The Journal of Microbiology, February 2007, p. 79-82
Vol. 45, No. 1

NOTE

Nucleotide Sequence and Secondary Structure of 5S rRNA from *Sphingobium chungbukense* DJ77

Hae-Ryong Kwon and Young-Chang Kim*

School of Life Sciences, Chungbuk National University, Cheongju, Chungbuk 361-763, Republic of Korea

(Received January 19, 2007 / Accepted February 14, 2007)

The 5S rRNA gene from *Sphingobium chungbukense* DJ77 was identified. The secondary structure of the 199-base-long RNA was proposed. The two-base-long D loop was the shortest among all of the known 5S rRNAs. The U19-U64 non-canonical pair in the helix II region was uniquely found in strain DJ77 among all of the sphingomonads.

Keywords: 5S ribosomal RNA, *Sphingobium chungbukense*, secondary structure, phylogenetic analysis, organization of rRNA operon

Sphingobium chungbukense DJ77 is a Gram-negative bacterium that is a very interesting organism due to its capacities to degrade monocyclic and polycyclic aromatic compounds, to synthesize glycosphingolipids as components of the cell envelope, and to produce exopolysaccharides as extracellular polymers (Kim *et al.*, 1986; Kim *et al.*, 2000; Lee *et al.*, 2005). This strain was isolated in Korea (Kim *et al.*, 1986), and was subsequently classified in the genus *Sphingomonas* as *Sphingomonas chungbukensis* sp. nov. The type strain is strain DJ77 (Kim *et al.*, 2000). Takeuchi *et al.* (2001) proposed that the genus formerly known as *Sphingomonas* can be divided into four clusters: *Sphingomonas* (cluster I), *Sphingobium* (II), *Novosphingobium* (III), and *Sphinogpyxis* (IV). According to phylogenetic analyses of 16S rRNA gene sequences, Lee *et al.* (2005) reported that *Sphingomonas chungbukensis* DJ77 was a member of the genus *Sphingobium*. Subsequently, Pal *et al.* (2005) reclassified *Sphingomonas chungbukensis* as *Sphingobium chungbukense* comb. nov.

Interest in the genomes of the sphingomonads has been rapidly increasing in recent years. Two whole-genomes were sequenced by the DOE Joint Genome Institute (JGI) in 2006: *Novosphingobium aromaticivorans* DSM11244 (JGI web site: http://genome.jgi-psf.org/finished_microbes/novar/novar.home.html), and *Sphingopyxis alaskensis* RB2256 (JGI web site: http://genome.jgi-psf.org/finished_microbes/sphal/sphal.home.html). In 2007, four whole-genome sequencing projects are currently in progress to sequence the genomes of *Sphingomonas* sp. SKA58 (J. Craig Venter Institute), *Sphingomonas elodea* (Hiram College), *Sphingomonas wittichii* RW1 (JGI), and *Sphingobium chungbukense* DJ77 (Chungbuk National University, Korea). The organization of the ribosomal RNA transcription units was conserved in two of the completed sphingomonad genomes. The order of the genes was 16S ribosomal RNA, $tRNA^{Ile}$, $tRNA^{Ala}$, 23S ribosomal RNA, 5S ribosomal RNA, and the $tRNA^{fMet}$ gene. There was only one unit in *Sphingopyxis alaskensis* and three units in *Novosphingobium aromaticivorans* that contained all of the same sequences in terms of both the genes themselves and the intergenic regions. This gene organization was found in many bacterial strains including *Zymomonas mobilis* ZM4 (Lee *et al.*, 2001). In *E. coli* strain K12, there were seven units that somehow differed in their sequences (*Escherichia coli* K12 MG1655, complete genome ACCESSION U00096). The 5S rRNA gene was found in the course of the *Sphingobium chungbukense* DJ77 genome project, and their primary and secondary structures and phylogenetic relationships among the sphingomonads are reported in this paper.

DNA fragments of 1-2 kb or 3-4 kb in size were collected from sonicated genomic DNA from strain DJ77 using agarose gel electrophoresis. Fragments of between 1 and 2 kb in size were ligated to a pBluescript II SK(-) vector that had been digested with *Eco*RV and treated with SAP (Promega, USA). Fragments of between 3 and 4 kb in size were ligated to a pUC19 vector that had been digested with *Hinc*II and treated with SAP. The recombinant plasmids were propagated in *E. coli* XLI-Blue strains. The nucleotide sequences were determined using the ABI 377 and 3700 automatic sequencers. Ninety percent of the whole-genome sequence (3 Mb in size) had been read by January 2007. The DNA sequences were assembled using the PHRED-PHRAP-CONSED contig assembly program (Ewing *et al.*, 1998; Ewing and Green, 1998; Gordon *et al.*, 1998). The sequences of the 5S rRNA genes of *E. coli* and the sphingomonads that were used in the phylogenetic study were obtained from the NCBI GenBank sequence database. The sequences were aligned using the CLUSTAL

* To whom correspondence should be addressed.
(Tel) 82-43-261-2302; (Fax) 82-43-268-2538
(E-mail) youngkim@chungbuk.ac.kr

