JOURNAL OF BACTERIOLOGY, Oct. 1990, p. 6165–6168
0021-9193/90/106165-04$02.00/0

Vol. 172, No. 10

Nucleotide Sequence Analysis of the *Leptospira biflexa* Serovar patoc *rpsL* and *rpsG* Genes

RICHARD L. ZUERNER[1,2]* AND NYLES W. CHARON[1]

Department of Microbiology and Immunology, School of Medicine, Health Sciences Center, West Virginia University, Morgantown, West Virginia 26506,[1] and Leptospirosis and Mycobacteriosis Research Unit, National Animal Disease Center, U.S. Department of Agriculture, Agricultural Research Service, P.O. Box 70, Ames, Iowa 50010[2]

Received 11 December 1989/Accepted 27 July 1990

The *Leptospira biflexa rpsL* and *rpsG* genes were sequenced. Although similar in many respects, proteins encoded by these *L. biflexa* genes had several unusual features when compared with homologous proteins of other organisms. Unlike the *rpsL* genes of other eubacteria, the *L. biflexa rpsL* gene is adjacent to a *rpoC*-like gene.

The genus *Leptospira* is part of a distinct and ancient branch of the eubacteria (5, 21, 30). This genus contains two genetically diverse species: *Leptospira biflexa* (saphrophytic) and *L. interrogans* (pathogenic) (1, 14, 15, 31). Little is known about the genetics of these bacteria. There are no known mechanisms for genetic exchange among members of the genus *Leptospira*, thus precluding the use of classical genetic techniques to study these bacteria. To circumvent these difficulties, we have used molecular cloning techniques to analyze *Leptospira* genes (32, 34). The results of those and other studies suggest that *Leptospira* genes may be organized differently from homologous genes in other bacteria (7, 25, 33, 34). In this report we present an analysis of the nucleotide sequences of two *L. biflexa* genes which exhibit significant sequence similarity to the *Escherichia coli* ribosomal protein genes *rpsL* (encoding ribosomal protein S12) and *rpsG* (encoding ribosomal protein S7).

(A preliminary version of this work was presented previously [R. L. Zuerner and N. W. Charon, Abstr. Annu. Meet. Am. Soc. Microbiol. 1987, H-151, p. 164].)

The *L. biflexa* serovar patoc strain Patoc I *rpsL* and *rpsG* genes were cloned on plasmids pZC300 and pZC301, which have been described previously (34). Both pZC300 and pZC301 were isolated by complementation of an *argE* mutation in *E. coli* AB1157 (34). The *L. biflexa* DNA cloned in pZC300 was subcloned into M13 bacteriophage vectors and sequenced by dideoxynucleotide termination reactions as described previously (34). Three open reading frames (ORFs) were detected within the sequenced DNA. The first ORF was identified as the *argE*-complementing gene (hereinafter referred to as the *argE* gene). The product of the *argE* gene shares amino acid sequence similarity with the *E. coli* β′ subunit of RNA polymerase (a product of the *rpoC* gene) (34). An analysis of the *L. biflexa* DNA comprising the second and third ORFs (ORF2 and ORF3) on pZC300 is presented here. ORF2 and ORF3 were located 3′ to and encoded by the same strand of DNA as *argE* (Fig. 1). Both ORFs were preceded by a potential ribosome-binding sequence, AAGG. ORF2 began 196 base pairs (bp) from the termination codon of *argE*, spanned 372 bp, and had the potential to code for a 13,932-dalton protein. ORF3 began 17 bp 3′ to the termination codon of ORF2, spanned 474 bp, and had the potential to encode an 18,288-dalton protein. Proteins corresponding to those predicted for ORF2 and ORF3 were synthesized in vitro from these cloned *L. biflexa* DNA sequences in coupled transcription-translation reactions (34).

