



HAL
open science

The integron integrase efficiently prevents the melting effect of Escherichia coli single-stranded DNA-binding protein on folded attC sites.

Céline Loot, Vincent Parissi, José Antonio Escudero, Jihane Amarir-Bouhram, David Bikard, Didier Mazel

► To cite this version:

Céline Loot, Vincent Parissi, José Antonio Escudero, Jihane Amarir-Bouhram, David Bikard, et al.. The integron integrase efficiently prevents the melting effect of Escherichia coli single-stranded DNA-binding protein on folded attC sites.. Journal of Bacteriology, 2014, 196 (4), pp.762-771. 10.1128/JB.01109-13 . hal-01101080

HAL Id: hal-01101080

<https://hal.science/hal-01101080>

Submitted on 1 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The integron integrase efficiently prevents the melting effect of *Escherichia coli* SSB**
2 **protein on folded *attC* sites**

3

4 Céline Loot¹, Vincent Parissi², José Antonio Escudero¹, Jihane Amarir-Bouhram¹, David
5 Bikard¹ and Didier Mazel^{1#}

6

7 ¹Institut Pasteur, Unité Plasticité du Génome Bactérien, CNRS UMR3525, 75724 Paris,
8 France.

9 ²Laboratoire MCMP, UMR 5234 CNRS-Université Victor Segalen Bordeaux 2, 33076
10 Bordeaux, France.

11

12 #Correspondent footnote: mazel@pasteur.fr

13 Tel : +33 1 4061 3284

14 Fax : +33 1 4568 8834

15

16 Running title: Regulation of *attC* site folding

17

18 **Abstract**

19 Integrons play a major role in the dissemination of antibiotic resistance genes among bacteria.
20 Rearrangement of gene cassettes occurs by recombination between *attI* and *attC* sites,
21 catalyzed by the integron integrase. Integron recombination uses an unconventional
22 mechanism involving a folded single-stranded *attC* site. This site could be a target for several
23 host factors and more precisely for the proteins able to bind single-stranded DNA. One of
24 these, the *Escherichia coli* Single-Stranded DNA-binding protein (SSB), regulates many
25 DNA processes. We studied the influence of this protein on integron recombination. Our
26 results show the ability of SSB to strongly bind folded *attC* sites and to destabilise them. This
27 effect is observed only in absence of the integrase. Indeed, we provided evidence that the
28 integrase is able to counterbalance the observed effect of SSB on *attC* site folding. We
29 showed that IntI1 possesses an intrinsic property to capture *attC* sites at the moment of their
30 extrusion, stabilizing them and recombining them efficiently. The stability of DNA secondary
31 structures in the chromosome must be restrained to avoid genetic instability (mutations or
32 deletions) and/or toxicity (replication arrest). SSB, which hampers *attC* site folding in the
33 absence of the integrase, likely plays an important role in maintaining the integrity, and thus

34 the recombinogenic functionality of the integron *attC* sites. We also tested the RecA host
35 factor and excluded any role of this protein in integron recombination.

36

37

38 **Introduction**

39 Integrons have been isolated on mobile elements responsible for the capture and
40 dissemination of antibiotic resistance genes among Gram negative bacteria (1). More recently,
41 their importance has been highlighted by the discovery of integrons in the chromosomes
42 (chromosomal integrons) of environmental strains of bacteria (2). This has led to the
43 extension of their role from acquisition of resistance genes to a wider role in the adaptation of
44 bacteria to different environments (3). The discovery that integron recombination was
45 controlled by the SOS response has further connected these elements to bacterial adaptive
46 responses (4). Integrons are composed of three key elements: an integrase gene (*intI*), which
47 encodes a tyrosine (Y) recombinase performing the site-specific recombination (SSR)
48 reaction, a primary recombination site (*attI*), where the incorporation of gene cassettes occurs,
49 and a strong resident promoter (Pc) (5). Gene cassettes are devoid of promoters and mostly
50 correspond to genes associated with an *attC* recombination site. The *attC* X *attI*
51 recombination event catalyzed by IntI leads to cassette integration at the *attI* site, downstream
52 of the Pc promoter allowing gene cassette expression (6). IntI can also catalyze recombination
53 between two consecutive *attC* sites leading to cassette excision.

54 The structures of *attI* and *attC* sites differ significantly from the canonical Y-recombinase
55 core sites (for review see (7)) and integron site-specific recombination is described as an
56 atypical mechanism, in comparison to the classical SSR largely used in bacteria for movement
57 of mobile genetic elements such as prophages or genomic islands (8).

58 The major distinctive characteristic of integron SSR comes from the properties of the *attC*
59 sites. Indeed, it has been previously shown that the *attC* site recombines as a single-stranded
60 (ss) DNA folded structure, in contrast to the *attI* site, which is recombined in a double-
61 stranded (ds) form (9). The length of natural *attC* sites varies from 57 to 141 bp. They include
62 two regions of inverted homology, R''-L'' and L'-R', separated by a central region that is
63 highly variable in length and sequence (Figure 1A) (10). In contrast with the heterogeneity of
64 their sequences, *attC* sites display a strikingly conserved palindromic organization (11) that
65 can form secondary structures through DNA strand self-pairing (Figure 1B). Upon folding, ss
66 *attC* sites show a structure resembling a canonical core site consisting of R and L boxes
67 separated by an unpaired central segment (UCS), two or three extrahelical bases (EHB) and a

68 variable terminal structure (VTS), (12). The VTS varies in length among various *attC* sites,
69 from 3 predicted unpaired nucleotides such as for *attC_{aadA7}*, to a complex branched secondary
70 structure in the larger sites such as VCR (*Vibrio cholerae* repeat) sites (Figure 1B).

71 It has been shown using a DNA binding assay that IntI1 binds strongly and specifically to the
72 bottom strand (bs) of ss *attC* DNA (13). *In vivo*, it has also been shown that only the bs of the
73 *attC* site is used as a substrate during gene cassette integration (9) thereby creating an atypical
74 Holliday junction (aHJ) between the ss *attC* and the ds *attI* site. We recently proved the
75 implication of replication in the resolution of the aHJ and definitively ruled out the
76 involvement of a second strand exchange of any kind in the *attC* x *attI* reaction (14).

77 We have previously shown that cellular bacterial processes delivering ssDNA, such as
78 conjugation and replication, favor the proper folding of *attC* sites (Figure 1C, “ssDNA
79 pathway” (1)), (15), (14). By developing a very sensitive *in vivo* assay, we also provided
80 evidence that *attC* sites can recombine as cruciform structures by extrusion from double-
81 stranded DNA (Figure 1C, “dsDNA pathway” (2)), (14). Moreover, we showed an influence
82 of DNA superhelicity on *attC* site extrusion and recombination *in vitro* and *in vivo* (14). We
83 demonstrated that the proper folding of *attC* sites depends on both the propensity to form non-
84 recombinogenic structures and the length of their variable terminal structures (VTS).

85 Since integron recombination uses *attC* sites in a single-stranded form, the host cell
86 factors which have affinity towards single-stranded DNA can potentially regulate integron
87 recombination by acting on *attC* sites, impairing their folding and/or impeding the integrase
88 activity.

89 In this paper, we examine the effect of several of these specific host factors on *attC* site
90 folding. Among them, we chose to investigate the effect of single-stranded DNA-binding
91 proteins SSB and RecA.

92 *Escherichia coli* Single-Stranded DNA-Binding Protein (SSB) is essential for cell viability
93 (16). During replication, SSB binds tightly and cooperatively to ssDNA to prevent premature
94 annealing. Indeed, the single-stranded DNA has a natural tendency to revert to the double-
95 stranded form, but SSB binds to the single strands, keeping them separate and allowing the
96 DNA replication machinery to proceed (for review see (17)). In addition, SSB is implicated in
97 several other cellular processes, such as homologous recombination, mismatch repair and
98 excision repair (for review see (16)). Moreover, recent studies found that SSB is not static
99 when bound to DNA but could migrate randomly on single-stranded DNA. Indeed, SSB
100 diffusion can melt short DNA hairpins and by consequence remove these secondary structures
101 spontaneously formed by DNA (18). This function is biologically important because RecA,

102 which is necessary for recombinational DNA repair, does not bind well to secondary
103 structures formed in ssDNA. SSB, through its ability to destabilize the intramolecular DNA
104 basepairing within ssDNA, allows RecA to form a contiguous filament on the DNA (19).

105 RecA, a multifunctional DNA-binding protein with ubiquitous distribution in nature,
106 plays a central role in both homologous recombination and post-replicative DNA repair
107 mechanisms (see above). In bacteria, it also plays a role in the SOS response as a co-protease
108 for the autocatalytic cleavage of the LexA repressor leading to de-repression of the SOS
109 regulon genes (20). The RecA protein binds strongly, and in long clusters, to ssDNA to form a
110 nucleoprotein filament.

111 Since integron recombination uses a folded single-stranded *attC* site, we hypothesized
112 that SSB could potentially act on the folded *attC* site and ultimately on integron
113 recombination. SSB could bind parts of the folded *attC* that remain as single-stranded DNA
114 (i.e. UCS and/or VTS, see Figure 1B), thus destabilizing/unpairing recombinogenic *attC*
115 structures and impeding integrase activity. We also studied the role of RecA in integron
116 recombination.

117 By using a gel retardation assay, we showed that SSB can bind folded *attC* sites. Moreover, a
118 replication slippage assay allowed us to conclude that this binding induces the destabilization
119 of *attC* site folding *in vivo*. Interestingly, this effect of SSB is not observed in the presence of
120 the integron integrase. We also provide evidence that, contrary to SSB, IntI1 possesses an
121 intrinsic property to capture *attC* sites at the moment of their extrusion, stabilizing and
122 recombining them efficiently.

