Vol. 33, No. 4, July 1994 BIOCHEMISTRY *and* MOLECULAR BIOLOGY INTERNATIONAL
Pages 651–660

MOLECULAR CLONING OF gyrA AND gyrB GENES OF MYCOBACTERIUM TUBERCULOSIS: ANALYSIS OF NUCLEOTIDE SEQUENCE.

Madhusudan, K., Ramesh, V., and Nagaraja, V*.
Centre for Genetic Engineering, Indian Institute of Science,
Bangalore, 560 012, INDIA.

Received May 11, 1994

SUMMARY: We have recently reported the cloning of gyrA and gyrB genes from Mycobacterium tuberculosis H37Ra [Curr. Science, (1994) 66, 664-667). Here, we present the complete nucleotide sequence of gyrB gene from M.tuberculosis H37Ra along with the flanking regions. The gyrA gene has been located 34 nucleotides downstream of gyrB and has been partially sequenced; both the genes seem to be transcribed from the promoter elements located upstream of gyrB coding sequence. The gyrB gene encodes a polypeptide of 714 amino acids. The deduced amino acid sequences of gyrB and a part of gyrA show extensive homology to the corresponding genes from other bacterial species. The DNA gyrase of M.tuberculosis could be utilised to develop new line of antitubercular drugs.

INTRODUCTION

Tuberculosis and other mycobacterial diseases have continued to be devastating health problems in the third world [1]. More recently with the increasing incidence of tuberculosis and the appearance of multi-drug resistant strains of M.tuberculosis in USA, the problem has attained global dimension [2,3]. Some of the new strains of M.tuberculosis are resistant to isoniazid, rifampicin and streptomycin - the front line drugs against the disease [3,4]. Hence there is an urgent need for new antimycobacterials to combat multi-drug resistant strains. A thorough understanding of potential molecular targets for antitubercular agents is necessary to

---

Abbreviations: aa-amino acid; MW-molecular weight; nt-nucleotides; ORF-Open Reading Frame; SD-Shine-Dalgarno;

*For correspondence:
V.NAGARAJA
Centre for Genetic Engineering
Indian Institute of Science
Bangalore, 560 012, INDIA
Fax: 91-80-3341683
Telex: 91-80-3344668
E-mail: vraj@cge.iisc.ernet.in

1039-9712/94/040651-10$05.00/0

develop new front line drugs. DNA gyrase, with its unique properties and structural features [5-8] becomes an ideal candidate in this direction.

The structure-function and mechanism of action of DNA gyrase has been studied extensively [5-8 and references therein]. The tetrameric protein (A_2B_2) catalyses the ATP dependent negative supercoiling of DNA. Two classes of compounds inhibit DNA gyrase from different bacteria. The quinolones interact with GyrA:DNA complex (gyrase:DNA) and trap the transient covalent enzyme-DNA intermediate [6,8]. The coumarins, on the other hand, bind to GyrB subunit and inhibit the energy dependent negative supercoiling step [6,8]. While the quinolones are proven antibacterials, the coumarins are yet to be developed as drugs [9].

We have reported recently the cloning of gyrA and gyrB genes from M.tuberculosis H37Ra as a first step to use them as targets for new antimycobacterials [10]. The sequence information provided here and strategies discussed could form a basis in the development of new therapeutic agents.

MATERIALS AND METHODS:

Streptomyces sphaeroides $gyrB^r$ probe, pLST182 [11] was a generous gift from A.S. Thiara and E. Cundliffe (Leicester Univ., U.K.). M.tuberculosis H37Rv was provided by the National Tuberculosis Institute, Bangalore. M.tuberculosis H37Ra was cultured as described earlier [10]. Restriction and modifying enzymes were from Boehringer Mannheim, Amersham, and New England Biolabs. Genescribe vectors pTZ18U, pTZ19U and Sequenase version 2.0 were from U.S.Biochemicals. Radionucleotides were from Bhabha Atomic Research Centre. All other reagents were purchased from Sigma.

