of cruciform extrusion and on the course of this transition.

It had been earlier shown that intra-strand base pairing aptitude (i.e. the palindromic structure), not primary sequence, conditions the *attC* site recombination (Bouvier *et al*, 2005, 2009). As modification of the structural properties of the *attC* site directly affects the recombination efficiency, *attC* site folding is likely the limiting step in the recombination process. Consequently, the frequencies obtained in the transformation assay (recombination from ds plasmid) suggest a low probability of *attC* cruciform formation. This corroborates the fact that cruciform structures were never observed for imperfect palindromes that are not particularly AT rich.

We propose a stabilizing effect of the integrase, which could capture *attC* sites on their extrusion and recombine them efficiently. We obtained preliminary results that support this hypothesis (C Loot, V Parissi, D Bikard and D Mazel, in preparation) and we are currently exploring the parameters of this stabilization process. This type of stabilizing effect on cruciform formation has been earlier observed. For instance, in a study of the effect of S1 nuclease on cruciform extrusion, Singleton and Wells (1982) obtained data supporting the fact that the nuclease may exert a transient stabilizing effect on cruciform formation. Similarly, Noirot *et al* (1990) showed the ability of the initiator RepC protein to enhance cruciform extrusion from the pT181 origin of replication.

In this study, we chose to develop an *in vivo* assay (recombination from non-replicative ds plasmid), which, contrary to the other classical assays earlier used, allowed us to observe low probability *attC* site cruciform extrusion (from 10^{-3} down to 10^{-5}), and directly showed the implication of cruciform structures in integron recombination. Moreover, the fact that only the *attC* bs can recombine allowed us to separate hairpin formation occurring on the lagging strand from hairpin formation in cruciform structures (see Figure 3C).

About the constraints of the natural sites

Sites such as $attC_{aadA7}$ seem evolutionarily optimized as they display very favourable folding. On the contrary, the VCR sites display large VTS and non-recombinogenic structures hindering recombination.

In addition to those constraints exerted on the *attC* sites for their efficient folding, we found that their propensity to accumulate mutations could also affect their upper length limit (141 bp for $attC_{qacE}$; Stokes *et al*, 1997). Indeed, during the construction of the VCR_{180} site, we obtained a much higher proportion of mutations than with the smaller sites. Long palindromes pose a threat to genome stability by hindering passage of the replication fork. It is known that cells have evolved a post-replicative mechanism for the elimination and/or repair of large DNA secondary structures using the SbcCD endonuclease (Leach, 1994). Thus, for the larger sites, if they can fold well, they would hinder replication and thus be unstable. Conversely, large sites that fold poorly would have recombination frequencies too low to be selected.

Folded attC sites as sensors of environmental stress

Cruciforms have lower thermodynamic stability than regular duplex DNA. They have been observed only in negatively supercoiled molecules, in which the unfavourable free energy of formation is offset by the topology of the torsionally stressed molecule. This can be a disadvantage, as cruciform structures can be observed only in relatively large supercoiled DNA circles, and are destabilized when a break is introduced at any location. In *E. coli*, superhelicity has been shown to vary considerably during cell growth and to change in different growth conditions (Balke and Gralla, 1987; Jaworski *et al*, 1991). The level of superhelicity can also vary between bacterial strains. Indeed, the average supercoil density of a pBR322 reporter plasmid extracted from mid-log cultures of *Salmonella* is 13% lower ($\sigma = -0.060$) than that from *E. coli* ($\sigma = -0.069$) (Champion and Higgins, 2007).

In addition, biological processes such as transcription may generate domains of supercoiling on circular DNA (Liu and Wang, 1987). On the other hand, changes in supercoiling in response to external and/or internal stimuli could have a significant function in the formation and stability of cruciform *attC* sites. Indeed, analysis of topology of reporter plasmids isolated from SOS+ and SOS− strains revealed higher levels of negative supercoiling in strains with the constitutively expressed SOS network, suggesting a link between the induction of bacterial SOS repair and changes in DNA topology (Majchrzak *et al*, 2006).