123 The stability of secondary structures in the cell must be limited to avoid genetic instability
124 (mutations or deletions) and/or toxicity (replication arrest). SSB is able to restrain the
125 formation of *attC* secondary structures in the absence of the integrase, ensuring the integrity
126 of integron *attC* sites when recombination is not happening. We also demonstrated that, when
127 expressed, the integrase is able to counterbalance this effect of SSB even when SSB is over-
128 expressed.

129 By using a set of similar experiments, we demonstrated that RecA does not regulate integron
130 recombination.

131 We conclude that the interplay between SSB and the *attC* sites represents another example of
132 host regulation of *attC* site folding.

133

134 **Materials and Methods**

135 *Bacterial strains, plasmids, oligonucleotides and media*

136 Bacterial strains, plasmids and oligonucleotide primer sequences are described in Tables S1,
137 S2 and S3. *Escherichia coli* strains were grown in Luria Bertani broth (LB) at 37°C.
138 Antibiotics were used at the following concentrations: ampicillin (Ap), 100 µg/ml,
139 chloramphenicol (Cm), 25 µg/ml, kanamycin (Km), 25 µg/ml, spectinomycin (Sp), 50 µg/ml
140 and tetracyclin (Tc), 15 µg/ml. Diaminopimelic acid (DAP) was supplemented when
141 necessary to a final concentration of 0.3 mM. Glucose and L-arabinose were added at a final
142 concentration of 10 and 2 mg/ml respectively. Chemicals were obtained from Sigma-Aldrich
143 (France).

144

145 ***Integron cassette excision assay***

146 SSB over-expression assay

147 The two plasmids pBAD::*intI1* (p7383, Km^R) and pHYD620 (p7305, Sp^R, which carries the
148 *ssb* gene on a pSC101 vector; approximately 5 copies/cell), were introduced by
149 transformation into a MG1655 Δ *dapA* derivative strain which contains an insertion (in the
150 attB lambda site) of a plasmid carrying a *dapA* gene interrupted by the two specific *attI1* and
151 *attC_{aadA7}* sites (pB117, ωB768, ωC021, see Table S1 and S2), *attC_{aadA7}* and *attC_{ereA2}* sites
152 (pB647, ωB661, ωB726) and *attI1* and *attI1* sites (pB119, ωB662, ωB735). These strains are
153 unable to synthesize DAP (2, 6-diaminopimelic acid), and as a result are not viable without
154 DAP supplemented in the medium. Recombination between *att* sites causes excision of the
155 synthetic cassette, restoring a functional *dapA* gene and allowing the strain to grow on DAP-
156 free medium. As controls, we also transformed this MG1655 Δ *dapA* strain with the previously
157 used pBAD::*intI1* plasmid and with the control vector pCL1920 devoid of the *ssb* gene
158 (p8008, Sp^R, ωC020, ωB725, ωB734). We also transformed MG1655 Δ *dapA* strains by
159 pCL1920 or pHYD620 but without pBAD::*intI1* (ωC018, ωC019, ωB723, ωB724, ωB732 and
160 ωB733).

161 After overnight growth in the presence of appropriate antibiotics (Sp, Cm and Km), DAP and
162 1% glucose, strains were cultivated for 6h in the presence of 0.2% arabinose to allow *intI1*
163 expression and plated on agar containing either LB or LB+DAP. Recombination activity is
164 calculated as the ratio of the number of cells growing in the absence of DAP over the total
165 number of cells.

166

167 RecA over-expression assay

168 In order to test RecA over-expression in the integron cassette excision assay, we processed
169 exactly as in the SSB over-expression assay but transformed pGB2-ara-RecA_{Eco} (pB981, Sp^R,

170 which carries the *recA* gene on a pSC101 vector) in place of the pHYD620 (ωC071, ωC069,
171 ωB720, ωB722, ωB729, ωB731) and the control vector pGB2 (pB495), in place of pCL1920
172 (ωC070, ωC068, ωB719, ωB721, ωB728, ωB730).

173

174 *ssb-200* and *recA269::tn10* assays

175 In order to test the effect of SSB and RecA inactivation in the integron cassette excision
176 assay, we proceeded exactly as in the SSB and RecA over-expression assays (described
177 above). pBAD::*intI1* (p7383, Km^R) was introduced by transformation into MG1655Δ*dapA* ,
178 MG1655Δ*dapA ssb-200 and recA269::tn10* derivative strains which contain an insertion (in
179 the attB lambda site) of a plasmid carrying a *dapA* gene interrupted by the two specific *attI1*
180 and *attC_{aadA7}* sites (pB117, ωB768, ωB769, ωB149, ωB848, ωB849, ωB645, see Table S1 and
181 S2), *attC_{aadA7}* and *attC_{ereA2}* sites (pB647, ωB661, ωB771, ωB660, ωB718, ωB851, ωB717)
182 and *attI1* and *attI1* sites (pB119, ωB662, ωB770, ωB150, ωB727, ωB850, ωB646).

183

184 **Replication slippage assay**

185 To determine reversion frequencies, 100 ml of *E. coli* GJ1885 and GJ1890 strains (ω6737 and
186 ω6738), transformed by the pBR derivative plasmids (p5087, p5088, p6671, p6672, p8558
187 and p8559), were grown up to OD₆₀₀=0.8 in LB medium containing ampicillin from 4 ml of
188 an overnight culture. Reversion frequencies are calculated as the number of Cm^R cells in the
189 population, which is determined by plating the equivalent of 90 ml of cells on medium
190 containing ampicillin and chloramphenicol. Total viable cells were determined by plating on
191 medium containing ampicillin. The frequency represents the mean of 4 independent
192 experiments. The partial or complete deletion of the *attC* site was confirmed for several Cm^R
193 reversion events by sequencing. To study the effect of the integrase on reversion frequency,
194 we performed the same experiment by employing the previously used GJ1885 and GJ1890
195 strains. We transformed them with the pBAD::*intI1Y312F* plasmid (Sp^R, p8741). To
196 determine reversion frequencies, 100 ml of these strains (ω8749, ω8752 and ω8755) were
197 grown up to OD₆₀₀=0.8 in LB medium containing ampicillin and kanamycin with 1% glucose
198 (integrase gene repression) or 0.2% arabinose (integrase gene expression) from 4 ml of an
199 overnight culture.

200

201 **Protein purification**

202 Bacterial strain BL21(DE3)plysS (21) was transformed with pET-derived plasmids
203 expressing the His-IntI integrase (p4634, ωC004). The protein purification was carried out as

204 described by Johansson and collaborators (22) and with the minor modifications described by
205 Demarre and collaborators (23). The *E. coli* Single-Strand Binding protein (SSB) is provided
206 from Sigma (product code **S3917**).

207

208 ***Electrophoretic mobility shift assay (EMSA)***

209 The SSB proteins were incubated with 0.6 pmol of P³²-labeled DNA oligonucleotides
210 corresponding to the bottom strands of the wild type (o32), the synthetic paired (o42) or the
211 synthetic unpaired VCR_{2/1} site derivative (o36, see Table S3). The EMSA was carried out as
212 described by Frumerie and collaborators (24). Quantifications were made using Image Gauge
213 version 4.0.0 (FUJI Photo Film Co. Ltd), manually defining the lanes and using the automatic
214 peak search function. All other functions were at default values.

215

216 ***attC site folding assay***

217

218 The 100 bp ds *attI1* containing fragment was generated by annealing the attI1-F
219 oligonucleotide with the complementary attI1-R oligonucleotide. The 120 bp ds *aadA7*
220 fragment was obtained by PCR amplification using the p4136 vector as template and aadA7-F
221 and aadA7-R primers. The 131 bp ds VCR fragment was obtained by PCR amplification
222 using the p1880 vector and VCR-F and VCR-R primers. The 119 bp ds *aadA7* mutant
223 fragment was obtained by PCR amplification using the p4250 vector as template and aadA7-F
224 and aadA7-R primers.

225 Under standard conditions the melting temperature of a 50 µg/ml solution of each fragment is
226 measured as the OD_{260nm} in a thermostatic spectrophotometer at 37°C. Fragments are pre-
227 chilled to 4°C and pre-heated to 95°C. To analyze the effect of IntI1, proteins are added to the
228 fragment solution 30 min after the beginning of the measurement. The OD_{260nm} measurement
229 of the proteins alone was previously performed and subtracted from the measurement
230 obtained in the mixture in order to determine the accurate values of the fragment solution.

231

232 **Results**

233

234 ***SSB binds folded attC sites***

235 It has been shown that single-stranded DNA-binding protein, SSB, destabilizes secondary
236 structures in ssDNA during major cellular processes such as DNA replication, repair, and
237 homologous recombination (25) (19), (26). We therefore hypothesized that SSB can bind

238 folded *attC* sites and destabilize them. To determine if this was the case, we investigated the
239 binding capacity of SSB on folded *attC* sites using an electrophoretic mobility shift assay
240 (EMSA). In order to mimic the *attC* site's bottom strand (bs), we used VCR_dbs substrate, a
241 single-stranded oligonucleotide of ~90 bases containing a shorter version of the VCR_{2/1}bs site
242 (23). In this substrate, the central part of the site (VTS) was removed and substituted with
243 GAA (VCR_dbs, Figure 2A), a modification previously described and shown to be fully
244 recombinogenic (23).

245 As controls, we constructed two other synthetic *attC* sites, namely the unpaired and paired
246 VCR_dbs sites derivative. To obtain fully paired VCR_dbs site derivative, we introduced specific
247 mutations in the unpaired strand parts of the folded VCR_dbs site in order to introduce
248 complementarity and pair them. Finally, to obtain the unpaired VCR_dbs derivative, we
249 introduced specific mutations in the double-stranded parts of the folded VCR_dbs site in order
250 to unpair them (Figure 2A).

