Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 4753–4757, May 1993
Biochemistry

Reverse gyrase: A helicase-like domain and a type I topoisomerase in the same polypeptide

(positive supercoiling/Archaebacteria/hyperthermophiles)

FABRICE CONFALONIERI*, CHRISTIANE ELIE†, MARC NADAL*, CLAIRE BOUTHIER DE LA TOUR*, PATRICK FORTERRE†, AND MICHEL DUGUET*‡

*Laboratoire d'Enzymologie des Acides Nucléiques and †Laboratoire des Archaebactéries, Institut de Génétique et Microbiologie, Unité de Recherche Associée 1354 Centre National de la Recherche Scientifique, Université Paris Sud 91 405 Orsay, France

Communicated by James C. Wang, January 19, 1993

ABSTRACT **Reverse gyrase is a type I DNA topoisomerase able to positively supercoil DNA and is found in thermophilic archaebacteria and eubacteria. The gene coding for this protein was cloned from *Sulfolobus acidocaldarius* DSM 639. Analysis of the 1247-amino acid sequence and comparison of it with available sequence data suggest that reverse gyrase is constituted of two distinct domains: (*i*) a C-terminal domain of ≈630 amino acids clearly related to eubacterial topoisomerase I (*Escherichia coli topA* and *topB* gene products) and to *Saccharomyces cerevisiae top3*; (*ii*) an N-terminal domain without any similarity to other known topoisomerases but containing several helicase motifs, including an ATP-binding site. These results are consistent with those from our previous mechanistic studies of reverse gyrase and suggest a model in which positive supercoiling is driven by the concerted action of helicase and topoisomerase in the same polypeptide: this constitutes an example of a composite gene formed by a helicase domain and a topoisomerase domain.**

It is today well established that topoisomerases play a crucial role in DNA structure and function (1, 2). These enzymes seem to act in two ways: (*i*) they are able to solve the topological problems intrinsic to the DNA double helix during replication, transcription, recombination, or chromatin condensation and decondensation; (*ii*) some topoisomerases have exploited the circular or pseudocircular (chromatin loops) structure of the genetic material to introduce stress into DNA by means of supercoiling. The enzymes specialized in the production of superhelical turns in a circular DNA were named "gyrases" (3, 4). Among these, the best characterized enzyme is the eubacterial gyrase, a type II topoisomerase that uses the energy of ATP hydrolysis to maintain the *in vivo* level of negative supercoiling of the bacterial chromosome (5). Several years ago, another kind of DNA supercoiling activity was discovered in hyperthermophilic archaebacteria (6) and was called "reverse gyrase." We have shown that reverse gyrase is widely distributed in various archaebacterial families and also in thermophilic eubacteria (7, 8). This enzyme has the specific ability to catalyze *in vitro* positive supercoiling of a closed circular DNA at high temperature (9). Surprisingly, reverse gyrase turned out to be a type I topoisomerase (9) and is the only topoisomerase I that depends on ATP and can catalyze DNA supercoiling. The biological function of reverse gyrase is still unclear, but we found that positive supercoiling occurs *in vivo* (10). The enzyme from *Sulfolobus acidocaldarius* is composed of a single polypeptide of apparent molecular mass of ≈130,000 (11). Mechanistic studies (12) indicated that reverse gyrase transiently cleaves a single DNA strand, forming a covalent link with the 5′ end of the broken strand, as described for eubacterial topoisomerase I (13). In addition, we have shown that stoichiometric binding of the enzyme changes the DNA conformation, presumably by unwinding the double helix (12). We report here the cloning and sequencing of the gene encoding reverse gyrase from the archaebacterium *S. acidocaldarius* strain DSM 639. Sequence data[§] support our previous mechanistic studies of reverse gyrase and a model in which positive supercoiling is driven by a helicase-plus-topoisomerase process.

MATERIALS AND METHODS

Construction and Screening of the Genomic Library of *S. acidocaldarius*. Total DNA from *S. acidocaldarius* was partially digested with *Sau3A*. DNA fragments with a length of 4–9 kb, isolated by sedimentation on a sucrose gradient, were linked to *Bam*HI/*Eco*RI adaptators and ligated to the *Eco*RI arms of the expression vector λgt11 (protoclone λgt11 system, Promega). Packaged phages were titered by using *Escherichia coli* Y 1090 R^- strain: a total of 10^6 total phages with 89% recombinants was obtained. The library was screened, as described (14), with polyclonal antibodies previously prepared against the purified protein (11). Fifty positive clones were isolated and purified. Some clones were tested in *E. coli* lysogen strain Y1089, as described (14), and all of them express a protein with an apparent molecular mass of 120 kDa (visualized by immunoblotting, see *Results*).