Nucleotide sequences spanning both ORFs were used to search the GenBank and EMBL nucleotide sequence data bases for genes with similar sequences. ORF2 (or its predicted protein product) was similar to all available *rpsL* (or S12) sequences (Fig. 2A). Likewise, ORF3 (or its predicted product) was similar to all deduced *rpsG* (or S7) sequences (Fig. 2B). Because of the similarity between these ORFs and all known *rpsL* (S12) or *rpsG* (S7) sequences, we concluded that ORF2 and ORF3 encoded proteins with functions homologous to those of *E. coli* S12 and S7, respectively. Additionally, Southern blot analysis of *L. biflexa* DNA (data not shown) showed that the cloned sequences are unique and that other *rpsL*- or *rpsG*-like sequences are not present in the *L. biflexa* genome. In *E. coli* the S12 protein is important in translational fidelity (9), whereas the S7 protein is important in ribosome assembly (27). On the basis of our conclusion, these data provide the first information on ribosomal protein gene organization and composition in any member of the genus *Leptospira*.

In all bacteria characterized to date, *rpsL* precedes*rpsG*. This arrangement was observed for the *L. biflexarpsL* and *rpsG* genes. However, the *L. biflexa rpsL* genewas closely linked to a *rpoC*-like gene (34). This genetic linkage was unlike those of other eubacteria but was simi-lar to the organization of the *rpoC* and *rpsL* genes of *Methanococcus vannielii* (17), a member of the archaebacteria. In *M. vannielii*, *rpoC* precedes *rpsL* by a 1,061-bp segment containing two ORFs (17). *Leptospira* spp. share several other unusual phenotypes with archaebacteria, including possession of rifampin-resistant RNA polymerases (12, 18) and an unusual pathway for isoleucine biosynthesis (3, 4).

Several features of the *L. biflexa* S12 and S7 proteinswere of particular interest (Fig. 2). The *L. biflexa* S12protein showed conservation of amino acids at positions44 and 88 (Fig. 2A), which in *E. coli* are altered in streptomycin-resistant mutants (9). The *L. biflexa* S7 protein contained

* Corresponding author.

