The EMBO Journal vol.3 no.4 pp.761–766, 1984

A linear DNA plasmid from *Streptomyces rochei* with an inverted terminal repetition of 614 base pairs

H.Hirochika[1], K.Nakamura and K.Sakaguchi*

Laboratory of Microbiological Chemistry, Mitsubishi-Kasei Institute of Life Sciences, Machida-shi, Tokyo 194, Japan

[1]Present address: Institute for Plant Virus Research, Tsukuba Science City, Yatabe, Ibaraki 305, Japan

*To whom reprint requests should be sent
Communicated by M.Salas

The terminal structure of a linear plasmid pSLA2, which was isolated from *Streptomyces rochei*, was analysed. The 5′ ends of pSLA2 DNA were blocked by the association of a protein probably covalently bonded with the DNA. This block is removed by alkali treatment and blunt ends with 5′-phosphate and 3′-hydroxy termini were released. The two terminal fragments of pSLA2 were cloned and the nucleotide sequence was determined. An inverted terminal repetition of 614 bp was found along with the presence of further interrupted homologous sequences beyond this area up to 800 bp. These are the first inverted terminal repeat sequences found in microbial linear plasmids.
Key words: 614-bp inverted terminal repetition/*Streptomyces*/ linear plasmids/terminal protein

Introduction

Streptomyces species are filamentous Gram-positive bacteria which produce various antibiotics as secondary metabolites. We have found two linear DNA plasmids designated pSLA1 and pSLA2 (Hayakawa *et al.*, 1979; Hirochika and Sakaguchi, 1982), and a circular transferrable plasmid pSLA3 (unpublished result) from *Streptomyces rochei* producing the lankacidin group of antibiotics. The two linear plasmids have almost the same size, 17 kb, and are close relatives: one is thought to be derived from the other by a small deletion or insertion (Hirochika and Sakaguchi, 1982). Curing experiments suggested that these linear plasmids might be involved in the production of antibiotics and the resistance to the produced antibiotics (Hayakawa *et al.*, 1979). Both linear plasmids have a unique terminal structure: the association of protein with their 5′ ends (Hirochika and Sakaguchi, 1982). A similar terminal structure has been found in adenovirus (Robinson and Bellett, 1974; Rekosh *et al.*, 1977) and *Bacillus subtilus* phage φ29 (Salas *et al.*, 1978; Harding *et al.*, 1978; Ito, 1978). The linear mitochondrial DNAs, named S1 and S2, present in some cytoplasmic male-sterile maize also have a protein bound at the 5′ ends (Kemble and Thompson, 1982). These proteins, named terminal proteins, appear to be involved in DNA replication, serving as a primer for DNA polymerase in adenovirus (Rekosh *et al.*, 1977; Challberg *et al.*, 1980) and φ29 (Watabe *et al.*, 1982; Penalva and Salas, 1982).

Another unique structure, terminal inverted repetition, is found in adenovirus (Steenberg *et al.*, 1977; Arrand and Roberts, 1979; Tolun *et al.*, 1979), φ29 (Yoshikawa *et al.*, 1981; Escarmis and Salas, 1981) and S2 (Levings and Sederoff, 1983). The length of the repetition ranges from 6 bp (φ29) to 208 bp (S2). The inverted terminal repetition was shown to be essential for adenovirus multiplication (Stow, 1982), but the mechanism by which the repetition functions is unclear.

Here we describe the cloning of the two terminal regions of pSLA2 after removal of the terminal protein with alkali, and the analysis of the terminal nucleotide sequences. The results show that pSLA2 has a long inverted terminal repetition of 614 bp, and a model for its secondary structure is presented. This is the first report of the presence of a terminal repetition in DNA sequences amongst microbial linear plasmids. Some of the results presented in this paper have been reported at the Fourth International Symposium on Genetics of Industrial Microorganisms, Kyoto, Japan, 6–11 June, 1982.

Results

Structure of the ends of pSLA2

We have shown that pSLA2 has a terminal protein which is associated, probably covalently, with the 5′ ends of the DNA (Hirochika and Sakaguchi, 1982). The 5′ ends of pSLA2 DNA are insensitive to 5′ exonuclease, even when treated with pronase. This is thought to be due to the covalent association of amino acid(s) which are not removed with pronase. The presence of this 5′ end block is also shown by 5′ end-labeling experiments with polynucleotide kinase (Table I). Pronase-treated pSLA2 was digested with *Bgl*II and labeled with [γ-^{32}P]ATP using polynucleotide kinase after treatment with alkaline phosphatase. If the 5′ ends are blocked with amino acid(s), only one end of the terminal fragments, *Bgl*II-D and -E, is expected to be labeled. As shown in Table I, the amount of radioactivity incorporated into the terminal fragment is ~60% of that of the internal fragments. A