251 The results from EMSA are presented in Figure 2B. In the presence of SSB, distinct shifted
252 bands (I and II), corresponding to SSB-DNA complexes, can be seen for both VCR_dbs and
253 unpaired VCR_dbs sites derivative. Bands I and II probably correspond to different SSB-
254 binding modes and/or different types of intertetramer cooperativity (see Discussion). As
255 expected, for the paired VCR_dbs derivative, we failed to observe distinct and significantly
256 shifted bands. The band intensities on the gels were quantified, confirming the ability of SSB
257 to bind single-stranded parts of *attC* sites very efficiently (Figure 2C). We observed, among
258 free DNA, higher electrophoretic migration bands in the VCR_dbs and paired VCR_dbs
259 derivative lanes. Note that these supplementary bands probably correspond to DNA dimer
260 formation favored by the presence of inverted repeats. The presence of these DNA dimers
261 does not alter the interpretation of the results.

262

263 ***SSB destabilizes folded attC sites***

264 We hypothesized that the consequence of SSB binding on the folded *attC* sites could be a
265 destabilization effect. Indeed, Reddy and collaborators have previously shown the ability of
266 SSB to efficiently and rapidly bind stem loops, melting these secondary structures and
267 facilitating the binding of RecA (26). To study this, we developed a replication slippage
268 assay. Indeed, a direct correlation between the stability of secondary structures and the
269 frequency of replication slippage has already been demonstrated (27). The model proposed is
270 that the formation of a stem-loop structure between a pair of flanking direct sites during
271 replication permits a slippage event. This results in precise or nearly precise loss of the

272 structured region. As SSB is essential in *E. coli*, the slippage assays were performed in the
273 wild type SSB context or in an *ssb-200* mutant. The *ssb-200* mutation corresponds to a G > A
274 substitution at nucleotide 41 resulting in an amino acid substitution at position 4 (G4D). This
275 mutation induces a *tex* (transposon excision) phenotype leading to an increase in the precise
276 excision frequency of a Tn10 derivative (Tn10dKan) (28). Interestingly, Reddy and
277 collaborators showed that *tex* mutant *ssb-200* was not affected in any other functions
278 mediated by SSB, such as replication, recombination, or repair. This suggests that the
279 observed *tex* phenotype of this mutant is probably only due to an incapability of the SSB-200
280 protein to destabilize stem loop structures formed by the inverted repeats of the Tn10dKan.
281 Thus if the previously observed binding of SSB on the *attC* folded sites induces their melting,
282 it is very likely that such an effect would not be observed with SSB-200.

283 We constructed a pBR325 derivative plasmid carrying an *attC_{aadA7}* site inserted in the *cat*
284 resistance gene, in such way that the loss of this *attC* site by “replication slippage” would lead
285 to the reconstitution of a functional *cat* gene (p5087, see Materials and Methods and Figure
286 3A). This plasmid was transformed into an *ssb-200* mutant strain (ω6738, from (28) and in its
287 *ssb-wt* parental strain GJ1885 (ω6737). The Cm^R reversion frequency obtained for the
288 parental strain was 5.5×10^{-11} . Interestingly, the frequency of reversion for the *ssb-200* strain
289 was approximately fourteen fold higher (7.6×10^{-10}) (Figure 3B). Note that several of the Cm^R
290 clones were sequenced in order to confirm deletion of the *attC* sites (see *attC* deletion
291 frequencies, Figure 3B). These results suggest that SSB protein reduces the *attC* site loss
292 mediated by replication slippage events, probably by binding and destabilizing the single-
293 stranded folded *attC* site. To confirm that the obtained frequency of replication slippage is not
294 influenced by the local context (transcription, replication, nucleotide sequences), we
295 performed a set of experimental controls. *attC* sites were cloned into further pBR derivatives
296 in both orientations within the *cat* resistance gene. Moreover, the *cat* gene was cloned in both
297 orientations on the plasmid (see Materials and Methods). In all cases we obtained the same
298 order of Cm^R reversion frequencies for the four tested constructions in the *wt* context and in
299 the *ssb-200* context (data not shown, see Materials and Methods, Table S1 and S2). As an
300 additional control, we constructed a synthetic derivative of *attC_{aadA7}* site presenting a lower
301 free energy of single-stranded *attC* site folding ($\Delta G = -5.1 \text{ kcal mol}^{-1}$) than the natural
302 *attC_{aadA7}* site ($\Delta G = -19.1 \text{ kcal mol}^{-1}$). For this, we introduced specific mutations in the double-
303 stranded parts of the natural *attC_{aadA7}* site in order to unpair them. We called this synthetic
304 *attC* site an unpaired *attC_{aadA7}* site derivative (Figure 4A). As with the *attC_{aadA7}* site, we also
305 observed a decrease in *attC* deletion frequencies in the *ssb-200* context (~13-fold, ω8609)

306 compared to the *wt* context (ω 8605) (Figure 4B). This result shows that the probable
307 destabilization effect of SSB is not only on the *attC* sites, but more generally on single-
308 stranded folded secondary structures. We also cloned the unpaired *attC*_{*aadA7*} site derivative in
309 the opposite orientation within the *cat* resistance gene and we obtained the same order of Cm^R
310 reversion and *attC* deletion frequencies for this construction in both the *wt* context (ω 8604)
311 and in the *ssb-200* context (ω 8608, data not shown).

312 Altogether, these assays confirmed that SSB can destabilize folded *attC* sites in a non-specific
313 manner.

314

315 ***The integrase stabilizes attC site folding***

316 We showed a destabilizing effect of SSB on folded *attC* sites, in the absence of integrase. As
317 the integrase likely stabilizes the folded *attC* structure, we could expect this effect to be
318 counterbalanced in presence of this protein. In order to test this hypothesis, we assessed the
319 effect of IntI1 expression on the slippage frequency in the previously described assay. We
320 introduced in the GJ1885 strain (ω 6791), the p8741 plasmid carrying the *intI1* gene under the
321 control of the inducible P_{bad} promoter (ω 8749). We chose to use an *intI1* gene mutated for the
322 catalytic tyrosine (Y312F) in order to observe the effect of IntI1 binding without being
323 influenced by potential recombination events that could occur between copies of the plasmid.
324 It has been previously shown that when substituting the reactive nucleophilic tyrosine with
325 the inert phenylalanine (IntI1Y312F), the ability of the integrase to attack the DNA was lost,
326 but the *attC* site binding capability was preserved (22). The results showed that in the
327 IntI1Y312F expression context (arabinose induction, +) the Cm^R reversion and *attC* deletion
328 frequencies are 25-fold higher than in the absence of the integrase (glucose repression, -)
329 (Figure 3B and 4C). This suggests that IntI1, contrary to SSB, increases the *attC* site loss
330 mediated by replication slippage events, probably by stabilizing folded *attC* sites. To confirm
331 that this observed stabilizing effect of IntI1 on *attC* folding is specific to *attC* sites, we
332 performed the same experiment by using the previously used unpaired *attC*_{*aadA7*} derivative
333 (Figure 4A) which theoretically lost its IntI1-binding boxes (ω 8752). As expected in this case,
334 we obtained a very low and probably non-significant stabilizing effect of IntI1 (6-fold higher)
335 compared to the wild type *attC*_{*aadA7*} site (25-fold higher, Figure 4C). We demonstrated here
336 that the integron integrase stabilizes the folding of *attC* sites in a specific manner.

337

338 ***In vitro study of IntI1 and SSB on attC site folding***

339 In order to specify the effect of the integrase on *attC* site folding, we performed additional *in*
340 *vitro* experiments. Indeed, it has been shown that IntI1 possesses the intrinsic property to
341 efficiently catalyze strand transfer *in vitro* between *attC* fragments in the absence of other
342 bacterial proteins, suggesting the ability of the integrase to favor the *attC* site folding or to
343 stabilize it (29), (30). To confirm this observation, we set up an *in vitro* folding assay
344 allowing us to easily monitor the single-stranded or double-stranded folding of short DNA
345 fragments containing *attC* sites. The OD variations, measured at 260 nm (OD₂₆₀), directly
346 reflect the DNA hybridization modifications. Indeed, OD₂₆₀ is comparatively lower when the
347 DNA fragment is in the double-stranded (ds) form rather than in the single-stranded (ss) form.
348 The pre-heated *attC_{aadA7}* fragment led to an OD₂₆₀ close to 1.2 corresponding to the fully
349 ssDNA form of the fragment (red curve, Figure 5A) and the pre-chilled *attC_{aadA7}* fragment led
350 to an OD₂₆₀ close to 0.6 corresponding to the fully dsDNA form (blue curve, Figure 5A). By
351 stabilizing the temperature to 37°C, the OD₂₆₀ of the pre-heated fragment decreased to 0.9
352 corresponding to the partially dsDNA form of the fragment. We also obtained the same
353 partially dsDNA structure from the pre-chilled fragment. These results were also reproduced
354 using the VCR_{2/1}-containing fragments (Figure 5B) but not for *attI1* sites (site not forming
355 hairpin structure, Figure 5C). This confirms that *attC* sites share a common partially dsDNA
356 form, in contrast to *attI1* which remains fully double-stranded. Addition of IntI1 to both pre-
357 chilled and pre-heated *attC_{aadA7}* (Figure 5A) and VCR_{2/1} (Figure 5B) fragments led to an
358 OD₂₆₀ close to 1.2, suggesting that the enzyme favors the formation of the fully single-
359 stranded folded *attC* site and stabilizes this structure. In contrast, no such effect was found
360 using the *attI1* site, suggesting that this activity of IntI1 is highly specific to *attC* sites. As a
361 supplementary control, we tested a mutated *attC_{aadA7}* site (Figure 5D) where the extrahelical G
362 base has been deleted in order to prevent the specific binding of IntI1 on the L box (Figure
363 1B) (22). As expected, by using this mutated site, we failed to observe any influence of IntI1
364 on *attC* site folding highlighting the specificity of the IntI1 properties for the *attC* sites.
365 Interestingly, when SSB is used instead of IntI1, perturbation of the fragments' folding was
366 only observed on pre-heated DNA. This suggests that the protein binds the single-stranded
367 fragments and keeps it in a single-stranded form (Figure 6A and 6B) and does not intervene in
368 the *attC* folding property as observed with IntI1. This also highlights the specificity of the
369 effect observed with IntI1 confirming that the folding effect was due to an intrinsic property
370 of the enzyme to favor and stabilize the folding of DNA strands, and not to an unspecific
371 DNA avidity.

