Nucleic Acids Research, 1997, Vol. 25, No. 15 **3009–3016**

A cruciform–dumbbell model for inverted dimer formation mediated by inverted repeats

Ching-Tai Lin, Yi Lisa Lyu and Leroy F. Liu*

Department of Pharmacology, UMDNJ–Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854, USA

Received April 29, 1997; Revised and Accepted June 22, 1997

ABSTRACT

Small inverted repeats (small palindromes) on plasmids have been shown to mediate a recombinational rearrangement event in *Escherichia coli* leading to the formation of inverted dimers (giant palindromes). This recombinational rearrangement event is efficient and independent of RecA and RecBCD. In this report, we propose a cruciform–dumbbell model to explain the inverted dimer formation mediated by inverted repeats. In this model, the inverted repeats promote the formation of a DNA cruciform which is processed by an endonuclease into a linear DNA with two hairpin loops at its ends. Upon DNA replication, this linear dumbbell-like DNA is then converted to the inverted dimer. In support of this model, linear dumbbell DNA molecules with unidirectional origin of DNA replication (ColE1 *ori*) have been constructed and shown to transform *E.coli* efficiently resulting in the formation of the inverted dimer. The ability of linear dumbbell DNA to transform *E.coli* suggests that the terminal loops may be important in bypassing the requirement of DNA supercoiling for initiation of replication of the ColE1 *ori*.

INTRODUCTION

In *Escherichia coli*, RecA is known to be central for homologous recombination (1–3). However, recent studies have demonstrated efficient RecA-independent homologous recombination on both plasmids and chromosomes (4–11). The RecA-independent homologous recombination is independent of the function of known recombination enzymes (9,11,12). Analysis of the RecA-independent recombination on plasmids has revealed complexity of the system. In addition to the expected deletion product of recombination between direct repeats, head-to-tail dimeric products have also been observed (6,8–11). The frequency of formation of these various recombination products also depends on the distance separating the homologous DNA sequences (13) and the presence of DNA sequences (possibly serving as spacers between the direct repeats and an unknown *cis*-element on the plasmid) distant to the homologous sequences (4). Models of sister-strand exchange during DNA replication have been proposed to explain the formation of these dimeric products (6,9,11).

Recent studies of RecA-independent recombination between inverted repeats have also revealed the formation of a dimeric recombination product (5). The studies on recombination between inverted repeats were facilitated by construction of an HPH/*tet* cassette on pBR322. The HPH/*tet* cassette functions as a genetic switch controlling expression of the *tet* gene depending on the orientation of the P fragment (promoter-containing fragment). Recombination between the two inverted H fragments, which changes the orientation of the P fragment and thereby activates expression of the functional *tet* gene, can be readily monitored by tetracycline selection (5). Unlike the dimeric recombination products of direct repeats which are head-to-tail dimers with 1+2 and 1+3 structures (the numbers refer to the number of the repeat units on the head-to-tail dimer; e.g. 1+2 has a total of three repeat units with one repeat located diagonally from the other two tandem repeat units) (6,8–11), the dimeric recombination product of inverted repeats is exclusively head-to-head with two pairs of giant inverted repeats, resembling certain gene amplification products in drug-resistant cells such as the double minute (DM) chromosomes in mammalian cells, the inverted dimer containing the *DFR1* gene in yeast and the H-circles in *Leishmania* (14–19).

Similar to the replication models proposed for direct repeat-mediated formation of dimeric recombination products, a reciprocal-strand-switching (RSS) model involving DNA replication has also been proposed for the formation of the inverted dimer from plasmids containing short inverted repeats (5). However, there has been no direct evidence supporting this model. In the present communication, we consider an alternative model (Fig. 1) which can also explain the formation of the head-to-head dimer satisfactorily. In this model, the inverted repeats are presumed to undergo a structural transition to form a DNA cruciform at a frequency depending on a variety of conditions known to favor this structural transition (20,21). The cruciform is then processed by an endonuclease which cuts diagonally at the Holliday junction. The resulting linear DNA, which is in the form of a dumbbell, can be further processed by replication to form the head-to-head dimer. The simplicity of this cruciform–dumbbell model has prompted us to test the aspects of this model using *in vitro* engineered dumbbell DNA containing the ColE1 *ori*. Theoretically, dumbbell DNA containing the ColE1 *ori* is not expected to replicate in *E.coli* for the following two reasons; first, ColE1 *ori* is known to require negative supercoiling for initiation of DNA replication (22–25). The linear dumbbell DNA is not expected to be supercoiled by gyrase. Second, ColE1 *ori* is unidirectional (26,27) and only one arm of the dumbbell DNA is expected to be replicated through initiation at ColE1 *ori*. We show in the present communication that dumbbell DNA can transform *E.coli* efficiently.

*To whom correspondence should be addressed. Tel: +1 732 235 4592; Fax: +1 732 235 4073; Email: lliu@umdnj.edu

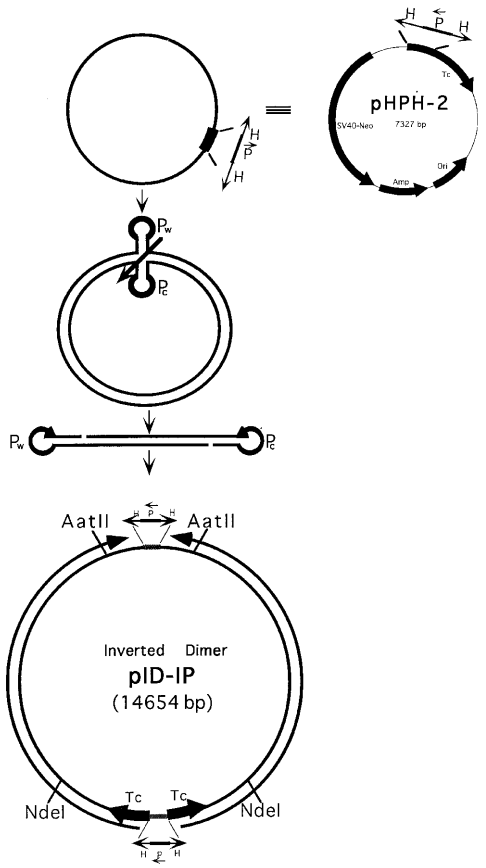

Figure 1. The cruciform–dumbbell model for inverted dimer formation mediated by inverted repeats. In this model, the inverted repeats are presumed to mediate the formation of a cruciform on a DNA molecule (a circular plasmid DNA is shown in this figure as an example). Processing of the cruciform by a junction cutting endonuclease results in the formation of a dumbbell DNA molecule. Following repair, the dumbbell DNA is replicated into an inverted dimeric DNA with two identical junctional fragments (P fragments) also positioned in an inverted orientation.