DNA Sequencing. The DNA of each clone was prepared, as described (14). The inserts of different independent clones were subcloned into the double-stranded vector pGEM-3Zf(+) (Promega) and sequenced by the dideoxynucleotide chain-termination method (15) using a T7 sequencing kit (Pharmacia) and synthetic oligodeoxynucleotides as primers. Because in *S. acidocaldarius* DNA the proportion of A + T is more important than that of G + C (16), we used long-termination mix and short-termination mix together in the ratio of 4:1.

Purification of Reverse Gyrase and Peptide Sequencing. Microsequence was analyzed as follows: reverse gyrase was partially purified as in ref. 11, except that the order of phenyl-Sepharose and phosphocellulose chromatographies has been inverted. The protein was isolated and concentrated by SDS/PAGE (17), then transferred to a poly(vinylidene difluoride) membrane, according to the method proposed by Rasmussen *et al.* (18). Reverse gyrase was digested overnight with Endolys-C at 37°C in 100 mM Tris·HCl, pH 9.0/10% acetonitrile, and the released peptides were separated on a C_{18} HPLC column with a 0–60% acetonitrile gradient in 0.1% trifluoroacetic acid. Microsequencing of each peptide was done on a microsequencer A 470 (Applied Biosystems).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

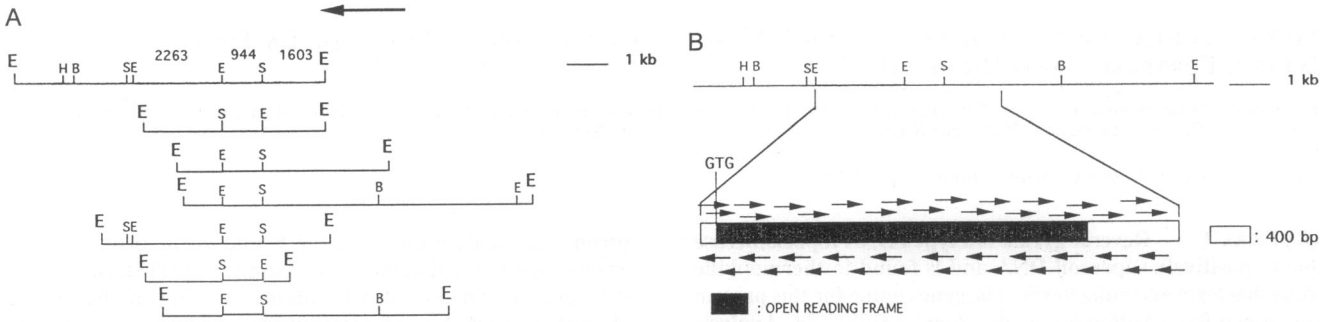
‡To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. L10651).

Analysis of Sequences. Alignments were realized by using the VIZZ program (19) and reorganized to reduce gaps.

RESULTS

To clone the gene coding for *S. acidocaldarius* reverse gyrase, we used polyclonal antibodies previously prepared against a pure preparation of the enzyme (11). We constructed a genomic library of *S. acidocaldarius* DNA in the expression vector λgt11 (see *Materials and Methods*). About 45,000 recombinant phages were screened with antibodies: 50 positive clones were purified, and their restriction maps were determined. All inserts share a common region, as checked