DNA Techniques: The isolation and restriction endonuclease digestion of genomic DNA from M.tuberculosis strains were as described [10]. After electrophoresis, DNA fragments were transferred on to nylon membrane [12]. The blot was prehybridised with 6X SSC containing 0.4% SDS, 5X Denhardt's solution [13] and 500 μg/ml yeast total RNA for 6 h. The 2.2 Kb BamHI fragment derived from pLST182 was radiolabelled by random priming [13] and hybridised to the blot at 65 ^{o}C for 12 h. The blot was washed at 65 ^{o}C with 0.5X SSC containing 0.1% SDS and autoradiographed. Cloning procedures and other experiments were essentially as described in Sambrook et al. [13]. Dideoxy sequencing was according to Sanger [14]. Fragments in the size range of 0.4-0.8 Kb were subcloned into Genescribe vectors using restriction enzymes which cut the mycobacterial genome frequently. Exonuclease III mediated nested deltions [13] were also created to obtain overlapping clones. Single strand sequencing was according to the suppliers manual using universal primer. The regions of secondary structures and compressions were resolved using 7-deaza dGTP and by sequencing both the strands. The nt sequence was analysed using UWGCG [15] software at the

Bioinformatics Centre, Indian Institute of Science. The sequence is deposited in EMBL data library (EMBL accession no. X78888).

RESULTS AND DISCUSSION:

The genomic Southern blots of M.tuberculosis H37Ra and H37Rv using S.sphaeroides $gyrB^r$ [11] are presented in Fig. 1A. The $gyrB^r$ hybridisable fragments from both the genomes are of same size in different restriction digests. Similar results were obtained with other restriction enzymes. The identical sized fragments (Fig. 1A) from both strains suggest that the

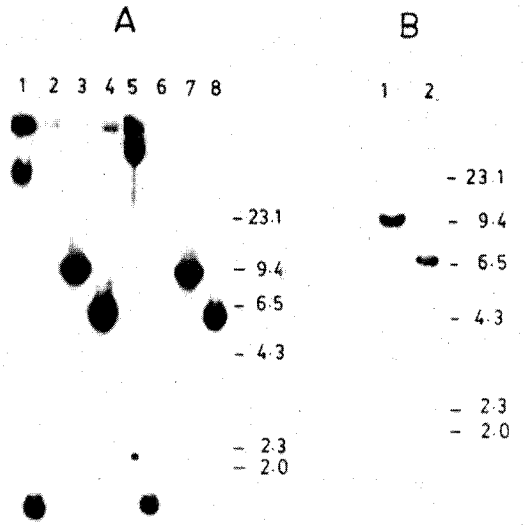


Figure 1: Representative autoradiogram of Southern blots.
A.The genomic DNA from M.tuberculosis H37Ra and H37Rv (lanes 1,5) were digested with BamHI (lanes 2,6), PstI (lanes 3,7), EcoRI (lanes 4,8) and hybridised with 2.2 Kb BamHI fragment from pLST182 [11]. A 4.2 Kb Eco RI fragment was also observed in lanes 4 and 8 at lower stringency [10]. **B**. M.tuberculosis H37Ra DNA digested with KpnI (lane 1) and KpnI+BsiWI (lane 2).

genes for DNA gyrase may occupy same location in the respective genomes. The genomic blot of highly virulent local strain also showed similar results (not shown).

Single hybridising fragment was obtained with KpnI or KpnI+BsiWI digestions of M.tuberculosis H37Ra DNA (Fig. 1B) The 7 Kb KpnI+BsiWI fragment was cloned into pUC19 to obtain pMN7KB. The plasmids pMN7KB, pMN13R and pMN6R [10] were used for further analysis. The restriction map of these plasmids and the sequencing strategy is given in Fig. 2. The gyrB gene was localised to XhoI-PstI fragment. The sequencing strategy shown includes subclones generated from SalI to NcoI (near the unique XhoI and PstI sites respectively), although partial sequencing from other regions were also carried out.

