

# Cloning and sequence analysis of DNA gyrase genes from *Mycobacterium tuberculosis*

K. Madhusudan, V. Ramesh and V. Nagaraja\*

Centre for Genetic Engineering, Indian Institute of Science, Bangalore 560 012, India

**Using a heterologous DNA gyrase probe from *Streptomycesphaeroides*, we have cloned the *gyrA* and *gyrB* genes of *Mycobacterium tuberculosis*. The *M. tuberculosis gyrB* gene has been sequenced and compared with the amino acid sequences of GyrB from other bacteria. There is a significant level (60–80%) of amino acid sequence homology between GyrB of *M. tuberculosis* and that of other bacterial species, indicating a high degree of conservation of this essential enzyme.**

DNA gyrase is a type II topoisomerase, present ubiquitously in bacteria. The enzyme has essential function in catalysing the introduction of negative superhelical turns into DNA<sup>1</sup>. The reaction involves ATP-dependent transient breakage and resealing of both strands of DNA<sup>2,3</sup>. The negative supercoiling thus generated is a prerequisite for replication and transcription. Apart from serving as a swivel during replication fork movement, the enzyme acts on positive supercoils generated ahead of the transcription complex. Along with topoisomerase I, DNA gyrase serves to maintain the homeostatic balance of chromosomal superhelical density<sup>4,5</sup>.

*E. coli* DNA gyrase is a tetrameric protein having two A and two B subunits, encoded by *gyrA* and *gyrB* respectively<sup>6</sup>. Both the genes of *E. coli* have been cloned, sequenced<sup>7–9</sup> and overexpressed<sup>10</sup>. The genes have also been isolated and characterized from other bacterial sources. Two classes of compounds—quinolones and coumarins and their derivatives inhibit DNA gyrase activity. The prototype compounds are nalidixic acid, oxolinic acid (quinolones) and coumermycin A1, novobiocin (coumarins). While quinolones interact with DNA–protein complex<sup>11,12</sup>, coumarins affect the ATPase activity of GyrB<sup>13</sup>, resulting in inhibition of supercoiling mediated by the enzyme.

The genus *Mycobacterium* has attained importance due to pathogenic species such as *Mycobacterium tuberculosis*, *Mycobacterium leprae* and other opportunistic members. Tuberculosis and leprosy have continued to be serious public health hazards in India and other developing countries over a long period of time. Only recently, the industrialized countries have shown enor-

mous interest in this problem due to the increased incidence of tuberculosis and development of resistance to known antimycobacterial drugs<sup>14,15</sup>. Though DNA gyrase has been well studied in several organisms, hardly any information is available from *Mycobacteria*. The enzyme could be an ideal target for the development of new line of antimycobacterial drugs. Although DNA gyrase activity was detected in *M. smegmatis* earlier<sup>16,17</sup>, details about the enzyme structure and function from the *Mycobacterial* species are yet to be elucidated. Cloning and expression of the gyrase genes should facilitate the isolation of the protein in large quantities to carry out *in vitro* drug screening experiments. Here we report the cloning of *gyrA* and *gyrB* genes from *M. tuberculosis*. Further, we have determined the nucleotide sequence of *gyrB* and compared with other known *gyrB* sequences.