C

```
131
TAATACTTTGACTCATGTGACATATTACTTTAACCTTTTTAAATAAGTCTAACTGTAATAAAAT
195
ACTGTGTGCAGTCTTTGAGTGATATTCCACCCTCAATATATTTATTCTCATGTCCTAATTGTGG
1     Q S L S D I P P S I Y L F S C P N C G
259
GAGATCTATTTCGACATATAGGTTACTTCTAGGATCA.....
20    R S I S T Y R L L L G S VCNICLEEDKEYKNIGDLIKDIEKQGNL
60    IKLKDIQRVLDDYESFVSVFRRLLGFPPFGPQKSWIYRLLSGESFAIIAPPGLGKTTFGLISSIY
125   LYLRGKKSILVFPTKSLVRQAIDKLSSYIQNLAEIKENPPKVIYYYSGMSASERKEADEGLQ
187   SKTFDIFISTNRFLIDNIDQISSTSYQFLFVDDVDTALKSSKSAQAILKLLGFTPSDQDKIKESL
252   KKYRENTQKNEQNEYIFEEIDKIRKDRLASKTVIFSSATLNRSNPILTSLVGFKPGSSVIYIRK
316   VYDMYVKQPDKEQETFNLIKSLLHRLGDGGLIFVPVDKKQEYIKRLQSELSNEFNVAAITSTSAT
381   KIDDFANGEIDVLIGSATHYGILVRGLDLPWRVKYSIFVGIPKFKFRLGEKVNLLTLSRLLSLI
445   ALITKDQEVIYISRRVKDKIRRLSPAALTMLSVQAKEGKLEDSILLKAYDLLNKYLSNQNVLKRI
510   SEIGDFVLSPDNDILIPDYLTYVQASGRTSRIYAGDVTTGLSILLVDDFNLFRLLNKKLQYILDD
575   IQWRELDVEKWTAGDVEIKNLISKINEERNEISKLKNEGNVAPALQKVKTVLLVVESPNKAKT
638   ISSFFSRPSIRQIGNMRVYETVLGDKVLMVTASGGHVYDLTTQDMGIYGVDIMKQNSSLVFIPIY
703   NSIKKCENNHQFTDFFESNKCPRCMTTKVRYDSLKSINVLRNLAVEADEVLIGTDPDTEGEKIAW
768   DLYLALRPYNSNIRRAEFHEVTRKAILQAINQPREFNVNLVKSQLVRRIEDRWIGFKLSSILQTR
833   FWPEYCKSLSSNKQLNCNENKNLSAGRVQTPVLSWIVDRYTEYQRNKSRVYYGKIDQLQDIVIY
897   VPKQDGVRKNSKIVVVFNEINQLEEEFGPLPPYTTDTLLSDSNNFFGLSAPETMRIAQDLFELGL
962   ITYHRTDSNRISNTGISVAENYLKDVLGDKYTNIFKPRSWGDGGAHEGIRPTKPIDVEQLRLLIE
1027  EGELELAKRLTNNHFKVYDIIFRRFISSQIIPLKVRKEIVKIELYGENKKEKINSNQNIIEVI
1090  TGITLPGIDTEISKFAYVPVRNVSRSVAERLKELGRSIPTDFSIEISNSFIKSTVNLYTQADLVM
1155  EMKNKKIGRPSTYATIIGTILRRGYVLESLKTKKLIPTRLGVEVNKYLNENYGRFVSEDRTRKLL
                3914
      .......CAAGTTTATGAAGAAATAAACGAAATACGGTGAAAAAAATGCTCATAGG
1220  QLMDMVEAGQEKYEEVLKQ V Y E E I N E I R STOP
```

FIG. 1. Cloning and sequencing of reverse gyrase. (*A*) Restriction map of seven λgt11 positive clones. Orientation of the clones is represented as in λgt11 DNA. The cloning site is in boldface letters. Direction of transcription of the β-galactosidase gene is indicated by the arrow. Size of the three different restriction fragments containing the coding sequence of reverse gyrase is mentioned in base pairs. Restriction sites used for subcloning and DNA sequencing are indicated: E, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; S, *Sac* I. (*B*) Restriction map of *S. acidocaldarius* genomic DNA in the region of reverse gyrase gene. Comparison of this map with that of all positive clones indicates that no recombination appears to occur in these clones (data not shown). The region covering the three restriction fragments of 2263 bp, 944 bp, and 1603 bp is represented by the box. The shaded box represents the open reading frame corresponding to reverse gyrase gene. Nucleotide sequences, illustrated by arrows, were acquired with a series of primers. (*C*) Amino acid sequence of reverse gyrase deduced from the nucleotide sequence. Only the 5′ and 3′ ends of the nucleotide sequence are represented. The origin of "numerotation" (italics) is the left *Eco*RI site of the 2263-bp fragment. The putative promoter sequence (TTTAAA), the Shine–Dalgarno sequence, and the stop signal (TGA) are boxed. The nine Endolys-C-generated peptides are represented in boldface letters and underlined. The methionine corresponding to the first ATG and the peptidic sequence composing the zinc-finger motif are doubly underlined.

by cross-hybridization. Southern blot profiles and analysis of protein products from *E. coli* lysogens (data not shown) suggested that the gene was located within a region covering partly or totally three restriction fragments: *Eco*RI–*Eco*RI (2263 bp), *Eco*RI–*Sac* I (944 bp), and *Sac* I–*Eco*RI (1603 bp) (Fig. 1*A*). The strategy of sequencing is shown in Fig. 1*B*, and the results are shown in Fig. 1*C*.