Fig.3 shows a 2723 nt sequence which includes the longest ORF of 2145 nt (EMBL accession no: X78888), encoding B subunit of DNA gyrase. The gyrB

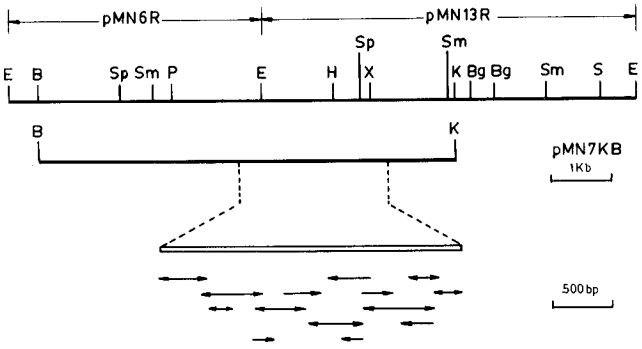


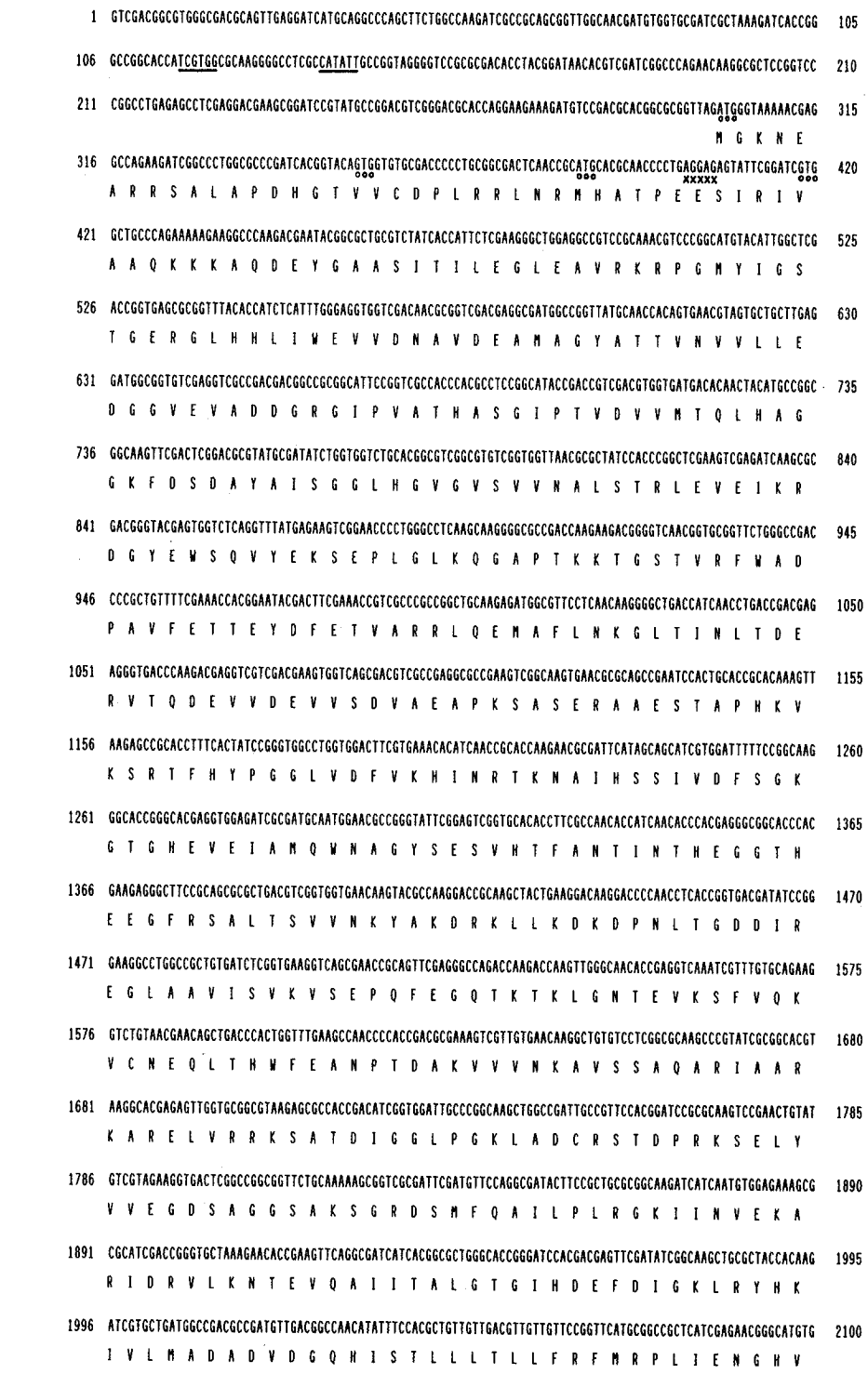
Figure 2: Sequencing strategy and restriction endonuclease maps of plasmids pMN13R, pMN6R and overlapping pMN7KB. Bidirectional arrows indicate sequencing of both the strands. B, BsiWI; Bg, BglII; E, EcoRI; H, HpaI; K, KpnI; P, PstI; S, SacI; Sm, SmaI; Sp, SphI; X, XhoI.