## Materials and methods

Bacterial strains *M. tuberculosis* H<sub>37</sub>R<sub>a</sub>, *E. coli* DH5 $\alpha$ , and plasmid vectors pUC19, pBR322, were from our laboratory stocks. *M. tuberculosis* H<sub>37</sub>R<sub>a</sub> was propagated on Youmans and Karlson's medium as described earlier<sup>17</sup>. All *E. coli* strains were grown on Luria–Bertani medium. The following plasmids and strains were generous gifts: pSLS447, *E. coli gyrA* (J. C. Wang, Harvard Univ., Cambridge, USA)<sup>7</sup>; pMK90, pMK47, *E. coli gyrA* and *gyrB* and *E. coli* N4177 *gyrB*<sup>ts</sup> *cou*<sup>r</sup> (M. Gellert, NIH, USA)<sup>4,8</sup>; *E. coli* LE316 *gyrB*<sup>ts</sup> (E. Orr, Leicester Univ., UK)<sup>18</sup>; *E. coli* NH647, KNK453, KNK402 are *gyrA*<sup>ts</sup> strains (N. R. Cozzarelli, Univ. of Chicago, Chicago, USA)<sup>19</sup>; pGPD80, *K. pneumoniae gyrA* (H. K. Das, Jawaharlal Nehru Univ., New Delhi, India)<sup>20</sup>; pAK1, *B. subtilis gyrA* and *gyrB* (K. F. Bott, Univ. of North Carolina, Chapel Hill, USA)<sup>21</sup>; pLS182, *S.phaeroides gyrB*<sup>r</sup> (A. S. Thiara and E. Cundliffe, Leicester Univ., UK)<sup>22</sup>.

Restriction and DNA modifying enzymes were from Pharmacia and Boehringer Mannheim and used according to their specifications. Nylon membrane (Gene screen plus) was from New England Nuclear, Du Pont. [ $\alpha$ <sup>32</sup>P]-dCTP was obtained from Bhabha Atomic Research Centre, India. 7-deaza dGTP was from Boehringer Mannheim and Sequenase version 2.0 and Genescribe vectors pTZ18U, pTZ19U were from U.S. Biochemicals. All other

\*For correspondence.

reagents were purchased from Sigma Chemical Co.

### Isolation of genomic DNA and other DNA manipulations

Genomic DNA from *M. tuberculosis* H<sub>37</sub>R<sub>a</sub> was isolated by the method of Srivastava *et al.*<sup>23</sup>. Briefly, the bacterial pellet was treated with lysozyme in GTE (Glucose 50 mM; Tris HCl pH 8.0, 25 mM; EDTA 10 mM) containing 0.1% Tween 80 at 37° C for 3 to 4 h and incubated with proteinase K (100 µg/ml) at 60° C for 2 h in presence of 1% SDS. After phenol and chloroform extractions the genomic DNA was precipitated with ethanol, spooled and dissolved in TE (Tris. HCl pH 8.0, 10 mM; EDTA 0.1 mM). RNase A digestion was followed by cetyl triethyl ammonium bromide treatment<sup>24</sup> and phenol, chloroform extractions. DNA was spooled out after adding ethanol, air dried and dissolved in TE.

### Southern analysis of genomic DNA

Genomic DNA (2 µg) was digested completely with several restriction enzymes, such as *Bam*HI, *Eco*RI and *Bgl*II. The digested DNA was electrophoresed in 0.8% agarose (4–5 V/cm) and transferred on to nylon membrane in presence of alkali. This blot was prehybridized for 4–6 h and then hybridized with radiolabelled 2.2 kb *Bam*HI fragment carrying *S. sphaeroides gyrB<sup>r</sup>* gene for 12–24 h. The experimental details were as described in Sambrook *et al.*<sup>25</sup>.

### Isolation of positive clone by colony hybridization

The DNA was electroeluted from the region of the signal and ligated to *Eco*RI digested, dephosphorylated pBR322. The transformants in *E. coli* DH5α were transferred on to nitrocellulose circles and hybridization was performed as described above. The colonies were screened and the positive clones were isolated and reanalysed by Southern hybridization<sup>25</sup>.

### DNA sequence determination

Different restriction enzyme fragments were subcloned into Genescribe vectors (pTZ18U and pTZ19U) to prepare ssDNA templates. Sequencing reactions were carried out following supplier's manual and the products were resolved on 6% Urea PAGE. The DNA sequence was analysed using UWGCG software, at the Distributed Information Centre, Indian Institute of Science. The sequence information has been communicated to EMBL Data Library.