372 In conclusion, while both SSB and IntI1 proteins bind the folded *attC* site, the consequences
373 of their binding would be antagonistic, since SSB seems to destabilize the folded *attC* site,
374 while IntI1 assists with *attC* site folding.

375

376 *SSB does not modulate integron recombination*

377 We tested the recombination of *attC* sites in conditions of SSB over-expression and in a *ssb-*
378 *200* mutant context. We used the excision assay described in (31). This assay allows us to
379 directly measure *attC* recombination through restoration of the *dapA* reading frame. We
380 inserted in a MG1655 Δ *dapA* strain a plasmid containing a *dapA* gene interrupted by two
381 specific recombination sites, *attI1* and *attC_{aadA7}* (Figure 7A, ω B768), *attC_{aadA7}* and *attC_{ereA2}*
382 (ω B661), or *attI1* and *attII* (ω B662, see Table S1). Integrase expression causes site-specific
383 recombination and excision of the synthetic cassette, restoring a functional *dapA* gene and
384 allowing these strains to grow on DAP-free medium (Figure 7A). We transformed these
385 strains with the pBAD::*intI1* multi-copy plasmid carrying the *intI1* gene under the control of
386 the P_{bad} promoter (p7383) and a second plasmid, p7305 (pHYD620, from (28) containing the
387 *ssb* gene ensuring the over-expression of the SSB protein. pCL1920, the pHYD620 progenitor
388 plasmid (p8008, devoid of *ssb* gene) was used as a control. Integrase driven recombination
389 frequencies are calculated as the ratio of the number of cells growing in the absence of DAP
390 over the total number of cells. Results are presented in Figure 7B.

391 The frequencies of excision in the absence (vector) and presence (*ssb+*) of SSB over-
392 expression are significantly similar for each tested *att* combinations. These results show that
393 SSB over-expression does not affect integron recombination.

394 We also tested the recombination of these three *att* combinations in the previously described
395 *ssb-200* context (see paragraph 2, ω B769, ω B771, ω B770). For these, we constructed the
396 MG1655 Δ *dapA ssb-200* strain (ω B704, see Table S1). Results are presented in Figure 7B.

397 Once more, excision frequencies in the *WT* and *ssb-200* strains are similar in each tested *att*
398 combinations. These results show that SSB did not affect integron recombination.

399 As a control, we also performed both previously described assays in the absence of integrase
400 expression (- IntI1, Figure 7B). We only observed a few recombination events only for the
401 *attII* x *attI1* reaction probably due to slippage events and/or homologous recombination
402 between both identical 58 bp long *attII* sites.

403

404 *Effect of RecA in integron recombination*

405 We tested the recombination of *att* sites in conditions of RecA over-expression and in a *recA*
406 deleted mutant. Tests were performed using the previously described *in vivo* excision assay
407 (see above). We also tested the recombination of the three previously used *att* combinations
408 (see Materials and Methods, Table S1 and S2). The frequencies of excision in the RecA over-
409 expression and *recA* deleted strains are significantly similar in each tested *att* combination
410 (Figure 8). These results show that RecA does not affect integron recombination.

411 We also performed these assays in the absence of integrase expression (- IntI1, Figure 8). As
412 in the SSB studies, we only observed recombination events for *attI1* x *attI1* combination,
413 probably due to slippage events and/or homologous recombination (only in *recA*+ strain)
414 between both *attI1* sites.

415

416

417 **DISCUSSION**

418

419 ***Biological functions of secondary structures***

420 It is now obvious that DNA is not always present in its canonical double-stranded form but
421 can also form secondary structures through intra-strand base pairing. These structures are
422 strongly favored in single-stranded DNA and can also be extruded from dsDNA at a lower
423 rate. These structures have been found to be involved in biological functions (15). Indeed,
424 hairpins play an essential role in initiation of replication and transcription, and are present in a
425 majority of origins of transfer (*oriT*) ensuring conjugation (15). Moreover, to date, there are
426 three examples of recombination systems using secondary structures as substrates; *attC*
427 integron recombination (9), CTX-phage integration (32) and IS608 transposition (33). In all
428 of the examples described above, secondary structures could represent a supplementary way
429 to expand information storage in DNA in addition to the primary base sequence.

430 It is also known that too long and stable palindromes pose a threat to genome stability.
431 Indeed, they cannot be maintained *in vivo* (for review, see (15)), either because they are
432 inviable; i.e., intrinsically toxic to the cell or because they are genetically unstable; i.e.,
433 partially mutated or deleted (34). It is assumed that *inviability* is caused by an arrest of the
434 replication fork, as it is unable to process these secondary structures, and *instability* is caused
435 by the presence of proteins such as SbcCD destroying these structures. This leads to
436 constraints on the size and perfection of the inverted repeats that can be maintained *in vivo*.

437

438 ***Single-strand DNA binding proteins***

439 Hairpin formation in the cell is most likely to occur in the presence of ssDNA in the cell. The
440 presence of ssDNA is not merely an intermediate state between its functional double-stranded
441 forms. Indeed, ssDNA can be found in the cell during many of its physiological processes,
442 such as lagging-strand synthesis during replication and DNA repair, but ssDNA is also tightly
443 connected to horizontal gene exchange in bacteria, *i.e.* in bacterial conjugation, natural
444 transformation or viral infections. When produced, cellular ssDNA is not left naked; several
445 proteins, such as RecA and SSB, bind it without sequence specificity. The best characterized
446 SSB is the one from the bacterium *E. coli*. SSB prevents premature annealing, stabilizes and
447 protects the single-stranded DNA, and also removes the secondary structures (16), (17). It acts
448 as a homotetramer that can migrate via random walk along ssDNA providing a mechanism by
449 which it can be repositioned along ssDNA, while remaining tightly bound (18). SSB can form
450 different complexes and bind ssDNA with several binding modes (SSB)_n depending on the
451 number of nucleotides bound per tetramer (n=35, 56 or 65). ssDNA interacts with only two
452 SSB subunits in the (SSB)₃₅ complex and with all four subunits in the (SSB)₆₅ complex,
453 resulting in different extents of ssDNA compaction. The binding mode defines several
454 intertetramer cooperativities: in the (SSB)₃₅-binding mode, cooperativity is “unlimited” and
455 tetramers can form long protein clusters, while in the (SSB)₆₅-binding mode, only dimers of
456 tetramers (octamers) can be formed (“limited” cooperativity) (17). Moreover, SSB directs
457 RecA binding to ssDNA and ensures the directionality of RecA polymerization across stem
458 loop structures in ssDNA (19), (26). Indeed, the RecA protein also binds ssDNA, forming a
459 nucleoproteic filament. Recent single-molecule studies have shown how SSB can
460 spontaneously migrate along ssDNA, melting unstable hairpins while stimulating RecA
461 filament elongation (18).

462 In this paper, we were interested in examining the role of these ssDNA-binding proteins, SSB
463 and RecA, in integron recombination, and more specifically the role of SSB in *attC* site
464 folding. Whereas SSB and RecA are not involved in integron recombination, we demonstrate
465 here the involvement of SSB in the modulation of *attC* site folding in the absence of
466 recombination.

467 ***The RecA protein does not regulate integron recombination***

468 In order to test the effect of RecA on integron recombination, *in vivo* tests were performed by
469 using the integron cassette excision assay where the recombination sites are carried by the
470 chromosome. Note that we also performed supplementary tests by using our suicide

471 conjugation and replicative assays (data not shown). While the conjugation assay permits the
472 delivery of *attC* sites as ssDNA, the test performed under replicative conditions supplies *attC*
473 sites that have to be extruded as cruciforms from double strand molecules. Our results show
474 that, whatever the delivery mode of *attC* sites, RecA does not influence integron
475 recombination. Interesting, even RecA over-expression did not affect integron recombination.
476 It's relevant since we know that expression of integron integrases depends on the induction of
477 the SOS response which in turn produces a large excess of RecA. Note that this result does
478 not exclude any regulatory role of RecA on *attC* site folding outside of recombination.
479 Nevertheless, it seems unlikely, since previous observations showed that the RecA protein
480 does not, by itself, bind well to ssDNA secondary structures and needs SSB to remove such
481 structures (19).

482 ***The SSB protein modulates attC site folding***

483 It was thought that SSB could have a role in integron recombination as it plays extensive
484 cellular roles in DNA replication, DNA repair, and DNA homologous recombination, where it
485 destabilizes secondary structures in ssDNA (25), (19), (26). Since SSB is essential for cell
486 viability, new experimental setups were developed to study its role in integron recombination
487 and *attC* folding. Especially, we developed a replication slippage assay using a strain with a
488 mutated form of SSB, SSB-200, which does not hamper cell viability but confers a transposon
489 excision phenotype (*tex*) (28). This *tex* phenotype is probably directly linked to SSB-200 not
490 destabilizing the stem-loop structures formed by the pair of inverted repeats of the transposon.
491 In the *ssb-200* mutated strain the frequency of reconstitution of a chloramphenicol resistance
492 gene induced by an excision of the folded *attC* site by replication slippage is increased only in
493 the absence of the integrase. This *in vivo* test confirmed that SSB can modify the *attC*-
494 mediated replication slippage, probably by destabilizing the folded *attC* sites. A DNA binding
495 assay allowed us to confirm the ability of SSB to strongly bind these sites. Note that, even if
496 the unpaired single-stranded parts contained in the folded *attC* site (and also the unpaired
497 derivative site) are smaller than the SSB tetramer binding site (n=35, see discussion above),
498 SSB is able to bind to these sites. Indeed, the folded *attC* sites are probably not fixed
499 structures and could “breathe” inducing a transient more unpaired state allowing SSB binding.
500

501 In conclusion, we demonstrated here that SSB, by “flattening” the folded *attC* sites, permits
502 the maintenance of their integrity in the cell avoiding their deletion or mutation and also
503 ensuring cell viability.