Table I. Presence of 5′ end block and its removal by alkali treatment

Alkali treatment[a]	*Bgl*II fragments[b]	[γ-^{32}P]ATP incorporated (c.p.m.)
–	A	1884
	B	1939
	C	2498
	D,E	1449/fragment
+	A	1867
	B	2010
	C	2584
	D,E	2456/fragment

[a]*Bgl*II fragments were treated with alkali. After neutralization and annealing, the fragments were precipitated with ethanol, treated with alkaline phosphatase, and labeled with polynucleotide kinase and [γ-^{32}P]ATP. The labeled fragments were precipitated with ethanol and fractionated by agarose gel (1%) electrophoresis. The radioactivity in each fragment was counted as described (Hirochika and Sakaguchi, 1982).
[b]*Bgl*II-D and -E fragments are not separable from each other, since these terminal fragments are within the inverted terminal repetition.

similar result was obtained with *Bam*HI-generated fragments. These results indicate that the 5′ ends of pSLA2 are blocked. Previously, the linkage between the terminal protein and DNA was shown to be alkali-labile in adenovirus (Rekosh *et al.*, 1977) and ϕ29 (Salas *et al.*, 1978). We have examined the susceptibility of the 5′ end block to alkali. When the *Bgl*II fragments were treated with alkali, the difference in the extent of labeling between the internal and terminal fragments disappeared (Table I). These results indicate that the 5′ end block is alkali-labile. When treatment with alkaline phosphatase was omitted, no radioactivity was incorporated into the terminal fragments or into the internal fragments, indicating that 5′-phosphate was released at the 5′ ends of pSLA2 after alkali treatment. Similar results were obtained with adenovirus (Carusi, 1977) and ϕ29 (Yoshikawa *et al.*, 1981; Escarmis and Salas, 1981).

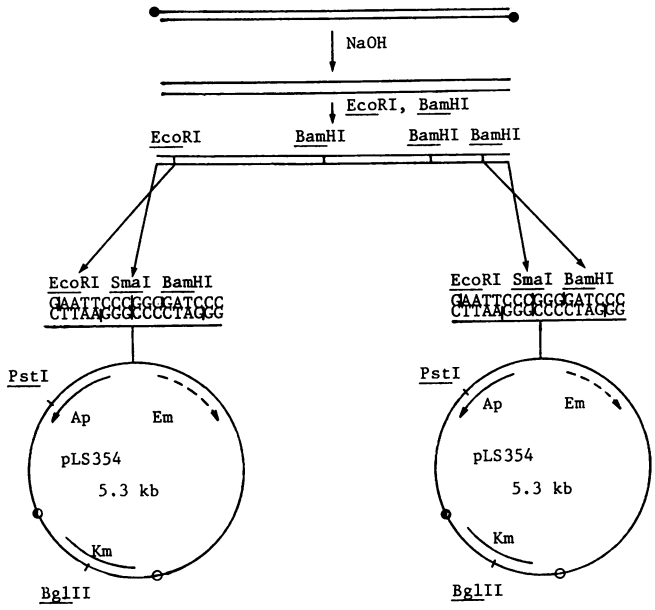

Fig. 1. Cloning of terminal fragments of pSLA2 (●). pSLA2 DNA was treated with NaOH in order to remove the 5′ end block. After neutralization with HCl and renaturation, pSLA2 DNA was treated with *Bam*HI and *Eco*RI, successively. The terminal fragments were separated by agarose gel electrophoresis and purified. These fragments were inserted into *Eco*RI-*Sma*I or *Bam*HI-*Sma*I sites of pLS354. The nucleotide sequence at the cloning site of pLS354 is shown. Ap, ampicillin; Km, kanamycin; Em, erythromycin; and ◐, ○, origins of replication of pBR322 and pUB110, respectively.

pSLA2 DNA is sensitive to exonuclease III (3′ exonuclease) (Hirochika and Sakaguchi, 1982). This suggests that 3′ ends of pSLA2 DNA are recessed or blunt, because this enzyme is double strand-specific. The ends of pSLA2 are not labeled with fill-in experiments using *Escherichia coli* DNA polymerase I (large fragment) and four [α-^{32}P]dNTPs, suggesting that the 3′ ends of pSLA2 DNA are protruding or blunt. We conclude that the ends of pSLA2 DNA are blunt. This is confirmed by the analysis of the nucleotide sequence described in the following paragraphs. 3′ Ends can be labeled with [α-^{32}P]ATP by using terminal deoxynucleotidyl transferase. This result, and the susceptibility to exonuclease III, indicate the presence of unblocked 3′-OH at the ends of pSLA2 DNA.