## Results and discussion

We have used different strategies to clone the DNA gyrase genes from *M. tuberculosis*. The genomic libraries of *M. tuberculosis* were prepared in pUC19 and pBR322 using different restriction enzymes. The libraries were transformed into temperature-sensitive *gyrA* and *gyrB* strains of *E. coli*<sup>4, 18, 19</sup> to obtain functionally complementing clones. This approach did not yield positive results. Probing the genomic DNA with gyrase genes of *E. coli*<sup>4, 7, 8</sup>, *B. subtilis*<sup>21</sup> and *K. pneumoniae*<sup>20</sup> also did not give hybridization signals in spite of several attempts using different conditions. The failure of these heterologous probing experiments could be due to the vast difference in the G + C content of the probes (≈ 50%) to that of Mycobacteria (≈ 70%). Hence, we used the cloned *gyrB<sup>r</sup>* gene<sup>22</sup> (a generous gift of A. S. Thiara and E. Cundliffe) of *S. sphaeroides*, a G + C rich organism closely related to Mycobacteria.

Southern hybridization of chromosomal DNA isolated from *M. tuberculosis* H<sub>37</sub>R<sub>a</sub> with *S. sphaeroides gyrB<sup>r</sup>* gene probe is shown in Figure 1. Single hybridizing fragment of different sizes was seen in each lane except

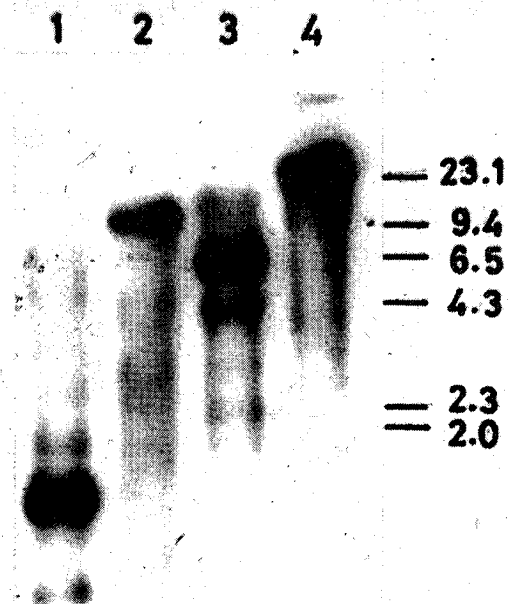


Figure 1. Southern analysis of *M. tuberculosis* H37Ra DNA. The details of the experiment are given in Methods. DNA was digested with *Bam*HI (lane 1), *Pst*I (lane 2), *Eco*RI (lane 3) and *Bgl*II (lane 4). Sizes of the Lambda DNA-*Hind*III fragments are indicated.

in *EcoRI* lane where two fragments 6.2 kb and 4.2 kb hybridized to the probe. DNA from *EcoRI* lane in the region corresponding to 6.2 kb and 4.2 kb was eluted and cloned into plasmid pBR322. Bacterial colonies harbouring these recombinant plasmids containing *gyrB*<sup>r</sup> hybridizable fragments were then picked up by colony hybridization with the same probe. Several recombinant plasmids were isolated and screened again by Southern hybridization to obtain pMN13R and pMN6R of sizes 6.2 kb and 4.2 kb respectively. The contiguous arrangement of these 6.2 kb and 4.2 kb fragments in the genome was confirmed by subsequent Southern analysis of genomic DNA using several restriction enzymes. The plasmids were digested with several restriction enzymes to obtain the restriction map. The combined restriction map of pMN13R and pMN6R is given in Figure 2. Enzymes such as *Bam*HI, *Sma*I, *Sac*II, *Pvu*II have multiple sites. This information was used to generate smaller subclones for sequencing. The *gyrB* gene was located by hybridizing the DNA from different subclones with the probe. The location of *gyrA* was identified by partial sequencing of plasmid pMN6R. We have determined the complete nucleotide sequence of *gyrB* gene by sequencing about 5.0 kb region of the cloned region. The sequencing strategy will be described in detail elsewhere. Sequence data obtained from double-stranded DNA sequencing were ambiguous. We therefore resorted to make single-stranded templates and use 7-deaza dGTP to relieve GC compressions. This modification provided unambiguous sequence data. Analysis of the sequence revealed the presence of an ORF of 714 amino acids (Figure 3). Examination of the derived amino acid sequence indicates the presence of several potential start codons and the exact N-terminus of the protein is not determined at this stage. Comparison of this derived amino acid sequence to known DNA GyrB subunit amino acid sequences using GAP (GCG software) showed a high degree of homology. At the amino acid level, the sequence shows about 55% identity and greater than 70% homology to *E. coli* GyrB subunit. DNA sequence comparison indicated poor homology between the *gyrB* genes of *E. coli* and *M. tuberculosis*. This result is not surprising considering the difference in the G + C content of these two unrelated species. However, the gene shows about 80% homology at the amino acid level and greater

