JOURNAL OF BACTERIOLOGY, Sept. 1994, p. 5835–5842
0021-9193/94/$04.00+0

Vol. 176, No. 18

Sequencing Analysis Reveals a Unique Gene Organization in the *gyrB* Region of *Mycoplasma hominis*

SØREN A. LADEFOGED* AND GUNNA CHRISTIANSEN

Institute of Medical Microbiology, University of Aarhus, DK-8000 Aarhus C, Denmark

Received 14 March 1994/Accepted 12 July 1994

The homolog of the *gyrB* gene, which has been reported to be present in the vicinity of the initiation site of replication in bacteria, was mapped on the *Mycoplasma hominis* genome, and the region was subsequently sequenced. Five open reading frames were identified flanking the *gyrB* gene, one of which showed similarity to that which encodes the LicA protein of *Haemophilus influenzae*. The organization of the genes in the region showed no resemblance to that in the corresponding regions of other bacteria sequenced so far. The *gyrA* gene was mapped 35 kb downstream from the *gyrB* gene.

DNA gyrase is a type II topoisomerase that catalyze the conversion of relaxed duplex DNA to a superhelical form. It is thought to have an essential role in DNA replication, to enhance transcription, and to be involved in DNA recombination and repair (6). DNA gyrase is a tetrameric molecule composed of two A and two B subunits, which are encoded by the *gyrA* and *gyrB* genes, respectively. DNA gyrase was first isolated from *Escherichia coli* and has since been detected in a number of bacterial species (24). Because of the apparently ubiquitous existence of DNA gyrases in microbial species and the high degree of sequence conservation among genes encoding them, much current research is focused on DNA gyrases as a target for antimicrobial agents (24). In *Bacillus subtilis* (17), *Salmonella typhimurium* (16), and probably *Pseudomonas putida* (5), the chromosomal replication origin is found in the *dnaA* region. In all of the prokaryotes analyzed so far, the genes of the *dnaA* region are linked to the *gyrB* gene (5, 7, 13, 14, 16, 17, 20, 26). All of the microorganisms that have been analyzed share a remarkable similarity both in the structure of the individual genes and in their relative organization. To test if this highly important region was conserved in *Mycoplasma hominis*, we cloned and sequenced the region surrounding the *gyrB* gene.

A *Bam*HI linking clone (pBMhHB-3) used in mapping studies of the *M. hominis* genome (12) was shown by sequence analysis to contain part of the *M. hominis* PG21 *gyrB* gene. This plasmid was used to identify clones containing the entire gene, including the flanking regions, from a *Bgl*II and *Eco*RI library of *M. hominis* PG21 DNA made in pBluescript SK+ (Stratagene, La Jolla, Calif.). Two clones, pBMhB507C and pE8, covering a span of 14 kb were picked for subcloning and sequenced as described by Hattori and Sakaki (9). A span of 7,246 bp in the *gyrB* region was sequenced and numbered as shown in Fig. 1. The sequence data were analyzed with the Genetics Computer Group Sequence Analysis Software Package, Version 7.1-UNIX (4). Possible open reading frames (ORFs) of more than 40 amino acids are shown in Fig. 2A. The small ORFs located within larger ORFs were excluded, and the remaining six ORFs were considered for further analysis. ORF219, ORF648, ORF249, ORF499, and ORF268 had the same direction of transcription, while the transcription direction of partially sequenced ORF445 was the opposite. Only the 3′ end of ORF445 was sequenced, since the 5′ end, including the initiation codon, was beyond the cloned fragments.

Amino acid sequences were deduced from the nucleotide sequences of all six ORFs, as shown in Fig. 1. Searches of the National Biomedical Research Foundation and European Molecular Biology Laboratory databases were performed. On the basis of similarity, ORF648 was found to correspond to the *E. coli gyrB* gene. The *gyrB* coding sequence is 1,944 bp long and encodes a protein of 648 amino acids, corresponding to a molecular mass of 72.7 kDa. The predicted protein shares 55% identity with *B. subtilis* gyrase B, 52% identity with the *Mycoplasma pneumoniae* protein, and 48.9% identity with *E. coli* gyrase B. The homology search also revealed considerable similarity between *M. hominis* GyrB and the ParE polypeptide of *E. coli* (39.9%). ParE is a subunit of the type IV topoisomerase, which is supposed to anchor chromosomes on membranes (10). The similarity to *E. coli* ParE was significantly less than that to *E. coli* GyrB, and some of the amino acid variations were at positions conserved in all known GyrB polypeptides. DNA hybridization analysis using probes specific for *M. hominis gyrB* on chromosomal DNA did not show hybridization to any additional fragments. This indicates that a *parE* gene in *M. hominis*, if present, has a low level of identity to *M. hominis gyrB*.

ORF249 exhibited 42% identity to a putative protein in *Mycoplasma capricolum* called LicA, which has an unknown function (15), and 28% identity to the LicA protein of *Haemophilus influenzae*. The latter is supposed to be necessary for expression of the outer membrane lipopolysaccharide of *H. influenzae* (30). The specific enzymatic function of the gene is unresolved, but it is involved in phase variation in the lipopolysaccharide. Mycoplasmas are closely related to gram-positive bacteria and have no cell wall, and ORF249 in *M. hominis* and *licA* in *M. capricolum* are, accordingly, not involved in lipopolysaccharide synthesis. The similarity may reflect a different but related enzymatic function in mycoplasmas.