An open reading frame of 3741 nt starting from a GTG was revealed. To confirm that the sequence shown in Fig. 1*C* truly encoded reverse gyrase, amino acid microsequencing of the protein purified from *S. acidocaldarius* was analyzed: we could not obtain the N-terminal sequence of the native protein that was possibly blocked; however, nine different peptides, totaling 80 amino acids, were generated by digestion of blotted reverse gyrase with Endolys-C protease. All peptide sequences confirm the data obtained from DNA sequencing (Fig. 1*C*).

Analysis of the 5′ region of the nucleotide sequence reveals a consensus motif present in many *Sulfolobus* promoters (consensus box A: TTTAAA) together with a Shine–Dalgarno sequence. The first ATG, located >500 nt downstream of box A, cannot be the initiator because it codes for an internal methionine belonging to the peptide VIYYYSGMSASERK, independently sequenced (Fig. 1*C*). However, previous works reported that GTG may be used along with ATG as initiation codon in many *Sulfolobus* genes (20, 21). Thus, we propose that the GTG found ≈25 nt downstream from box A serves as an initiation triplet. As for the gene encoding aspartate aminotransferase from *Sulfolobus solfataricus*, the GTG triplet is enclosed within the Shine–Dalgarno sequence (21).

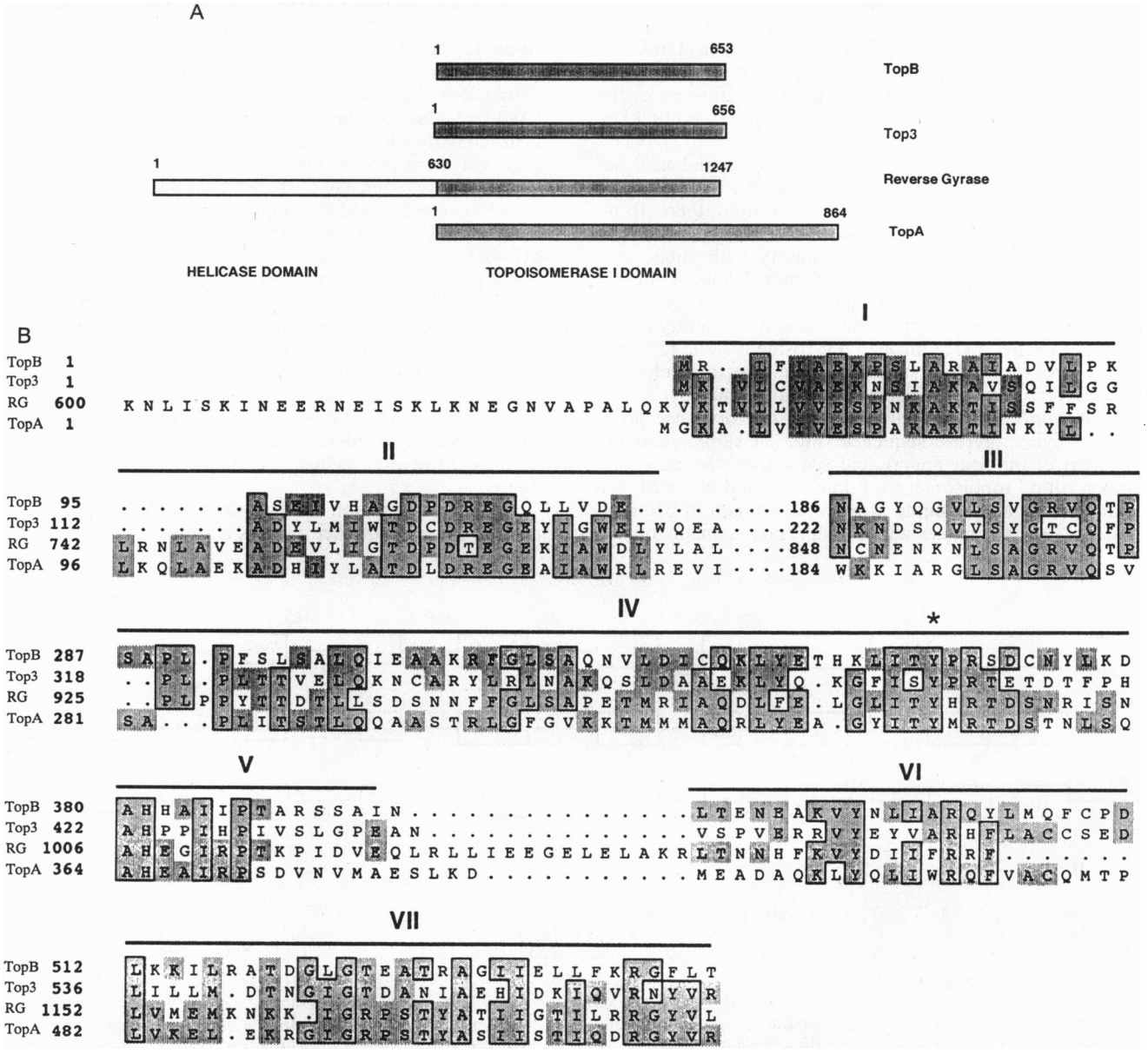


FIG. 2. Comparison of reverse gyrase with type I topoisomerases. (*A*) Schematic alignment of reverse gyrase with the other type I topoisomerases. TopA, *E. coli* topoisomerase I; TopB, *E. coli* topoisomerase III; Top3, *S. cerevisiae* topoisomerase III. Topoisomerase domains are represented by light gray boxes, and the helicase domain of reverse gyrase is represented by an open box. The C-terminal domain of reverse gyrase is clearly related to other eubacterial type I topoisomerases. (*B*) Alignment of the four amino acid sequences from TopB, Top3, TopA, and reverse gyrase (RG) genes in seven regions of similarity. Alignments were made by using the VIZZ program (19) and reorganized to reduce gaps. Identical amino acids were shaded. Strict identity for at least three of four amino acids is surrounded by a solid line. The putative tyrosine of the active site is indicated by a star. The percentage of identity between TopA and reverse gyrase varies between 30% and 65% in the different regions of similarity. There is no significant similarity with eukaryotic type I topoisomerases.