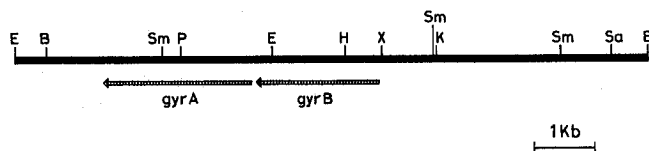


Figure 2. Restriction map of the cloned region. A contiguous stretch of 10.4 kb from plasmids pMN13R and pMN6R is shown. The location and orientation of both genes are indicated by the arrow mark. Abbreviations: B, *Bst*WI; E, *Eco*RI; H, *Hpa*I; K, *Kpn*I; P, *Pst*I; Sa, *Sac*I and Sm, *Sma*I.

```

1  . M G K N E A R R S A L A P D H G T V V C D P L R R L N R H H .
31 A T P E E S I R I V A A Q K K K A Q D E Y G A A S I T I L E
61 G L E A V R K R P G M Y I G S T G E R G L H H L I W E V V D
91 N A V D E A M A G Y A T T V N V V L L E D G G V E V A D D G
121 R G I P V A T H A S G I P T V D V V M T Q L H A G G K F D S
151 D A Y A I S G G L H G V G V S V V N A L S T R L E V E I K R
181 D G Y E W S Q V Y E K S E P L G L K Q G A P T K K T G S T V
211 R F W A D P A V F E T T E Y D F E T V A R R L Q E M A F L N
241 K G L T I N L T D E R V T Q D E V V D E V V S D V A E A P K
271 S A S E R A A E S T A P H K V K S R T F H Y P G G L V D F V
301 K H I N R T K N A I H S S I V D F S G K G T G H E V E I A M
331 Q W N A G Y S E S V H T F A N T I N T H E G G T H E E G F R
361 S A L T S V V N K Y A K D R K L L K D K D P N L T G D D I R
391 E G L A A V I S V K V S E P Q F E G Q T K T K L G N T E V K
421 S F V Q K V C N E Q L T H W F E A N P T D A K V V V N K A V
451 S S A Q A R I A A R K A R E L V R R K S A T D I G G L P G K
481 L A D C R S T D P R K S E L Y V V E G D S A G G S A K S G R
511 D S M F Q A I L P L R G K I I N V E K A R I D R V L K N T E
541 V Q A I I T A L G T G I H D E F D I G K L R Y H K I V L M A
571 D A D V D G Q H I S T L L L T L L F R F M R P L I E N G H V
601 F L A Q P P L Y K L K W Q R S D P E F A Y S D R E R D G L L
631 E A G L K A G K K I N K E D G I Q R Y K G L G E M D A K E L
661 W E T T M D P S V R V L R Q V T L D D A A A A D E L F S I L
691 M G E D V D A R R S F I T R N A K D V R F L D V *
    
```

Figure 3. The derived amino-acid sequence of *gyrB* of *M. tuberculosis*. Potential translational starts are indicated by dots.

than 70% DNA sequence homology to *S. sphaeroides gyrB* gene.