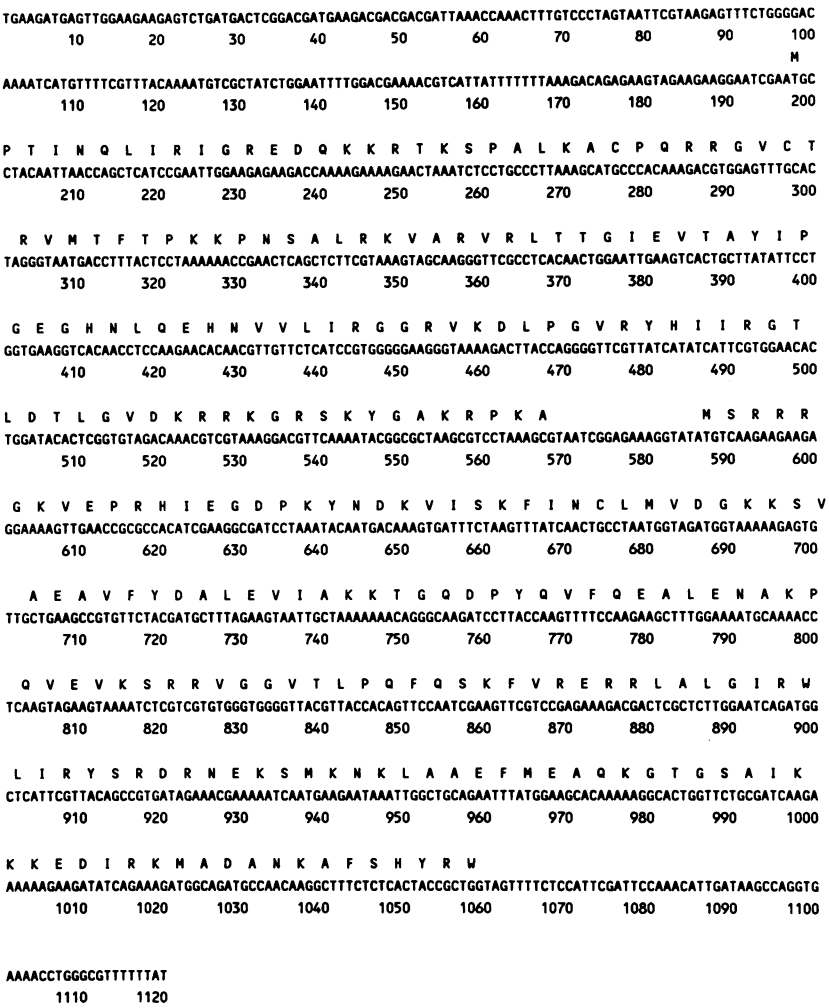
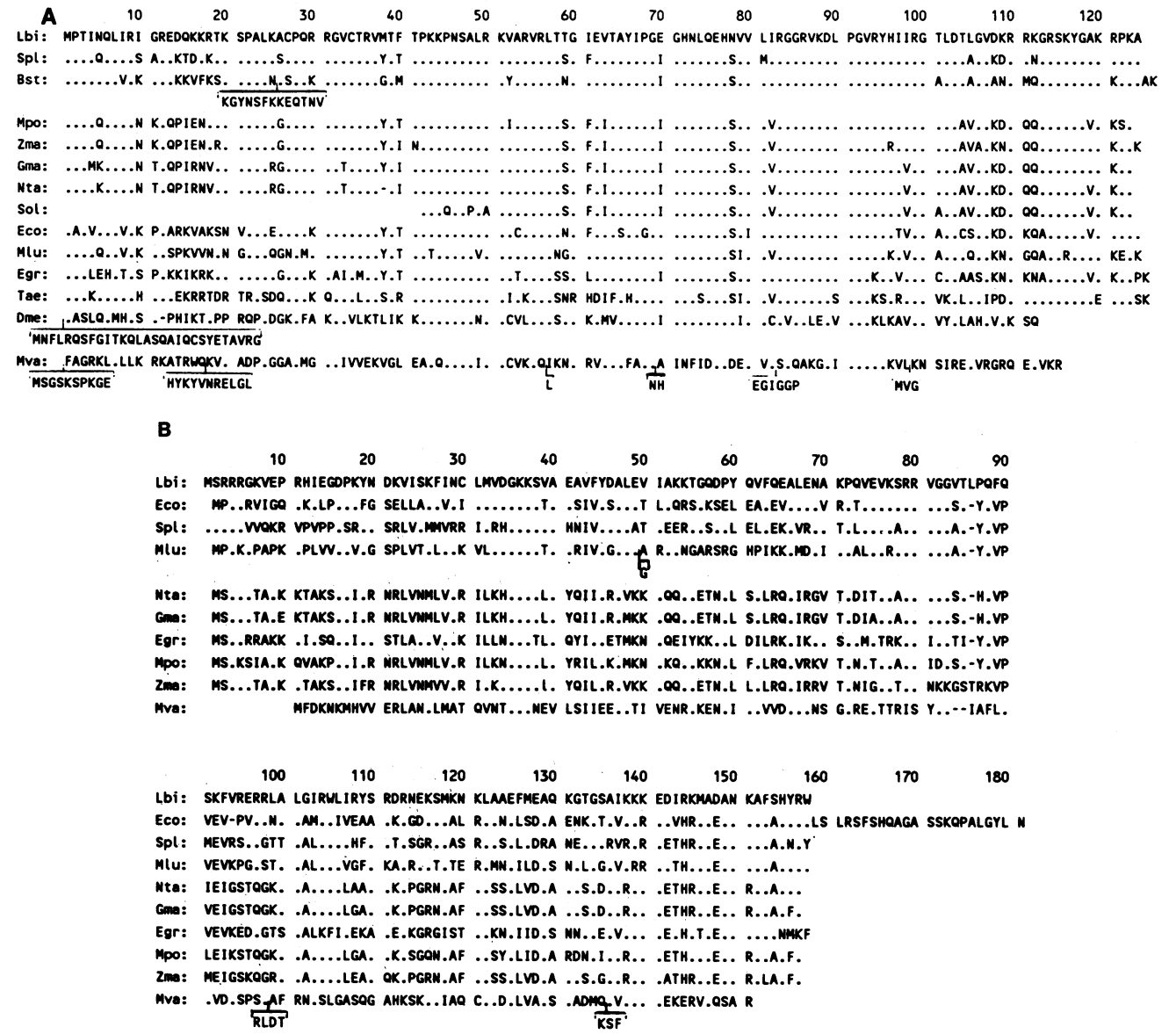