SDS/PAGE of the protein purified from *S. acidocaldarius* revealed a specific polypeptide with an apparent molecular mass of ≈130 kDa (11). Expression of reverse gyrase in *E. coli* lysogens results in a protein with a lower apparent molecular mass of 120 kDa (visualized by immunoblotting; data not shown). It is likely that, in *E. coli*, translation of the gene started from the ATG, producing a protein that lacked the first 172 amino acids. This fact would explain our negative tests for reverse gyrase activities with crude or semipurified extracts from lysogens.

The molecular mass of the 1247-amino acid polypeptide deduced from the nucleotide sequence is not 130 kDa but is, instead, 143 kDa. This discrepancy may be partly explained by the relatively low resolution of electrophoresis in this high-molecular-mass range. Moreover, by direct N-terminal microsequencing, we have shown that the protein migrating slightly below reverse gyrase in SDS/PAGE is the β chain of *S. acidocaldarius* RNA polymerase: this polypeptide, which migrated almost like the β-galactosidase marker (116 kDa), has a molecular mass of 126 kDa, based on the DNA sequence (22). When this value is used as a reference, the molecular mass of reverse gyrase would be ≈140 kDa, consistent with the sequence data.

Comparative analysis of reverse gyrase amino acid sequence with available sequence data indicates that, although the protein performs a reaction of gyration, there is no similarity with eubacterial gyrase or with any type II topoisomerase. In contrast, a clear similarity with eubacterial topoisomerase I appears in the C-terminal half of the sequence, consistent with the ability of reverse gyrase to transiently break single strands and remain covalently bound to the 5′ DNA end (12) through a tyrosine residue (23). We have compared the three available sequences that belong to the eubacterial type I topoisomerase family—namely, *E. coli topA* and *topB*, together with *S. cerevisiae top3* gene products, to reverse gyrase sequence. Fig. 2*A* summarizes the alignment of the four amino acid sequences: the canonical "prokaryotic" topoisomerase I domain would be ≈620–650 residues, with an extruding N-terminal domain of 620 residues for reverse gyrase and a C-terminal domain of 200 residues for TopA protein. Fig. 2*B* shows some of the regions of similarity shared by the four sequences. Unexpectedly, the highest score of similarity is found between TopA and reverse gyrase, with 55%, 44%, and 64% identity in regions I, IV, and VII, respectively. By comparison, TopA and TopB show only 30%, 42%, and 35% identity in the same regions. The similarity observed around the active tyrosine of TopA (region IV) suggests that the active site of reverse gyrase is Tyr-964. Finally, the large extent of similarity of reverse gyrase in the vicinity of the amino end of TopA (region I) allows us to propose that the topoisomerase I domain of reverse gyrase starts about residue 620.