The sequences of the terminal loops appear to be important for transformation. The potential roles of the terminal loops in replication of linear dumbbell DNA are discussed.

MATERIALS AND METHODS

Enzymes and reagents

Klenow polymerase (large DNA polymerase fragment) was purchased from GIBCO-BRL. T4 DNA ligase was from NEB. Restriction enzymes were from several commercial sources. *Escherichia coli* DNA gyrase was a gift from Dr Martin Gellert (NIH, MD). *Escherichia coli* DNA topoisomerase I was a gift from Dr James C.Wang (Cambridge, MA).

Construction of dumbbell DNA *in vitro*

As shown in Figure 1, the fragment containing the SV40 origin and the neomycin-resistant gene on pHPH-2 was obtained from *Bam*HI digestion of pMAMneo (Clontech, CA). After polymerase fill-in, the blunted fragment was cloned into the *Ssp*I site of pHPH (5). Plasmid pHPH was derived from pBR322 and contained the HPH/*tet* inverted repeats cassette. As reported previously, the HPH/*tet* cassette, which consists of a flipped P*tet* promoter fragment including part of the *tet* gene (the P fragment) flanked by inverted repeats (the two H fragments) can mediate efficient RecA-independent recombination/rearrangement resulting in the exclusive formation of a special inverted dimer (5). The HPH/*tet* cassette is basically a genetic switch controlling transcription of the functional tetracycline gene, depending on the orientation of P fragment. The inverted dimer, pID-IP, was generated by transforming pHPH-2 into *E.coli* DH5α (*rec*A$^-$) followed by selection with tetracycline. This inverted dimer contains a functional tetracycline gene due to the inversion of the flipped P fragment and therefore cells containing the inverted dimer can be readily selected for by resistance to tetracycline.

pID-IP* was constructed from pID-IP by destroying one of the two identical *Nde*I sites on the inverted dimer. This was accomplished by partial digestion with *Nde*I, followed by gel purification of the full length linear DNA, and religation after Klenow polymerase fill-in of the cohesive ends. The resulting plasmid, pID-IP*, therefore, contains only a single *Nde*I site. The dumbbell molecules were generated by linearization of 5 μg of pID-IP* with *Nde*I, followed by alkaline denaturation (0.1 N NaOH) and rapid renaturation (neutralization with 0.1 N HCl plus 100 mM Tris–HCl pH 7.6) at 37°C for 10 min.

To flip the direction of one P fragment in pID-IP*, the pID-IP* DNA was digested with *Aat*II (Fig. 6). Following phenol–$CHCl_3$ extraction and ethanol precipitation, the digested DNA was ligated and transformed into *E.coli* DH5α. The clone with the P fragment flipped was identified by restriction enzyme analysis and designated pID-PP* (inverted dimer with parallel P fragments). Dumbbell DNA molecules derived from pID-PP* were prepared in the same manner as described above for pID-IP*.

Transforming dumbbell molecules into *E.coli*

Dumbbell molecules were digested with *Eco*RV to reduce the residual linear or supercoiled molecules and then transformed into *E.coli* DH5α by heat-shock at 42°C for 30 s following incubation of DNA with competent cells at 4°C for 30 min. The competent cells were made by incubation in 100 mM $CaCl_2$ at 4°C for 20 min. Transformation frequency was obtained from the colony number after plating the transformation mixture on LB (10 g/l Bacto-tryptone, 5 g/l Bacto-yeast extract and 10 g/l NaCl) plates supplemented with 100 μg/ml ampicillin. Plasmid DNAs isolated from transformants were analyzed by restriction enzyme digestion. The inverted dimers recovered from dumbbell transformants have both *Nde*I sites inactivated and are referred to as pID-IP** and pID-PP**, respectively.

Generation of supercoiled dumbbell DNA *in vitro*

The dumbbell DNA molecules derived from pID-IP* were circularized by renaturation at 65°C for 10 min. Circularization was achieved through an intramolecular interaction between the two complementary hairpin loops. As shown in Figure 5A, the circularized dumbbell DNA (CDB) contained a two-base gap which was filled in by Klenow polymerase and ligase *in vitro*. The closed-circular dumbbell DNA (CCDB) was then treated with 40 U *E.coli* DNA gyrase and/or 5 U *E.coli* topoisomerase I (ω protein) in a reaction mixture (containing 100 mM KCl, 40 mM Tris pH 7.5, 10 mM $MgCl_2$, 0.5 mM EDTA and 30 μg/ml BSA) at 37°C for 60 min. The reaction was terminated by addition of Proteinase K and SDS (final concentrations of 200 mg/ml and 1%, respectively) and further incubated at 37°C for 15 min.