Next, we aligned one of the most conserved regions amongst the known GyrB proteins. This region includes Lys-103 and Tyr-109 residues of *E. coli* GyrB which contribute to the ATP-binding pocket<sup>26</sup>. The conservation to near identity of this 27 amino acid stretch reflects the importance of this region in the structure and function of the protein across diverse bacterial genera (Figure 4). It should be noted here that crystal structure studies of the N-terminal fragment of *E. coli* DNA GyrB protein,

<i>M.tuberculosis</i>	137	V V M T Q L H A G G K F D S D A Y A I S G G L H G V
<i>E.coli</i>	93	V I M T V L H A G G K F D D N S Y K V S G G L H G V
<i>B.subtilis</i>	95	V I M T V L H A G G K F D G S G Y K V S G G L H G V
<i>M.pneumoniae</i>	97	T V F T V L H A G G K F D N D S Y K V S G G L H G V
<i>P.putida</i>	95	V I M T V L H A G G K F D D N S Y K V S G G L H G V
<i>P.mirabilis</i>	93	V I M T V L H A G G K F D D N S Y K V S G G L H G V
<i>N.gonorrhoeae</i>	97	V I M T V L H A G G K F D N S Y K I S G G L H G V
<i>S.sphaeroides</i>	114	V V L T V L H A G G K F G G G G Y S V S G G L H G V

Figure 4. Alignment of the conserved N-terminus region of different GyrB proteins.

complexed with ATP analogue, ADPNP have revealed important features of cofactor-protein interactions<sup>26</sup>. Apart from Lys-103 and Tyr-109, Tyr-5, Asn-46, Asp-73, Gln-335 and Lys-337 are shown to be involved in interacting with the cofactor. These amino acids are also found to be conserved in *gyrB* gene of *M. tuberculosis*.

While there are structural and possible functional similarities between *E. coli* and *M. tuberculosis* GyrB proteins, the following differences may be noted.

(1) An unusually long N-terminal stretch of 46 amino acids (considering the first ATG of the ORF) when compared to GyrB of *E. coli*; the significance, if any, of this region is yet to be ascertained.

(2) Absence of a stretch of about 165 amino acids towards the C-terminal region of *M. tuberculosis* GyrB which is present in *E. coli*. This might explain the failure of our earlier functional complementation experiments using temperature-sensitive *E. coli* *gyrA* and *gyrB* strains. It should be mentioned here that *B. subtilis* *gyrB* which lacks this stretch also fails to functionally complement its homologue of *E. coli*<sup>27,28</sup>. This result is in general agreement with our observation. We have noted that the same stretch of amino acids is absent in GyrB of *M. pneumoniae*<sup>29</sup>, and *S. sphaeroideis*<sup>30</sup>.

The DNA gyrase of Mycobacteria could be a very effective target for new antimycobacterials, either derivatives of known inhibitors of DNA gyrase or novel ones. There is an urgent need for new drugs for the treatment of tuberculosis due to the resurgence of the disease<sup>14</sup> and alarming increase in resistance<sup>15</sup> to proven antimycobacterials such as Rifampicin and INH. The cloning and sequencing of the gyrase genes is a first step in this direction. Expression and characterization of the protein should enable us to screen a variety of quinolone and coumarin compounds.

*Note added in proof:* Takiff *et al.*<sup>31</sup> have recently cloned and sequenced *gyrA* and *gyrB* genes from *M. tuberculosis* H<sub>37</sub> R<sub>v</sub>.

- Gellert, M., Mizuuchi, K., O'Dea, M. H. and Nash, H. A., *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 3872-3875.
- Maxwell, A. and Gellert, M., *Adv. Protein Chem.*, 1986, **38**, 69-107.
- Mizuuchi, K., Fisher, L. M., O'Dea, M. H. and Gellert, M., *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 1847-1851.
- Menzel, R. and Gellert, M., *Cell*, 1983, **34**, 105-113.
- Reece, R. J. and Maxwell, A., *Crit. Rev. Biochem. Mol. Biol.*, 1991, **26**, 335-375.
- Sugino, A., Peebles, C. L., Kreuzer, K. N. and Cozzarelli, N. R.,

*Proc. Natl. Acad. Sci. USA*, 1977, **74**, 4767-4771.