FIG. 1. Nucleotide sequence of the *L. biflexa* serovar patoc *rpsL* and *rpsG* genes and amino acid sequence of the corresponding gene products. The sequence starts with the termination codon of *argE*. The first translated sequence is ORF2 (*rpsL*), which is followed by ORF3 (*rpsG*).

two amino acid insertions (arginine at position 3 and leucine at position 86) which were not common to other S7 proteins (Fig. 2B). An insertion at position 3 was shared only with S7 from *Spirulina platensis*, and an insertion at position 86 was shared only with S7 proteins from *Zea mays* chloroplasts and *M. vannielii*. Valine was found at position 23 of the *L. biflexa* S7 protein, while leucine was found at this position in other S7 proteins (Fig. 2B). Likewise, instead of a conserved isoleucine at position 43, the *L. biflexa* S7 protein contained valine (Fig. 2B). Several amino acids (positions 102, 128, 144 and 148) are normally conserved among S7 proteins; however, both the *L. biflexa* and *M. vannielii* S7 proteins diverged from the consensus sequence at these positions (Fig. 2B). These amino acid changes may reflect adaptations to different rRNA and protein structures in ribosomes of evolutionarily diverse organisms.

Nucleotide sequence accession number. The nucleotide sequence data presented here have been entered in the EMBL, GenBank, and DDBJ nucleotide sequence data bases with the accession number M30695.

FIG. 2. Comparisons of *L. biflexa* serovar patoc ribosomal proteins S12 (A) and S7 (B) with homologous proteins from various organisms. Proteins are aligned for the best fit to the *L. biflexa* sequence and are arranged in order of decreasing levels of similarity. Amino acids identical to the *L. biflexa* sequence are represented by dots. Amino acid deletions are represented by dashes. Amino acid insertions are indicated with brackets. Abbreviations: Lbi, *L. biflexa*; Spl, *S. platensis* (2); Bst, *Bacillus stearothermophilus* (16); Mpo, *Marchantia polymorpha* (8); Zma, *Z. mays* (chloroplast) (11); Gma, *Glycine max* (29); Nta, *Nicotiana tabacum* (6, 28); Sol, *Spirodela oligorhiza* (22); Eco, *E. coli* (9, 10, 23, 24); Mlu, *Micrococcus luteus* (20); Egr, *Euglena gracilis* (19); Tae, *Triticum aestivum* (13); Dme, *Drosophila melanogaster* (26); Mva, *M. vannielii* (17). The amino acid sequences for mitochondrial S12 proteins of *T. aestivum* and *Z. mays* are identical; therefore, only the *T. aestivum* mitochondrial sequence is shown. Since the amino acid sequences of mitochondrial and chloroplast forms of the *Z. mays* S12 are different, the chloroplast form is shown.



We appreciate the suggestions of R. Hendrix, J. Strobl, and D. Yelton and thank T. Casey, S. Halling, C. Minion, and D. Yelton for reviewing the manuscript.

Part of this research was supported by Public Health Service grant DE04645 from the National Institute of Dental Research (to N.W.C.).

