JOURNAL OF BACTERIOLOGY, Mar. 1997, p. 1721–1726 Vol. 179, No. 5
0021-9193/97/$04.00+0
Copyright © 1997, American Society for Microbiology

Characterization of the Reverse Gyrase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*†

KIM M. BORGES,[1,2] AGNES BERGERAT,[3] ALLISON M. BOGERT,[1] JOCELYNE DIRUGGIERO,[1]* PATRICK FORTERRE,[3] AND FRANK T. ROBB[1]

Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, Maryland[1]; Institut de Génétique et Microbiologie, CNRS, Université Paris-Sud, Paris, France[3]; and Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut[2]

Received 23 August 1996/Accepted 17 December 1996

The reverse gyrase gene *rgy* from the hyperthermophilic archaeon *Pyrococcus furiosus* was cloned and sequenced. The gene is 3,642 bp (1,214 amino acids) in length. The deduced amino acid sequence has relatively high similarity to the sequences of the *Methanococcus jannaschii* reverse gyrase (48% overall identity), the *Sulfolobus acidocaldarius* reverse gyrase (41% identity), and the *Methanopyrus kandleri* reverse gyrase (37% identity). The *P. furiosus* reverse gyrase is a monomeric protein, containing a helicase-like module and a type I topoisomerase module, which resembles the enzyme from *S. acidocaldarius* more than that from *M. kandleri*, a heterodimeric protein encoded by two separate genes. The control region of the *P. furiosus rgy* gene contains a typical archaeal putative box A promoter element which is located at position −26 from the transcription start identified by primer extension experiments. The initiating ATG codon is preceded by a possible prokaryote-type ribosome-binding site. Purified *P. furiosus* reverse gyrase has a sedimentation coefficient of 6S, suggesting a monomeric structure for the native protein. The enzyme is a single polypeptide with an apparent molecular mass of 120 kDa, in agreement with the gene structure. The sequence of the N terminus of the protein corresponded to the deduced amino acid sequence. Phylogenetic analysis indicates that all known reverse gyrase topoisomerase modules form a subgroup inside subfamily 1A of type I DNA topoisomerases (sensu Wang [J. C. Wang, Annu. Rev. Biochem. 65:635–692, 1996]). Our results suggest that the fusion between the topoisomerase and helicase modules of reverse gyrase occurred before the divergence of the two archaeal phyla, *Crenoarchaeota* and *Euryarchaeota*.

The occurrence and recent isolation of hyperthermophilic microorganisms, which grow above 90°C, have introduced new concepts into the field of high-temperature biochemistry. Reverse gyrase is an enzyme unique to hyperthermophilic prokaryotes (members of domains *Archaea* and *Bacteria*) and was first discovered in the archaeon *Sulfolobus acidocaldarius* (22). This enzyme catalyzes the ATP-dependent introduction of positive supercoils into DNA (13). Since positive supercoiling corresponds to an increase in the number of topological links between the two strands of a closed DNA molecule (linking number [Lk]), reverse gyrase could help to counteract the effects of high temperature on DNA structure (17). Indeed, plasmids from hyperthermophilic archaea exhibit higher Lks than plasmids from mesophilic archaea or bacteria (8).

In marked contrast to gyrase, which is a type II DNA topoisomerase, reverse gyrase is a type I DNA topoisomerase (19), an enzyme which changes the linking number by increments of 1, making transient single-stranded breaks in DNA. This was surprising at first, since all type I topoisomerase enzymes previously discovered were ATP independent. However, the cloning and sequencing of the *S. acidocaldarius* gene encoding reverse gyrase has resolved this paradox, revealing an ATP-independent type I topoisomerase module and an ATP-dependent helicase-like module associated in a single polypeptide (9). It is believed that the helicase module causes positive supercoils to accumulate ahead of the migrating reverse gyrase, and the topoisomerase module releases the negative supercoils behind the enzyme, resulting in accumulation of positive supercoils (13).

Reverse gyrases from several archaea and one bacterium have been purified and characterized (2, 19, 23, 28, 31). All of the enzymes but one consist of one polypeptide with a size similar to that of the *S. acidocaldarius* enzyme. The exception is the reverse gyrase from the hyperthermophilic archaeon *Methanopyrus kandleri* (a methanogen), which is composed of two polypeptides (24). Krah and coworkers (24) suggested that this dimeric structure could be a primitive feature of reverse gyrase, since *M. kandleri* is the first lineage branching in the euryarchaeal kingdom, at the base of the archaeal 16S rRNA tree (29). The two genes for the *M. kandleri* reverse gyrase, *rgyA* and *rgyB*, have recently been cloned and sequenced (24). The helicase-like module corresponds to the N-terminal half of *rgyA*, whereas the topoisomerase module is divided between the C-terminal part of *rgyA* and the entire *rgyB* gene (24). However, the sequence of another euryarchaeal *rgy* gene, from *Methanococcus jannaschii*, has been recently published (7), predicting a single polypeptide structure of the enzyme, similar to that of *S. acidocaldarius.*

We describe here the cloning and sequencing of the gene encoding reverse gyrase from the hyperthermophile *Pyrococcus furiosus*, the activity of the corresponding enzyme, and the phylogenetic implications of the discovery of this new gene. *P. furiosus* is a strictly anaerobic heterotrophic archaeon which grows optimally at 100°C (15) and branches after *M. kandleri* in the 16S rRNA tree of the euryarchaeal kingdom (29). Our results suggest that the ancestral archaea had a monomeric reverse gyrase.

MATERIALS AND METHODS

Materials. Restriction enzymes were from Promega (Madison, Wis.). All other chemicals were of the highest purity available and were used without further purification.

* Corresponding author. Phone: (410) 234-8890. Fax: (410) 234-8896. E-mail: Jocelyne@UMBI.UMD.EDU.

† Contribution number 280 from the Center of Marine Biotechnology.

Cloning and sequencing. A cDNA clone from *P. furiosus* was sequenced and shown to have high homology to the reverse gyrase gene from *S. acidocaldarius* by BLASTX analysis (4). PCR primers for the gene were designed from the 500-bp cDNA sequence and used to amplify a 450-bp reverse gyrase fragment from *P. furiosus* genomic DNA. A lambda Zap mixed partial genomic library of *P. furiosus* (4) was screened by using this PCR fragment labeled with [α-^{32}P] dATP by random priming. Four positive plaques were rescued into the pBluescript KS$^+$ plasmid (Stratagene, La Jolla, Calif.) and purified by cesium chloride gradient sedimentation (3). The positive clones were sequenced by the dideoxy-chain termination method with T3 and T7 primers, as well as primer walking methodology (3). Two overlapping clones, pKMB12 and pKMB14, together contain the complete sequence of the reverse gyrase gene.