- Swanberg, S. L. and Wang, J. C., *J. Mol. Biol.*, 1987, **197**, 729-736.
- Mizuuchi, K., Mizuuchi, M., O'Dea, M. H. and Gellert, M., *J. Biol. Chem.*, 1984, **259**, 9199-9201.
- Adachi, T., Mizuuchi, M., Robinson, E. A., Apella, E., O'Dea, M. H., Gellert, M. and Mizuuchi, K., *Nucleic Acids Res.*, 1987, **15**, 771-781.
- Hallet, P., Grimshaw, A. J., Wigely, D. B. and Maxwell, A., *Gene*, 1990, **93**, 139-142.
- Drlica, K. and Franco, R. J., *Biochemistry*, 1988, **27**, 2253-2259.
- Shen, L. L., Kohlbrenner, W. E. and Baranowski, J., *J. Biol. Chem.*, 1989, **264**, 2973-2978.
- Sugino, A., Higgins, N. P., Brown, P. O., Peebles, C. L. and Cozzarelli, N. R., *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 4838-4842.
- Bloom, B. R. and Murray, L. J. C., *Science*, 1992, **257**, 1055-1064.
- Young, D. B. and Cole, S. T., *J. Bacteriol.*, 1993, **175**, 1-6.
- Nagaraja, V. and Gopinathan, K. P., *FEBS Lett.*, 1981, **127**, 57-62.
- Nagaraja, V., Ph D thesis, Indian Institute of Science, India, 1981.
- Orr, E. and Staudenbauer, W. L., *Mol. Gen. Genet.*, 1981, **181**, 52-57.
- Kreuzer, K. N. and Cozzarelli, N. R., *J. Bacteriol.*, 1979, **140**, 425-435.
- Dimri, G. P. and Das, H. K., *Nucleic Acids Res.*, 1990, **18**, 151-156.
- Lampe, M. F. and Bott, K. F., *Nucleic Acids Res.*, 1984, **12**, 6307-6323.
- Thiara, A. S. and Cundliffe, E., *Embo J.*, 1988, **7**, 2255-2259.
- Srivastava, R., Gopinathan, K. P. and Ramakrishnan, T., *J. Bacteriol.*, 1981, **148**, 716-719.
- Srivastava, R., Ph D thesis, Indian Institute of Science, India, 1983.
- Langridge, J., Langridge, P. and Berquist, P. L., *Anal. Biochem.*, 1980, **103**, 264-271.
- Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, Second edition, 1989.
- Wigely, D. B., Davies, G. J., Dodson, E. J., Maxwell, A. and Dodson, G., *Nature*, 1991, **351**, 624-629.
- Lampe, M. F. and Bott, K. F., *J. Bacteriol.*, 1985, **162**, 78-84.
- Moriya, S., Ogaswara, N. and Yoshikawa, H., *Nucleic Acids Res.*, 1985, **13**, 2251-2265.
- Colman, S. D., Hu, P. C. and Bott, K. F., *Mol. Microbiol.*, 1990, **4**, 1129-1134.
- Thiara, A. S. and Cundliffe, E., *Mol. Microbiol.*, 1993, **8**, 495-506.
- Takiff, H. E., Salazar, L., Guerrero, C., Philipp, W., Huang, W. M., Kreiswirth, B., Cole, S. T., Jacobs, W. R. and Telenti, A., *Antimicrobial Agents and Chemotherapy*, 1994, **38**, 773-780.

**ACKNOWLEDGEMENTS.** We thank N. R. Cozzarelli, M. Gellert and J. C. Wang for temperature sensitive *E. coli* *gyrA* and *gyrB* strains and DNA probes; K. F. Bott for *B. subtilis* gyrase clones; H. K. Das for *K. pneumoniae* *gyrA* plasmid; A. S. Thiara and E. Cundliffe for *gyrB*<sup>f</sup> of *S. sphaeroideis*. Thanks are due to H. Sharat Chandra for his interest and critical reading of the manuscript. The work is supported by a grant from Council for Scientific and Industrial Research (CSIR), India. MK and VR are senior research fellows of CSIR.

Received 7 April 1994; accepted 12 April 1994