The remaining ORFs (ORF219, ORF499, ORF268, and ORF445) were compared with the sequence databases by using the FASTA and TFASTA programs. No significant similarity to any of the sequences was found. In an 8-kb region upstream from ORF219, a total of 3 kb of noncontinuous DNA fragments were sequenced (data not shown). ORFs were likewise identified and compared to the sequence databases, and no similarities were found.

In all of the bacteria so far analyzed, the genetic organiza-

* Corresponding author. Mailing address: Institute of Medical Microbiology, Bartholin Building, University of Aarhus, DK-8000 Aarhus C, Denmark. Phone: 4589421746. Fax: 4586191277. Electronic mail address: sorenl@biobase.aau.dk.

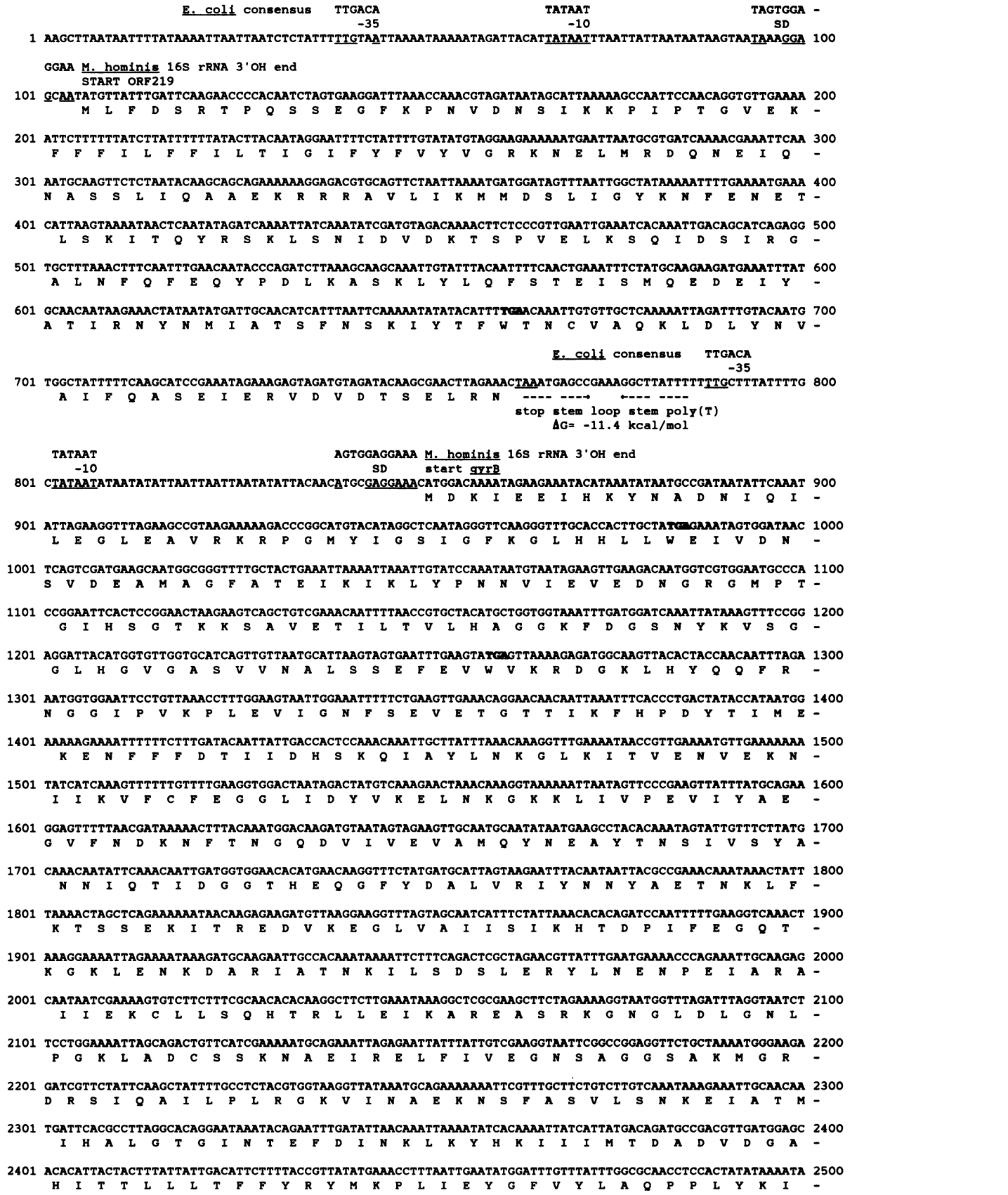

FIG. 1. Nucleotide sequence of the *gyrB* region. The deduced amino acid sequence is shown below the DNA sequence. Possible promoter and Shine-Dalgarno (SD) sequences are underlined and compared with the corresponding sequences in *E. coli* and the 16S rRNA of *M. hominis* (8), respectively. Tentative stem-and-loop structures are shown by arrows, and the free energy calculated as described by Tinoco et al. (28) is indicated. UGA_{Trp} codons are in boldface. Over the deduced amino acid sequence for ORF445, the noncoding strand is shown.