504

505 *Stabilization constraints and recombination of hairpins*

506 Although ssDNA is present in many contexts within the cell, hairpin formation is strongly
507 constrained by SSB binding. Proteins that function through hairpin binding are thus in
508 competition with SSB for substrate availability. It could be the case for the integrase since,
509 during recombination, folded *attC* sites need to be stable enough to resist SSB melting and
510 coating. Therefore, we proposed that the integrase could capture *attC* sites at the moment of
511 their extrusion, efficiently stabilizing and recombining them. Here, we confirmed this
512 hypothesis and demonstrated that, when expressed, the integrase is able to counterbalance the
513 effect of SSB even when SSB is over expressed.

514 Finally, integron recombination efficiency is directed by a precise regulation of the stability of
515 the folded *attC* sites. Bacteria have to find a subtle balance between the benefit provided by
516 biological functions of secondary structures in DNA, and the avoidance of such structures that
517 could be detrimental. Here, this balance seems to have been reached through the production of
518 SSB and integrase, even though other host factors may have a role in *attC* site folding. Both
519 proteins are capable of recognizing and acting upon these secondary structures, mediating an
520 antagonistic effect.

521 We previously demonstrated the influence of cellular processes such as replication and/or
522 horizontal gene transfer events (such as conjugation) on *attC* site folding and also the
523 implication of superhelicity in cruciform extrusion. The multiple means of regulating *attC* site
524 folding by the host show once again the network of cell processes that regulate integron
525 recombination.

526

527 **ACKNOWLEDGMENTS**

528

529 This work was supported by the Institut Pasteur, the Centre National de la Recherche
530 Scientifique (CNRS-UMR3525), the European Union Seventh Framework Programme (FP7-
531 HEALTH-2011-single-stage) and the “Evolution and Transfer of Antibiotic Resistance”
532 (EvoTAR).

533 The authors acknowledge Jason Bland, Dr Zeynep Baharoglu, Dr Alfonso Soler and
534 Aleksandra Nivina for critical reading of the manuscript, and Dr Clara Frumerie for
535 performing the EMSA quantification, and Dr Bénédicte Michel, Dr Ivan Matic and Dr Robert
536 Lloyd for providing bacterial strains. This work is dedicated to the memory of Guy Duval-
537 Valentin.

538

539

540 REFERENCES

541

- 542 1. **Hall RM, Collis CM.** 1998. Antibiotic resistance in gram-negative bacteria: the role
543 of gene cassettes and integrons. *Drug Resistance Updates* **1**:109-119.
- 544 2. **Rowe-Magnus DA, Guerout A-M, Ploncard P, Dychinco B, Davies J, Mazel D.**
545 2001. The evolutionary history of chromosomal super-integrons provides an ancestry
546 for multi-resistant integrons. *P Natl Acad Sci USA* **98**:652-657.
- 547 3. **Rowe-Magnus DA, Mazel D.** 2002. The Role of Integrons in Antibiotic Resistance
548 Gene Capture. *International Journal of Medical Microbiology* **292**:115-125.
- 549 4. **Guerin E, Cambray G, Sanchez-Alberola N, Campoy S, Erill I, Da Re S,**
550 **Gonzalez-Zorn B, Barbe J, Ploy MC, Mazel D.** 2009. The SOS response controls
551 integron recombination. *Science* **324**:1034.
- 552 5. **Mazel D.** 2006. Integrons: agents of bacterial evolution. *Nat Rev Microbiol* **4**:608-
553 620.
- 554 6. **Jove T, Da Re S, Denis F, Mazel D, Ploy MC.** 2010. Inverse correlation between
555 promoter strength and excision activity in class 1 integrons. *PLoS Genet* **6**:e1000793.
- 556 7. **Grindley ND, Whiteson KL, Rice PA.** 2006. Mechanisms of site-specific
557 recombination. *Annu Rev Biochem* **75**:567-605.
- 558 8. **Hallet B, Sherratt DJ.** 1997. Transposition and site-specific recombination: adapting
559 DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS*
560 *Microbiol Rev* **21**:157-178.
- 561 9. **Bouvier M, Demarre G, Mazel D.** 2005. Integron cassette insertion: a recombination
562 process involving a folded single strand substrate. *Embo J* **24**:4356-4367.
- 563 10. **Stokes HW, O'Gorman DB, Recchia GD, Parsekhian M, Hall RM.** 1997. Structure
564 and function of 59-base element recombination sites associated with mobile gene
565 cassettes. *Molecular Microbiology* **26**:731-745.
- 566 11. **Hall RM, Brookes DE, Stokes HW.** 1991. Site-specific insertion of genes into
567 integrons: role of the 59-base element and determination of the recombination cross-
568 over point. *Mol Microbiol* **5**:1941-1959.
- 569 12. **Bouvier M, Ducos-Galand M, Loot C, Bikard D, Mazel D.** 2009. Structural features
570 of single-stranded integron cassette attC sites and their role in strand selection. *PLoS*
571 *Genet* **5**:e1000632.
- 572 13. **Francia MV, Zabala JC, de la Cruz F, Garcia-Lobo JM.** 1999. The Int11 integron
573 integrase preferentially binds single-stranded DNA of the attC site. *Journal of*
574 *Bacteriology* **181**:6844-6849.
- 575 14. **Loot C, Bikard D, Rachlin A, Mazel D.** 2010. Cellular pathways controlling
576 integron cassette site folding. *EMBO J* **29**:2623-2634.

- 577 15. **Bikard D, Loot C, Baharoglu Z, Mazel D.** 2010. Folded DNA in action: hairpin
578 formation and biological functions in prokaryotes. *Microbiol Mol Biol Rev* **74**:570-
579 588.
- 580 16. **Meyer RR, Laine PS.** 1990. The single-stranded DNA-binding protein of *Escherichia*
581 *coli*. *Microbiol Rev* **54**:342-380.
- 582 17. **Lohman TM, Ferrari ME.** 1994. *Escherichia coli* single-stranded DNA-binding
583 protein: multiple DNA-binding modes and cooperativities. *Annu Rev Biochem*
584 **63**:527-570.
- 585 18. **Roy R, Kozlov AG, Lohman TM, Ha T.** 2009. SSB protein diffusion on single-
586 stranded DNA stimulates RecA filament formation. *Nature* **461**:1092-1097.
- 587 19. **Kowalczykowski SC, Krupp RA.** 1987. Effects of *Escherichia coli* SSB protein on
588 the single-stranded DNA-dependent ATPase activity of *Escherichia coli* RecA protein.
589 Evidence that SSB protein facilitates the binding of RecA protein to regions of
590 secondary structure within single-stranded DNA. *J Mol Biol* **193**:97-113.
- 591 20. **Walker GC.** 1996. The SOS response of *Escherichia coli*. *Escherichia coli and*
592 *Salmonella*. , p. 1400-1416, Neidhardt, FC Washington DC American Society of
593 Microbiology, vol. 1.
- 594 21. **Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW.** 1990. Use of T7 RNA
595 polymerase to direct expression of cloned genes. *Methods Enzymol* **185**:60-89.
- 596 22. **Johansson C, Kamali-Moghaddam M, Sundstrom L.** 2004. Integron integrase
597 binds to bulged hairpin DNA. *Nucleic Acids Res* **32**:4033-4043.
- 598 23. **Demarre G, Frumerie C, Gopaul DN, Mazel D.** 2007. Identification of key
599 structural determinants of the IntI1 integron integrase that influence attC x attI1
600 recombination efficiency. *Nucleic Acids Res* **35**:6475-6489.
- 601 24. **Frumerie C, Ducos-Galand M, Gopaul DN, Mazel D.** 2010. The relaxed
602 requirements of the integron cleavage site allow predictable changes in integron target
603 specificity. *Nucleic Acids Res* **38**:559-569.
- 604 25. **Muniyappa K, Shaner SL, Tsang SS, Radding CM.** 1984. Mechanism of the
605 concerted action of recA protein and helix-destabilizing proteins in homologous
606 recombination. *Proc Natl Acad Sci U S A* **81**:2757-2761.
- 607 26. **Reddy MS, Vaze MB, Madhusudan K, Muniyappa K.** 2000. Binding of SSB and
608 RecA protein to DNA-containing stem loop structures: SSB ensures the polarity of
609 RecA polymerization on single-stranded DNA. *Biochemistry* **39**:14250-14262.
- 610 27. **Sinden RR, Zheng GX, Brankamp RG, Allen KN.** 1991. On the deletion of
611 inverted repeated DNA in *Escherichia coli*: effects of length, thermal stability, and
612 cruciform formation in vivo. *Genetics* **129**:991-1005.
- 613 28. **Reddy M, Gowrishankar J.** 1997. Identification and characterization of *ssb* and *uup*
614 mutants with increased frequency of precise excision of transposon Tn10 derivatives:
615 nucleotide sequence of *uup* in *Escherichia coli*. *J Bacteriol* **179**:2892-2899.
- 616 29. **Dubois V, Debreyer C, Litvak S, Quentin C, Parissi V.** 2007. A new in vitro strand
617 transfer assay for monitoring bacterial class 1 integron recombinase IntI1 activity.
618 *PLoS One* **2**:e1315.
- 619 30. **Dubois V, Debreyer C, Quentin C, Parissi V.** 2009. In vitro recombination
620 catalyzed by bacterial class 1 integron integrase IntI1 involves cooperative binding
621 and specific oligomeric intermediates. *PLoS One* **4**:e5228.
- 622 31. **Baharoglu Z, Bikard D, Mazel D.** 2010. Conjugative DNA transfer induces the
623 bacterial SOS response and promotes antibiotic resistance development through
624 integron activation. *PLoS Genet* **6**:e1001165.