Cloning of terminal DNA fragments

We have examined the nucleotide sequence of both terminal regions of pSLA2. Because preparation and purification of pSLA2 are laborious for the sequence analysis, we have cloned these regions in *E. coli*. The cloning strategy is shown in Figure 1. As shown above, blunt ends with 5′-phosphate and

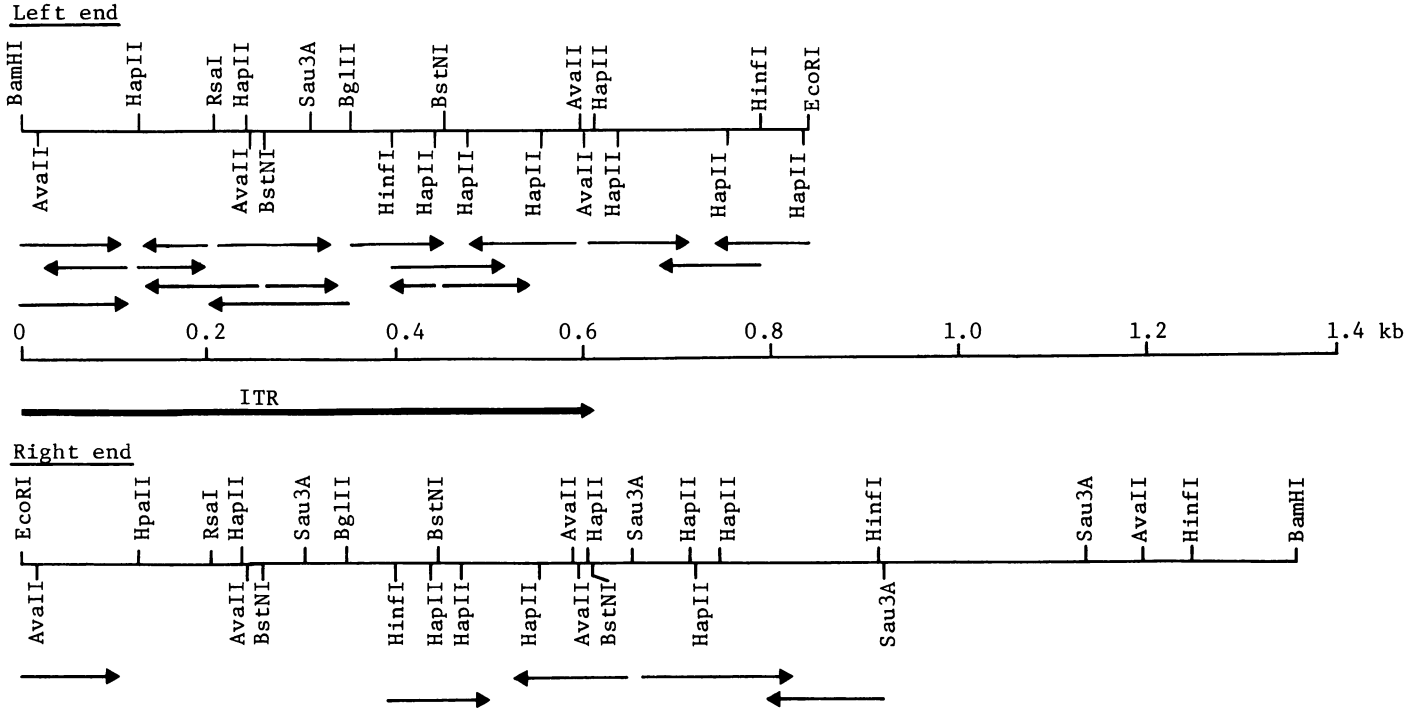

Fig. 2. Physical maps of the cloned left and right terminal fragments. The restriction sites were determined by restriction analysis and confirmed by computer analysis of the nucleotide sequences. The *Bst*NI and *Hap*II sites were not examined in the region where the nucleotide sequence was not determined. *Bam*HI and *Eco*RI sites on the left ends of the maps are derived from the cloning vector. The DNA sequence strategy is shown just below the restriction maps. The location of the inverted terminal repetition (ITR) is shown by the arrow.

3′-OH were released after alkali treatment and renaturation. These ends and the blunt ends generated by *Sma*I digestion can be covalently linked by T4 DNA ligase. The left-end fragment, *Eco*RI-B (0.8 kb), was inserted into *Eco*RI-*Sma*I sites of pLS354. The right-end fragment, *Bam*HI-D (1.4 kb), was inserted into *Bam*HI-*Sma*I sites of pLS354. The recombinant plasmids thus obtained are designated pSL2-E and pSL2-B, respectively. Since the restriction sites used for cloning are close to each other, both cloned fragments are recovered by *Eco*RI-*Bam*HI double digestion with a small addition (seven bases) at the native ends of pSLA2 (Figures 1 and 3). Fine maps of cloned fragments were constructed (Figure 2) which agree well with those of the native terminal fragments. It is concluded that the terminal fragments of pSLA2 are cloned in *E. coli*.