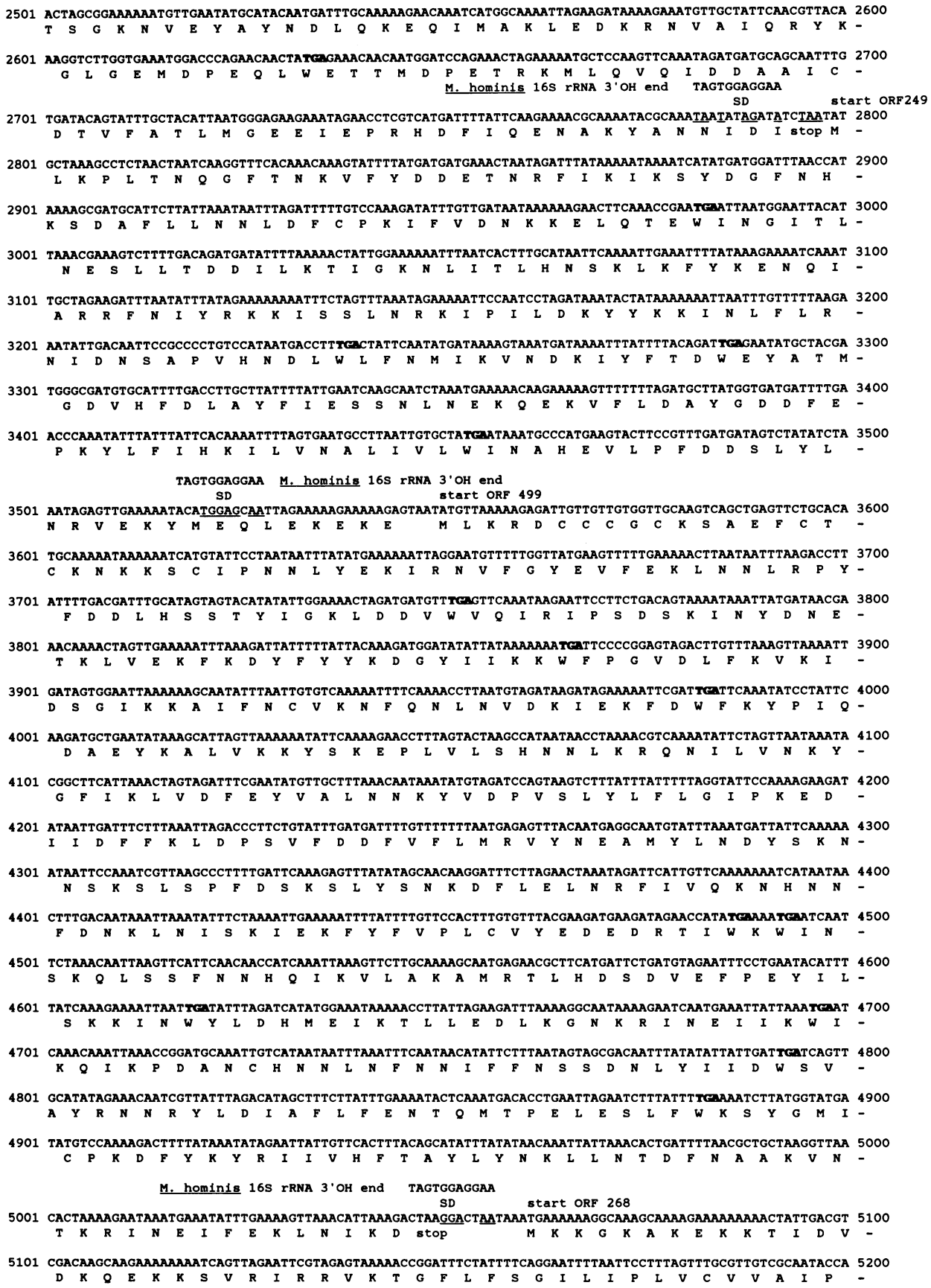

FIG. 1—*Continued.*

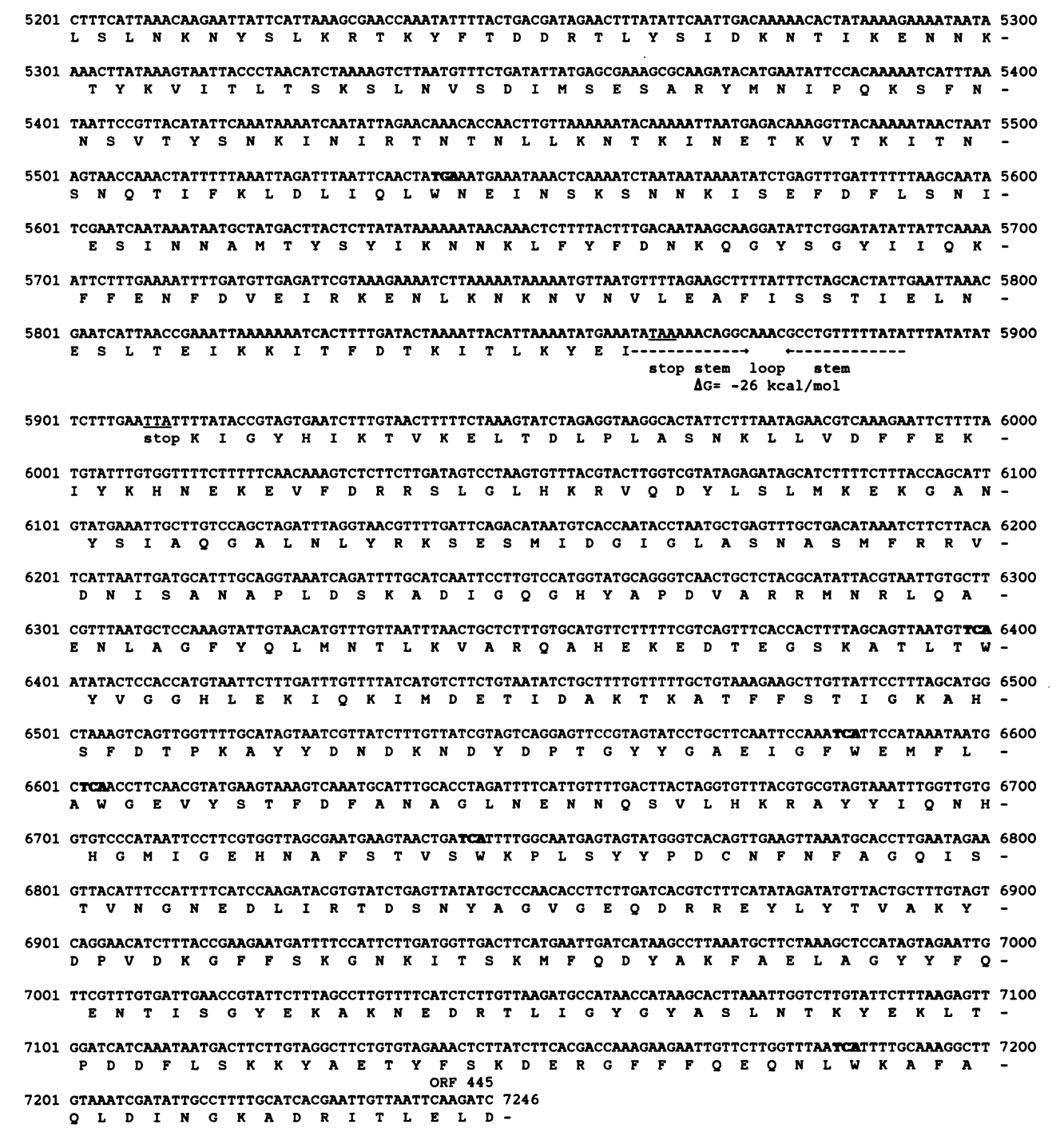

FIG. 1—*Continued.*

tion upstream of the *gyrB* gene is highly conserved and the genes are organized in the following order, in *E. coli* (7), *B. subtilis* (17), *P. putida* (5), *Pseudomonas mirabilis* (26), *S. typhimurium* (16), *Staphylococcus aureus* (14), and *Buchnera aphidicola* (13): *rnpA-rpmH-dnaA-dnaN-recF-gyrB*. An exception is present in *Borrelia burgdorferi*, where *dnaA* and *dnaN* are reversed. In *M. capricolum*, the arrangement *rnpA-rpmH-dnaA-dnaN* is identical to that of other bacteria but the *gyrB* gene has not been identified (15). To localize the *dnaA* gene relative to *gyrB* on the chromosomal map of *M. hominis*, clones from *E. coli* (29) and *Mycoplasma genitalium* (22) containing the *dnaA* gene were used in Southern hybridizations under low-stringency conditions but it was not possible to obtain any signal. An *M. genitalium gyrA* probe (see below) was applied as a positive control.

The *gyrB* and *gyrA* genes in *E. coli* (16) and *P. putida* (21) are localized several kilobases apart on the chromosome, but in most bacteria, including *S. aureus*, *B. subtilis*, *Haloferax* sp.