Primer extension analysis. Total RNA was isolated as described by DiRuggiero and Robb (11). Primer extension (3) was performed with a 5′-end ^{32}P-labeled synthetic oligonucleotide, 5′ CATCCAAACATGCATCAC 3′, which is complementary to positions +80 to +97 on the *rgy* gene. Three reaction mixtures containing 2.5, 5, and 10 μg of total RNA were used. Extension products were separated by 7 M urea–8% polyacrylamide gel electrophoresis and analyzed by using a PhosphorImager and ImageQuant analysis software (Molecular Dynamics, Sunnyvale, Calif.).

Reverse gyrase purification. Purification of *P. furiosus* reverse gyrase was performed by phenyl-Sepharose and heparin-Sepharose chromatography as described by Nadal et al. (28), followed by sucrose gradient centrifugation. A partially purified fraction from phenyl-Sepharose and heparin-Sepharose columns was loaded onto a 5 to 20% sucrose gradient in purification buffer (50 mM potassium phosphate, 1 mM EDTA, 1 mM dithiothreitol, 0.01% Triton X-100, 0.25 M KCl) and centrifuged at 40,000 rpm for 15 h in an SW41 rotor. Fractions of 0.4 ml were collected. Reverse gyrase was detected by determination of relaxation activity in the presence of ATP and by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (see below). Protein markers centrifuged in different tubes at the same time were detected by the Bradford assay (5).

***P. furiosus* reverse gyrase assays.** The standard reaction mixture (20 μl) contained 35 mM Tris-HCl (pH 8.8, 25°C), 0.1 mM Na_2-EDTA, 30 mM $MgCl_2$, 2 mM dithiothreitol, 1 mM ATP, 0.2 μg of plasmid pTZ18 DNA, and a 2-μl enzyme sample. Each reaction mixture was incubated for 2 min at 90°C, and then the reaction was stopped and the Lk was analyzed by electrophoresis as previously described (21). One unit of *P. furiosus* reverse gyrase activity is defined as the amount of enzyme required for positive supercoiling of 0.2 μg of negatively supercoiled plasmid pTZ18 DNA under the standard conditions described above.

Protein gel electrophoresis. SDS–8% PAGE was performed as described by Laemmli (26). Silver staining was used to visualize protein bands. The molecular mass markers were myosin (205 kDa), β-galactosidase (116 kDa), phosphorylase *b* (97.4 kDa), and bovine serum albumin (66 kDa).

Peptide sequencing. Microsequencing was performed with 2 μg of reverse gyrase obtained from the sucrose gradient. The 120-kDa protein was separated from other contaminant proteins by SDS-PAGE and then transferred to a polyvinylidene difluoride membrane by the method proposed by Kyhse-Anderson (25). Proteins were visualized by staining with Coomassie blue R-250 (0.1% in 50% methanol) and then destaining with several washes with 50% methanol–10% acetic acid–40% deionized, distilled water. Microsequencing was performed by cyclic Edman degradation.

Phylogenetic analysis. The three reverse gyrase topoisomerase domains and all available complete type I DNA topoisomerases of subfamily 1A were aligned by using CLUSTAL and corrected by hand. Distance and parsimony trees were constructed from a compact alignment of clearly homologous regions (336 amino acids) by using three different programs from the PHYLIP package (version 3.4): Fitch and Margoliash (distance method), neighbor joining (distance method), and the PROTPARS program (parsimony).

Nucleotide sequence accession number. The sequence reported in this paper has been deposited in the GenBank database and assigned accession number U66557.

RESULTS

Cloning and sequencing of the *P. furiosus* reverse gyrase gene. A fragment of the *P. furiosus* reverse gyrase gene had been previously identified by random sequencing of a cDNA clone (4). A PCR fragment derived from this cDNA clone was used to screen a lambda Zap library of *P. furiosus* genomic DNA. Several clones were identified by plaque hybridization and rescued into pBluescript KS$^+$. The *P. furiosus* reverse gyrase gene was sequenced by using two of these subclones, pKMB12 and pKMB14, which contain the entire *rgy* gene. Hybridization of genomic DNA from *P. furiosus*, digested with several restriction enzymes (data not shown), with either clone showed a single band corresponding to a single copy of the reverse gyrase gene in the archaeal genome.

The complete nucleotide sequence of *P. furiosus* reverse gyrase gene *rgy* consists of a single open reading frame of 3,642 nucleotides. The deduced amino acid sequence of the reverse gyrase is composed of 1,214 residues, similar in length to the *S. acidocaldarius* protein, which is 1,247 amino acids (9). The deduced *P. furiosus* reverse gyrase amino acid sequence is 48, 41, and 37% identical to the deduced reverse gyrase sequences of *M. jannaschii*, *S. acidocaldarius*, and *M. kandleri*, respectively. The *P. furiosus* protein has an amino-terminal helicase-like module and a C-terminal topoisomerase I module, so the enzyme possesses the hallmark modular structure of a reverse gyrase. The two modules are arranged within a single polypeptide structure, as in the *S. acidocaldarius* reverse gyrase, rather than the heterodimeric structure of the *M. kandleri* enzyme (Fig. 1).

Comparative analysis of amino acid identities for each of the modules among the three enzymes indicated that *P. furiosus* reverse gyrase is more closely related to the *M. jannaschii* enzyme than those from the *S. acidocaldarius* and *M. kandleri* enzymes, especially in the topoisomerase domain, whereas the *S. acidocaldarius* and *M. kandleri* reverse gyrases are most distantly related to each other (Table 1). This is especially clear when the identities are calculated based on the fraction of variable amino acid residues. The helicase modules have evolved more rapidly than the topoisomerase modules, with 32% identity between *P. furiosus* and *M. jannaschii*, 23% identity between *P. furiosus* and *S. acidocaldarius*, and 15% identity between *P. furiosus* and *M. kandleri.*

Figure 1 compares the structures of the three archaeal reverse gyrase genes. The helicase module of *P. furiosus* reverse gyrase possesses all of the helicase-like motifs (a to h) present in the other reverse gyrases in the same arrangement, as well as the conserved zinc finger motif with four cysteines near the amino terminus. Interestingly, 4 of the 10 insertions of at least 4 nucleotides present in the *M. kandleri* helicase-like module (compared to the corresponding region in the *S. acidocaldarius* enzyme) are also present in the *P. furiosus* gene, in agreement with an intermediate position of the *P. furiosus* enzyme between the *S. acidocaldarius* and *M. kandleri* reverse gyrases. The topoisomerase domain of the *P. furiosus* enzyme also possesses all of the motifs (I to VII) present in the other two reverse gyrases, again in the same arrangement. The putative active-site tyrosine is located at position 955. Two large insertions present in the *M. kandleri* enzyme are missing in *P. furiosus* reverse gyrase, in particular, a C-terminal extension in *rgyA* that could be involved in the stabilization of the *M. kandleri* reverse gyrase dimeric structure (24).