Sequence analysis of cloned terminal fragments

Figure 2 indicates that the cleavage maps from the end to ~600 bp are the same between the left- and the right-end fragments. These results indicate the presence of an inverted terminal repetition of ~600 bp in length. This was confirmed by the following experiment. pSLA2 was alkali-denatured and renatured at low DNA concentration. Under this condition, intrastrand-hybrid formation predominated over interstrand-hybrid formation. Renatured DNA was digested with single strand-specific nuclease S1. The length of S1 nuclease-resistant DNA was estimated by gel electrophoresis. The DNA fragment obtained was ~600 bp long and had a cleavage pattern identical to that of the terminal fragments. These results indicate that an ~600 bp long inverted repetition is present at the two termini of pSLA2.

The structure and the exact length of the inverted repetition were examined by determining the nucleotide sequence using the method of Maxam and Gilbert (1980). The sequencing strategy is shown in Figure 2. The gel showing the sequence around the junction between the end of pSLA2 and the vector plasmid is shown in Figure 3. The nucleotide sequences at both ends are identical. To examine whether an intact end was cloned or not, the nucleotide sequence of the ends of native pSLA2 was examined. Native 5′ and 3′ ends were labeled with [γ-^{32}P]ATP and [α-^{32}P]ATP, respectively, and subjected to the sequence analysis. Both right and left ends are shown to be blunt and to have the same sequence: 5′ CCCGCGGAGCGGTCC 3′. This sequence is identical to that of the cloned terminal fragments (Figure 3). These results indicate that intact ends were cloned. The nucleotide sequence of the terminal regions of pSLA2 is shown in Figure 4. An inverted terminal repetition of 614 bp is detected. The right-end repetitive region is not entirely sequenced, since the hybridization experiments formerly described showed that the repetition is identical.

In the region beyond the 614 bp of complete terminal homology, there are still 11 sites of homologous sequences which are indicated by boxes on Figure 4. The total length of homologous sequences in the region from 615 to 800 bp from both ends reached 65%.

Discussion

Inverted terminal repetition

Various specific structures such as cohesive ends, direct or inverted repeats, or hairpins, are usually present at the terminal region of linear DNA genomes (Kornberg, 1980). We have shown here, and previously (Hirochika and Sakaguchi, 1982), that the structure of the ends of pSLA2 is similar to that of adenovirus DNA and *Bacillus subtilis* phage ϕ29 DNA: (i) these genomes have inverted terminal repetitions with the terminal protein covalently linked to each 5′ end; (ii) the linkage between the genome and the terminal protein is alkali-labile, and 5′-phosphate is released after alkali treatment; (iii) the

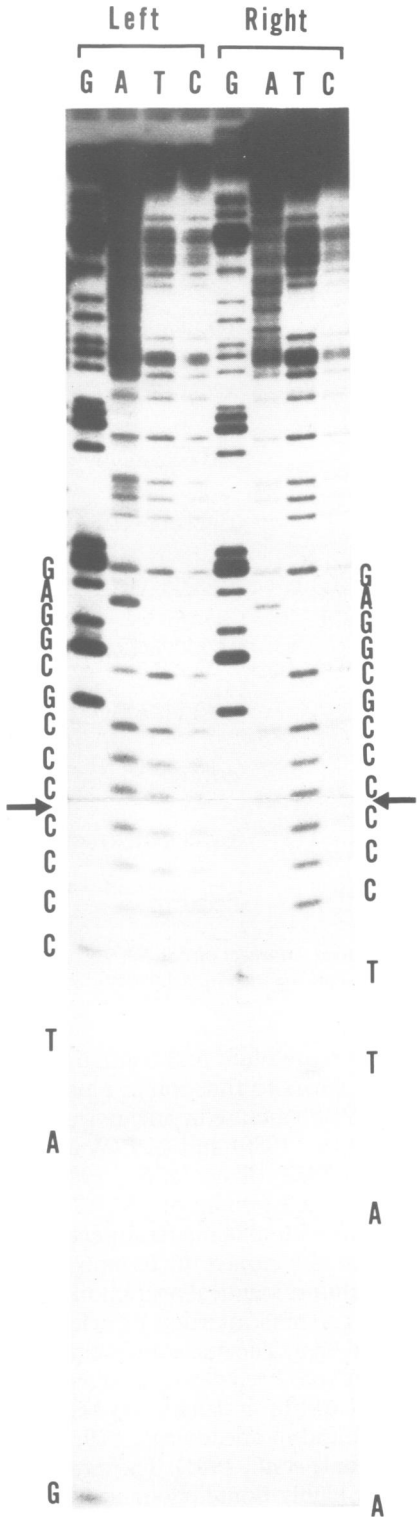

Fig. 3. A 20% sequence gel showing the end sequence of the cloned left and right at the 5′ ends of *Bam*HI and *Eco*RI sites, respectively, labeled with [γ-^{32}P]ATP. The sequences shown under the arrows are derived from the cloning vector.