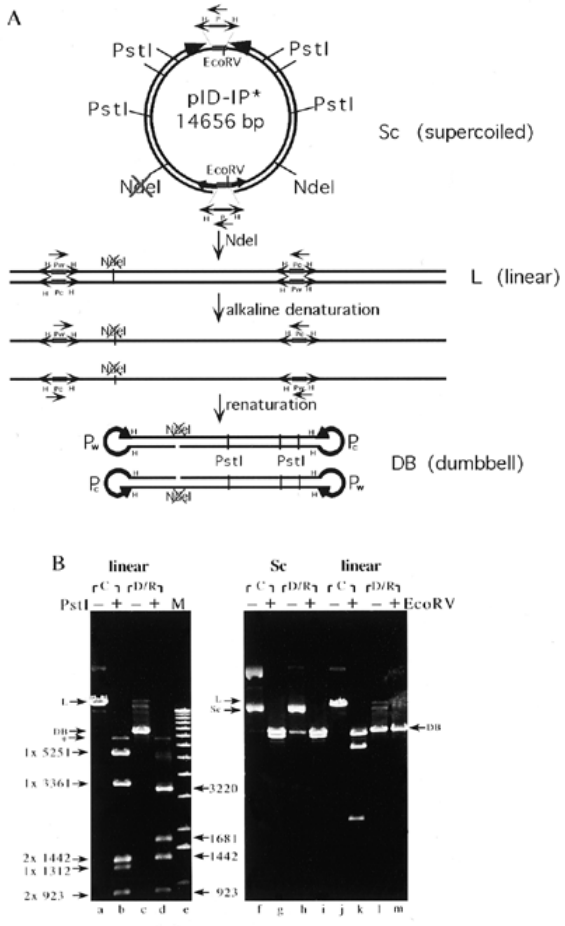

Figure 2. Construction of dumbbell DNA *in vitro*. (**A**) The inverted dimer plasmid pID-IP* with one *Nde*I site destroyed (see Materials and Methods) was linearized by *Nde*I. The linearized pID-IP* DNA was then alkali-denatured and rapidly renatured (R/D) at 37°C for 10 min. The resulting DNA under such conditions is the dumbbell DNA (DB). (**B**) The identity of the dumbbell DNA was verified by restriction enzyme digestion and electrophoresis. The double-stranded region and the single-stranded hairpin loops of the dumbbell DNA were confirmed by *Pst*I and *Eco*RV, respectively. Lanes a and j: the *Nde*I-linearized full length pID-IP* DNA (marked L). Lanes c and l: dumbbell DNA molecules (marked DB) prepared as described in (A) following denaturation/renaturation (D/R) of the linearized pID-IP* DNA. Lane b: linearized pID-IP* DNA digested with *Pst*I. Lane d: the dumbbell DNA (DB) digested with *Pst*I. As expected, four restriction fragments were observed. An extra band (marked *) was generated from residual supercoiled and nicked pID-IP*. Lane e: the 1 kb ladder used as molecular weight markers. Lanes f and g: supercoiled pID-IP* without and with *Eco*RV digestion, respectively. Lanes h and i: supercoiled pID-IP* DNA without and with *Eco*RV digestion following denaturation/renaturation (D/R), respectively. Lanes j and k: *Nde*I-linearized pID-IP* DNA without and with *Eco*RV digestion, respectively. Lanes l and m: dumbbell DNA prepared from pID-IP* without and with *Eco*RV digestion, respectively. The symbol C above the lanes indicates control, meaning no denaturation/renaturation treatment. The symbol D/R above the lanes indicates denaturation/renaturation treatment of the DNA.

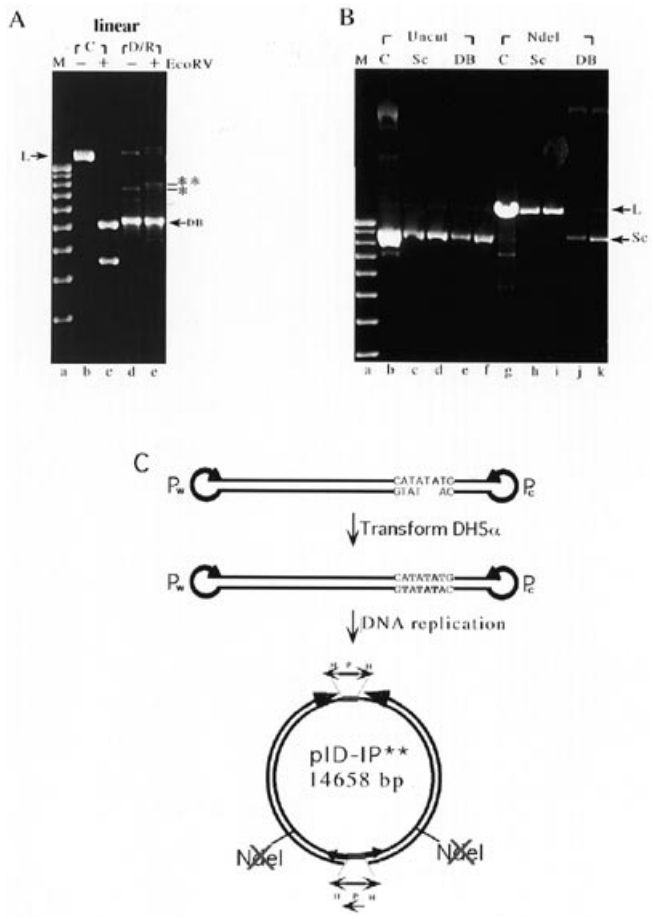

Figure 3. Transformation of *E.coli*. with the dumbbell DNA. (**A**) The dumbbell (DB) DNA molecules prepared from pID-IP* DNA following denaturation/renaturation (D/R) were analyzed by gel electrophoresis. Lane a: the 1 kb ladder used as molecular weight marker. Lanes b and c: *Nde*I-linearized pID-IP* (L) without and with *Eco*RV digestion, respectively. Lanes d and e: the dumbbell (DB) DNA without and with *Eco*RV digestion, respectively. (**B**) The inverted dimers designated pID-IP** isolated from dumbbell transformants have both *Nde*I sites inactivated. Lane a: the 1 kb ladder. Lane b: supercoiled pID-IP* DNA (labeled C on top of the lane). Lanes c and d: supercoiled plasmid DNAs isolated from cells transformed by supercoiled pID-IP* (labeled Sc on top of the lanes) DNA. Out of the 18 transformants, 16 contained inverted dimers (two were monomers). Only two out of the 16 dimers were shown in lanes c and d. Lanes e and f: supercoiled plasmid DNAs isolated from dumbbell transformants (labeled DB on top of the lanes). Lanes g–k: same as DNAs in lanes b–f except that the DNAs were digested by *Nde*I. (**C**) A schematic diagram showing the inactivation of the *Nde*I site following transformation of cells by the dumbbell DNA.

RESULTS

In vitro engineering of linear dumbbell DNA

Our strategy for preparing dumbbell DNA is schematically shown in Figure 2A. We took advantage of the special head-to-head dimer (pID-IP) (inverted dimer with inverted P fragments) generated due to recombinational rearrangement of the HPH/*tet* cassette-containing plasmids. The pID-IP used in our current studies was isolated from tetracycline-resistant clones of pHPH-2, an HPH/*tet* cassette-containing plasmid (see Fig. 1 and Materials and Methods for details). In order to prepare dumbbell DNA, we eliminated one of the two *Nde*I sites on the pID-IP DNA (see Materials and Methods). The resulting plasmid, pID-IP*, was then converted to full-length linear DNA by *Nde*I digestion. Upon alkali denaturation and rapid renaturation, the majority of the linear pID-IP* DNA was converted to the dumbbell (DB) form as shown in Figure 2A. Because the two HPH/*tet* cassettes on the pID-IP* DNA are in the inverted orientation, the two hairpin

loops of the dumbbell DNA are expected to be complementary in their DNA sequences. The structure of dumbbell molecule was confirmed by restriction enzyme analysis (Fig. 2B). *Pst*I digestion of the linear pID-IP* DNA resulted in six bands, and five bands (two of them have twice the amount of DNA due to repeated DNA sequences) are expected based on the restriction map of *Nde*I linearized pID-IP* (Fig. 2B, lane b). The extra band (marked with * to the left of the gel) in Figure 2B was generated from residual supercoiled and nicked pID-IP*. The dumbbell DNA was expected to be digested by *Pst*I into four fragments. Upon denaturation/renaturation, the presumed dumbbell DNA indeed gave only four major bands (Fig. 2B, lane d). The presence of the hairpin loops at the ends of the dumbbell DNA was further confirmed by *Eco*RV digestion which was expected to cut within the double-stranded P segment of the HPH/*tet* cassette of pID-IP*. Since the P segment of the HPH cassette of the dumbbell DNA was located at the single-stranded hairpin loops, no digestion was expected. Indeed, as shown in Figure 2B (compare lanes l and m), *Eco*RV did not digest the dumbbell DNA. As a positive control, *Eco*RV cut the *Nde*I-linearized pID-IP* DNA into three bands (Fig. 2B, compare lanes j and k). Similarly, supercoiled pID-IP* DNA was digested by *Eco*RV into two bands under identical conditions (Fig. 2B, compare lanes h and i). These experiments confirmed the structure of the dumbbell DNA.