The start site of the reverse gyrase transcript was at position −14 upstream from the putative ATG initiation site (Fig. 2). Transcription starts at a purine following a pyrimidine; this is typical of archaeal promoters, which usually start at a pyrimidine-purine dinucleotide motif (30). A putative box A promoter element (TTTTAA) at position −26 from the transcription start site was detected (Fig. 3). This box A resembles the eucaryotic TATA box and conforms to the archaeal consensus sequence TTTAA/TA, which was determined by analysis of over 80 archaeal promoters (30). The box A sequence is located within the proposed consensus distance (30). The box B sequence, described for methanogens by Wich et al. (33) and later modified by Reiter et al. (30), includes the motif TGCA located at the transcription start site. No box B sequence was found in the *P. furiosus rgy* control region, as has been the case for several other genes from hyperthermophilic archaea (12). Immediately upstream of the open reading frame at position

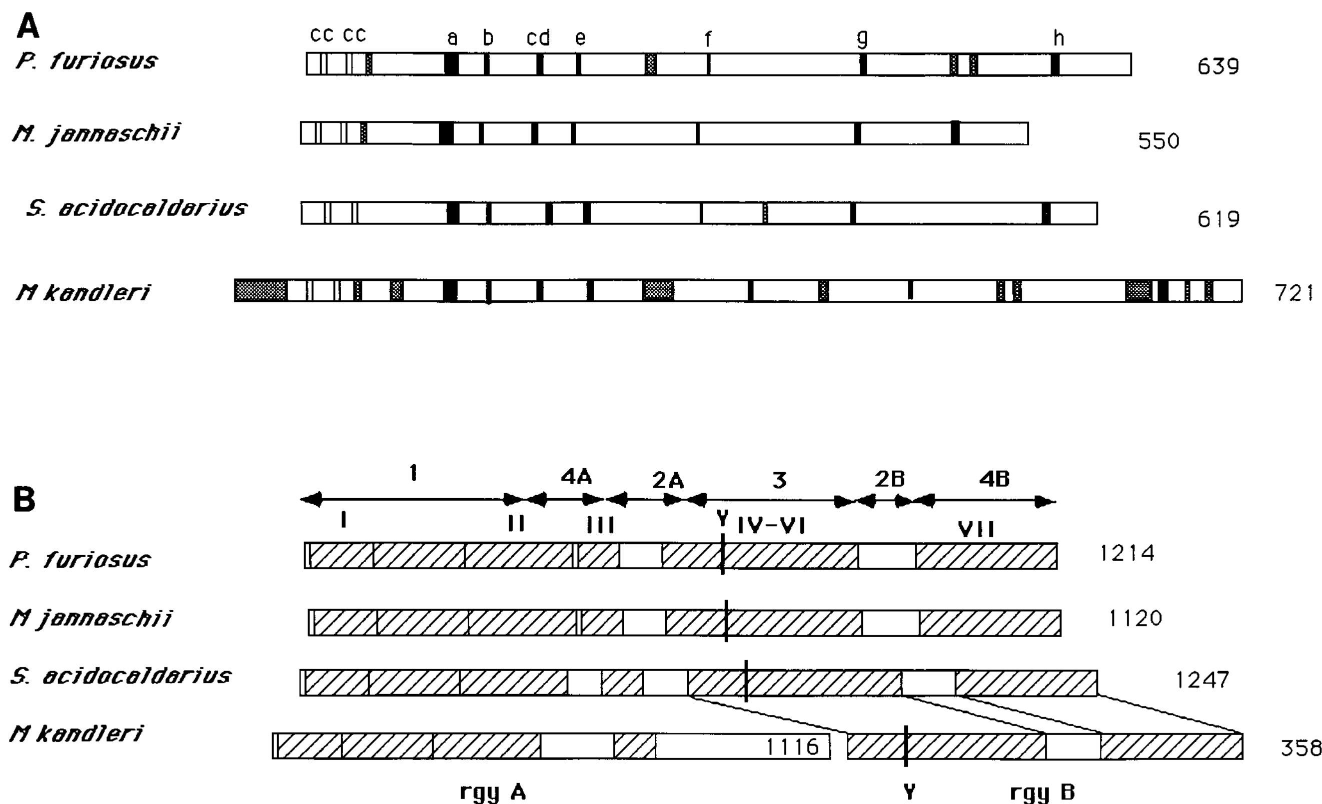

FIG. 1. Schematic alignments of helicase-like and topoisomerase modules of archaeal reverse gyrases. (A) Helicase-like modules. c, conserved cysteines of a putative zinc finger motif (13); black boxes, helicase signatures (9); grey boxes, insertions of at least 4 nucleotides (compared to the *S. acidocaldarius* sequence). (B) Topoisomerase modules. Hatched boxes, regions that can be aligned with precision; arrows and arabic numbers, domain structures deduced from the three-dimensional structure of *E. coli* TopA (27); roman numbers, conserved regions in enzymes of the TopIA family, according to Confalonieri et al. (9); Y, active-site tyrosine.

−6, a putative ribosome-binding site, GGAGA, has been found (Fig. 3), which corresponds to many Shine-Dalgarno sequences from archaea (6).

Size and activity of the *P. furiosus* reverse gyrase enzyme. To identify the protein encoded by our cloned gene, we purified the *P. furiosus* reverse gyrase to near homogeneity on the basis of the ATP-dependent introduction of positive supercoiling in a negatively supercoiled plasmid. During sucrose gradient centrifugation in 250 mM KCl, this reverse gyrase activity cosedimented with a polypeptide of 120 kDa in fraction 16 (Fig. 4). The sedimentation coefficient of this polypeptide was estimated to be 6S, suggesting that this protein is a monomer. The sucrose gradient fractions were concentrated 10-fold and incubated at 95°C in the presence of ATP with negatively supercoiled plasmid DNA for various times (Fig. 5). This fraction was shown to relax negative superturns and to introduce pos-

TABLE 1. Percentages of amino acid identity between topoisomerase[a] and helicase-like modules[b] of archaeal reverse gyrases[c]

Species	% Identity with:			
	Pfu	Mja	Sac	Mka
Pfu	100	36	33	24
Mja	32	100	26	20
Sac	23	20	100	15
Mka	15	21	15	100

[a] Top right.
[b] Bottom left.
[c] Percentages of amino acid identity were determined in each module by considering only regions homologous among all four proteins. The identities were calculated by taking into account only variable positions, i.e., subtracting amino acids identical in the four sequences. Pfu, *P. furiosus*; Mja, *M. jannaschii*; Sac, *S. acidocaldarius*; Mka, *M. kandleri.*