ORF codes for a protein of 714 aa with a calculated MW of 78.4 kDa and a pI of 6.6. Analysis of the ORF shows the presence of two ATG at positions 301, 385 and two GTG at positions 352, 418 towards N-terminus. One of these may serve as initiation codon. A near perfect SD sequence at 401 makes the GTG at 418 a strong candidate. This would result in a protein of 675 aa. The N-terminal analysis of GyrB protein is necessary to confirm the translational start point.

Examination of 5' end of gyrB shows the presence of regulatory elements. The putative -10 and -35 sequences similar to other characterised promoters from this organism [16] are located about 150 base pairs upstream of the first ATG. The gyrA ORF begins 34 nt downstream of gyrB stop codon. A potential SD sequence for gyrA is located 5 nt upstream of the ATG. We have not found terminator stem-loop structure beyond the gyrB stop codon. Further, no promoter elements have been detected within the 3' end of the gyrB gene or in the short intergenic region. These observations suggest that both the genes are transcribed as a single transcription unit initiating upstream of gyrB.

The codon usage for gyrB gene is given in Table I. The overall codon usage pattern is similar to other genes characterised from this organism [16]. Among the 714 codons used, CCA (P) is not represented at all, while TTA (L), TCA (S), CCT (P), ACT (T), AAT (N), TGT (C), AGG (R), GGA (G) are represented only once. Codons such as CTT (L), GTA (V), ACA (T) are poorly represented. The most abundant ones are CTG (L), ATC (I), ACC (T) CGC (R) and GGC (G).

The nt sequence of gyrB of M.tuberculosis was compared with that of E.coli [17] and S.sphaeroides [18] (Fig. 4). The gyrB nt sequence has a high degree of homology with S.sphaeroides but less homology with E.coli. This result is consistent with different codon usage between the genes and G+C



```
2101 TTTTTGGCACAACCGCCGCTGTACAAACTCAAGTGGCAGCGCAGTGACCCGGAATTCGCATACTCCGACCGCGAGCGCGACGGTCTGCTGGAGGCGGGGCTGAAG  2205
     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  E  A  G  L  K

2206 GCCGGGAAGAAGATCAACAAGGAAGACGGCATTCAGCGGTACAAGGGTCTAGGTGAAATGGACGCTAAGGAGTTGTGGGAGACCACCATGGATCCCTCGGTTCGT  2310
     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  W  E  T  T  M  D  P  S  V  R

2311 GTGTTGCGTCAAGTGACGCTGGACGACGCCGCCGCCGCCGACGAGTTGTTCTCCATCCTGATGGGCGAGGACGTCGACGCGCGGCGCAGCTTTATCACCCGCAAC  2415
     V  L  R  Q  V  T  L  D  D  A  A  A  A  A  D  E  L  F  S  I  L  M  G  E  D  V  D  A  R  R  S  F  I  T  R  N

2416 GCCAAGGATGTTCGGTTCCTGGATGTCTAACGCAACCCTGCGTTCGATTGCAAACGAGGAATAGATGACAGACACGACGTTGCCGCCTGACGACTCGCTCGACCG  2520
     A  K  D  V  R  F  L  D  V  *                                   M  T  D  T  T  L  P  P  D  D  S  L  D  R

2521 GATCGAACCGGTTGACATCGAGCAGGAGATGCAGCGCAGCTACATCGACTATGCGATGAGCGTGATCGTCGGCCGCGCGCTGCCGGAGGTGCGCGACGGGCTCAA  2625
     I  E  P  V  D  I  E  Q  E  M  Q  R  S  Y  I  D  Y  A  M  S  V  I  V  G  R  A  L  P  E  V  R  D  G  L  K

2626 GCCCGTGCATCGCCGGGTGCTCTATGCAATGTTCGATTCCGGCTTCCGCCCGGACCGCAGCCACGCCAAGTCGGCCCGGTCGGTTGCCGAGACCATGG  2723
     P  V  H  R  R  V  L  Y  A  M  F  D  S  G  F  R  P  D  R  S  H  A  K  S  A  R  S  V  A  E  T  M
```