Integron recombination could not only be controlled by all processes implicated in the variation of DNA topology, but also we showed that it is controlled by the processes, which generate ssDNA (conjugation, replication…). Interestingly, ssDNA can also be produced in response to external stimuli such as environmental stress. For example, recently, it has been shown that in *Streptococcus pneumoniae*, a Gram+ bacterium, competence and, therefore, single-strand production is induced by an antibiotic stress response (Prudhomme *et al*, 2006). In *Bacillus subtilis*, the competence state has been found to be required for the cell to revert point mutations in auxotrophic alleles when grown on minimal medium (Robleto *et al*, 2007). These are two examples in which ssDNA production is triggered in response to stress conditions.

This novel concept for regulating the recombination of gene cassettes in integrons seems to be of considerable importance, as filamentous phages (Val *et al*, 2005) and insertion sequences (Ton-Hoang *et al*, 2005) also use this mobilization mechanism. It was recently shown that integron-mediated recombination is integrated with the SOS response (Guerin *et al*, 2009), which in turn is activated by the production of the substrate for integron-mediated recombination—ssDNA. The use of this unconventional form of DNA as substrate allows another level of regulation in the integron recombination process. The results presented here confirm the position of integrons as an integrated adaptative system and strengthen the model in which folded ssDNA can 'serve' as sensor of the environmental stress triggering bacterial adaptation.

Materials and methods

Bacterial strains and media

See Supplementary data.

Plasmids

Plasmids used in these studies are described in Supplementary Table S1. Primers were obtained from Sigma-Aldrich (France) and are listed in Supplementary Table S2.

In vivo recombination assays

Suicide-conjugation assay and recombination assay with a replicative ds substrate are described in Supplementary data.

Recombination assay with unidirectional-replicative substrate. In this assay, the two natural $attC_{aadA7}$ and $VCR_{2/1}$ sites were used to analyse the effect of replication on integron recombination. To this end, we constructed two unidirectional replicating pTSC plasmids carrying the *attC* sites in either of the two orientations. A thermosensitive origin was chosen (oripSC101ts). Each pTSC::*attC* plasmid was introduced in a UB5201 strain containing the pBAD::*IntI1* plasmid and the pSU38Δ::*attI1* plasmid. The transformed cells were grown for 6 h at 30°C (to allow pTSC::*attC* replication) in the presence of the respective antibiotics: Cm (pTSC::*attC* marker), Ap (pBAD::IntI1 marker) and Km (pSU38Δ ::*attI1* marker). The integrase was expressed by addition of 0.2% arabinose. Then, cells were plated at 42°C on Cm so that only the cells containing recombined pTSC::attC could grow. A total of 1% glucose was added to the plates to repress the pBAD promoter and prevent residual recombination events. The integration activity was calculated as the ratio of cells expressing the Cm^R marker to the total number of Ap^R Km^R clones.

Recombination assay with a non-replicative substrate. This assay supplies the *attC* site on a ds plasmid that cannot replicate once introduced into the recipient cell by transformation. For these, we transformed a *pir*− strain (UB5201) containing the pBAD::*IntI1* and the pSU38Δ::*attI1* plasmids with 200 ng of the pSW::*attC* plasmids. Competent cells were prepared in the presence of 0.2% arabinose to allow integrase expression. Transformants were selected on Cm^R (the pSW::*attC* marker). As pSW::*attC* cannot replicate in the UB5201 strain, Cm^R clones correspond to *attC* × *attI* recombination events. To establish the recombination activity, we in parallel transformed a *pir*+ strain (UB5201-Pi), which allows the replication of the pSW::*attC* plasmids. Transformants were selected for Cm^R (the pSW::*attC* marker). The recombination activity corresponds to the ratio of Cm^R clones obtained in *pir*− conditions to those obtained in *pir*+ conditions. Note that the efficiency of transformation of each strain was determined beforehand and used to adjust the final ratio and normalize the results. To study the superhelicity effect, we performed the same assay by transforming the JTT1 and SD7 strains with the same quantity (200 ng) of supercoiled and relaxed (topoisomerase I-treated) plasmids.