LITERATURE CITED

1. **Brendle, J. J., M. Rogul, and A. D. Alexander.** 1974. Deoxyribonucleic acid hybridization among selected leptospiral serotypes. Int. J. Syst. Bacteriol. **24**:205–214.
2. **Buttarelli, F. R., R. A. Calogero, O. Tiboni, C. O. Gualerzi, and C. L. Pon.** 1989. Characterization of the *str* operon genes from *Spirulina platensis* and their evolutionary relationship to those of other prokaryotes. Mol. Gen. Genet. **217**:97–104.
3. **Charon, N. W., R. C. Johnson, and D. Peterson.** 1974. Amino acid biosynthesis in the spirochete *Leptospira*: evidence for a novel pathway of isoleucine biosynthesis. J. Bacteriol. **117**:203–211.
4. **Ekiel, I., I. C. P. Smith, and G. D. Sprott.** 1984. Biosynthesis of isoleucine in methanogenic bacteria: a ^{13}C NMR study. Biochemistry **23**:1683–1687.
5. **Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonene, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. M. Chen, and C. R. Woese.** 1980. The phylogeny of prokaryotes. Science **209**:457–463.
6. **Fromm, H., M. Edelman, B. Koller, P. Goloubinoff, and E. Galun.** 1986. The enigma of the gene coding for ribosomal protein S12 in the chloroplasts of *Nicotiana*. Nucleic Acids Res. **14**:883–898.
7. **Fukanaga, M., and I. Mifuchi.** 1989. Unique organization of *Leptospira interrogans* rRNA genes. J. Bacteriol. **171**:5763–5767.
8. **Fukuzawa, H., T. Kohchi, H. Shirai, K. Ohyama, K. Umesono, H. Inokuchi, and H. Ozeki.** 1986. Coding sequences for chloroplast ribosomal protein S12 from the liverwort, *Marchantia polymorpha*, are separated far apart on the different DNA strands. FEBS Lett. **198**:11–15.
9. **Funatsu, G., and H. G. Wittman.** 1972. Ribosomal proteins. XXXIII. Location of amino-acid replacements in protein S12 isolated from *Escherichia coli* mutants resistant to streptomycin. J. Mol. Biol. **68**:547–550.