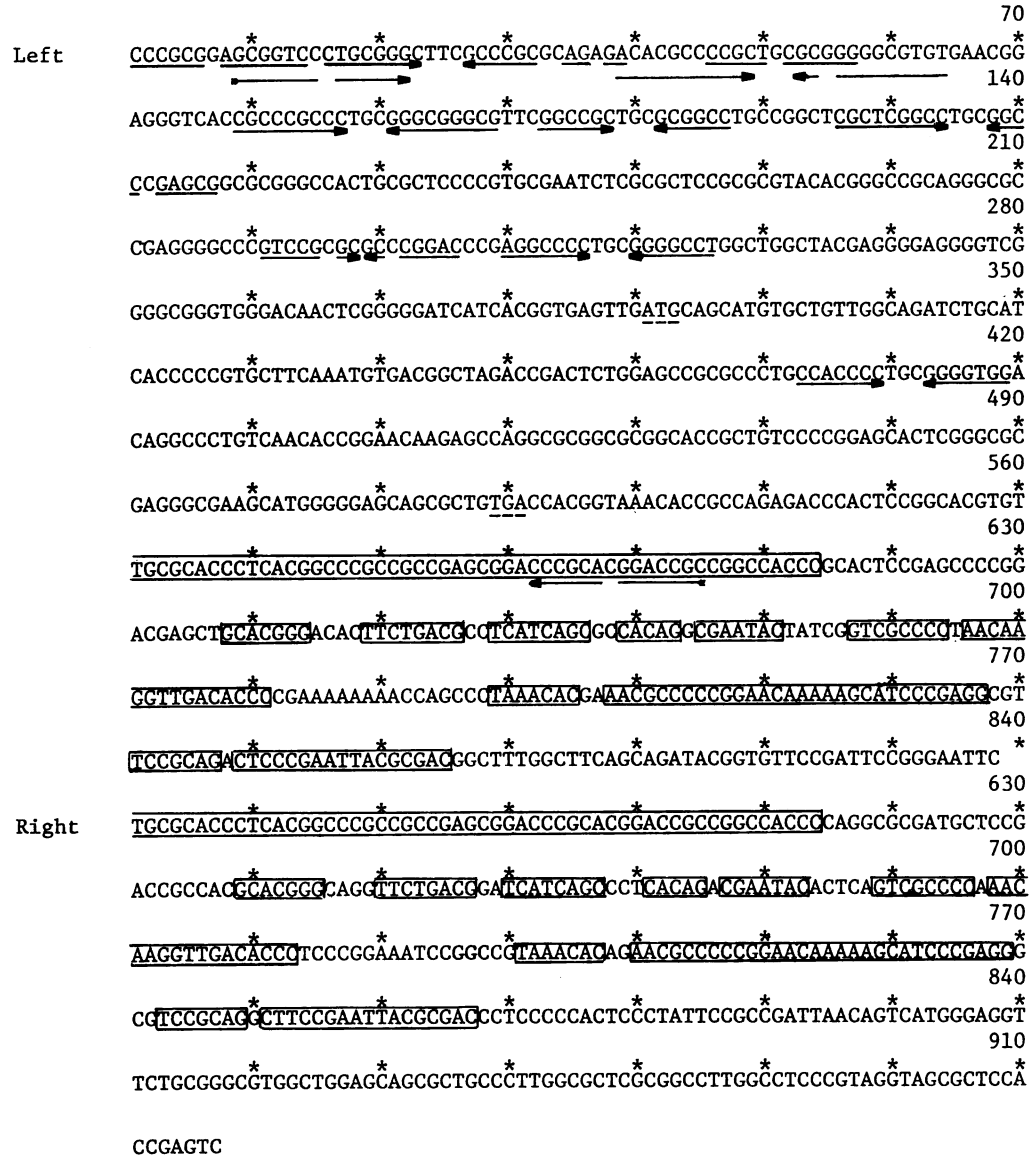

Fig. 4. Terminal nucleotide sequences of pSLA2. The inverted terminal repetition extends to nucleotide 614. The sequences homologous between the two ends are boxed. The palindromic sequences are indicated by the arrows. The initiation and termination codons of a possible coding sequence are indicated by the broken underlining.

ends of the genomes are blunt and contain free 3′-OH. The terminal protein appears to function as a primer for 'protein priming' the DNA polymerase in adenovirus (Rekosh *et al.*, 1977; Challberg *et al.*, 1980) and ϕ29 (Watabe *et al.*, 1982; Peñalva and Salas, 1982; Blanco *et al.*, 1983).