Contrasting to the C-terminal half, the N-terminal part of the reverse gyrase sequence exhibits no similarity to any topoisomerase I or II. This domain contains a putative ATP-binding site common to many ATP-binding proteins (24–26). This motif (AXXGXGKT) was intriguing because it is not found in the other ATP-dependent topoisomerases (type II topoisomerases). Further comparison of reverse gyrase sequence with those of ATP-binding protein families shows that the highest similarity is found with the DEAD (Asp-Glu-Ala-Asp) family of helicases (24): eight motifs characteristic of this family were found in the N-terminal region of our sequence (Fig. 3*A*). Two types of considerations support the relevance of the helicase motifs in reverse gyrase: (*i*) we have compared the sequences of the eight motifs of a canonical RNA helicase, the translation initiation factor eIF-4AI (27), with those of a DNA helicase involved in bacterial primosome function PriA (28, 29) and with the N-terminal half of reverse gyrase; Fig. 3*B* shows that, in most cases, the similarity extends outside the box, and this is especially true for PTRELA, TPGR, and HRIGR boxes; (*ii*) the various motifs are not distributed randomly within the reverse gyrase sequence but appear in the same order as in eIF-4AI and PriA and are separated approximately by the same number of residues, except for the last box (HRIGR) in the case of reverse gyrase and ARGXD box for PriA (Fig. 3*A*).

Finally, examination of the position of cysteine residues in reverse gyrase sequence reveals a series of 22 amino acids

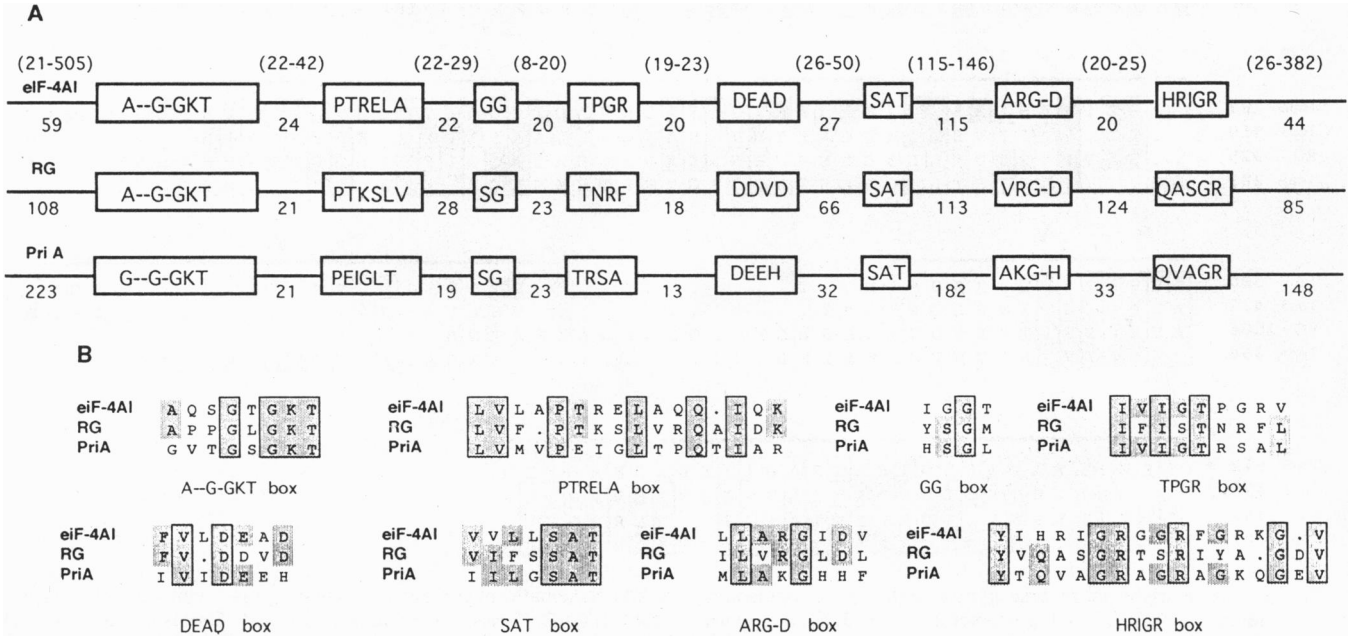


FIG. 3. Helicase motifs in reverse gyrase. Comparison of reverse gyrase (RG) with RNA helicase eIF-4AI and DNA helicase PriA. (*A*) Succession of the eight helicase motifs in eIF-4AI, reverse gyrase, and PriA. The nomenclature of the motifs is taken from RNA helicases (24). The number of amino acid residues between each box is indicated, and its variation within the DEAD family is under brackets. (*B*) Comparison of the helicase motifs in eIF-4AI, reverse gyrase, and PriA. Identical residues are shaded. Strict identities between the three sequences are surrounded by a solid line.

that fit the consensus for a metal binding or "zinc-finger" domain (30) identified in a number of prokaryotic or eukaryotic proteins. The putative reverse gyrase zinc-finger domain extends from amino acid 15 to 36 and contains the sequence Cys-Xaa_2-Cys-Xaa_{14}-Cys-Xaa_2-Cys (Fig. 1*C*). This motif is similar to that found within the primase domain of the T7 gene *4* primase–helicase protein (31) and identical to one of the motifs described in PriA (28, 29).