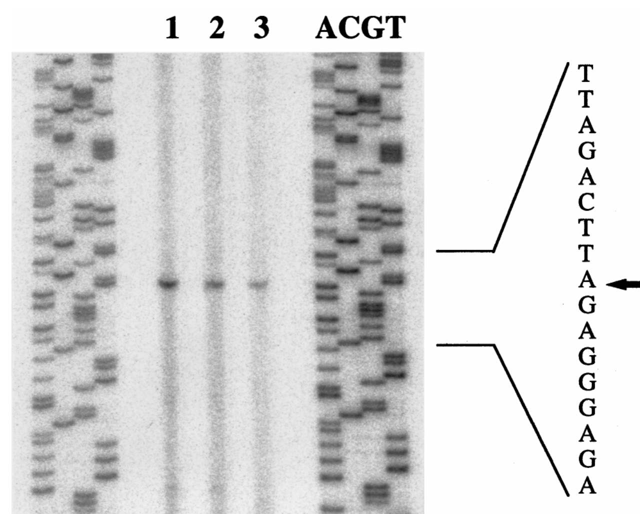

FIG. 2. Primer extension analysis of the *P. furiosus rgy* transcript. Lanes 1, 2, and 3 contained total RNA (10, 5, and 2.5 μg, respectively). Also shown are the sequencing ladders obtained with the primer used in the primer extension experiment (ACGT). The relevant nucleotide sequence is shown, and the transcription start is indicated by the arrow.

TATGAATTGCAAGGAATATAATACTATTAAGAAAAAGAGATAGAAATAGCTTATTTCTAAAATACGGGGA 70
GCTAGAAACCTTTTTAAGAACAAAAGTAGTGAACTTAGACTTAGAG**GGAGA**GCATTATG AAG GCA 135
M K A 3
ATA TAT AGG GAT ATG TGT CCC AAT TGT AGA GGG GCA ATA ACT GAT GAA AGA 186
I Y R D M C P N C R G A I T D E R 20
CTT GCA GCA AAA AAT CCC TGT GAT GCA TGT TTG GAT GAA CCA ATA AGT ATG 237
L A A K N P C D A C L D E P I S M 37
GAT GAT TAC TTT CAA CTT GTA ACT GCC ATT AGG AAG GCT CTC AAA TTA AAG 288
D D Y F Q L V T A I R K A L K L K 54
GGA GTG CTT AAA GAG TGG GAG GAA ATT TAT GCT CTT AAC AAA GAA GTT AAA 339
G V L K E W E E I Y A L N K E V K 71
GAA ATT GAA GAA CTC TTT AAA AAG GCA ACT GGA TTT ACA TTT TGG AGC GCC 390
E I E E L F K K A T G F T F W S A 88
CAG CGC TCC TGG GTG AAG AGA ATA ATA AAA GGA AAG AGC TTT TCT ATT ATA 441
Q R S W V K R I I K G K S F S I I 105
GCT CCA ACA GGA ATG GGT AAG AGT ACC TTC GGA GCT TTT ATG TCC ATT TAC 492
A P T G M G K S T F G A F M S I Y 122
TTT GCA TTA AAA GGA AAG AAG AGT TAT ATC GTA GTT CCT ACA ACT CCT CTA 543
F A L K G K K S Y I V V P T T P L 139
GTT GTC CAA ACT GTA AAG AAA ATT AAA GCG ATG ATG GAA AAA GCG GAA GTT 594
V V Q T V K K I K A M M E K A E V 156
AAA GTT AAC CTC GTT TAT TAC CAT GGA AAC CTA AAG AAG AAA GAA AAA GAT 645
K V N L V Y Y H G N L K K K E K D 173
GAA GCC CTC GAA AAA ATA AAA TCA GGA AAC TTT GAT ATC TTG GTA ACT TCA 696
E A L E K I K S G N F D I L V T S 190
AGC CAA TTC TTG GCC ACA AGG TTT GAT GAA TTG CTA AAG GAT AAG CGT ATA 747
S Q F L A T R F D E L L K D K R I 207
GAC TTC ATG TTT GTT GAC GAT GTG GAT GCA TTT CTT AAG GCA AGC AAA AAC 798
D F M F V D D V D A F L K A S K N 224
ATT GAC CGC TCT TTA ATG ATG CTT GGA TTT AAT GAA GAG ATA ATC AAA AAG 849
I D R S L M M L G F N E E I I K K 241
GCC TGG GAA ATT ATA AAG CTC AAG AAA CAA CTC GCC AAG CTT CTT CAA AAT 900
A W E I I K L K K Q L A K L L Q N 258
GAA AGT GGT AAT GAA GAA GTG GAA AAG TTA AAT AGG GAG ATA GAA AGA CTT 951
E S G N E E V E K L N R E I E R L 275
GAG AGA GAA ATT GAC AGG TAC AAG AAA GAA AAT AAA ATA GGA ATC TTG ATC 1002
E R E I D R Y K K E N K I G I L I 292
GTT GCG TCT GCT ACG GGA AGT GCC AGG GGA GAT AGA ATA AAA CTA TAC AGA 1053
V A S A T G S A R G D R I K L Y R 309
GAG CTC CTT GGA TTT GAA GTC GGA AGT GGA AGG AGC GTT CTG AGA AAT ATT 1104
E L L G F E V G S G R S V L R N I 326
GTT GAC ACT TAC ATT ATT CCC GAG AAA AGC ATG GAG GAA CAT GTC GAA GAA 1155
V D T Y I I P E K S M E E H V E E 343
CTA CTA AAG ACC CTT GGA AGA GGT GGC TTG ATT TTC GTT CCA ATT GAT AAG 1206
L L K T L G R G G L I F V P I D K 360
GGA ATT GAG TAT GCC GAA CAA CTT ACA AAT TAC CTA AAG TCC AAA GGA TTT 1257
G I E Y A E Q L T N Y L K S K G F 377
AAA GTT GAG CTT GTC TCA GCT AGA GAC AAG AAA GGA CTA GAA CTT TTC GAA 1308
K V E L V S A R D K K G L E L F E 394
AAA GAA GAA GTC GAT TAT CTC GTT GGA GTG GCC ACA TAT TAT GGG ACA ATT 1359
K E E V D Y L V G V A T Y Y G T I 411