We have cloned intact termini of pSLA2 after removal of the terminal protein with alkaline treatment. Using the same procedure, we have also cloned the termini of pSLA1 DNA, which showed an almost identical restriction pattern with that of pSLA2, but has a small insertion or deletion of < 100 bp in the non-terminal area. The nucleotide sequence of the cloned termini of pSLA2 has been determined and an inverted terminal repetition of 614 bp in length has been found. This is longer than that found in adenovirus, ϕ29, and maize S1 or S2 plasmids (Lonsdale *et al.*, 1981). The terminal sequence of pSLA2 plasmid is highly homologous up to ~ 800 bp from the ends, forming an interrupted homologous region as shown on Figure 4. The 11 boxed sequences are identical.

The functions of inverted terminal repetitions still remain unclear. However, Lechner and Kelly (1977), have proposed a model in which the single-stranded DNA molecules, which are formed during the replication of adenovirus DNA by a strand-displacement mechanism, form a panhandle structure as a result of annealing of the complementary repetitive sequences at both ends. Molecules with this structure have not yet, however, been detected.

Palindromes in the terminal sequences

In pSLA2, nine palindromic structures are detected at the terminal half of the repetition sequences (see Figure 4). One of the functions of such structures could be the binding of specific proteins, which might act by sticking together two double-stranded DNAs. An example of such a cohesive protein is the terminal protein of the adenovirus DNA, molecules of which stick to each other (Challberg *et al.*, 1980). Also, the protein-mediated juxtaposition of many DNA strands is visible in the salivary chromosome of *Drosophila*.

If such a four-stranded structure exists, DNA possessing terminal inverted repeats would assume a racket frame-like form (Figure 5). Inverted terminal repetition sequences are common in transposable elements, in the proviral structure of

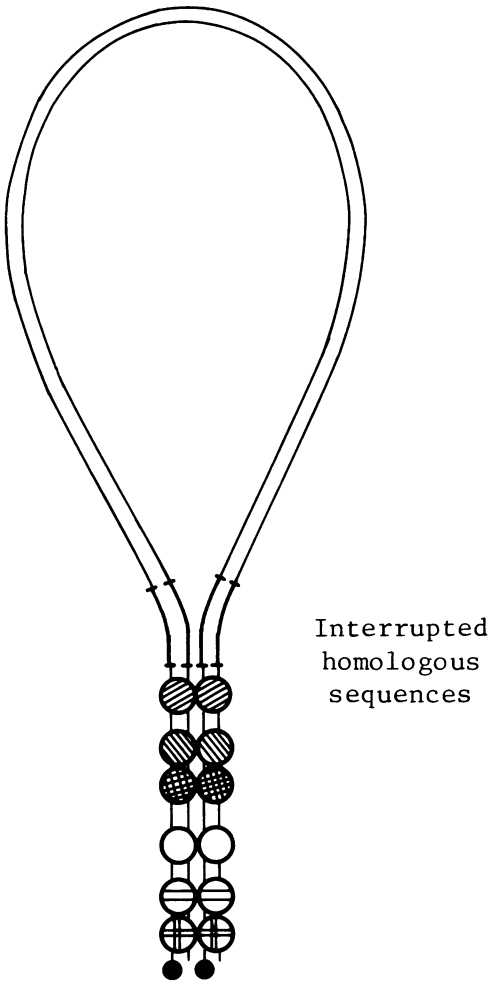

Fig. 5. Racket frame model.

● The terminal protein which initiates DNA replication and has itself a cohesive nature.

The proteins which recognize and bind to the specific DNA sequences and stick together.

some viruses (Varmus, 1982), in the transposable element of fold-back type in *Drosophila* cells (Potter, 1982) and in the maize S1 and S2 linear plasmids (Weissinger *et al.*, 1982). A common feature of these DNA elements is the ability to insert into chromosomal or mitochondrial DNA and, in the case of pSLA2, attempts to detect such integration are under way.

Hairpin structures at termini of inverted repeats

Computer analysis of the possible secondary structure of pSLA2 DNA termini, from 1 to 70 bp from the two 5′ termini, reveals several hairpin structures on the basis of low free-energy requirement, which might be easily formed, perhaps by binding of proteins. For this analysis, matrices were constructed to pair each nucleotide of a sequence with every other nucleotide in order. To calculate free energies of base pairing, we used the rules of Tinoco *et al.* (1971), as modified to account for effects of neighbouring pairs (Tinoco *et al.*, 1973). The two most energetically favorable structures formed from the 5′-terminal repeats are depicted in Figure 6. Relative to an unpaired single DNA strand, the free energies