Topoisomers production

See Supplementary data.

In vitro detection of cruciform

Potential cruciform loops on the pSW::$attC_{aadA7}$, pSW::$attC_{aadA7Mut3}$, pSW::$VCR_{2/1}$ and pSW::VCR_{GAA} plasmids were detected by S1 nuclease sensitivity and digested with NcoI or BglII for the pSW::$attC_{aadA7}$ and pSW::$attC_{aadA7Mut3}$ plasmids and NcoI or ScaI for the pSW::$VCR_{2/1}$ and pSW::VCR_{GAA} plasmids (see Supplementary data).

Mapping of S1-cleavage position

To precisely map the S1-cleavage positions in the $attC_{aadA7}$-folded site, we purified the two fragments arising from molecules cleaved by S1 nuclease and BglII. A C nucleotides tail was added to the 3′ terminus of purified fragments using the Terminal transferase (TdT) (5′RACE Invitrogen kit, version 2.0). TdT is active on both ssDNA (protruding and recessing 3′ends) and dsDNA. The C-tailed products were then used as templates for PCR using the Abridged Anchor primer (containing a stretch of G nucleotides) and either the SW23beg primer for the larger band, or the SW23end primer for the smaller band. Amplified products were then blindly cloned using the TOPO TA cloning kit (Invitrogen). In total, 91 resulting clones were sequenced by using the MFD primer (see Supplementary data).

Analysis of recombination events and point localization

See Supplementary data.

Calculation of the probability to fold a recombinogenic attC site from ssDNA and of the free energy of cruciform formation

See Supplementary data.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We acknowledge Dean Rowe-Magnus for critical reading of the paper and Bénédicte Michel for helpful discussions. This study was carried out with financial assistance from the Institut Pasteur, the Centre National de la Recherche Scientifique (CNRS-URA 2171), the French National Research Agency (ANR-08-MIE-016), the EU (NoE EuroPathoGenomics, LSHB-CT-2005-512061) and the Fondation pour la Recherche Médicale (équipe FRM 2007).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Azaro MA, Landy A (2002) Chapter 7—λ integrase and the λ Int family. In *Mobile DNA II*, Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) pp 118–148. Washington, DC: ASM Press

Bacolla A, Wells RD (2004) Non-B DNA conformations, genomic rearrangements, and human disease. *J Biol Chem* **279:** 47411–47414

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

Benham CJ, Savitt AG, Bauer WR (2002) Extrusion of an imperfect palindrome to a cruciform in superhelical DNA: complete determination of energetics using a statistical mechanical model. *J Mol Biol* **316:** 563–581

Bierne H, Vilette D, Ehrlich SD, Michel B (1997) Isolation of a dnaE mutation which enhances RecA-independent homologous recombination in the Escherichia coli chromosome. *Mol Microbiol* **24:** 1225–1234

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

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 Genet* **5:** e1000632

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

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

Courey AJ (1999) Analysis of altered DNA structures: cruciform DNA. *Methods Mol Biol* **94:** 29–40

d'Alencon E, Petranovic M, Michel B, Noirot P, Aucouturier A, Uzest M, Ehrlich SD (1994) Copy-choice illegitimate DNA recombination revisited. *EMBO J* **13:** 2725–2734

Dai X, Rothman-Denes LB (1999) DNA structure and transcription. *Curr Opin Microbiol* **2:** 126–130

Demarre G, Guerout AM, Matsumoto-Mashimo C, Rowe-Magnus DA, Marlière P, Mazel D (2005) A new family of mobilizable suicide plasmids based on the broad host range R388 plasmid (IncW) or RP4 plasmid (IncPα) conjugative machineries and their cognate *E. coli* host strains. *Res Microbiol* **156:** 245–255

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

Frumerie C, Ducos-Galand M, Gopaul DN, Mazel D (2010) The relaxed requirements of the integron cleavage site allow predictable changes in integron target specificity. *Nucleic Acids Res* **38:** 559–569