DISCUSSION

The finding of helicase motifs in reverse gyrase is consistent with some properties of the enzyme that are reminiscent of helicases: (*i*) reverse gyrase binds to DNA duplexes with a preference for single-stranded DNA (32, 33); (*ii*) a number of enzyme molecules are bound to DNA; and (*iii*) the binding induces a change in DNA structure that may be interpreted as left-handed DNA wrapping or, more likely, as DNA unwinding (12). Even if reverse gyrase contains a functional helicase domain, it is possible that the complete enzyme does not exhibit helicase activity. If so, it would be interesting to express separately the two domains of the protein, to see whether the N-terminal domain presents an intrinsic helicase activity while the C-terminal domain exhibits an ATP-independent topoisomerase.

Comparison of our previous mechanistic data concerning reverse gyrase with the present sequence information led us to propose a model of positive gyration. This model is reminiscent of the early models for gyrase (34) and is based on the theory of twin DNA supercoiling during transcription (35). As for RNA polymerase, the progression of a helicase through a circular DNA duplex produces two waves of stress into DNA: an overwound (sc^+) DNA region downstream from the helicase and an underwound region (sc^-) upstream of the helicase. Specific relaxation of the negatively supercoiled (sc^-) region by the topoisomerase I domain of reverse gyrase would increase the DNA linking number and result in net positive supercoiling of the whole DNA circle (see Fig. 4). Indeed, production of positive DNA supercoils has been reported *in vitro* with simian virus 40 large T antigen (36) and, more recently, by an artificial mixture containing ATP, extracts from *top1*$^-$ strains of *Saccharomyces cerevisiae*, and bacterial topoisomerase I (37).

The finding of a bifunctional structure in reverse gyrase has several implications of wide interest; indeed, an association of helicase and topoisomerase activities is of potential importance in many enzymatic mechanisms where DNA is involved: replication, transcription, or repair/recombination processes. Accordingly, in addition to its putative role in DNA stabilization at high temperature, reverse gyrase could be involved in one or several of these mechanisms.

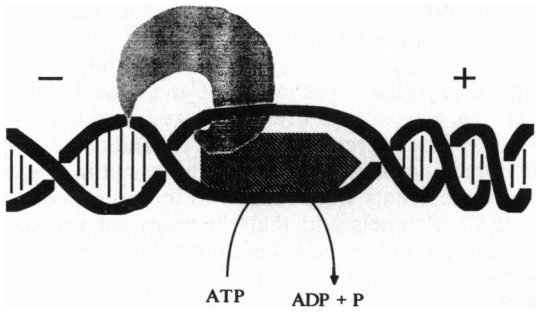


FIG. 4. Tentative model of reverse gyration. The helicase domain is hatched, and its arrowhead indicates the movement of the enzyme along the DNA double helix. Topoisomerase domain is represented by a light gray crescent. − and + indicate, respectively, underwound (sc^-) and overwound (sc^+) domains in the DNA molecule.

The cloning and sequencing of reverse gyrase, an archaebacterial topoisomerase I, opens the field of phylogenetic studies. Comparison with the other available topoisomerase sequences led us to define common motifs useful to identify the homologous genes in the various archaebacterial or eubacterial branches. Finally, the finding of a composite structure for reverse gyrase addresses a fundamental question about the emergence of such a gene: is it an ancestral gene, further split in the course of evolution into a helicase and a topoisomerase, or, conversely, was it formed by gene fusion or recombination?

We thank J. P. Lecaer, Institut de Neurobiologie Centre National de la Recherche Scientifique, Gif sur Yvette, and J. D'Alayer, Institut Pasteur, Paris, for advice and help in protein microsequencing. We also thank the Centre Universitaire d'Information à Orientation Biomédicale (Paris) for their data. This work has been partly supported by a grant "Evolution" of the Ministère de l'Education Nationale (France).