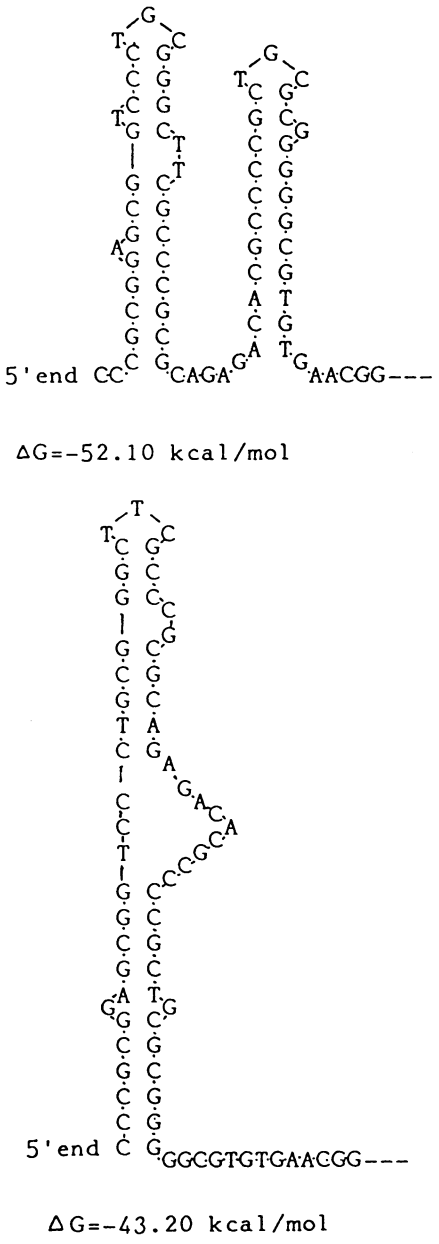

Fig. 6. Hairpin structures at termini of inverted repeats. See text for explanation.

of formation of these base-paired structures are −52.1 kcal/mol and −43.2 kcal/mol, respectively. Given the number of nucleotides in the repeat sequences and the numbers of G-C pairs within the folded structures, the two energies indicate a high degree of stability of the structures shown, and they could be expected to form spontaneously in single-stranded sequences at the termini.

Materials and methods

Enzymes and nucleoside triphosphates

Restriction endonucleases were purchased from New England Biolabs, Bethesda Research Laboratories or Takara Shuzo. T4 polynucleotide kinase was obtained from P-L Biochemicals. Bacterial alkaline phosphatase, *E. coli* DNA polymerase I (large fragment), T4 DNA ligase, nuclease S1 and terminal deoxynucleotidyl transferase were from Bethesda Research Laboratories. [γ-^{32}P]ATP (3000 − 5000 Ci/mmol; 1 Ci = 3.7 x 10^{10} Be), [γ-^{32}P]dNTPs (3000 Ci/mmol) and [α-^{32}P]ATP (3000 Ci/mmol) were purchased from Amersham.

Preparation of plasmid DNAs

pSLA2 was prepared from *S. rochei* 7434-AN4 after pronase treatment as described (Hirochika and Sakaguchi, 1982) and purified by CsCl equilibrium centrifugation. The recombinant plasmids carrying the terminal fragments of pSLA2 were identified by a rapid isolation procedure (Birnboim and Doly, 1979) and prepared on a large scale for the sequence analysis as described (Tanaka and Kawano, 1980).

Labeling of DNA

The cloned pSLA2 fragments were labeled by using [γ-^{32}P]ATP and T4 polynucleotide kinase as described (Maxam and Gilbert, 1980). The native 5′ ends of pSLA2 DNA were labeled after alkali treatment and renaturation. Alkali treatment and renaturation were performed as follows. pSLA2 DNA or its restriction fragments were treated with 0.1 M NaOH at 37°C for 90 min. The DNA samples were neutralized with HCl, and annealed by incubation at 65°C for 2 h in 2 x SSC. After ethanol precipitation, the samples were used for end-labeling. 3′ Ends of pSLA2 DNA were labeled by using [α-^{32}P]ATP and terminal deoxynucleotidyl transferase as described (Maxam and Gilbert, 1980). Fill-in experiments were performed by using four [α-^{32}P]dNTPs and DNA polymerase I (large fragment) as described (Yoshikawa *et al.*, 1981).

DNA sequence analysis

5′ or 3′ ends of DNA were labeled as described above and subjected to sequence analysis according to Maxam and Gilbert (1980).