Figure 3: Nt sequence of gyrB gene and flanking regions. Translated aa sequence is shown below the DNA sequence. The possible start and stop codons are indicated by open (ooo) and closed (●●●) circles respectively. The SD sequences are marked with xxxxx. Promoter consensus is underlined.

Table 1. Codon usage frequency in GyrB subunit of M.tuberculosis.

CODON	n	n/t	CODON	n	n/t	CODON	n	n/t	CODON	n	n/t
TTT-F	6	0.27	TCT-S	4	0.10	TAT-Y	6	0.40	TGT-C	1	0.34
TTC-F	16	0.73	TCC-S	10	0.24	TAC-Y	9	0.60	TGC-C	2	0.66
TTA-L	1	0.01	TCA-S	1	0.02	TAA-*	1	1.00	TGA-*	0	0.00
TTG-L	11	0.20	TCG-S	15	0.36	TAG-*	0	0.00	TGG-W	7	1.00
CTT-L	2	0.03	CCT-P	1	0.04	CAT-H	5	0.25	CGT-R	7	0.15
CTC-L	9	0.16	CCC-P	9	0.40	CAC-H	15	0.75	CGC-R	21	0.46
CTA-L	5	0.09	CCA-P	0	0.00	CAA-Q	10	0.50	CGA-R	2	0.04
CTG-L	28	0.50	CCG-P	13	0.56	CAG-Q	10	0.50	CGG-R	13	0.28
ATT-I	8	0.23	ACT-T	1	0.02	AAT-N	1	0.04	AGT-S	3	0.07
ATC-I	24	0.70	ACC-T	32	0.71	AAC-N	22	0.96	AGC-S	9	0.21
ATA-I	3	0.05	ACA-T	3	0.07	AAA-K	10	0.20	AGA-R	2	0.04
ATG-M	13	1.00	ACG-T	9	0.20	AAG-K	39	0.80	AGG-R	1	0.02
GTT-V	9	0.14	GCT-A	6	0.09	GAT-D	12	0.24	GGT-G	17	0.29
GTC-V	21	0.33	GCC-A	31	0.45	GAC-D	38	0.76	GGC-G	29	0.50
GTA-V	2	0.03	GCA-A	10	0.14	GAA-E	21	0.39	GGA-G	1	0.02
GTG-V	32	0.50	GCG-A	22	0.32	GAG-E	33	0.61	GGG-G	11	0.19

n= usage of individual codon
t= total usage of synonymous codons

content of the two organisms. Next, the derived polypeptide sequences of gyrB from M.tuberculosis, S.sphaeroides and E.coli were aligned. Dot matrix comparison at the aa level reveals significant homology with that of S.sphaeroides (79%) and E.coli (70%). An aa sequence comparison of known GyrB sequences shows significant sequence similarity. The conservation is more towards the C-terminus, but a stretch of 27 aa towards the N-terminus