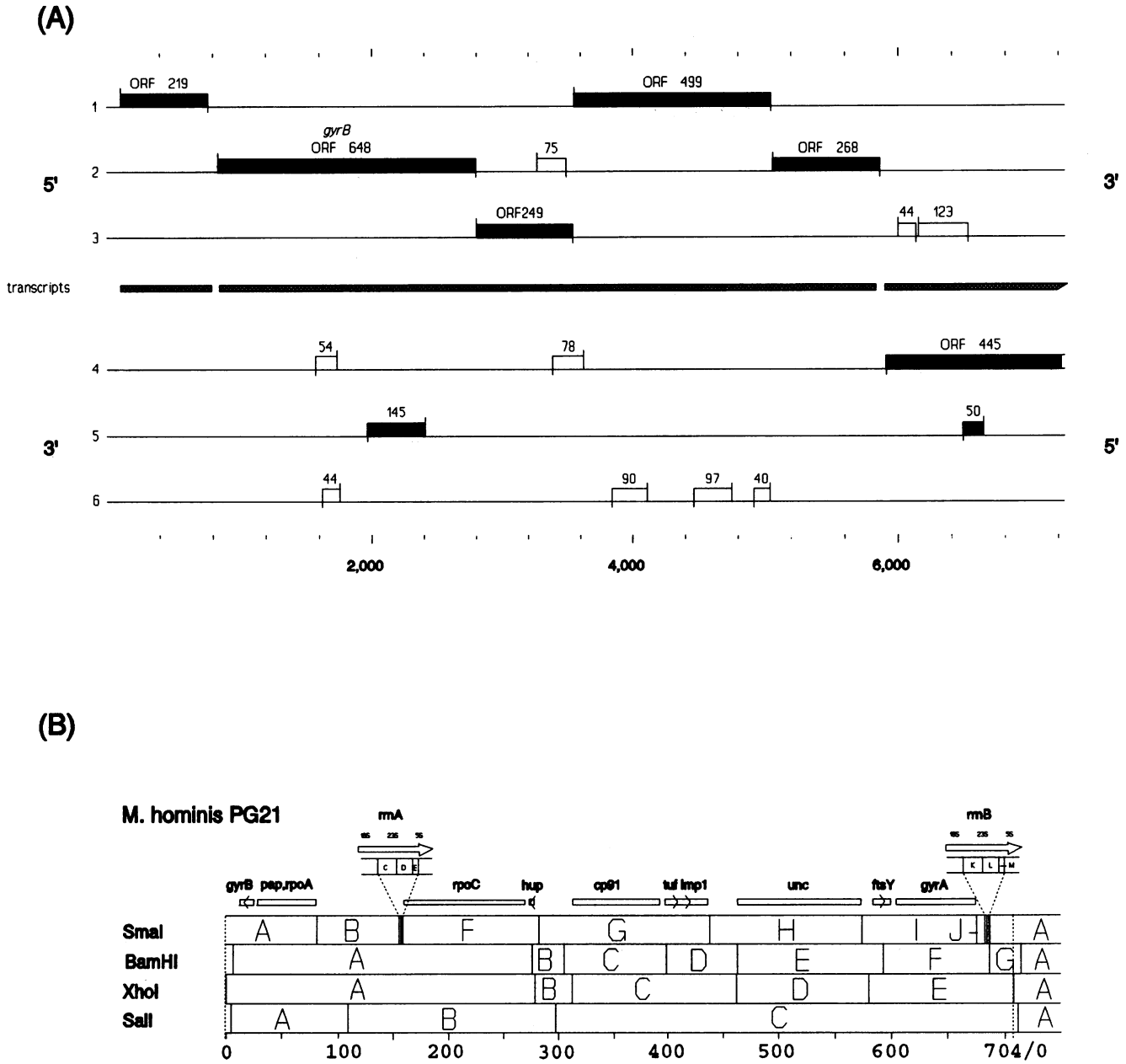

FIG. 2. (A) ORFs in the *gyrB* region. All possible coding frames which code for more than 40 amino acids are shown. The frames with typical Shine-Dalgarno sequences are indicated by black blocks. The numbers of amino acids in the ORFs are shown with the numbers of amino acids giving the names of the ORFs. Transcript lengths are shown by hatched bars. (B) Physical and genetic map of *M. hominis* PG21. *Sma*I, *Bam*HI, *Xho*I, and *Sal*I restriction sites are indicated. Map units are in kilobases. The positions of the functional loci are indicated as precisely as allowed by the resolution of the map.

strain AA 2.2, *B. burgdorferi*, *M. pneumoniae* (24), and *M. genitalium* (22), *gyrA* and *gyrB* are coupled. DNA sequence analysis of the region downstream of the *gyrB* gene in *M. hominis* indicated that it was not coupled with *gyrA*. *gyrB* has been mapped on the *M. hominis* PG21 genome by pulsed-field gel electrophoresis (12). By using a DNA fragment containing most of the *M. genitalium gyrA* gene (23) as a probe in hybridization analysis, the *gyrA* gene in *M. hominis* PG21 was mapped and shown to be located at least 35 kb upstream of the *gyrB* gene (Fig. 2B).