Table 1. The transformation frequency of linear, dumbell and supercoiled DNA

Conditions		Colonies[a]		
		Sc pID-IP*[b]	linear pID-IP*[b]	linear pID-PP*[c]
Control	–*Eco*RV	11 300 ± 70	400 ± 120	100 ± 50
	+*Eco*RV	40 ± 10	2 ± 1	4 ± 3
D/R	–*Eco*RV	10 180 ± 210	1370 ± 360	200 ± 140
	+*Eco*RV	370 ± 20	1020 ± 40	100 ± 40

Sc, supercoiled; D/R, denaturation/renaturation.
[a]Transformation efficiency was the average from four independent experiments using 100 ng DNA.
[b]The structure of pID-IP* is shown in Figure 1A. Linear pID-IP* was obtained by digestion with *Nde*I.
[c]The structure of pID-PP* is shown in Figure 6A.

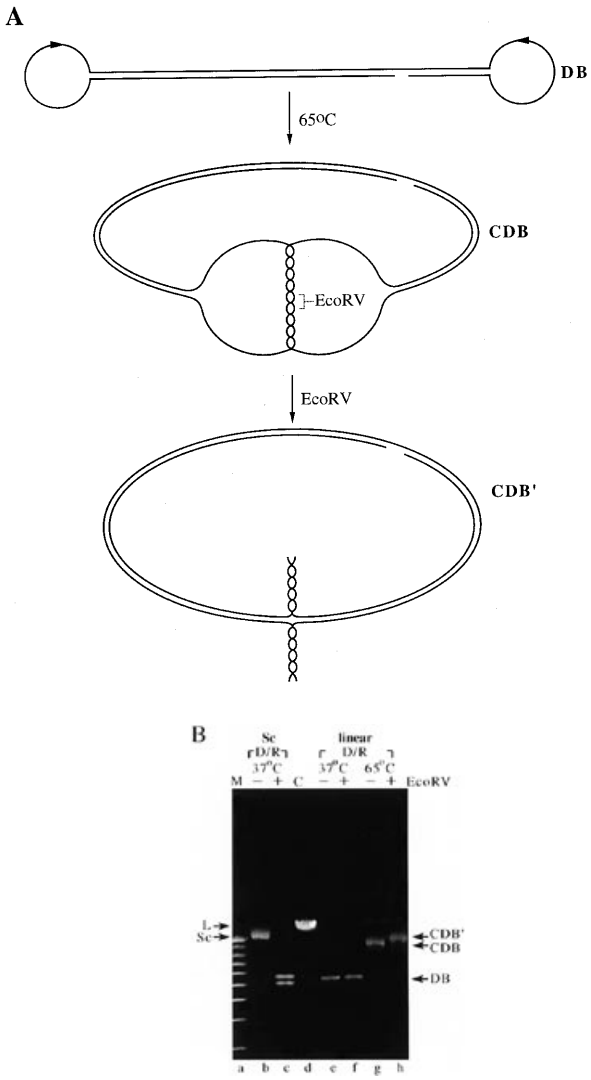

Figure 4. Circularization of the dumbbell DNA through intramolecular interaction between the two complementary hairpin loops. (**A**) A schematic diagram showing the circularization of the dumbbell DNA through intramolecular interaction between the two complementary hairpin loops. The circularized dumbbell DNA (CDB) regenerates a single *Eco*RV site in the duplex region of the two interacting loops. *Eco*RV-cutting of the CDB can then generate an isoform of the circularized dumbbell DNA (CDB′). CDB′ is topologically equivalent to a gapped circular DNA. If the gap is filled in by polymerase/ligase, CDB′ can be considered as an isoform of closed-circular DNA. (**B**) Circularization of the dumbbell DNA. Circularization of the dumbbell DNA was achieved by renaturation of the dumbbell DNA at 65°C for 10 min. Lane a: the 1 kb ladder. Lanes b and c: supercoiled pID-IP* DNA treated without and with *Eco*RV, respectively. Lane d: *Nde*I-linearized pID-IP* DNA. Lane e: *Nde*I-linearized DNA was denatured and renatured (D/R) at 37°C for 10 min to form the dumbbell DNA. Lane f: the dumbbell DNA from lane e was digested with *Eco*RV. Lanes g and h: the same as in lanes e and f, respectively, except that renaturation was performed at 65°C for 10 min.

Dumbbell DNA can efficiently transform *E.coli*

The dumbbell molecules, generated as described above and schematically shown in Figures 2A and 3C, were used to transform *E.coli* DH5α (*rec*A$^-$). *Eco*RV digestion was performed before transformation to reduce residual supercoiled pID-IP* DNA which contaminated the dumbbell preparation. As shown in Table 1, based on four independent transformation experiments, the transformation efficiency of the dumbbell DNA was ~9% of that of supercoiled DNA. On the other hand, linear pID-IP* DNA gave a transformation frequency of only 0.02%, as expected. To ascertain that it was indeed the dumbbell DNA that transformed *E.coli*, we have characterized the plasmid DNAs isolated from the transformants. As shown in Figure 3C, plasmid DNAs isolated from the transformants are expected to be resistant to *Nde*I digestion. This is due to the gapped nature (a two-base gap) of the dumbbell DNA at the site of *Nde*I (one of the *Nde*I sites was destroyed by polymerase fill-in). Upon transformation, the gap was supposed to be filled in in *E.coli* and the resultant dumbbell DNA was expected to be converted into supercoiled DNA with both *Nde*I sites destroyed (Fig. 3C). As expected, plasmid DNAs isolated from all analyzed transformants arising from supercoiled pID-IP* DNA were digestible by *Nde*I, while plasmid DNAs isolated from all analyzed transformants arising from dumbbell DNA were resistant to *Nde*I digestion (18 plasmid