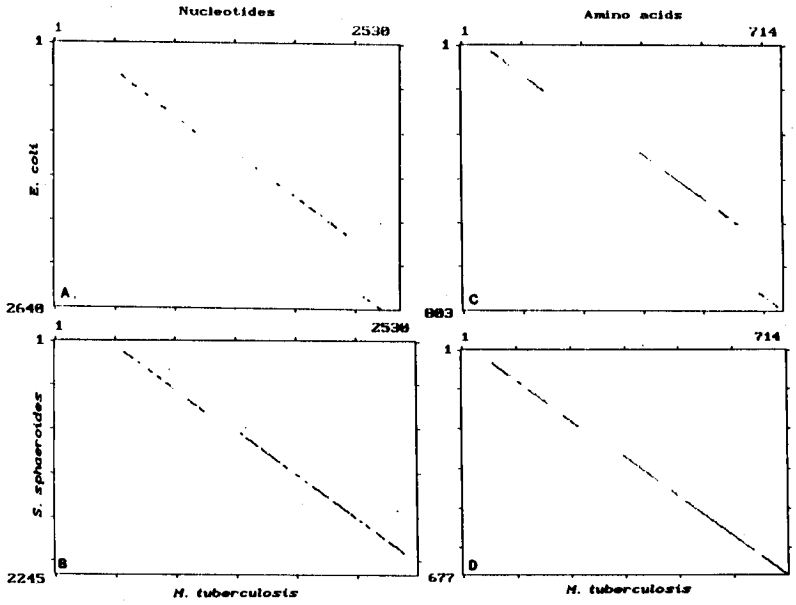


Figure 4: Dot Plot analysis of gyrB nt and aa sequences from M.tuberculosis, S.sphaeroides and E.coli. Comparisons were made with a window size of 30 nt at a stringency of 21 (A and B) and a window size of 20 aa at a stringency of 12 (C and D) using DNASIS and PROSIS.

from diverse species are conserved to near identity [10]. However, a non-homologous region exists towards the middle (closer to the N-terminus) in all the three genes.

The N-terminus regions of various GyrA polypeptides [19-24] are aligned in Fig. 5. The 29 aa stretch from different species shows a high degree of conservation. Very few conservative substitutions are observed. This near identical stretch in diverse species of bacteria underlines the remarkable structural similarity of the enzyme. The importance of this sequence is not known as the active site region is further downstream in all gyrA polypeptides. It may as well be a part of the protein, retained to keep the structural framework. We have not completed the sequencing of gyrA as yet, but anticipate homology in other regions including the active site.

The essential role of DNA gyrase and the information available on struc-

Organism	Sequence	Reference
E.coli	22 LDYAMSVIVGRALPDVRDGLKPVHRRVLY	[19]
K.pneumoniae	22 LDYAMSVIVGRALPDVRDGLKPVHRRVLY	[20]
M.pneumoniae	34 MEYAMSVIVARALPDARDGLKPVHRRVLY	[21]
S.aureus	23 LDYAMSVIVARALPDVRDGLKPVHRRILY	[22]
B.subtilis	23 LDYAMSVIVSRALPDVRDGLKPVHRRILY	[23]
R.prowazekii	23 LDYAMSVIVSRAIPDVRDGLKPVHRRIIY	
Haloferax sps.	33 IDYAMSVIAGRALPDVRDGLKPVHRRILY	[24]
M.tuberculosis	29 IDYAMSVIVGRALPEVRDGLKPVHRRVLY	[this work]

Figure 5: Alignment of conserved amino terminus region of the GyrA polypeptides from eight different species. R.prowazekii gyrA sequence was obtained from GenBank (accession no. U02931).

ture and function makes it an ideal target for new antibacterial drugs. Further, among all the known topoisomerases only DNA gyrase can introduce negative superhelical turns into DNA, a feature not shared by any eukaryotic enzyme. Apart from the well known quinolone and coumarin compounds, the E.coli DNA gyrase is poisoned by two 'foreign' DNA encoded proteins. While the mini F plasmid coded CcdB interacts with GyrA, the microcin B17 produced by plasmid MccB17 binds to GyrB. It has been suggested that these proteins inhibit DNA gyrase activity by direct interaction. The potential of both the polypeptides in the development of new drug antibacterials has been recognised [25, 26]. Thus peptide fragments engineered from these proteins could also serve as powerful inhibitors of DNA gyrase from other bacteria along with quinolones and coumarins. Overexpression of M.tuberculosis DNA gyrase to produce the enzyme in sufficient quantities would allow us to carry out enzyme inhibition studies.