Comparison of amino acid sequences deduced from the DNA sequence indicates that the chromosomal organization of the *dnaA* region is highly conserved in prokaryotes (5, 14, 16, 26). This apparent conservation is remarkable and indicative of biological significance. One element of the conservation in the *dnaA* region could be the need for coordinated expression of the genes, most of which are essential for DNA metabolism. From the presented sequencing data and mapping studies of the *gyrB* region of *M. hominis*, it seems possible that the *dnaA* region in some mycoplasmas is not as strictly conserved as in other organisms analyzed so far. Conservation of gene order has been recognized in the *spc* operon of *E. coli*,

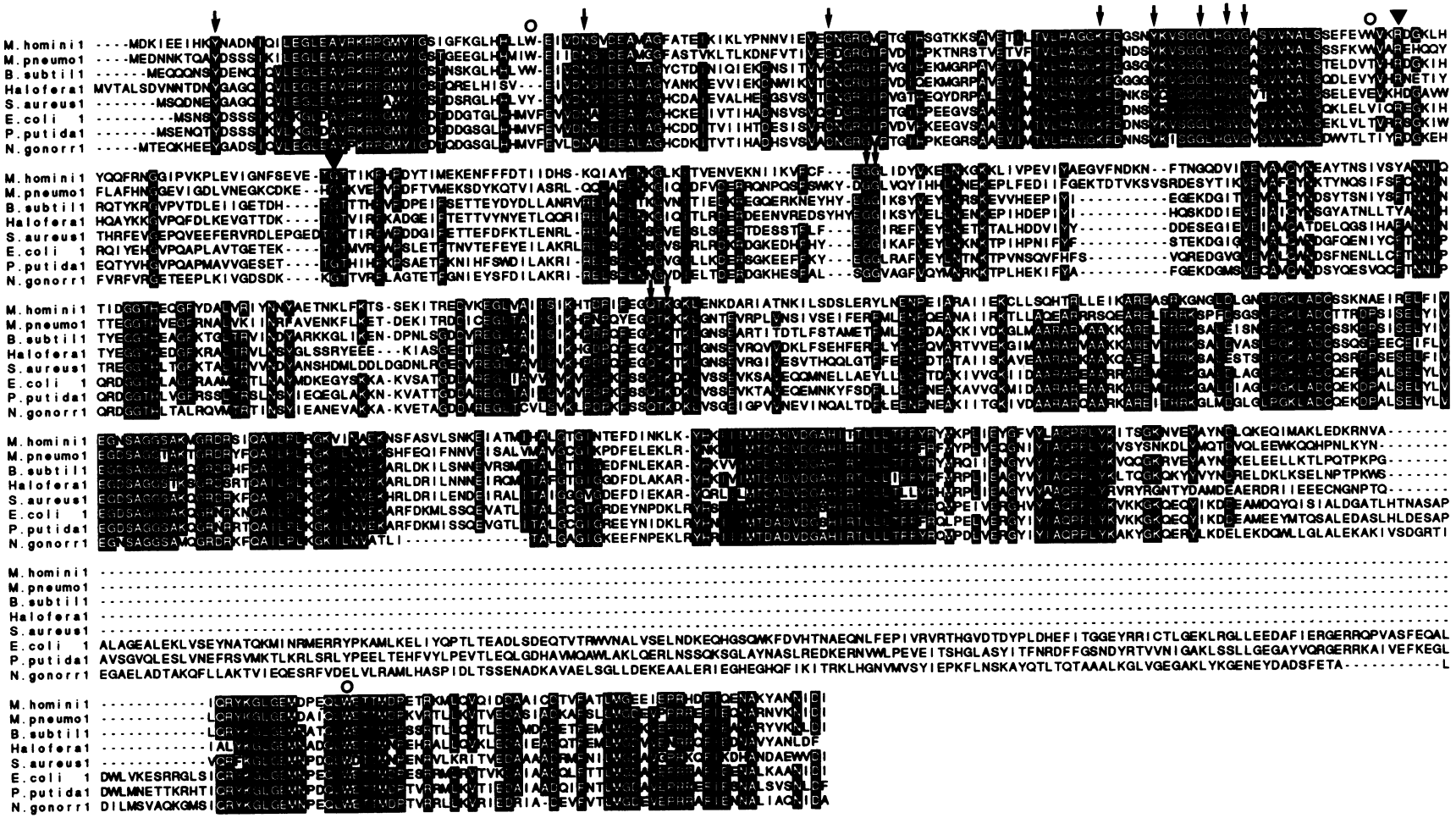

FIG. 3. Multiple alignment of gyrase B sequences from *M. hominis* (this report), *M. pneumoniae*, *B. subtilis*, *Haloferax* sp., *S. aureus*, *E. coli*, *P. putida*, and *N. gonorrhoeae* (14, 24). Reverse-printed letters indicate residues identical in more than five of the eight organisms. The open circles denote the locations of tryptophan (W) residues in *M. hominis*. The arrows indicate residues important in ATP binding and hydrolysis in *E. coli* (27, 31). The triangle indicates the two residues shown to be involved in coumarin resistance in *E. coli*.

B. subtilis, *M. capricolum*, and *Micrococcus luteus* (16, 18, 19). Unlike in these bacteria, in *M. hominis* the *tuf* gene is not part of the *spc* operon (11). Gene rearrangement may thus be a more pronounced phenomenon in *M. hominis* than in the organisms analyzed so far.