10. **Funatsu, G., M. Yaguchi, and B. Wittmann-Liebold.** 1977. Primary structure of protein S12 from the small *Escherichia coli* ribosomal subunit. FEBS Lett. **73**:12–17.
11. **Giese, K., A. R. Subramanian, I. M. Larrinua, and L. Bogorad.** 1987. Nucleotide sequence, promoter analysis, and linkage mapping of the unusually organized operon encoding ribosomal proteins S7 and S12 in maize chloroplast. J. Biol. Chem. **262**:15251–15255.
12. **Gropp, F., W. D. Reiter, A. Sentenac, W. Zillig, R. Schnabel, M. Thomm, and K. O. Stetter.** 1986. Homologies of components of DNA-dependent RNA polymerases of archaebacteria, eukaryotes and eubacteria. Syst. Appl. Microbiol. **7**:95–101.
13. **Gualberto, J. M., H. Wintz, J.-H. Weil, and J.-M. Grienenberger.** 1988. The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are co-transcribed. Mol. Gen. Genet. **215**:118–127.
14. **Haapala, D. K., M. Rogul, L. B. Evans, and A. D. Alexander.** 1969. Deoxyribonucleic acid base composition and homology studies of *Leptospira*. J. Bacteriol. **98**:421–428.
15. **Johnson, R. C., and S. Faine.** 1984. Genus I. *Leptospira* Noguchi 1917, 755.^{AL}, p. 62–67. *In* N. R. Krieg and J. G. Holt (ed.), Bergey's manual of systematic bacteriology, vol. 1. The Williams & Wilkins Co., Baltimore.
16. **Kimura, M., and J. Kimura.** 1987. The complete amino acid sequence of ribosomal protein S12 from *Bacillus stearothermophilus*. FEBS Lett. **210**:91–96.
17. **Lechner, K., G. Heller, and A. Bock.** 1989. Organization and nucleotide sequence of a transcriptional unit of *Methanococcus vannielii* comprising genes for protein synthesis elongation factors and ribosomal proteins. J. Mol. Evol. **29**:20–27.
18. **Leschine, S. B., and E. Canale-Parola.** 1986. Rifampin resistant RNA polymerase in spirochetes. FEMS Microbiol. Lett. **35**: 199–204.
19. **Montandon, P.-E., and E. Stutz.** 1984. The genes for the ribosomal proteins S12 and S7 are clustered with the gene for the EF-Tu protein on the chloroplast genome of *Euglena gracilis*. Nucleic Acids Res. **12**:2851–2859.
20. **Ohama, T., F. Yamao, A. Muto, and S. Osawa.** 1987. Organization and codon usage of the streptomycin operon in *Micrococcus luteus*, a bacterium with a high genomic G+C content. J. Bacteriol. **169**:4770–4777.
21. **Paster, B. J., E. Stackebrandt, R. B. Hespell, C. M. Hahn, and C. R. Woese.** 1984. The phylogeny of the spirochetes. Syst. Appl. Microbiol. **5**:337–351.
22. **Posno, M., W. R. Verweij, I. C. Dekker, P. M. de Waard, and G. S. P. Groot.** 1986. The genes encoding chloroplast ribosomal proteins S7 and S12 are located in the inverted repeat of *Spirodela oligorhiza* chloroplast DNA. Curr. Genet. **11**:25–34.
23. **Post, L. E., and M. Nomura.** 1980. DNA sequences from the *str* operon of *Escherichia coli*. J. Biol. Chem. **255**:4660–4666.
24. **Reinbolt, J., D. Tritsch, and B. Wittmann-Liebold.** 1979. The primary structure of ribosomal protein S7 from *E. coli* strains K and B. Biochimie **61**:501–522.
25. **Richaud, C., D. Margarita, G. Baranton, and I. Saint-Girons.** 1990. Cloning of genes required for amino acid biosynthesis from *Leptospira interrogans* serovar *icterohaemorrhagiae*. J. Gen. Microbiol. **136**:651–656.
26. **Royden, C. S., V. Pirrotta, and L. Y. Jan.** 1987. The *tko* locus, site of a behavioral mutation in *D. melanogaster*, codes for a protein homologous to prokaryotic ribosomal protein S12. Cell **51**:165–173.
27. **Stern, S., T. Powers, L.-M. Changshien, and H. F. Noller.** 1989. RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. Science **244**:783–790.
28. **Torazawa, K., N. Hayashida, J. Obokata, K. Shinozaki, and M. Sugiura.** 1986. The 5′ part of the gene for ribosomal protein S12 is located 30 kbp downstream from its 3′ part in tobacco chloroplast genome. Nucleic Acids Res. **14**:3143.
29. **von Allmen, J.-M., and E. Stutz.** 1987. Complete sequence of 'divided' *rps*12 (r-protein S12) and *rps*7 (r-protein S7) gene in soybean chloroplast DNA. Nucleic Acids Res. **15**:2387.
30. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51**:221–271.
31. **Yasuda, P. H., A. G. Steigerwalt, K. R. Sulzer, A. F. Kaufmann, F. Rogers, and D. J. Brenner.** 1987. Deoxyribonucleic acid relatedness between serogroups and serovars in the family *Leptospiraceae* with proposals for seven new *Leptospira* species. Int. J. Syst. Bacteriol. **37**:407–415.
32. **Yelton, D. B., and N. W. Charon.** 1984. Cloning of a gene required for tryptophan biosynthesis from *L. biflexa* serovar *patoc* into *Escherichia coli*. Gene **28**:147–152.
33. **Yelton, D. B., and S. L. Peng.** 1989. Identification and nucleotide sequence of the *Leptospira biflexa* serovar *patoc trpE* and *trpG* genes. J. Bacteriol. **171**:2083–2089.
34. **Zuerner, R. L., and N. W. Charon.** 1988. Nucleotide sequence analysis of a gene cloned from *Leptospira biflexa* serovar *patoc* which complements an *argE* defect in *Escherichia coli*. J. Bacteriol. **170**:4548–4554.