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

KIM M. BORGES,1,2 AGNES BERGERAT,3 ALLISON M. BOGERT,1 JOCELYNE DI RUGGIERO,1* PATRICK FORTERRE,3 AND FRANK T. ROBB1

Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, Maryland1; Institut de Génétique et Microbiologie, CNRS, Université Paris-Sud, Paris, France3; and Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut2

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, Crenarchaeota 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 Lk 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 archaean 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 \emph{P. furiosus} was sequenced and shown to have high homology to the reverse gyrase gene from \emph{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 \emph{P. furiosus} genomic DNA. A lambda Zap II partial genomic library of \emph{P. furiosus} (4) was screened by using this PCR fragment labeled with [α-32P]dATP by random priming. Four positive plaques were rescued into the plasmid 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.

Total RNA was extracted as described by Druggingero and Robb (11). Primer extension (3) was performed with a 5′-end 32P-labeled synthetic oligonucleotide, 5′-ACATCAAACTGATCATCAC-3′, which is complementary to positions +80 to +97 on the \textit{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.).

RESULTS

Cloning and sequencing of the \emph{P. furiosus} reverse gyrase gene. A fragment of the \emph{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 II library of \emph{P. furiosus} genomic DNA. Several clones were identified by plaque hybridization and rescued into pBluescript KS+ (4). The \emph{P. furiosus} reverse gyrase gene was sequenced by using two of these subclones, pKMB12 and pKMB14, which contain the entire \textit{rgy} gene. Hybridization of genomic DNA from \emph{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 \emph{P. furiosus} reverse gyrase gene \textit{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 \emph{S. acidocaldarius} protein, which is 1,247 amino acids (9). The deduced \emph{P. furiosus} reverse gyrase amino acid sequence is 48, 41, and 37% identical to the deduced reverse gyrase sequences of \emph{M. jannaschii}, \emph{S. acidocaldarius}, and \emph{M. kandleri}, respectively. The \emph{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 \emph{S. acidocaldarius} reverse gyrase, rather than the heterodimeric structure of the \emph{M. kandleri} enzyme (Fig. 1).

Comparative analysis of amino acid identities for each of the modules among the three enzymes indicated that \emph{P. furiosus} reverse gyrase is more closely related to the \emph{M. jannaschii} enzyme than those from the \emph{S. acidocaldarius} and \emph{M. kandleri} enzymes, especially in the topoisomerase domain, whereas the \emph{S. acidocaldarius} and \emph{M. kandleri} reverse gyrase genes 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 \emph{P. furiosus} and \emph{M. jannaschii}, 23% identity between \emph{P. furiosus} and \emph{S. acidocaldarius}, and 15% identity between \emph{P. furiosus} and \emph{M. kandleri}.

Figure 1 compares the structures of the three archaeal reverse gyrase genes. The helicase module of \emph{P. furiosus} reverse gyrase possesses all of the helicase-like motifs (a to h) present in the other reverse gyrase genes 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 \emph{M. kandleri} helicase-like module (compared to the corresponding region in the \emph{S. acidocaldarius} enzyme) are also present in the \emph{P. furiosus} gene, in agreement with an intermediate position of the \emph{P. furiosus} enzyme between the \emph{S. acidocaldarius} and \emph{M. kandleri} reverse gyrase genes. The topoisomerase domain of the \emph{P. furiosus} enzyme also possesses all of the motifs (1 to VII) present in the other two reverse gyrase genes, again in the same arrangement. The putative active-site tyrosine is located at position 955. Two large insertions present in the \emph{M. kandleri} enzyme are missing in \emph{P. furiosus} reverse gyrase, in particular, a C-terminal extension in \textit{rgyA} that could be involved in the stabilization of the \emph{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 (TTTATAA) 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 TTATAA/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 sequences were identified 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 \emph{P. furiosus} \textit{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 1722 BORGES ET AL. J. BACTERIOL.
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 positive superhelical turns (Fig. 5).

---

**TABLE 1. Percentages of amino acid identity between topoisomerase and helicase-like modules of archaeal reverse gyrase**

<table>
<thead>
<tr>
<th>Species</th>
<th>% Identity with:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pfu</td>
</tr>
<tr>
<td>Pfu</td>
<td>100</td>
</tr>
<tr>
<td>Mja</td>
<td>32</td>
</tr>
<tr>
<td>Sac</td>
<td>23</td>
</tr>
<tr>
<td>Mka</td>
<td>15</td>
</tr>
</tbody>
</table>

* a Top right.
* b Bottom left.
* 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.
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. 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.

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

one's. 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. (A) 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). (B) The same fraction was 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. 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.

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. (A) 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). (B) The same fraction was 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. *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. (A) 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). (B) The same fraction was 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.
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 gyrase, we aligned the topoisomerase modules of the four known reverse gyrase 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 crenarchaeon S. acidocaldarius and two from the euryarchaeon 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 crenarchaeon, 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 gyrase (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-
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. janinachii 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 Eurarchaeota 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 archaean 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 unravelling 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), and 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-PK02-92ER20283 (F.T.R.), the National Science Foundation (BES 9410087) (F.T.R.), the European Union Biotechnology Program (Bio-CF93-02734) (P.F.), the Association de la Recherche sur le Cancer (P.F.), and the Ligue Française de Recherche sur le Cancer (P.F.).

REFERENCES