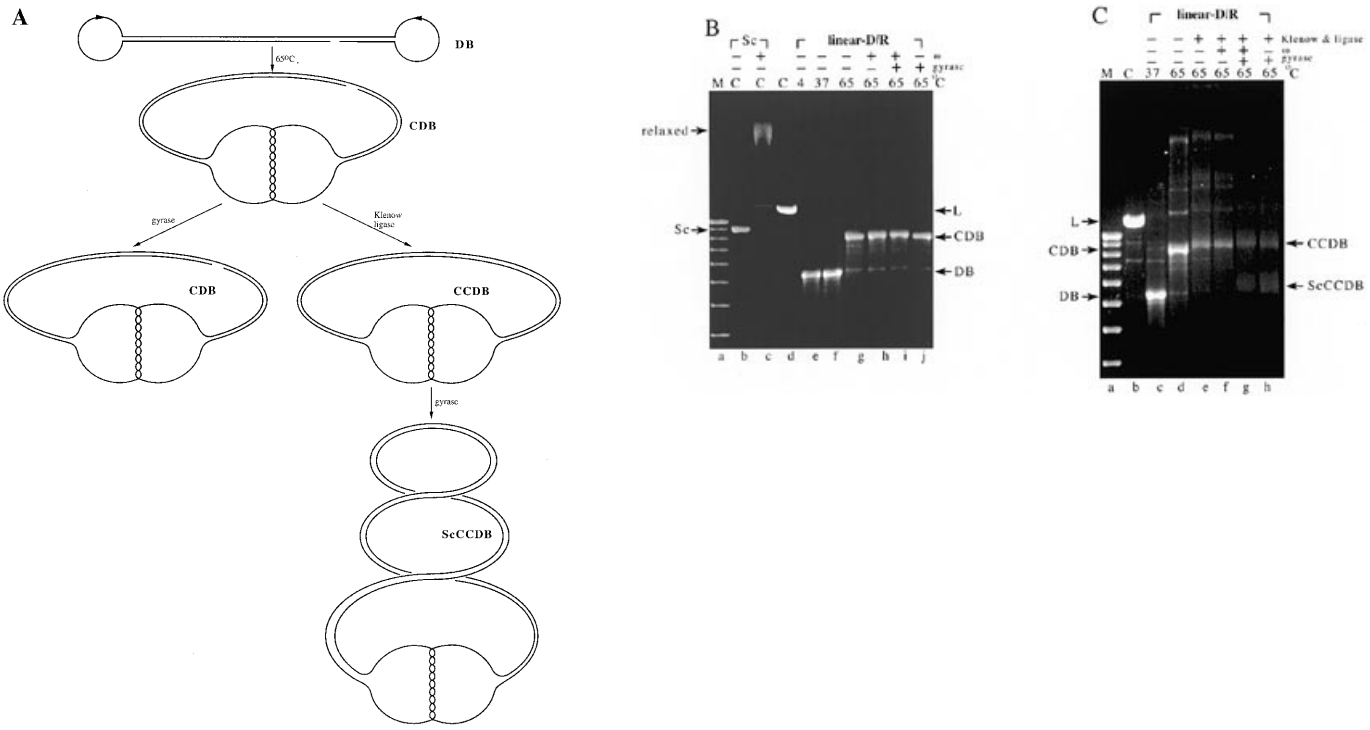

Figure 5. The circularized dumbbell DNA can be negatively supercoiled by DNA gyrase following repair of the gap. (**A**) The interaction of the two complementary hairpin loops (a paranemic joint) of the dumbbell DNA can lead to circularization of the dumbbell (CDB) DNA. *In vivo*, the gap in the circularized dumbbell is expected to be filled in by polymerase/ligase activities. The closed-circular dumbbell (CCDB) can then be negatively supercoiled by DNA gyrase to form the supercoiled closed-circular dumbbell (ScCCDB). ScCCDB can serve as the template for initiation of DNA replication to produce pID-IP**. (**B**) Circularized dumbbell DNA with a unfilled gap cannot be supercoiled by DNA gyrase. Lane a: the 1 kb ladder. Lanes b and c: supercoiled pID-IP* DNA without and with treatment with *E.coli* DNA topoisomerase I (ω). Lane d: *Nde*I-linearized pID-IP* DNA. Lanes e–j: *Nde*I-linearized pID-IP* DNA was renatured for 10 min at various temperatures as indicated above each lane. At 4°C (lane e) and 37°C (lane f), the predominant form of the DNA was the dumbbell DNA (DB). At 65°C (lanes g–j), the predominant form of the DNA was the circularized dumbbell DNA (CDB). Treatment of CDB with ω (lanes h and i) and/or gyrase (lanes i and j) had no effect on the mobility of the CDB. (**C**) Circularized dumbbell DNA with the gap filled in can be supercoiled by DNA gyrase. Lanes b–h: same as lanes d–j in (B), respectively, except that the circularized dumbbell (CDB) was converted to closed-circular dumbbell (CCDB) by polymerase/ligase fill-in of the gap. As shown in lanes g and h, CCDB was converted to supercoiled CCDB (ScCCDB) by gyrase treatment.

DNAs from each transformation were analyzed; only two plasmid DNAs from each transformation are shown in Fig. 3B). These experiments established that dumbbell DNA molecules can indeed efficiently transform *E.coli*.

Interaction between the two complementary hairpin loops of the dumbbell molecule can generate supercoiled DNA

As shown in Figure 3A, in addition to dumbbell DNA molecules (marked DB), denaturation/renaturation of the linear pID-IP* DNA also generated additional species which migrated slower than dumbbell DNA (DB) (e.g. the band marked * in Fig. 3A). Different from the dumbbell DNA, which is resistant to digestion with *Eco*RV, these slower migrating bands were digestible by *Eco*RV (e.g. see slight mobility upshift of band marked ** from the band marked * in Fig. 3A). One possible explanation for the presence of these additional species is that they represent molecules in which the two complementary hairpin loops interact (intra- or inter-molecularly) with each other to form DNA duplexes. As shown in Figure 4A (see CDB, circular dumbbell), intra-molecular interaction between the two complementary hairpin loops of the dumbbell DNA molecule can result in the formation of a circular DNA with a duplex stem formed between the two loops. While the duplex region formed between the two loops is schematically drawn as a regular right-handed duplex (plectonemic joint) in Figure 4A, the topological constraint for such an interaction (a paranemic joint, similar to the interaction between two complementary single-stranded DNA circles with zero linking number) would demand an equal number of left-handed DNA in the form of either Z-DNA or DNA writhe. *Eco*RV is expected to cut at this duplex region (the P segment of the HPH/*tet* cassette) generating a circular DNA with two duplex stems (see CDB′ in Fig. 4A). To demonstrate that such an interaction indeed can occur, we renatured the denatured linear pID-IP* DNA at 65°C for 10 min. As shown in Figure 4B, renaturation at 65°C resulted in the formation of a slower migrating species (labeled CDB in Fig. 4B, compare lanes e and g). This new species can be further converted into an even slower migrating species by digestion with *Eco*RV (labeled CDB′ in Fig. 4B, compare lanes g and h). These results are consistent with our interpretation that the species CDB represents the circular form of the dumbbell DNA through an intra-molecular interaction between the two complementary hairpin loops. The CDB′ species presumably represents a circular form of the dumbbell DNA molecule with the interacting duplex region linearized by *Eco*RV digestion (Fig. 4A).