GTG AGA GGT CTT GAT CTG CCT CAT TTA GTG AGG TTT GCA ATC TTT ACA GGA 1410
V R G L D L P H L V R F A I F T G 428
GTG CCA AAA TTC AGG TTC TCC CTC GAT TTA GAG CAA CCC ACA ATA TAT AGA 1461
V P K F R F S L D L E Q P T I Y R 445
GTC CTC GGA TTA ATG AGT GAA ATA CTG GAG TTT TTA CCT GAA GAA AAG AGA 1512
V L G L M S E I L E F L P E E K R 462
ACA GAA GGG GAG AAA TTA TAT GCA AGG CTT AGG CGT CTA ATA AGG AAC ATT 1563
T E G E K L Y A R L R R L I R N I 479
CCG CAA TAT GAA CTG ATG AAA ATA GAG GAA GCC CTT GCA GAG GGA CTA GAA 1614
P Q Y E L M K I E E A L A E G L E 496
CTA GAA GGA TTT TAC AAT CAT GTG TTA GAG GTA TTT AAG CAG GCT GTT GAA 1665
L E G F Y N H V L E V F K Q A V E 513
TTC TTA AGA GAG GCA TTA AAG GAT GAA GAA GTC CTA AAG AAA ATA GCA GAA 1716
F L R E A L K D E E V L K K I A E 530
AAT CCA TTC TTA AGT CTC AAG CAG ATA GAA GGT AAG TGG TAC ATT GAG ATA 1767
N P F L S L K Q I E G K W Y I E I 547
CCA GAT GTA AGA ACT TAT ATT CAG GCC AGC GGC AGG ACA TCT AGA CTC TTT 1818
P D V R T Y I Q A S G R T S R L F 564
GCG GGG GGC ATT ACC AAG GGT CTT AGC GTC ATA TTG GTA GAC GAT CAG AAA 1869
A G G I T K G L S V I L V D D Q K 581
GTG TTC AAC GGA CTT ATA AGG CAG ATG AGA TGG AGA TTT GTT GAG TTC GAG 1920
V F N G L I R Q M R W R F V E F E 598
ATA AAG CCT CTT GGC GAG ATA AAC ATA GAA GAA GTC CTT AAG GAA ATT GAT 1971
I K P L G E I N I E E V L K E I D 615
AGG GAT AGG GAA AAA GTT AGG TTA GTC TTA GAA GGC AAG ATA AGT GAA CAG 2022
R D R E K V R L V L E G K I S E Q 632
GTT AAA GAC CTT GTT AAG TCT GCA CTA ATG ATT GTG GAG AGC CCC AAC AAG 2073
V K D L V K S A L M I V E S P N K 649
GCC AGG ACA ATA GCT AAC TTC TTT GGC CAG CCA AGT AAA AGA AGA ATT GGT 2124
A R T I A N F F G Q P S K R R I G 666
GAT TTA GTT GCA TAT GAA GTT TCC ATT GGA GAC AAA ATG TTG ACG ATT CTA 2175
D L V A Y E V S I G D K M L T I L 683
GCA AGT GGA GGA CAT ATG TTT GAT CTA GTC ACT ACT GAA GGA TAT CAC GGA 2226
A S G G H M F D L V T T E G Y H G 700
GTT CTT CTT CTA GAA AAG GGG GGT AAA AGA TAC TTT GTC CCA GTC TAC GAC 2277
V L L L E K G G K R Y F V P V Y D 717
ACT ATA AAA AGG TGT AGA GAC TGT GGC CAT CAA TTC GTA GAT TGG GAG GAG 2328
T I K R C R D C G H Q F V D W E E 734
AAA GGG GTT TGT CCA AGA TGT GGA AGC AGG AAC GTT TAC GAT GCC CTA GAG 2379
K G V C P R C G S R N V Y D A L E 751
AAT GTA AAG GCA ATG AGA GAT CTT GCC CAG GAA GTT GAT GAA ATA TTA ATT 2430
N V K A M R D L A Q E V D E I L I 768
GGT ACG GAT CCA GAT ACA GAG GGA GAG AAG ATA GCT TGG GAC ATT AGG AAC 2481
G T D P D T E G E K I A W D I R N 785
GTT CTC TCA CCT TAT GCC CCC GTG ATT AAG AGA ATA GAA TTT CAC GAG GTT 2532
V L S P Y A P V I K R I E F H E V 802
ACA AGA CCT GCA ATC CTA AGA GCC ATC AAC GAG GCC AGG GAT ATA AAC GAA 2583
T R P A I L R A I N E A R D I N E 819
GAT AGA GTC AAT GCC CAA CTT GTG AGA AGG ATA GAA GAT AGG TGG ATT GGA 2634
D R V N A Q L V R R I E D R W I G 836
TTT GAG CTT AGC CAA AAG TTA TGG GAA GTG TTT GAG AAC TAC AAC TTA TCG 2685
F E L S Q K L W E V F E N Y N L S 853
GCT GGA AGA GTT CAG ACT CCA GTT CTT GGT TGG ATT GTC CAA AGA TAT AAG 2736
A G R V Q T P V L G W I V Q R Y K 870
GAG TTT ACG GAG AGT GAA ACT GAC TTC CTG GGA CTT ACG TTA GAG AAT GAT 2787
E F T E S E T D F L G L T L E N D 887
ATA ACT ATA ATC CTG GAG GGA GTC TCT GGA GAT GTT CAG GAA GTT TAC GTT 2838
I T I I L E G V S G D V Q E V Y V 904
AAG GAA GTA ACA ATT GAG GAG AGA GAG ATT AAT CCA CTT CCT CCC TAC ACC 2889
K E V T I E E R E I N P L P P Y T 921
ACA GAT GCC ATG CTA CAA GAT GCA TCT AGG TTC CTG GGG TTC TCA CCA ACT 2940
T D A M L Q D A S R F L G F S P T 938
CAC ACA ATG CAA TTA GCT CAA GAT CTC TTC GAA CTA GGT TTG ACA ACT TAT 2991
H T M Q L A Q D L F E L G L T T Y 955
CAC AGA ACT GAC AGT ATA CAT GTG AGT AAC ACT GGA ATT GAG ATA GCA AAA 3042
H R T D S I H V S N T G I E I A K 972
GAA TAC ATA ACC CAG GAA ATT GGG GAA GAA TAC TTT GCT CCG AGA AAG TGG 3093
E Y I T Q E I G E E Y F A P R K W 989
GGA GAA GAA GGA GCT CAC GAA GCG ATA AGA CCC ACT AGA CCA ATA GAT ACT 4144
G E E G A H E A I R P T R P I D T 1006
GGA AGA TTG ATC CAG CTC ATT AGA GAT GGA ATA ATA ACA TTG CCC AGA AAC 3195
G R L I Q L I R D G I I T L P R N 1023
TTA ACA AAA GAT CAC TTT AAG CTT TAT GAT CTC ATA TTC AGG AGA TTC ATG 3246
L T K D H F K L Y D L I F R R F M 1040
GCT AGT CAA ATG AAG CCT GCC AAG GTT TTG TAT GAA AGA GCA GTT ATT GGA 3297
A S Q M K P A K V L Y E R A V I G 1057
ACT CCT TAT GGA GAG GTT GAA ATA GAG GGA TAC ATT GAT GTT ATA TAT GAT 3348
T P Y G E V E I E G Y I D V I Y D 1074
GGA TGG TCT AGA ATT AAG CCT CTT CCA CTG AAG AAA CTT CCA AGG CTA GAA 3399
G W S R I K P L P L K K L P R L E 1091
AAA GAC CAA ATA TTG AAG GTC ACT GAA ATT AGA AAA TGG AGA GCT CCA AAA 3450
K D Q I L K V T E I R K W R A P K 1108
GTT TCA CTT TAT ACC CAG GGT GAC GTT ATA GCA TTA ATG AAA GAG AGG GGA 3501
V S L Y T Q G D V I A L M K E R G 1125
ATT GGT AGA CCA TCA ACT TAT GCG AAG ATT GTG CAA ACC CTA TTA CAG AGG 3552
I G R P S T Y A K I V Q T L L Q R 1142
GGA TAC GTT ATT GAA ACA AAG AGT AAG AAG AAA TTA GTT CCA ACG GAG AAG 3603
G Y V I E T K S K K K L V P T E K 1159
GGA ATA AAA GTA TAT CAC TAC CTT GTG TCT AAA TAC AAA GAC CTC GTT AGT 3654
G I K V Y H Y L V S K Y K D L V S 1176
GAA GAG AGA ACA AGA CAG CTA GAA AAG TTA ATG GAT ATG GTG GAG GAG GCG 3705
E E R T R Q L E K L M D M V E E A 1193
AAG GCA AAC TAT CAA GAA GTT CTC AAT GAA CTC TAT GAG GAA ATC TTA AGA 3756
K A N Y Q E V L N E L Y E E I L R 1210
TAC GTT AAA TCT TGA AGGGACAGTTGTTTATTTACTTTACCTCCTTCTATAACAATTTTTCTGAG 3821
Y V K S * 1214
ATCTATTTGTGGTGAAGAAAATCTTTTTAAGTTTAATTATAATGCATATCTATGCACTTCAATAAAAAT 3891
ATAAGAAGTATGTGATTTTCCTCCATGTAAGAGAAAAGCTT 3832