Cloning of terminal fragments

pSLA2 DNA was treated with alkali and renatured as described above. This DNA was digested with *Bam*HI and *Eco*RI, successively. Terminal fragments, *Eco*RI-B and *Bam*HI-D (see Figure 1), were fractionated by agarose gel electrophoresis, electrophoretically eluted from the gel slices, and purified by phenol and ether extractions. The vector plasmid pLS354 (Figure 1; H.Hirochika, K.Nakamura and K.Sakaguchi, in preparation) was digested with *Bam*HI and *Sma*I or *Eco*RI and *Sma*I to produce two fragments. The large fragments of *Bam*I-*Sma*I and *Eco*RI-*Sma*I-treated pLS354 were mixed with *Bam*HI-D and *Eco*RI-B fragments, respectively, and ligated with T4 DNA ligase. The ligated DNA was used to transform *E. coli* C600 rk$^-$mk$^-$ *thi thr leuB trpB* (Nagahari *et al.*, 1977) to ampicillin resistance (50 μg/ml). Transformation was carried out as described (Norgard *et al.*, 1978). The vector plasmid pLS354 is a *B. subtilis-E.coli* shuttle promoter probe vector. By using this vector, promoter activity of the fragments cloned in *Eco*RI, *Sma*I or *Bam*HI site can be assayed by monitoring the expression of erythromycin resistance gene.

Acknowledgements

The authors are thankful to Professor N. Otake of the University of Tokyo for generously providing the *Streptomyces* strains.

References

Arrand,J.R. and Roberts,R.J. (1979) *J. Mol. Biol.*, **128**, 577-594.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.
Blanco,L., Garcia,J.A., Peñalva,M.A. and Salas,M. (1983) *Nucleic Acids Res.*, **11**, 1309-1323.
Carusi,E.A. (1977) *Virology*, **71**, 380-394.
Challberg,M.D., Desiderio,S.V. and Kely,T.J.,Jr. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5105-5109.
Escarmis,C. and Salas,M. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1446-1450.
Harding,N., Ito,J. and David,G.S. (1978) *Virology*, **84**, 279-292.
Hayakawa,T., Tanaka,T., Sakaguchi,K., Otake,N. and Yonehara,H. (1979) *J. Gen. Appl. Microbiol.*, **25**, 225-260.
Hirochika,H. and Sakaguchi,K. (1982) *Plasmid*, **7**, 59-65.
Ito,J. (1978) *J. Virol.*, **28**, 895-904.
Kemble,R.J. and Thompson,D. (1982) *Nucleic Acids Res.*, **10**, 8181-8190.
Kornberg,A. (1980) *DNA Replication*, published by Freeman, San Francisco.
Lechner,R.L. and Kelly,T.J.,Jr. (1977) *Cell*, **12**, 1007-1020.
Levings,C.S. and Sederoff,R.R. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4055-4059.
Lonsdale,D.M., Thompson,R.D. and Hodge,T.P. (1981) *Nucleic Acids Res.*, **9**, 3657-3669.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Nagahari,K., Tanaka,T., Hishinuma,F., Kuroda,M. and Sakaguchi,K. (1977) *Gene*, **1**, 141-152.
Norgard,M.V., Keem,K. and Monahan,J.J. (1978) *Recomb. Technol. Bull.*, **1**, 4.
Peñalva,M. and Salas,M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5522-5526.
Potter,S.S., Truen,M., Phillips,M. and Maher,A. (1980) *Cell*, **20**, 639-647.
Rekosh,D.M.K., Rusell,W.C., Bellet,A.J.D. and Robinson,A.J. (1977) *Cell*, **11**, 283-295.
Robinson,A.J. and Bellett,A.J.D. (1974) *Cold Spring Harbor Symp. Quant. Biol.*, **39**, 523-531.
Salas,M., Mellado,R.P., Vinuela,E. and Sogo,J.M. (1978) *J. Mol. Biol.*, **119**, 269-291.
Steenberg,P.H., van Ormondt,H.M. and Sussenbach,J.S. (1977) *Nucleic Acids Res.*, **4**, 4371-4389.
Stow,N.D. (1982) *Nucleic Acids Res.*, **10**, 5105-5119.
Tanaka,T. and Kawano,N. (1980) *Gene*, **10**, 131-136.
Tinoco,I.,Jr., Uhlenbeck,O.C. and Levine,M.D. (1971) *Nature*, **230**, 362-367.
Tinoco,I.,Jr., Borer,P.N., Dengler,B., Levine,M.D., Uhlenbeck,O.C., Crothers,D.M. and Gralla,J. (1973) *Nature New Biol*, **246**, 40-41.
Tolun,A., Alestrom,P. and Petterson,U. (1979) *Cell*, **17**, 705-713.
Tonegawa,S. (1983) *Nature*, **302**, 575-581.
Varmus,H.A. (1982) *Science (Wash.)*, **216**, 812-820.
Watabe,K., Smith,M.F., Sugino,A. and Ito,J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5245-5248.
Weissinger,A.K., Timothy,D.H., Levings,C.S.,III, Hu,W.W.L. and Goodman,M.M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 1-5.
Yoshikawa,H., Friedmann,T. and Ito,J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1336-1340.

Received on 13 December 1983; revised on 30 January 1984