The CDB molecule represents a novel form of circular DNA. It seems possible that the CDB molecule may be supercoiled by

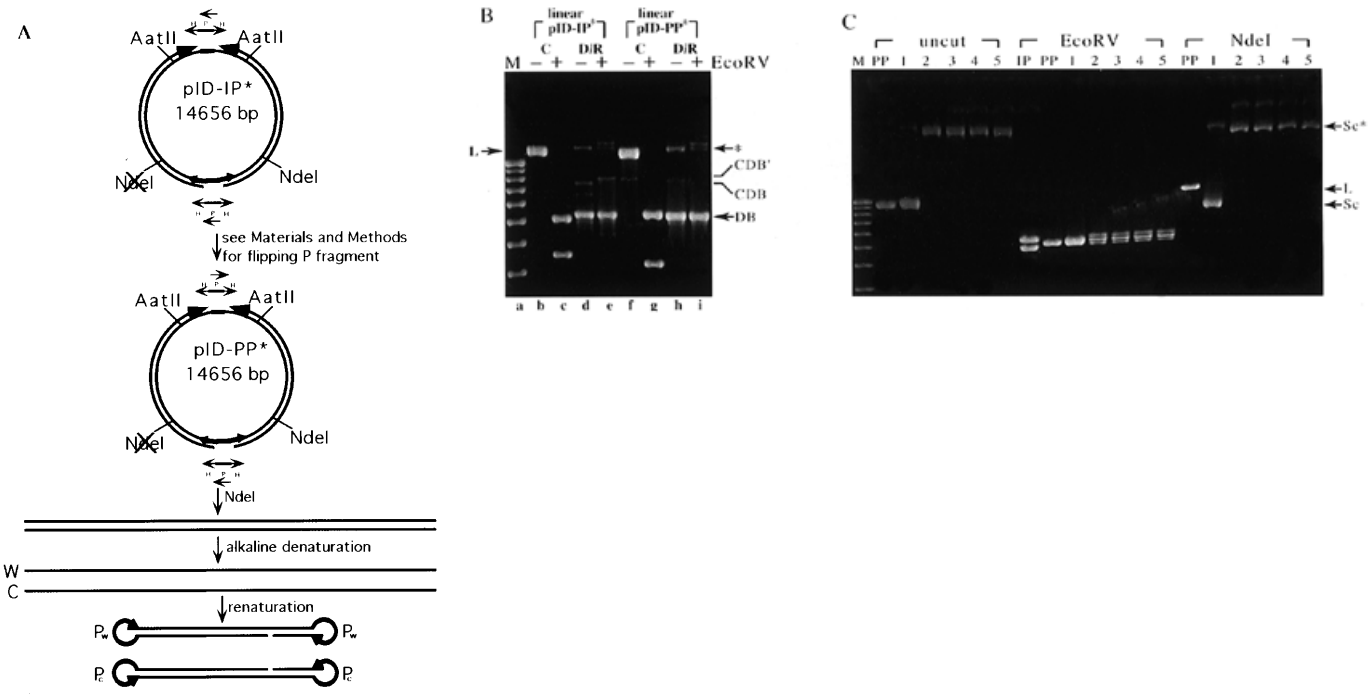

Figure 6. Dumbbell DNA with identical but non-complementary loops can transform *E.coli* to produce both inverted dimeric and inverted tetrameric plasmids. (**A**) A schematic diagram showing the procedure used to prepare dumbbell DNA with identical (and therefore non-complementary) hairpin loops. (**B**) Identification of the dumbbell DNA with identical hairpin loops by *Eco*RV digestion. Lane a: the 1 kb ladder. Lane b: *Nde*I-linearized pID-IP* DNA. Lane c: *Nde*I-linearized pID-IP* DNA digested with *Eco*RV. Lane d: dumbbell DNA prepared from *Nde*I-linearized pID-IP* DNA by denaturation/renaturation (D/R). Lane e: dumbbell DNA in lane d treated with *Eco*RV. Lanes f–i: the same as lanes b–e, respectively, except that pID-PP* DNA was used instead of pID-IP* DNA. The band marked by * on the right-hand side of (B) probably represents a circularized form of dumbbell DNA through an intermolecular interaction between two dumbbell molecules. (**C**) Analysis of the plasmids in cells transformed by dumbbell with identical hairpin loops. About 20% (seven out of 36 colonies) of the transformants contained a tetrameric species. Four DNA samples (labeled 2–5 on top of the lanes) isolated from transformants containing the tetrameric species, and one DNA sample (labeled 1 on top of the lane) isolated from a transformant with the expected dimeric species, were analyzed by restriction digestion. Supercoiled pID-PP* DNA was used as a control (labeled PP on top of the lanes). The lane labeled IP contained pID-IP* digested with *Eco*RV.

DNA gyrase *in vivo* for initiation of DNA replication. To test whether the CDB molecules may be a substrate of DNA gyrase, we first converted the CDB molecules into closed circular dumbbell (CCDB) molecules by filling in the two-base gap using Klenow polymerase and DNA ligase (Fig. 5A). The CCDB molecules were then treated with DNA gyrase. As shown in Figure 5C (lanes g and h), treatment of the CCDB molecules with DNA gyrase converted them into faster migrating species (labeled ScCCDB; Sc stands for supercoiled). As a control, CDB molecules without polymerase/ligase fill-in were shown not to be affected by DNA gyrase treatment (Fig. 5B lanes i and j). These results are consistent with our interpretation that ScCCDB molecules can be considered as supercoiled DNA.