FIG. 3. Nucleotide sequence of *P. furiosus rgy* and the flanking regions. The deduced amino acid sequence is given in one-letter code below the nucleotide sequence. The transcription start site is identified by the arrow, the sequence of the putative box A promoter element is underlined, and the putative ribosome-binding site is in boldface type.

itive ones. After 30 s of incubation (Fig. 5, lane b), plasmid pTZ18 exhibited both negatively (dark bands) and positively (light bands) supercoiled topoisomers. After 1 min of incubation (Fig. 5, lane c), the plasmid was already completely positively supercoiled.

N-terminal microsequencing has shown that the 120-kDa protein with the reverse gyrase activity corresponds to the 1,214-amino-acid polypeptide deduced from sequencing of the

FIG. 4. *P. furiosus* reverse gyrase activity is carried by a protein of 120 kDa. A partially purified fraction from a heparin-Sepharose column was sedimented through a linear 5 to 20% (vol/vol) sucrose gradient in the presence of 200 mM KCl. Fractions were collected and numbered from the bottom of the gradient. (B) SDS-PAGE analysis of collected fractions. Lane M contained the following molecular mass markers (whose masses are given in kilodaltons on the left): catalase (11.2S), aldolase (7.4S), phosphorylase *b* (6.7S), and bovine serum albumin (4.6S). (A) The same fractions were analyzed for relaxation and positive supercoiling of negatively supercoiled plasmid DNA in the presence of ATP as described in Materials and Methods. M, without enzyme; SC−, negatively supercoiled DNA; N, nicked DNA.

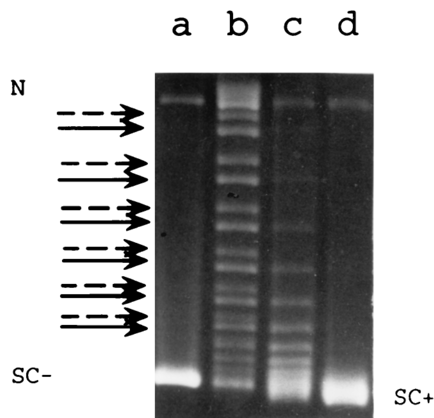

FIG. 5. Kinetics of positive supercoiling. A 0.2-μg sample of negatively supercoiled plasmid pTZ18 DNA was incubated at 95°C with 2 U of *P. furiosus* reverse gyrase under standard conditions. a, without incubation; b, 30 s of incubation; c, 1 min of incubation; d, 2 min of incubation; SC−, negatively supercoiled DNA; SC+, positively supercoiled DNA; N, nicked DNA; --→, negative superturns; →, positive superturns.

cloned *rgy* gene. The molecular mass deduced from the sequence of the reverse gyrase was 134 kDa. This discrepancy could be explained either by the relatively low resolution of electrophoresis in the high-molecular-weight range or by abnormal migration of *P. furiosus* reverse gyrase.

Phylogenetic analysis of reverse gyrases. To examine the origin and evolution of reverse gyrases, we aligned the topoisomerase modules of the four known reverse gyrases with the corresponding sequences of all other available type I DNA topoisomerases of the TopIA family (32b). The most conserved regions, corresponding to 366 amino acid residues, were combined to construct phylogenetic trees by using both distance and parsimony programs. All of the methods used produced the same tree topology. As shown in Fig. 6, the TopIA family can be divided into three subfamilies, one corresponding to bacterial ω-like proteins, another one grouping *Escherichia coli* type III topoisomerase and *Saccharomyces cerevisiae* Top3, and a third one corresponding to the three reverse gyrase sequences. Since this tree is unrooted, one cannot determine if the reverse gyrase subfamily is the outgroup for the other two subfamilies or if it is a sister group of one of them. However, in terms of distance, reverse gyrase is slightly more closely related to ω-like proteins. In all of the trees obtained, the *M. kandleri* reverse gyrase branched first in the reverse gyrase subtree.