- 625 32. **Val ME, Bouvier M, Campos J, Sherratt D, Cornet F, Mazel D, Barre FX.** 2005.
626 The single-stranded genome of phage CTX is the form used for integration into the
627 genome of *Vibrio cholerae*. *Mol Cell* **19**:559-566.
- 628 33. **Guynet C, Hickman AB, Barabas O, Dyda F, Chandler M, Ton-Hoang B.** 2008.
629 In vitro reconstitution of a single-stranded transposition mechanism of IS608. *Mol*
630 *Cell* **29**:302-312.
- 631 34. **Collins J, Volckaert G, Nevers P.** 1982. Precise and nearly-precise excision of the
632 symmetrical inverted repeats of Tn5; common features of recA-independent deletion
633 events in *Escherichia coli*. *Gene* **19**:139-146.

634

635

636 **Supporting information**

637 Table S1: Bacterial strains used in this study

638 Table S2: Plasmids used and constructed in this study

639 Table S3: Oligonucleotides

640

641 **FIGURE LEGENDS**

642 **Figure 1: *attC* recombination sites and a model for *attC* folding.**

643 **A)** Schematic representation of a double-stranded (ds) *attC* site. Inverted repeats R'', L'', L'
644 and R' are indicated by grey boxes. The dotted line represents the variable central part.
645 Conserved nucleotides are indicated. Asterisks (*) show the conserved G nucleotide which
646 corresponds to the extrahelical base (EHB) in the folded *attC* site bottom strand. Black arrow
647 shows the cleavage point.

648 **B)** Secondary structures of VCR_{2/1} and *attC_{aadA7}* sites bottom strands (bs). Structures were
649 determined by the UNAFOLD online interface at the Pasteur Institute. The four structural
650 features of *attC* sites, namely, the UCS (Unpaired Central Segment), the EHB, the stem and
651 the VTS (Variable Terminal Structure) are indicated. Black arrows show the cleavage points
652 and asterisks, the extrahelical G bases. Primary sequences of the *attC* sites are shown (except
653 for the VTS of the VCR_{2/1} site).

654 **C)** Representation of both "ssDNA" and "dsDNA" pathways for *attC* site folding. In the
655 "ssDNA pathway" (1, red), horizontal gene transfer (conjugation, transformation and phage
656 infection) and replication lead to the production of ssDNA favoring *attC* hairpin formation. In
657 the "dsDNA pathway" (2, green), an increase of supercoiling ensures cruciform extrusion.
658 Origins of replication are shown as blue ovals and replication complexes as yellow ovals.

659

660 **Figure 2: Comparison of SSB binding to several VCR_{2/1} site derivatives**

661 **A)** Nucleotide sequences and secondary structures of the bottom strands (bs) of VCR_{2/1} site
662 derivatives. Structures were determined by the UNAFOLD online interface at the Pasteur
663 Institute. VCR_{dbs} is a VCR derivative used in EMSA experiments in which the central and
664 imperfect part of the palindrome (VTS) was replaced by GAA. The black arrow shows the
665 cleavage point in the R box and the asterisk (*), the extrahelical G nucleotide. We introduced
666 mutations (bold letters) in this VCR_{dbs} site in order to generate, after folding, paired and
667 unpaired (pd and unpd) VCR_{dbs} site derivatives. The ΔG (free energy) of VCR_{dbs}, pd
668 VCR_{dbs} and unpd VCR_{dbs} folding are -21.7, -37.2 and -5.3 kcal mol⁻¹ respectively.

669 **B)** EMSA analysis of the binding of SSB to VCR_{2/1} site derivatives. Equal quantities of radio-
670 labeled VCR site derivative DNA fragments (0.6pmol) were incubated with increasing
671 amounts (0, 13, 26, 52 and 104 nM) of SSB. Complexes I and II probably represent several
672 degrees of SSB multimerization.

673 **C)** Densitometry analysis of EMSA. Binding of SSB was measured by monitoring the
674 fraction ($\times 100$) of the labeled DNA substrate bound by SSB in complexes I and II/labeled
675 (bound + free) DNA, detected by automatic peak search of Image Gauge.

676

677 **Figure 3: *In vivo* effect of SSB and IntI1 on *attC* site folding**

678 **A)** Experimental setup of the replication slippage assay. The model proposed is that folding of
679 the *attC* site between the pair of flanking direct repeats (EcoRI restriction sites, black
680 rectangles) permits a replication slippage event. This results in precise or nearly precise *attC*
681 deletion thus reconstituting a functional *cat* resistance gene (Cm^R reversion frequencies). The
682 precise *attC* deletion is measured by sequencing (*attC* deletion frequencies).

683 **B)** SSB and IntI1 effect on Cm^R reversion and *attC* deletion frequencies (see Materials and
684 Methods).

685 The replication slippage assay was used to study the IntI1Y312F and SSB effects on *attC_{aadA7}*
686 site folding. The Cm^R reversion frequencies and *attC* deletion frequencies are represented in
687 the absence (-) and in the presence of IntI1 (+) in the GJ1885 wild type strain. The Cm^R
688 reversion frequencies and *attC* deletion frequencies are represented in the GJ1885 wild type
689 and GJ1885 *ssb-200* (GJ1890) strains. The results represent a mean of four independent
690 experiments (see Materials and Methods). Error bars show the standard deviations.

691

692 **Figure 4: Specificity of the *in vivo* effect of SSB and IntI1 on *attC* site folding**

693

694 **A)** Secondary structures of the *attC_{aadA7}* site bottom strand (bs) and the unpaired *attC_{aadA7}*bs
695 site derivative. Structures were determined by the UNAFOLD online interface at the Pasteur
696 Institute. The black arrow shows the cleavage point. Primary sequences of the *attC* sites are
697 shown. The ΔG (free energy) of *attC_{aadA7}*bs and unpaired *attC_{aadA7}*bs folding are respectively -
698 19.1, -5.1 kcal mol⁻¹.

699 **B)** The replication slippage assay was used to study the effect of SSB on folding of the
700 *attC_{aadA7}* site (black bars) and unpaired *attC_{aadA7}* site derivative (grey bars). The relative
701 deletion rates obtained in *ssb-200* versus GJ1885 strains (*ssb-200/ssb-WT*) are indicated. The
702 results represent a mean of four independent experiments.

703 **C)** The replication slippage assay was used to study the effect of IntI1Y312F on folding of the
704 *attC_{aadA7}* site (black bars) and unpaired *attC_{aadA7}* site derivative (grey bars). The relative
705 deletion rates obtained in the presence of integrase versus in its absence (intI1+/intI1-) are
706 indicated. The results represent a mean of four independent experiments.

707

708 **Figure 5: *In vitro* effect of IntI1 on *attC_{aadA7}* (A), VCR_{2/1} (B), *attI1* (C) and *attC_{aadA7}*
709 *mutated* (D) site folding.** The OD_{260nm} of each fragment (50 μ g/ml) was measured at 37°C
710 either after pre-chilling (pre-chilled *attC_{aadA7}*, VCR_{2/1}, *attI1*, *attC_{aadA7m}*), or after pre-heating
711 for three minutes at 95°C (pre-heated *attC_{aadA7}*, VCR_{2/1}, *attI1* and *attC_{aadA7m}*) in the absence of
712 IntI1, as well with the addition of IntI1 (2 pmoles) 30 min after the beginning of the
713 measurement with pre-chilling (pre-chilled *attC_{aadA7}* + IntI1, VCR_{2/1} + IntI1, *attI1* + IntI1,
714 *attC_{aadA7m}* + IntI1) or after 95°C pre-heating (pre-heated *attC_{aadA7}* + IntI1, VCR_{2/1} + IntI1, *attI1*
715 + IntI1, *attC_{aadA7m}* + IntI1). The OD_{260nm} measurement of the proteins alone was previously
716 performed and subtracted from the measurements obtained in the mixture in order to
717 determine the accurate values of the ODN solution. Results represent a mean of three
718 independent experiments. Error bars show the standard deviations.

719

720 **Figure 6: *In vitro* effect of SSB on *attC_{aadA7}* (A) and VCR_{2/1} (B) site folding.** The OD_{260nm}
721 of each fragment (50 μ g/ml) was measured at 37°C either after pre-chilling (pre-chilled
722 *attC_{aadA7}*, VCR_{2/1}), or after pre-heating for three minutes at 95°C (pre-heated *attC_{aadA7}*,
723 VCR_{2/1}) in the absence of SSB, as well as after addition of SSB (2 pmoles) 30 min after the
724 beginning of the measurement with pre-chilling (pre-chilled *attC_{aadA7}* + SSB, VCR_{2/1} + SSB)
725 or after 95°C pre-heating (pre-heated *attC_{aadA7}* + SSB, VCR_{2/1} + SSB). The OD_{260nm}
726 measurement of the proteins alone was previously performed and subtracted from the
727 measurements obtained in the mixture in order to determine the accurate values of the

728 fragment solution. Results represent a mean of three independent experiments. Error bars
729 show the standard deviations.

730

731

732 **Figure 7: *In vivo* effect of SSB on cassette excision**

733 **A)** Experimental setup of cassette excision assay. The *dapA* gene is interrupted by a synthetic
734 cassette (black line) encountered by the *attI1* (black triangle) and *attC_{aadA7}* (white triangle)
735 sites. The recombination leads to the excision of the integron cassette and restores a
736 functional *dapA* gene.