1. Wang, J. C. (1987) *Biochim. Biophys. Acta* **909**, 1–9.
2. Wang, J. C., Caron, P. R. & Kim, R. A. (1990) *Cell* **62**, 403–406.
3. Gellert, M., Mizuuchi, K., Odea, M. H. & Nash, H. A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3872–3876.
4. Reece, R. J. & Maxwell, A. (1991) *Crit. Rev. Biochem. Mol. Biol.* **26**, 335–375.
5. Drlica, K. (1990) *Trends Genet.* **6**, 433–437.
6. Kikuchi, A. & Asai, K. (1984) *Nature (London)* **309**, 677–681.
7. Bouthier de la Tour, C., Portemer, C., Nadal, M., Stetter, K. O., Forterre, P. & Duguet, M. (1990) *J. Bacteriol.* **172**, 6803–6808.
8. Bouthier de la Tour, C., Portemer, C., Nadal, M., Huber, R., Forterre, P. & Duguet, M. (1991) *J. Bacteriol.* **173**, 3921–3923.
9. Forterre, P., Mirambeau, G., Jaxel, C., Nadal, M. & Duguet, M. (1985) *EMBO J.* **4**, 2123–2128.
10. Nadal, M., Mirambeau, G., Forterre, P., Reiter, W.-D. & Duguet, M. (1986) *Nature (London)* **321**, 256–258.
11. Nadal, M., Jaxel, C., Portemer, C., Forterre, P., Mirambeau, G. & Duguet, M. (1988) *Biochemistry* **27**, 9102–9108.
12. Jaxel, C., Nadal, M., Mirambeau, G., Forterre, P., Takahashi, M. & Duguet, M. (1989) *EMBO J.* **8**, 3135–3139.
13. Depew, R. E., Liu, L. F. & Wang, J. C. (1978) *J. Biol. Chem.* **253**, 511–518.
14. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, p. 49.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Zillig, W., Stetter, K. O. & Janekovic, D. (1979) *Eur. J. Biochem* **96**, 597–604.
17. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
18. Rasmussen, H. H., Van Damme, J., Baum, G., Puype, M., Gesser, B., Celis, J. E. & Vandekerckhove, J. (1991) in *Methods in Protein Sequence Analysis*, eds, Jörnval, H., Höög, J. O. & Gustavsson, A. M. (Birkhäuser, Basel), p. 103.
19. Saurin, W. & Marliere, P. (1986) *C.R. Acad. Sci. Ser. 3* **13**, 541–546.
20. Bokranz, M. & Klein, A. (1987) *Nucleic Acids Res.* **15**, 4350–4351.
21. Cubellis, M. V., Rozzo, C., Nitte, G., Arnone, M. I., Marino, G. & Sannia, G. (1989) *Eur. J. Biochem.* **186**, 375–381.
22. Pühler, G., Lottspeich, F. & Zillig, W. (1989) *Nucleic Acids Res.* **17**, 4517–4534.
23. Nadal, M. (1990) Ph.D. thesis (University of Paris), p. 98.
24. Schmid, S. R. & Linder, P. (1992) *Mol. Microbiol.* **6**, 283–292.
25. Gorbalenya, A. E., Koonin, A. E., Donchenko, A. P. & Blinov, V. M. (1989) *Nucleic Acids Res.* **17**, 4713–4730.
26. Higgins, C. F., Gallagher, M. P., Mimmarck, M. L. & Pearce, S. R. (1988) *BioEssays* **8**, 111–116.
27. Nielsen, P. J. & Trachsel, H. (1988) *EMBO J.* **7**, 2097–2105.
28. Nurse, P., DiGate, R. J., Zavitz, K. H. & Marians, K. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4615–4619.
29. Lee, E. H., Masai, H., Allen, G. C., Jr., & Kornberg, A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4620–4624.
30. Berg, J. M. (1986) *Science* **232**, 485–487.
31. Berstein, J. A. & Richardson, C. C. (1984) *Proc. Natl. Acad. Sci. USA* **85**, 396–400.
32. Shibata, T., Nakasu, S., Yasui, K. & Kikuchi, A. (1987) *J. Biol. Chem.* **263**, 10419–10421.
33. Slesarev, A. I. & Kozyavkin, S. A. (1990) *J. Biomol. Struct. Dyn.* **7**, 935–942.
34. Gellert, M. (1981) *Annu. Rev. Biochem.* **50**, 879–910.
35. Liu, L. F. & Wang, J. C. (1987) *Proc. Natl. Acad. Sci. USA* **87**, 7024–7027.
36. Zhang, H., Jessee, C. B. & Liu, L. F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9078–9082.
37. Koo, H. B., Lou, K., Wu, H. Y. & Liu, L. F. (1992) *Nucleic Acids Res.* **20**, 5067–5072.