DISCUSSION

Reverse gyrase activity has been detected in all of the hyperthermophiles tested so far (members of domain *Archaea* or *Bacteria*), suggesting that this enzyme is a prerequisite for life at extremely high temperatures (13, 18). The origin of the reverse gyrase gene is thus a crucial question in current discussions about the origin and nature of hyperthermophiles (16). Three reverse gyrase genes have been previously described, one from the crenoarchaeon *S. acidocaldarius* and two from the euryarchaea *M. jannaschii* and *M. kandleri*. All are made by the association of a helicase-like module and a topoisomerase module belonging to the TopIA family (9, 24). In *S. acidocaldarius*, and probably *M. jannaschii*, the enzyme is a monomer, whereas in *M. kandleri*, the topoisomerase module is divided between two linked genes, *rgyA* and *rgyB* (Fig. 1). Since *M. kandleri* branches first in the euryarchaeal branch of the 16S rRNA tree (29), it was unclear which of these two arrangements was the ancestral one in *Archaea.*

Here we report the sequence of another euryarchaeal reverse gyrase, from *P. furiosus*. Despite being from a euryarchaeon, this enzyme exhibits the same monomeric structure as the *S. acidocaldarius* enzyme. The same structure can also be deduced from the *rgy* gene sequence of *M. jannaschii*. This suggests that the common ancestor of all archaea contained a monomeric reverse gyrase and that the dual structure of the *M. kandleri* reverse gyrase is a derived character. In the reverse gyrase subtree, the *M. jannaschii*, *S. acidocaldarius*, and *P. furiosus* enzymes are grouped together (Fig. 6). In fact, comparative analyses of amino acid identities and patterns of insertion-deletions strongly suggest that the *P. furiosus* and *M. jannaschii* enzymes occupy an intermediate position between the other two reverse gyrases (Table 1 and Fig. 1). These observations apparently contradict rRNA phylogenies in which *M. kandleri* is grouped with *P. furiosus* and *M. jannaschii* among *Euryarchaeota* and occupies an intermediate position between *P. furiosus* and *S. acidocaldarius* (29). Either the rooting of the rRNA tree is erroneous and the actual root lies between meth-

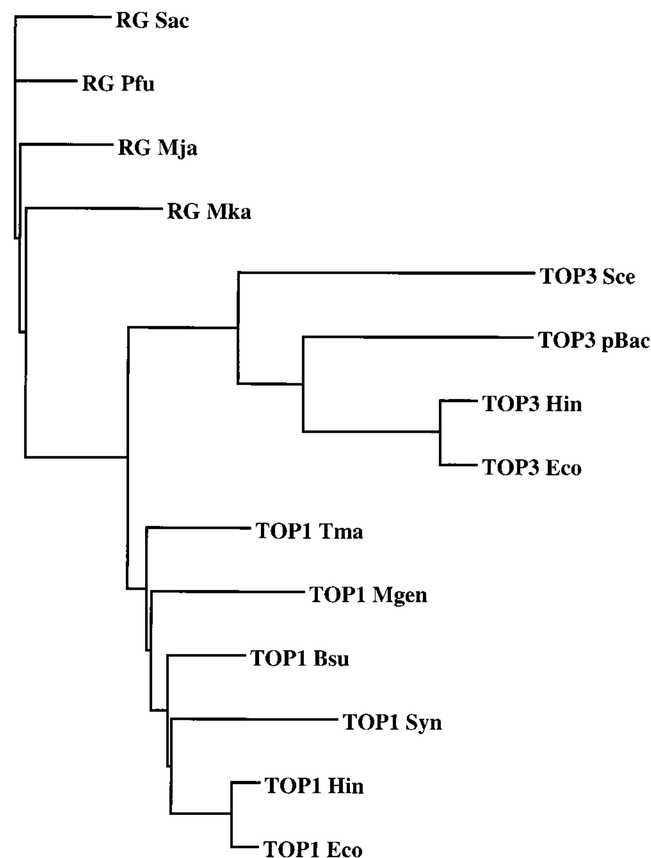

FIG. 6. Unrooted phylogenetic tree of type I DNA topoisomerases of subfamily 1A. This tree was constructed by using the Fitch and Margoliash algorithm in the PHYLIP package (version 3.4) from an alignment of the most-conserved regions (336 amino acids) of all available members of this family, including the four reverse gyrase topoisomerase modules. The analysis involved *S. acidocaldarius* reverse gyrase (RG Sac), *P. furiosus* reverse gyrase (RG Pfu), *M. jannaschii* reverse gyrase (RG Mja), *M. kandleri* reverse gyrase (RG Mka), *Saccharomyces cerevisiae* Top3 (TOP3 Sce), a *Bacillus anthracis* plasmid (TOP3 pBac), *Haemophilus influenzae* TopA (TOP1 Hin) and TopB (TOP3 Hin), *E. coli* TopA (TOP1 Eco) and TopB (TOP3 Eco), *T. maritima* TopA (TOP1 Tma), *Mycoplasma genitalium* TopA (TOP1 Mgen), *Bacillus subtilis* TopA (TOP1 Bsu), and *Synechococcus* sp. TopA (TOP1 Syn).

anogens and sulfothermophiles, or the reverse gyrase tree is biased by some artifact. Since *M. kandleri* reverse gyrase exhibits a novel dimeric structure, it is indeed possible that this structural modification has altered the rate of reverse gyrase evolution in its lineage.

The monomeric structure of the *P. furiosus* and *M. jannaschii* reverse gyrases suggests that the fusion between the helicase and topoisomerase modules to produce this unusual activity occurred before the diversification of the archaeal domain, since both organisms branch after *M. kandleri* in the *Euryarchaeota* domain of the 16S rRNA tree. Interestingly, the reverse gyrase from the bacterium *Thermotoga maritima* also exhibits a monomeric structure, suggesting either lateral gene transfer between the two domains or an even more ancient fusion event, predating the divergence of *Archaea* and *Bacteria* (4a).

A reverse gyrase-like activity in eucaryotes has recently been described, i.e., the functional association in vivo of a DNA helicase (the product of the *SGS* gene) and a type I DNA topoisomerase of the TopIA family (yeast Top3) (20). However, since the SGS protein is much more closely related to the *E. coli* RecQ protein than to the reverse gyrase helicase-like module (data not shown), it is unlikely that this eucaryotic association originated from archaeal reverse gyrases (or vice versa).

The isolation of the *P. furiosus* reverse gyrase gene, identification of its transcript and promoter, and purification of the protein should help in unraveling the exact role played by this enzyme in hyperthermophiles. Members of the order *Pyrococcales* constitute a good model for such studies, since they are used in many laboratories as model organisms for the study of hyperthermophiles, plasmids have been recently isolated (14), an in vitro transcription system is available (32), and genetic tools are under construction (1). It is also possible to overexpress *P. furiosus* enzymes in *E. coli* (10) and study their important features by using site-directed mutagenesis (32a).

Finally, our preliminary results indicate that the reverse gyrase from *P. furiosus* is more thermophilic and thermostable than the *S. acidocaldarius* enzyme. Since the two enzymes are structurally closely related, they could constitute an interesting two-model system to study adaptation to hyperthermophily in this fascinating class of enzymes.

ACKNOWLEDGMENTS

This research was supported by grants from the Department of Energy (DE-F902-92ER20083) (F.T.R.), the National Science Foundation (BES 9410687) (F.T.R.), the European Union Biotechnology Program (BIO-CT93-02734) (P.F.), the Association de la Recherche sur le Cancer (P.F.), and the Ligue Française de Recherche sur le Cancer (P.F.).