Gamper HB, Hearst JE (1982) A topological model for transcription based on unwinding angle analysis of *E. coli* RNA polymerase binary, initiation and ternary complexes. *Cell* **29:** 81–90

Guasch A, Lucas M, Moncalian G, Cabezas M, Perez-Luque R, Gomis-Ruth FX, de la Cruz F, Coll M (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10:** 1002–1010

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

Guynet C, Hickman AB, Barabas O, Dyda F, Chandler M, Ton-Hoang B (2008) *In vitro* reconstitution of a single-stranded transposition mechanism of IS608. *Mol Cell* **29:** 302–312

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

Hall RM, Collis CM (1995) Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. *Mol Microbiol* **15:** 593–600

Hartvig L, Christiansen J (1996) Intrinsic termination of T7 RNA polymerase mediated by either RNA or DNA. *EMBO J* **15:** 4767–4774

Heitman J, Zinder ND, Model P (1989) Repair of the *Escherichia coli* chromosome after *in vivo* scission by the EcoRI endonuclease. *Proc Natl Acad Sci USA* **86:** 2281–2285

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

Khan SA (2005) Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* **53:** 126–136

Langston LD, O'Donnell M (2006) DNA replication: keep moving and don't mind the gap. *Mol Cell* **23:** 155–160

Leach DR (1994) Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* **16:** 893–900

Lilley DM (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. *Proc Natl Acad Sci USA* **77:** 6468–6472

Lilley DM, Gough GW, Hallam LR, Sullivan KM (1985) The physical chemistry of cruciform structures in supercoiled DNA molecules. *Biochimie* **67:** 697–706

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

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

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

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

Mazel D (2006) Integrons: agents of bacterial evolution. *Nat Rev Microbiol* **4:** 608–620

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

Napierala M, Bacolla A, Wells RD (2005) Increased negative superhelical density *in vivo* enhances the genetic instability of triplet repeat sequences. *J Biol Chem* **280:** 37366–37376

Noirot P, Bargonetti J, Novick RP (1990) Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. *Proc Natl Acad Sci USA* **87:** 8560–8564

Pages V, Fuchs RP (2003) Uncoupling of leading- and lagging-strand DNA replication during lesion bypass *in vivo*. *Science* **300:** 1300–1303

Panayotatos N, Wells RD (1981) Cruciform structures in supercoiled DNA. *Nature* **289:** 466–470

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

Phillips GJ (1999) New cloning vectors with temperature-sensitive replication. *Plasmid* **41:** 78–81

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

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

Robleto EA, Yasbin R, Ross C, Pedraza-Reyes M (2007) Stationary phase mutagenesis in *B. subtilis*: a paradigm to study genetic diversity programs in cells under stress. *Crit Rev Biochem Mol Biol* **42:** 327–339

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

Singleton CK, Wells RD (1982) Relationship between superhelical density and cruciform formation in plasmid pVH51. *J Biol Chem* **257:** 6292–6295

Spies M, Dillingham MS, Kowalczykowski SC (2005) Translocation by the RecB motor is an absolute requirement for {chi}-recognition and RecA protein loading by RecBCD enzyme. *J Biol Chem* **280:** 37078–37087

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

Ton-Hoang B, Guynet C, Ronning DR, Cointin-Marty B, Dyda F, Chandler M (2005) Transposition of ISHp608, member of an unusual family of bacterial insertion sequences. *EMBO J* **24:** 3325–3338

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

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

Wang TC (2005) Discontinuous or semi-discontinuous DNA replication in *Escherichia coli*? *Bioessays* **27:** 633–636

Wolfson J, Dressler D (1972) Regions of single-stranded DNA in the growing points of replicating bacteriophage T7 chromosomes. *Proc Natl Acad Sci USA* **69:** 2682–2686

Zheng GX, Kochel T, Hoepfner RW, Timmons SE, Sinden RR (1991) Torsionally tuned cruciform and Z-DNA probes for measuring unrestrained supercoiling at specific sites in DNA of living cells. *J Mol Biol* **221:** 107–122

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31:** 3406–3415