737 **B)** Excision frequencies of the *attI1* x *attI1*, *attC_{aadA7}* x *attC_{ereA2}* and *attI1* x *attC_{aadA7}* cassettes
738 in WT or *ssb-200* strain (top panel) and in the presence of the pCL1920 control vector (WT)
739 or the pHYD620 plasmid (*ssb+*) which over-expresses the SSB protein (bottom panel). Black
740 bars represent the results obtained in presence of IntI1 and grey bars, in absence.

741 “nd” means “non-detected” recombination events. Results represent a mean of three
742 independent experiments. Error bars show the standard deviations.

743

744 **Figure 8: *In vivo* effect of RecA on cassette excision**

745 **A)** Experimental setup of cassette excision assay. The *dapA* gene is interrupted by a synthetic
746 cassette (black line) encountered by the *attI1* (black triangle) and *attC_{aadA7}* (white triangle)
747 sites. The recombination leads to the excision of the integron cassette and restores a
748 functional *dapA* gene.

749 **B)** Excision frequencies of the *attI1* x *attI1*, *attC_{aadA7}* x *attC_{ereA2}* and *attI1* x *attC_{aadA7}* cassettes
750 in WT or *recA-* strain (top panel) and in the presence of the pGB2 control vector (WT) or the
751 pGB2-*recA* plasmid (*recA+*) which over-expresses the RecA protein (bottom panel). Black
752 bars represent the results obtained in presence of IntI1 and grey bars, in absence. “nd” means

753 “non-detected” recombination events. Results represent a mean of three independent
754 experiments. Error bars show the standard deviations.

755

756

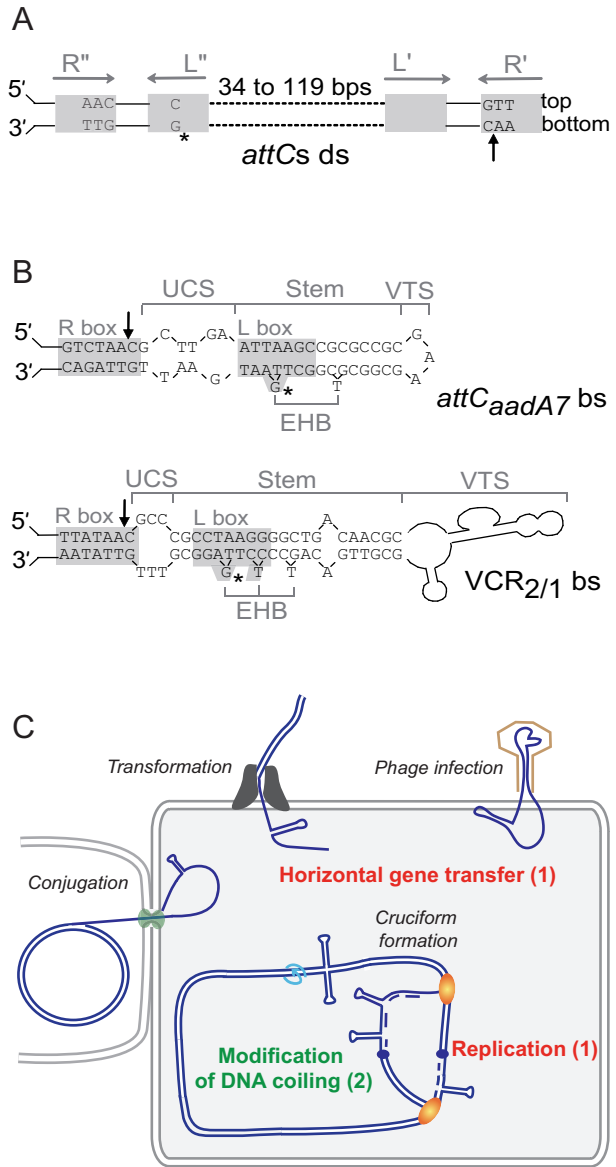


Figure 1

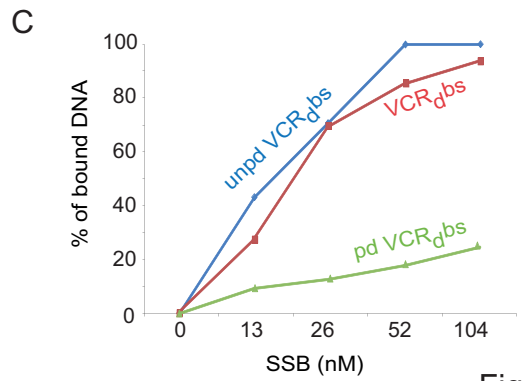
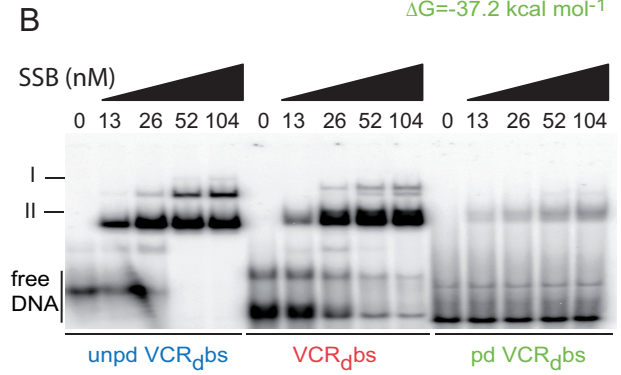
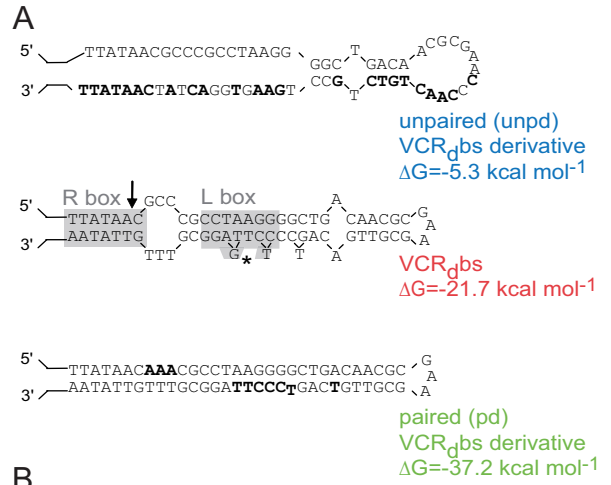
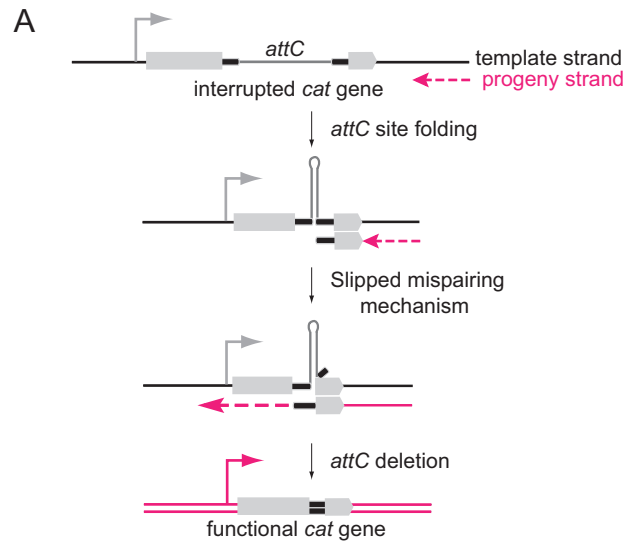


Figure 2



B

		<i>int11</i> expression	<i>Cm^R</i> reversion frequencies	<i>attC</i> deletion frequencies
Strains	GJ1885	-	$5.5 \times 10^{-11} \pm 2.3 \times 10^{-11}$	5.5×10^{-11}
		+	$1.4 \times 10^{-9} \pm 5.7 \times 10^{-10}$	1.4×10^{-9}
ssb-200		-	$7.6 \times 10^{-10} \pm 1.5 \times 10^{-10}$	7.6×10^{-10}
		+	$1.3 \times 10^{-9} \pm 2.0 \times 10^{-10}$	8.7×10^{-10}

Figure 3

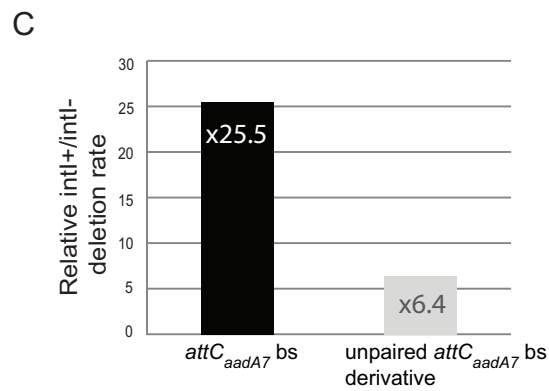
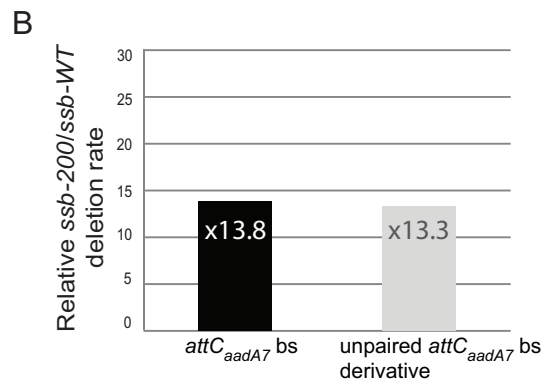
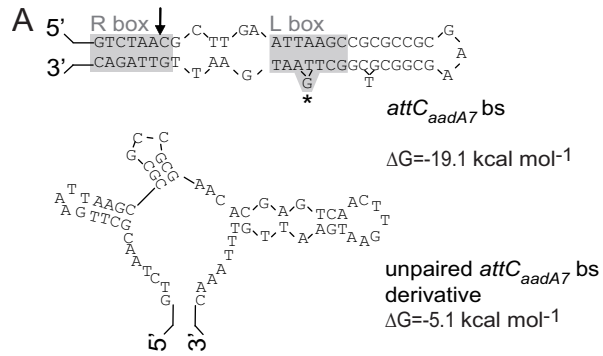