Supercoiling of the dumbbell DNA may be essential for DNA replication

To test whether supercoiling of the dumbbell DNA is important for DNA replication, we constructed dumbbell DNA molecules whose two hairpin loops are not complementary but identical in their DNA sequences. The strategy used for constructing this non-complementary dumbbell DNA is outlined in Figure 6A. A plasmid pID-PP* was constructed by flipping one of the two HPH/*tet* cassettes in pID-IP*. The resulting plasmid pID-PP* is identical to pID-IP* except that the P segments within the two HPH/*tet* cassettes are in a direct-repeat (as opposed to inverted-repeat) orientation. Following linearization with *Nde*I and denaturation/renaturation, two types of dumbbell DNA molecules are produced, both of which have identical DNA sequences in their two hairpin loops (Fig. 6A). While the two hairpin loops from each dumbbell molecule are identical and non-complementary, the hairpin loops between the two types of dumbbell DNA molecules are complementary in their DNA sequences (Fig. 6A). As shown in Figure 6B, the dumbbell DNA molecules prepared from pID-PP* are also resistant to *Eco*RV digestion, consistent with the presence of two single-stranded hairpin loops. However, the dumbbell molecules prepared from pID-PP* had a much lower (~10-fold) transformation efficiency than the dumbbell molecules prepared from pID-IP* (Table 1). Furthermore, when the 36 plasmid DNAs from the transformants were analyzed by *Eco*RV digestion, ~20% (seven out of the 36 colonies analyzed) contained tetrameric plasmids (Fig. 6C), two were monomers (data not shown) and the rest were dimers (one plasmid DNA was shown in Fig. 6C). None of the 36 plasmids could be digested by *Nde*I (five plasmid DNAs were shown in Fig. 6C), suggesting that all of them were derived from the linear dumbbell. In the case of dumbbell DNA molecules prepared from pID-IP*, none of the transformants (a total of 18 analyzed) was a tetramer (data not shown). Two of the 18 were monomers and the rest were dimers (data not shown). Restriction enzyme analysis of the dimers with *Eco*RV resulted in one band, consistent with the head-to-head dimeric structure of the dimer (Fig. 6C, plasmid DNA #1). On the other hand, digestion of the tetramers with *Eco*RV resulted in

three bands with the middle band having twice the amount of DNA (Fig. 6C, plasmid DNAs #2–5), suggesting that the tetrameric plasmid DNA can be considered as a fusion between two pID-PP* DNA molecules.

DISCUSSION

The cruciform–dumbbell (CD) model for the formation of inverted dimers (giant palindromes) from plasmids containing short inverted repeats (small palindromes) (Fig. 1) is different from the proposed reciprocal-strand-switching (RSS) model (5) in the initial steps of forming the dumbbell-like intermediates. According to the RSS model, reciprocal switching of the leading and lagging strands at the repeated sequences during DNA replication, followed by cutting of the Holliday junction, results in the dumbbell-like replication intermediate. In the current CD model, we propose that a palindrome–cruciform transition, followed by cutting of the Holliday junction, results in the dumbbell-like DNA intermediate. In both models, replication of the dumbbell-like DNA intermediate leads to the formation of the inverted dimer (the giant palindrome). While the CD model appears to be at least equally attractive, there are two conceptually difficult steps; one is the formation of the cruciform from the inverted repeats, and the other is initiation of DNA replication from the linear dumbbell-like DNA intermediate. The present work is aimed at addressing the latter question.

In the RSS model, the dumbbell-like DNA intermediate is already pre-initiated. In the CD model, the dumbbell-like DNA is most likely not pre-initiated. This is based on the following considerations; if the palindrome/cruciform transition occurs on the pre-initiated (replicating) DNA molecules, the cruciforms are most likely located behind the replication forks (the positive supercoiling in front of the replication forks would prevent palindrome/cruciform transition). Cutting of the cruciforms located behind the replication forks eventually leads to the formation of linear dumbbell DNA which cannot be supercoiled by gyrase. It has been well documented that linear or relaxed DNA with the ColE1 origin of DNA replication cannot initiate DNA replication *in vivo* and *in vitro* (22–25). Efficient initiation of DNA replication of the ColE1 origin requires a negatively supercoiled DNA template. It is therefore questionable whether the linear dumbbell-like DNA intermediate proposed in the cruciform–dumbbell model is able to initiate DNA replication. In order to test whether the linear dumbbell-like DNA can replicate in *E.coli*, we have engineered these dumbbell DNA molecules *in vitro* and used them for transformation studies. Surprisingly, while normal linear DNA does not transform *E.coli*, the linear dumbbell DNA (derived from pID-IP*) transforms *E.coli* ~9% as efficiently as supercoiled DNA (Table 1). It is unclear why the linear dumbbell DNA transforms *E.coli* efficiently. However, it appears that the DNA sequence and/or the complementarity of the two hairpin loops of the dumbbell DNA are important for efficient transformation (Table 1, compare linear pID-IP* with linear pID-PP* following D/R). Based on our *in vitro* studies, the dumbbell DNA can circularize via intra-molecular loop–loop interaction forming a paranemic joint between the two complementary loops (Fig. 4). The circularized dumbbell DNA can also be negatively supercoiled by DNA gyrase *in vitro* following covalent closure of the two-base gap by polymerase/ligase (Fig. 5). It seems possible that these reactions can also readily occur *in vivo* to convert the dumbbell DNA into a supercoiled DNA for initiation of DNA replication. It is also possible that *E.coli* DNA topoisomerase I can further stabilize the supercoiled circular dumbbell DNA by converting the paranemic joint into a plectonemic joint. Such a conversion (paranemic to plectonemic conversion) is not expected to affect segregation of replicated DNA since the plectonemic joint can be converted into double-stranded interlocks and further resolved by topoisomerase IV and to a lesser extent by DNA gyrase (28–34).

The dumbbell DNA prepared from pID-PP* contains two identical (non-complementary) hairpin loops and thus cannot undergo intramolecular circularization. As expected, the transformation efficiency is ~10 times lower. We have also constructed a plasmid, pID-PP′*, in which the two P fragments have completely different sequences. Again, the dumbbell DNA prepared this plasmid had a seven times lower transformation efficiency (unpublished results). While these results support the supercoiling explanation for efficient dumbbell transformation, we were puzzled by the rather high residual transformation frequency of these dumbbell DNA molecules derived from pID-PP* which cannot circularize. Furthermore, analysis of the transformants has revealed an abundance of a tetrameric plasmid DNA product (0% for pID-IP, 20% for pID-PP* and 35% for pID-PP′*).

While the mechanism for the formation of the tetrameric plasmid remains unclear, we can offer two possible explanations for the residual transformation efficiency of dumbbell DNA with non-complementary or different loops. These dumbbell DNAs can undergo intermolecular loop–loop interactions to form pseudo-supercoiled DNA substrates for initiation of DNA replication. Alternatively, the single-stranded loops can initiate DNA replication thereby bypassing the requirement of negative supercoiling. Priming of single-stranded DNA, especially at the regions with stem–loops, has been well documented (35–37). At present, we are uncertain about the exact mode of initiation of DNA replication by the dumbbell DNA. However, the ability of dumbbell DNA to transform *E.coli* provides partial support to the proposed CD model for the formation of giant palindromes from small palindromes in *E.coli*. The CD model may also be applied to explain the formation of inverted dimers (giant palindromes) during gene amplification in eukaryotic systems (14,16–19, 38–45). Further studies are necessary to verify aspects of the proposed model.