While this manuscript was in preparation, sequencing of the same genes was reported [27]. Although the same probe was used to detect the genes, Takif et al. [27] reported the isolation of genes from a cosmid library while we have cloned from genomic DNA restriction digests. The results presented

here agree with their findings in general. However, there are a few minor differences in the sequences reported. This variation could be due to sequencing artifacts or the differences in strains employed to clone the genes. Takiff et al. have isolated the genes from M.tuberculosis H37Rv while we report from M.tuberculosis H37Ra.

Acknowledgements: We thank E. Cundliffe and A.S Thiara for providing us pLST182 carrying $gyrB^r$ gene of S.sphaeroides; C. Mahesh Babu, H. V. Jayashree and H.S. Rajeshwari for help in some of the experiments. The data were analysed at the Bioinformatics Centre, Indian Institute of Science. Thanks are also due to U. Varshney for helpful discussions and K.P. Gopinathan for comments on the manuscript. Supported by a grant from Council for Scientific and Industrial Research (CSIR). K.M. and V.R. are Senior Research Fellows of CSIR.

REFERENCES:

1. Bloom, B.R., and Murray, L.J.C. (1992) Science 257, 1055-1064.
2. Young, D.B., and Cole, S.T. (1993) J. Bacteriol. 175, 1-6.
3. Collins, M.F.(1993) Critical Rev. Microbiol. 19, 1-16.
4. Snider, D.E., and Roper, W.L. (1992) New Engl. J. Med. 326, 703-705.
5. Cozzarelli, N.R. (1980) Science 207, 953-960.
6. Maxwell, A. and Gellert, M.(1986) Adv.Protein Chem. 38,69-107.
7. Wang, J.C., (1991) J. Biol. Chem. 266, 6659-6662.
8. Reece, R.J., and Maxwell, A. (1991) Critical Rev. Biochem. Mol. Biol. 26, 335-375.
9. Maxwell, A. (1993) Mol. Microbiol. 9, 681-686.
10. Madhusudan,K., Ramesh,V., and Nagaraja,V. (1994) Curr. Science,66,664-667.
11. Thiara, A.S., and Cundliffe, E. (1988) EMBO J. 7, 2255-2259.
12. Reed, K.C., and Mann, D.A., (1985) Nucleic Acids Res. 13, 7207-7221.
13. Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular Cloning, Cold Spring Harbor Laboratory Press, New York.
14. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467.
15. Devereux, J., Haeberli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395.
16. Dale, J.W., and Patki, A. (1990) in Molecular Biology of Mycobacteria (McFadden, J., Ed.), pp.173, Academic Press, London.
17. Adachi, T., Mizuuchi, M., Robinson, E. A., Apella, E., O'Dea, M.H., Gellert, M., and Mizuuchi, K. (1987) Nucleic Acids Res. 15, 771-775.
18. Thiara, A.S., and Cundliffe, E. (1993) Mol. Microbiol. 8, 495-506.
19. Swanberg,S.L., and Wang,J.C.(1987) J.Mol.Biol. 197, 729-736.
20. Dimri, G.P., and Das, H.K. (1990) Nucleic Acids Res. 18, 151-156.
21. Colman,S.D., Hu,P.C., and Bott, K.F.(1990) Mol. Microbiol. 4, 1129-1134.
22. Hopewell,R., Oram,M., Briesewitz,R., and Fisher,L.M. (1990) J.Bacteriol. 172, 3481-3484.
23. Moriya, S., Ogaswara, N., and Yoshikawa,H. (1985) Nucleic Acids Res. 13, 2251-2255.
24. Holmes, M.L., and Dyall-Smith,M. (1991) J.Bacteriol. 173, 642-648.
25. Bernard,P., Kezdy,E.K., Melderen,V.L. Steyaert, J., Wyns,L., Pato, M.L., Higgins,P.N., and Couturier,M. (1993) J.Mol.Biol. 234, 534-541.
26. Yorgey,P., Davagnino,J., and Kolter,R. (1993) Mol.Microbiol.9, 897-905.
27. Takiff,H.E., Salazar,L., Guerrero,C., Philipp,W., Huang, W.M., Kreiswirth, B., Cole, S.T., Jacobs Jr, W.R., and Telenti, A. (1994) Antimicrob. Agents Chemother. 38, 773-780.