Northern (RNA) blot hybridization of *M. hominis* RNA was performed by standard methods (2, 25) to identify the corresponding transcripts of the identified ORFs. A DNA probe specific for ORF219 hybridized to a 0.64-kb RNA fragment. Probes specific for ORF648 (*gyrB*), ORF249, ORF499, and ORF268 all hybridized to a 5.2-kb fragment. ORF445 hybridized to a 3.2-kb fragment. On the basis of these results, the six genes in the *gyrB* region are divided into three transcriptional units. Putative promoters and termination signals were found for each transcript, as shown in Fig. 1.

A characteristic feature of several mycoplasma species is the use of UGA to encode tryptophan (Trp) rather than translational termination (32). We found one UGA codon within ORF219, three in ORF648 (*gyrB*), four in ORF249 (*licA*), nine in ORF499, one in ORF268, and five in ORF445. None of the tryptophans was encoded by the universal Trp codon (UGG). As shown in Fig. 3, which compares gyrase B sequences, UGA codons occur at Trp sites in the corresponding proteins of other bacteria. One occurs where all of the other organisms have Trp (Trp-595), one occurs where *B. subtilis* and *M. pneumoniae* have Trp (Trp-45), and one occurs where *M. pneumoniae* has Trp (Trp-137). All of these organisms use the universal Trp codon UGG at the indicated locations. This indicates that UGA is a codon for Trp in *M. hominis*. Whether *M. hominis* has a $tRNA^{Trp}$ (CCA) in accordance with the universal genetic code cannot be interpreted from these data.

The *E. coli gyrB* sequence was published in 1987 (1), and now several genes have been sequenced, which makes a comprehensive comparison of *gyrB* sequences possible. In Fig. 3, the amino acid sequences deduced from the nucleotide sequences of eight *gyrB* genes are compared. Identical amino acid residues are printed in reverse. Like *M. pneumoniae*, *B. subtilis*, *S. aureus*, and *Haloferax* sp. gyrases B, the *M. hominis* protein lacks an internal 180-amino-acid stretch found toward the C-terminal end of the gram-negative bacterial GyrB protein (Fig. 3). The function of this region is unknown. Recent crystallographic analysis of the 393 N-terminal amino acids of *E. coli gyrB* has shown that this region hydrolyzes ATP, and furthermore, a number of critical residues that interact with the nucleotides have been identified (31). These sites are all conserved in the sequenced gyrase genes, including that of *M. hominis* (Fig. 3), indicating the importance of these residues. DNA gyrase is the target for a number of antibacterial agents, including coumarins. Mutations in the *E. coli gyrB* gene at positions 136 (Arg→Cys/His/Ser) and 164 (Gly→Val) confer resistance to high concentrations of coumarins (3). The amino acids in the corresponding positons of *M. hominis* are identical to those in the *E. coli* wild-type sequence (Fig. 3).

Nucleotide sequence accession number. The sequence data presented here will appear in the EMBL data library under accession number X77529.

This work was supported by the Aarhus University Research Foundation, the Danish Medical Research Council (grant 12-0166-1), and the Danish Natural Science Research Council (grant 11-9061-1).

We thank I. Andersen and K. Skovgaard for technical support.