Figure 4

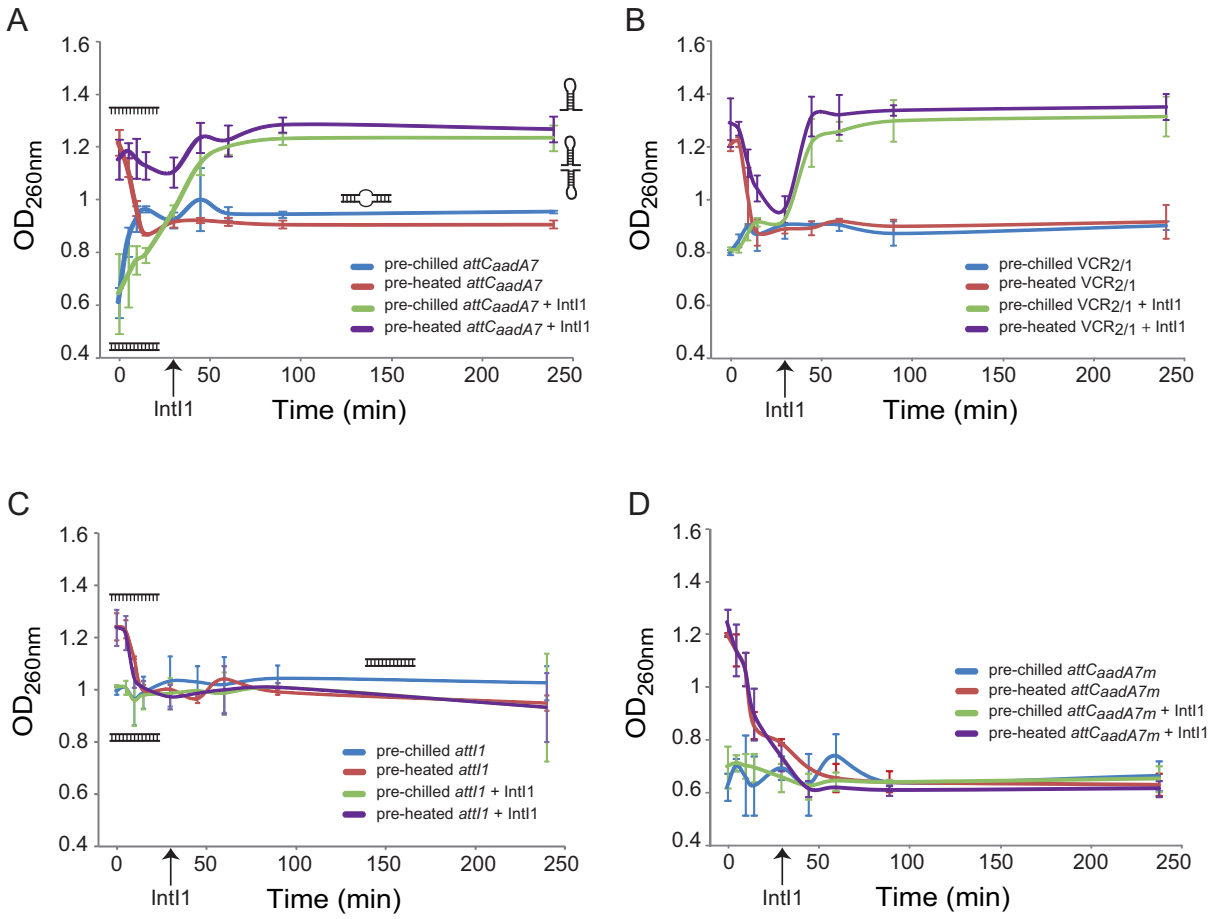


Figure 5

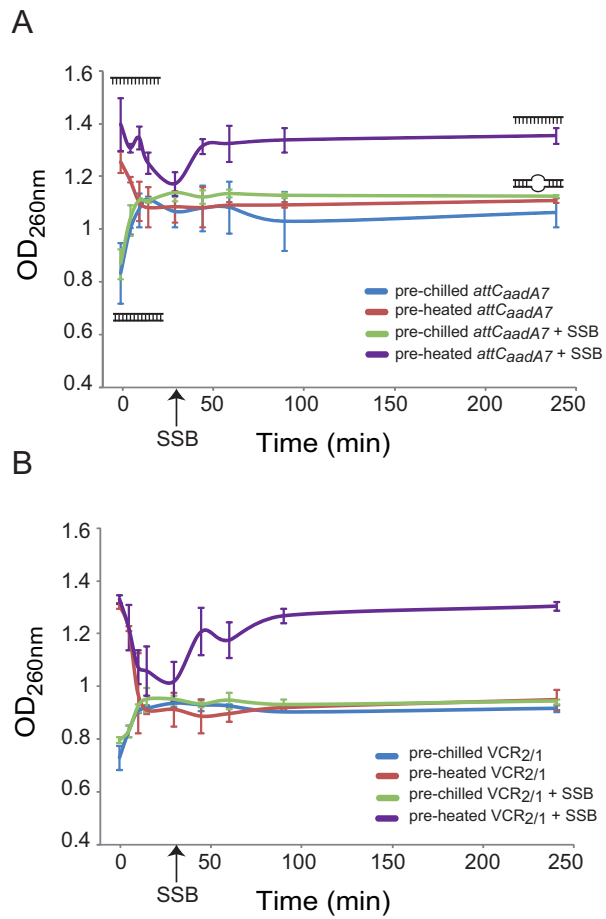


Figure 6

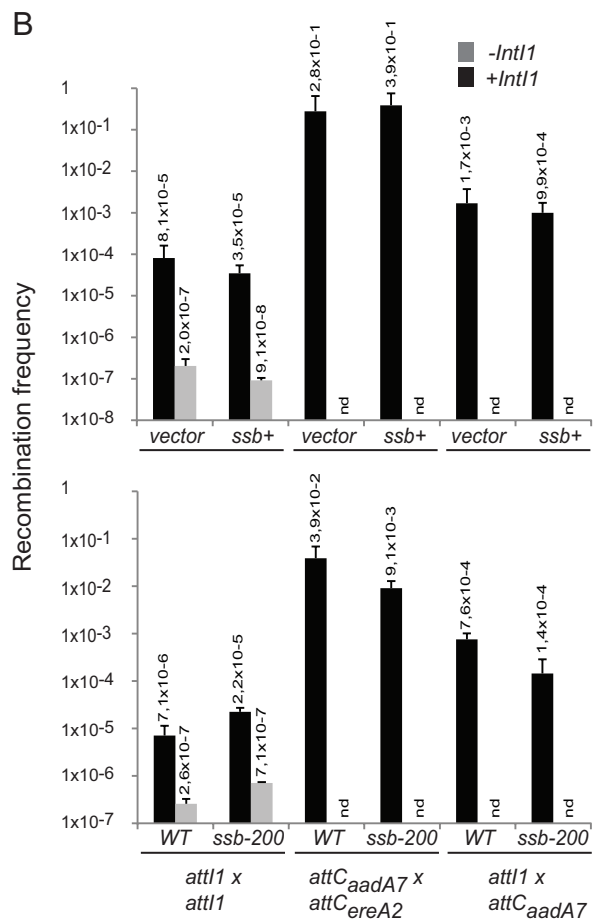
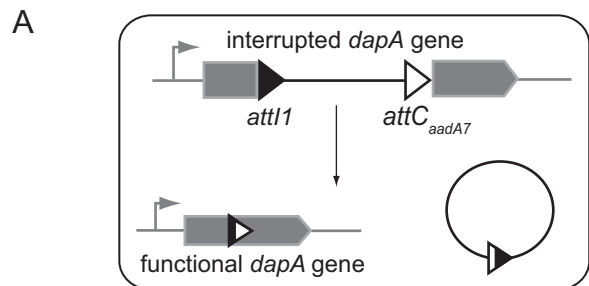


Figure 7

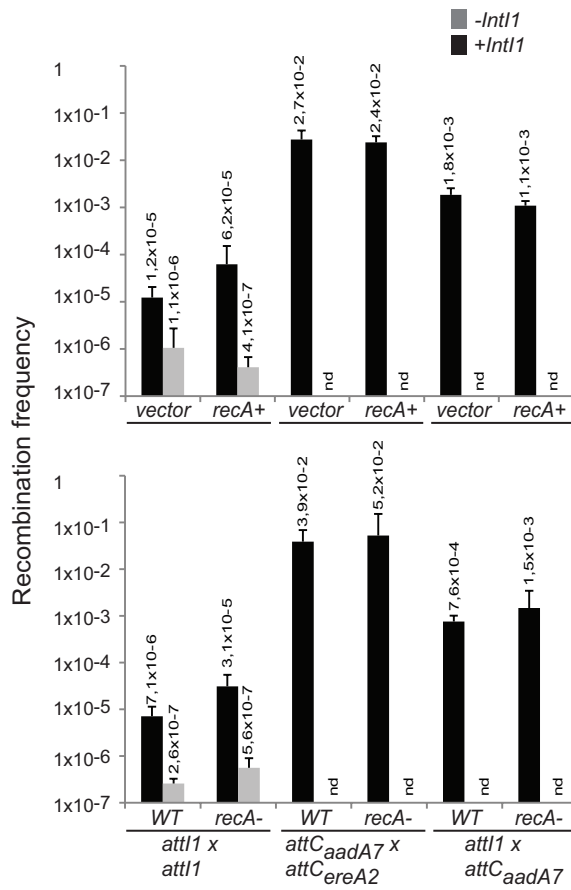


Figure 8