REFERENCES

1. **Aagaard, C., I. Leviev, R. N. Aravalli, P. Forterre, D. Prieur, and R. A. Garrett.** 1996. General vector for archaeal hyperthermophiles: strategies based on a mobile intron and a plasmid. FEMS Microbiol. Rev. **18:**93–104.
2. **Andrea, L., K. Mikulik, and N. D. Savelyeva.** 1993. Characterization of a reverse gyrase from the extremely thermophilic hydrogen-oxidizing eubacterium *Calderobacterium hydrogenophilum*. FEMS Microbiol. Lett. **110:**107–112.
3. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1987. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
4. **Borges, K. M., S. R. Brummet, A. P. Bogert, M. C. Davis, K. M. Hujer, S. T. Domke, J. Szasz, J. Ravel, J. DiRuggiero, C. Fuller, J. W. Chase, and F. T. Robb.** 1996. A survey of the genome of the hyperthermophilic archaeon, *Pyrococcus furiosus*. Genome Sci. Technol. **1:**37–46.

4a. **Bouthier de la Tour, C., and M. Duguet.** Personal communication.

5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–252.
6. **Brown, J. W., C. J. Daniels, and J. N. Reeve.** 1989. Gene structure, organization and expression in archaebacteria. Crit. Rev. Microbiol. **16:**287–338.
7. **Bult, C. J., O. White, G. J. Olson, L. Zhou, R. D. Fleishmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, et al.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science **273:**1058–1073.
8. **Charbonnier, F., and P. Forterre.** 1994. Comparison of plasmid DNA topology among mesophilic and thermophilic eubacteria and archaebacteria. J. Bacteriol. **176:**1252–1259.
9. **Confalonieri, F., C. Elie, M. Nadal, C. de La Tour, P. Forterre, and M. Duguet.** 1993. Reverse gyrase: a helicase-like domain and a type I topoisomerase in the same polypeptide. Proc. Natl. Acad. Sci. USA **90:**4753–4757. (Erratum, **91:**3478, 1994.)
10. **DiRuggiero, J., and F. T. Robb.** 1995. Expression and in vitro assembly of recombinant glutamate dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl. Environ. Microbiol. **61:**159–164.
11. **DiRuggiero, J., and F. T. Robb.** 1995. RNA extraction from sulfur-utilizing thermophilic archaea, p. 97–99. *In* F. T. Robb, A. R. Place, S. DasSarma, H. J. Schreier, and E. M. Fleinschmann (ed.), Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. **DiRuggiero, J., F. T. Robb, R. Jagus, H. H. Klump, K. M. Borges, M. Kessel, X. Mai, and M. W. W. Adams.** 1993. Characterization, cloning, and in vitro expression of the extremely thermostable glutamate dehydrogenase from the hyperthermophilic archaeon, ES4. J. Biol. Chem. **268:**17767–17774.
13. **Duguet, M.** 1995. Reverse gyrase, p. 84–114. *In* F. Eckstein and D. M. J. Lilley (ed.), Nucleic acids and molecular biology. Springer-Verlag KG, Berlin, Germany.
14. **Erauso, G., S. Marsin, N. Benbouzid-Rollet, M. Baucher, T. Barbeyron, Y. Zivanovic, D. Prieur, and P. Forterre.** 1996. Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. J. Bacteriol. **178:**3232–3237.
15. **Fiala, G., and K. O. Stetter.** 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch. Microbiol. **145:**56–61.
16. **Forterre, P.** 1996. A hot topic: the origin of hyperthermophiles. Cell **85:**789–792.
17. **Forterre, P., A. Bergerat, and P. Lopez-Garcia.** 1996. The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea. FEMS Microbiol. Rev. **18:**237–248.
18. **Forterre, P., F. Confalonieri, F. Charbonnier, and M. Duguet.** 1995. Speculations on the origin of life and thermophily. Origins Life Evol. Biosphere **25:**235–249.
19. **Forterre, P., G. Mirambeau, C. Jaxel, M. Nadal, and M. Duguet.** 1985. High positive supercoiling in vitro catalyzed by an ATP and polyethylene glycol-stimulated topoisomerase from *Sulfolobus acidocaldarius*. EMBO J. **4:**2123–2128.
20. **Gangloff, S., J. P. McDonald, C. Bendixen, L. Arthur, and R. Rothstein.** 1994. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. Mol. Cell. Biol. **14:**8391–8398.
21. **Kato, J. I., H. Suzuki, and H. Ikeda.** 1992. Purification and characterization of DNA topoisomerase IV. J. Biol. Chem. **267:**25676–25684.
22. **Kikuchi, A., and K. Asai.** 1984. Reverse gyrase, a topoisomerase which introduces positive superhelical turns into DNA. Nature **309:**677–681.
23. **Kikuchi, A., T. Shibata, and S. Kakasu.** 1986. Reverse gyrase and DNA supercoiling in *Sulfolobus*. Syst. Appl. Microbiol. **7:**72–78.
24. **Krah, R., S. A. Kozyavkin, A. I. Slesarev, and M. Gellert.** 1996. A two-subunit type I DNA topoisomerase (reverse gyrase) from an extreme hyperthermophile. Proc. Natl. Acad. Sci. USA **93:**106–110.
25. **Kyhse-Anderson, J.** 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. J. Biochem. Biophys. Methods **10:**203–209.
26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
27. **Lima, C. D., J. C. Wang, and A. Mondragon.** 1994. Three-dimensional structure of the 67K N-terminal fragment of E. coli DNA topoisomerase I. Nature **367:**138–146.
28. **Nadal, M., C. Jaxel, C. Portemer, P. Forterre, G. Mirambeau, and M. Duguet.** 1988. Reverse gyrase of sulfolobus: purification to homogeneity and characterization. Biochemistry **27:**9102–9108.
29. **Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
30. **Reiter, W. D., U. Hudepohl, and W. Zillig.** 1990. Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection *in vitro*. Proc. Natl. Acad. Sci. USA **87:**9509–9513.
31. **Slesarev, A. I.** 1988. Positive supercoiling catalysed in vitro by ATP-dependent topoisomerase from *Desulfurococcus amylolyticus*. Eur. J. Biochem. **173:**395–399.
32. **Thomm, M.** 1996. Archaeal transcription factors and their role in transcription initiation. FEMS Microbiol. Rev. **18:**159–171.

32a. **Vetriani, C., N. Tolliday, and F. T. Robb.** Personal communication.

32b. **Wang, J. C.** 1996. DNA topoisomerases. Annu. Rev. Biochem. **65:**635–692.

33. **Wich, G., H. Hummel, M. Jarsch, U. Bar, and A. Bock.** 1986. Transcription signals for stable RNA genes in *Methanococcus*. Nucleic Acids Res. **14:**2459–2479.