REFERENCES

1. **Adachi, T., M. Mizuuchi, E. A. Robinson, E. Appella, M. H. O'Dea, M. Gellert, and K. Mizuuchi.** 1987. DNA sequence of the *E. coli gyrB* gene: application of a new sequencing strategy. Nucleic Acids Res. **15**:771–784.
2. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
3. **Contreras, A., and A. Maxwell.** 1992. gyrB mutations which confer coumarin resistance also affect DNA supercoiling and ATP hydrolysis by *Escherichia coli* DNA gyrase. Mol. Microbiol. **6**:1617–1624.
4. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
5. **Fujita, M. Q., H. Yoshikawa, and N. Ogasawara.** 1989. Structure of the *dnaA* region of *Pseudomonas putida*: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli* and *P. putida*. Mol. Gen. Genet. **215**:381–387.
6. **Gellert, M.** 1981. DNA topoisomerases. Annu. Rev. Biochem. **50**:879–910.
7. **Hansen, E. B., T. Atlung, F. G. Hansen, O. Skovgaard, and K. von Meyenburg.** 1984. Fine structure genetic map and complementation analysis of mutations in the *dnaA* gene of *Escherichia coli*. Mol. Gen. Genet. **196**:387–396.
8. **Harasawa, R., T. Uemori, K. Asada, I. Kato, and N. Shiragami.** 1992. 'boxA'-like sequence between the 16 S/23 S spacer in *rRNA* operon of mycoplasmas. FEBS Lett. **297**:209–211.
9. **Hattori, M., and Y. Sakaki.** 1986. Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. **152**:232–238.
10. **Kato, J., H. Suzuki, and H. Ikeda.** 1992. Purification and characterization of DNA topoisomerase IV in *Escherichia coli*. J. Biol. Chem. **267**:25676–25684.
11. **Ladefoged, S. A., and G. Christiansen.** 1991. Analysis of the nucleotide sequence of the *Mycoplasma hominis tuf* gene and its flanking region. FEMS Microbiol. Lett. **79**:133–140.
12. **Ladefoged, S. A., and G. Christiansen.** 1992. Physical and genetic mapping of the genomes of five *Mycoplasma hominis* strains by pulsed-field gel electrophoresis. J. Bacteriol. **174**:2199–2207.
13. **Lai, C. Y., and P. Baumann.** 1992. Genetic analysis of an aphid endosymbiont DNA fragment homologous to the *rnpA-rpmH-dnaA-dnaN-gyrB* region of eubacteria. Gene **113**:175–181.
14. **Margerrison, E. E., R. Hopewell, and L. M. Fisher.** 1992. Nucleotide sequence of the *Staphylococcus aureus gyrB-gyrA* locus encoding the DNA gyrase A and B proteins. J. Bacteriol. **174**: 1596–1603.
15. **Miyata, M., K. I. Sano, R. Okada, and T. Fukumura.** 1993. Mapping of replication initiation site in *Mycoplasma capricolum* genome by two-dimensional gel-electrophoretic analysis. Nucleic Acids Res. **21**:4816–4823.
16. **O'Brien, S. J.** 1990. Genetic maps; locus maps of complex genomes. Book 2: bacteria, algae, and protozoa, p. 2.54–2.70. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
17. **Ogasawara, N., S. Moriya, K. von Meyenburg, F. G. Hansen, and H. Yoshikawa.** 1985. Conservation of genes and their organization in the chromosomal replication origin region of *Bacillus subtilis* and *Escherichia coli*. EMBO J. **4**:3345–3350.
18. **Ohama, T., A. Muto, and S. Osawa.** 1989. Spectinomycin operon of *Micrococcus luteus*: evolutionary implications of organization and novel codon usage. J. Mol. Evol. **29**:381–395.
19. **Ohkubo, S., A. Muto, Y. Kawauchi, F. Yamao, and S. Osawa.** 1987. The ribosomal protein gene cluster of *Mycoplasma capricolum*. Mol. Gen. Genet. **210**:314–322.
20. **Old, I. G., D. Margarita, and I. Saint Girons.** 1993. Unique genetic arrangement in the *dnaA* region of the *Borrelia burgdorferi* linear chromosome: nucleotide sequence of the *dnaA* gene. FEMS Microbiol. Lett. **111**:109–114.
21. **Parales, R. E., and C. S. Harwood.** 1990. Nucleotide sequence of the *gyrB* gene of *Pseudomonas putida*. Nucleic Acids Res. **18**:5880.
22. **Peterson, S. N., P. C. Hu, K. F. Bott, and C. A. Hutchison.** 1993. A survey of the *Mycoplasma genitalium* genome by using random sequencing. J. Bacteriol. **175**:7918–7930.
23. **Peterson, S. N., N. Schramm, P. C. Hu, K. F. Bott, and C. A. Hutchison.** 1991. A random sequencing approach for placing markers on the physical map of *Mycoplasma genitalium*. Nucleic Acids Res. **19**:6027–6031.
24. **Reece, R. J., and A. Maxwell.** 1991. DNA gyrase: structure and function. Crit. Rev. Biochem. Mol. Biol. **26**:335–375.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular

cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

26. **Skovgaard, O.** 1990. Nucleotide sequence of a *Proteus mirabilis* DNA fragment homologous to the *60K-rnpA-rpmH-dnaA-dnaN-recF-gyrB* region of *Escherichia coli*. Gene **93:**27–34.
27. **Tamura, J. K., and M. Gellert.** 1990. Characterization of the ATP binding site on *Escherichia coli* DNA gyrase. Affinity labeling of Lys-103 and Lys-110 of the B subunit by pyridoxal 5′-diphospho-5′-adenosine. J. Biol. Chem. **265:**21342–21349.
28. **Tinoco, I., O. Uhlenbec, and M. D. Levine.** 1971. Estimation of secondary structure in ribonucleic acids. Nature (London) **230:** 362–367.
29. **von Meyenburg, K., F. G. Hansen, T. Atlung, L. Boe, I. G. Clausen, B. van Deurs, E. B. Hansen, B. B. Jørgensen, F. Jørgensen, L. Koppes, O. Michelsen, J. Nielsen, P. E. Pedersen, K. V. Rasmussen, E. Riise, and O. Skovgaard.** 1985. Facets of the chromosomal origin of replication, *oriC*, of *Escherichia coli*, p. 260–281. *In* M. Schaechter, F. C. Neidhardt, J. L. Ingraham, and N. O. Kjeldgaard (ed.), Molecular biology of bacterial growth. Jones & Bartlett, Boston.
30. **Weiser, J. N., J. M. Love, and E. R. Moxon.** 1989. The molecular mechanism of phase variation of *H. influenzae* lipopolysaccharide. Cell **59:**657–665.
31. **Wigley, D. B., G. J. Davies, E. J. Dodson, A. Maxwell, and G. Dodson.** 1991. Crystal structure of an N-terminal fragment of the DNA gyrase B protein. Nature (London) **351:**624–629.
32. **Yamao, F., A. Muto, Y. Kawauchi, M. Iwami, S. Iwagami, Y. Azumi, and S. Osawa.** 1985. UGA is read as tryptophan in *Mycoplasma capricolum*. Proc. Natl. Acad. Sci. USA **82:**2306–2309.