ACKNOWLEDGEMENTS

We would like to thank Dr Carton W.Chen for helpful discussions and Mr Pu Duann and Ms Maarika Liivak for their help in various stages of this work. This work was supported by NIH grant GM27731.

REFERENCES

1 Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. and Rehrauer, W. M. (1994) *Microbiol. Rev.*, **58**, 401–465.
2 Smith, G. R. (1988) *Microbiol. Rev.*, **52**, 1–28.
3 Rao, B. J., Chiu S. K., Bazemore, L. R. and Radding C. M. (1995) *Trends Biochem. Sci.*, **20**, 109–113.
4 Bi, X. and Liu, L. F. (1994) *J. Mol. Biol.*, **235**, 414–423.
5 Bi, X. and Liu, L. F. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 819–823.
6 Bi, X. and Liu, L. F. (1996) *J. Mol. Biol.*, **256**, 849–858.
7 Chedin, F., Dervyn, E., Dervyn, R., Ehrlich, S. D. and Noirot, P. (1994) *Mol. Microbiol.*, **12**, 561–569.
8 Dianov, G. L., Kuzminov, A. V., Mazin, A. V. and Salganik, R. I. (1991) *Mol. Gen. Genet.*, **228**, 153–159.

9 Lovett, S. T., Drapkin, P. T., Sutera, V. A., Jr and Gluckman-Peskind, T. J. (1993) *Genetics*, **135**, 631–642.
10 Lovett, S. T., Gluckman, T. J., Simon, P. J., Sutera, V. A., Jr and Drapkin, P. T. (1994) *Mol. Gen. Genet.*, **245**, 294–300.
11 Mazin, A. V., Kuzminov, A. V., Dianov, G. L. and Salganik, R. I. (1991) *Mol. Gen. Genet.*, **228**, 209–214.
12 Matfield, M., Badawi, R. and Brammar, W. J. (1985) *Mol. Gen. Genet.*, **199**, 518–523.
13 Bi, X., Lyu, Y. L. and Liu, L. F. (1995) *J. Mol. Biol.*, **247**, 890–902.
14 Fried, M., Feo, S. and Heard, E. (1991) *Biochim. Biophys. Acta*, **1090**, 143–155.
15 Grondin, K., Roy, G. and Ouellette, M. (1996) *Mol. Cell. Biol.*, **16**, 3587–3595.
16 Hahn, P. J. (1993) *Bioessays*, **15**, 477–484.
17 Huang, T. and Campbell, J. L. (1995) *J. Biol. Chem.*, **270**, 9607–9614.
18 Ouellette, M., Hettema, E., Wust, D., Fase-Fowler, F. and Borst, P. (1991) *EMBO J.*, **10**, 1009–1016.
19 White, T. C., Fase-Fowler, F., van Luenen, H., Calafat, J. and Borst, P. (1988) *J. Biol. Chem.*, **263**, 16977–16983.
20 Lilley, D. M., Gough, G. W., Hallam, L. R. and Sullivan, K. M. (1985) *Biochimie*, **67**, 697–706.
21 Bowater, R. P., Chen, D. and Lilley, D. M. (1994) *Biochemistry*, **33**, 9266–9275.
22 Marians, K. J. (1992) *Annu. Rev. Biochem.*, **61**, 673–719.
23 Marians, K. J., Minden, J. S. and Parada, C. (1986) *Prog. Nucleic Acids Res. Mol. Biol.*, **33**, 111–140.
24 Ogasawara, N., Seiki, M. and Yoshikawa, H. (1979) *Nature*, **281**, 702–704.
25 Wood, D. C. and Lebowitz, J. (1984) *J. Biol. Chem.*, **259**, 11184–111847.
26 Lovett, M. A., Katz, L. and Helinski, D. R. (1974) *Nature*, **251**, 337–340.
27 Martin-Parras, L., Hernandez, P., Martinez-Robles, M. L. and Schvartzman, J. B. (1991) *J. Mol. Biol.*, **220**, 843–853.
28 Adams, D. E., Shekhtman, E. M., Zechiedrich, E. L., Schmid, M. B. and Cozzarelli, N. R. (1992) *Cell*, **71**, 277–288.
29 Hiasa, H. and Marians, K. J. (1994) *J. Biol. Chem.*, **269**, 16371–16375.
30 Minden, J. S. and Marians, K. J. (1985) *J. Biol. Chem.*, **260**, 9316–9325.
31 Minden, J. S. and Marians, K. J. (1986) *J. Biol. Chem.*, **261**, 11906–11917.
32 Peng, H. and Marians, K. J. (1995) *J. Biol. Chem.*, **270**, 25286–25290.
33 Zechiedrich, E. L. and Cozzarelli, N. R. (1995) *Genes Dev.*, **9**, 2859–2869.
34 Peng, H. and Marians, K. J. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8571–8575.
35 Harrington, C. and Perrino, F. W. (1995) *Nucleic Acids Res.*, **23**, 1003–1009.
36 Langer, U. and Alonso, J. C. (1994) *FEMS Microbiol. Lett.*, **119**, 123–128.
37 Miao, D. M., Honda, Y., Tanaka, K., Higashi, A., Nakamura, T., Taguchi, Y., Sakai, H., Komano, T. and Bagdasarian, M. (1993) *Nucleic Acids Res.*, **21**, 4900–4903.
38 Ouellette, M. and Borst, P. (1991) *Res. Microbiol.*, **142**, 737–746.
39 Hyrien, O., Debatisse, M., Buttin, G. and de Saint Vincent, B. R. (1988) *EMBO J.*, **7**, 407–417.
40 Passananti, C., Davies, B., Ford, M. and Fried, M. (1987) *EMBO J.*, **6**, 1697–1703.
41 Ruiz, J. C. and Wahl, G. M. (1988) *Mol. Cell. Biol.*, **8**, 4302–4313.
42 Ruiz, J. C. and Wahl, G. M. (1990) *Mol. Cell. Biol.*, **10**, 3056–3066.
43 Butler, D. K., Yasuda, L. E. and Yao, M. C. (1995) *Mol. Cell. Biol.*, **15**, 7117–7126.
44 Yasuda, L. F. and Yao, M. C. (1991) *Cell*, **67**, 505–516.
45 Butler, D. K., Yasuda, L. E. and Yao, M. C. (1996) *Cell*, **87**, 1115–1122.