W multiple sequence alignment program (Thompson *et al.*, 1994). The phylogenetic trees were constructed using the neighbor-joining and maximum-parsimony methods, and the reliability of the tree was evaluated by a PHYLIP package (Felsenstein, 1993). The secondary structures of the 5S rRNA genes in the sphingomonads were constructed from the secondary structure model of *E. coli* 5S rRNA. This model was constructed from the sequence of EMBL accession number V00336 (http://www.rna.icmb.utexas.edu). The 5S rRNA sequence of DJ77 was deposited in GenBank under accession number EF221603. The accession numbers or GeneIDs of the 5S rRNA sequence data for the species examined in this study were as follows: *Novosphingobium aromaticivorans* F199, 3917044; *Sphingomonas sp.* SKA58, AAQG01000001 (regions: 4554-4672); *Sphingopyxis alaskensis* RB2256, NC_008048 (regions: complement, 2847886-2848004); *Sphingomonas wittichii* RW1, NZ_AAVK01000005 (regions: complement, 391800-391918); *Sphingomonas ursincola*, L37447.

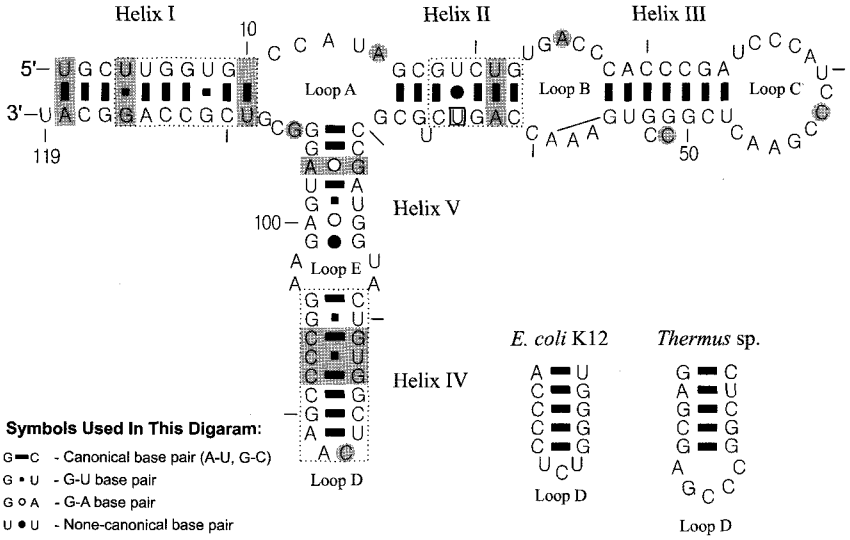

Fig. 1. Primary sequence and secondary structure model for *Sphingobium chungbukense* DJ77 5S rRNA. Dotted boxes in the helix I, II, and IV regions represent the highly variable regions between *Sphingobium chungbukense* and *E. coli* 5S rRNAs. The nucleotides in the gray boxes or circles are variable among the sphingomonads. Every 10th nucleotide is marked with a tick mark.

```
                                  |         |         |         |         |
DJ77          1  TGCTTGGTGACCATAGCGTCTGTGACCCACCCGATCCCATCCCGAACTCG  50
SKA58            TGCTTGGTGACCATAGCGTCAGTGACCCACCCGATCCCATCCCGAACTCG
RB2256           TGCTTGGTGACCATGGCGTCAGTGACCCACCCGATCCCATCTCGAACTCG
RW1              AGCTTGGTGGCCATAGCGTCTGTGACCCACCCGATCCCATCCCGAACTCG
F199             TGCCTGGTGGCCATAGCGTCGGTGCCCCACCCGATCCCATCTCGAACTCG
                  ** ***** **** ***** *** ********************* ********

                                  |         |         |         |         |
DJ77         51  GCCGTGAAACCAGTCTGCGCCGATGGTACTGTGGCTCAAGCGCGGGAAGA  100
SKA58            GACGTGAAACCTGACTGCGCCGATGGTACTGTGGCTCAAGCGCGGGAAGA
RB2256           GCCGTGAAACCTGACAGCGCCGATGGTACTGTGGCTCAAGCGCGGGAAGA
RW1              GCCGTGAAACCAGACTGCGCCGATGGTACTAGCGCTTAAGCGGTGGAAGA
F199             GCCGTGAAACCGGACAGCGCCAATGGTACTAACGCTCAAGCGTTGGAAGA
                 * ********* * * ***** ********   *** ****   ******

                                  |
DJ77         101 GTAGGGCGTCGCCAGGCAT  119
SKA58            GTAGGGCGTCGCCAGGCAT
RB2256           GTAGGACGTCGCCAGGCAT
RW1              GTAGGGCGTCGCCAGGCTT
F199             GTAGGTCGTCGCCAGGCAT
                 ***** ********** *
```

Fig. 2. Multiple sequence alignment of the 5S rRNA genes from *Sphingobium chungbukense* DJ77 and four other sphingomonads. Gray regions represent the highly variable sequences among the sphingomonads. A "T" enclosed with a black box indicates one of the signature sequences of *Sphingobium chungbukense* DJ77 5S rRNA. The residues TG and AT were inserted into positions 1-2 and 118-119 of the annotated 5S rRNA genes of *Novosphingobium aromaticivorans* F199 (GenBank accession number 3917044) and *Sphingopyxis alaskensis* RB2256 (GeneID, NC_008048), based on the sequence homology with the *Sphingobium chungbukense* DJ77 5S rRNA model. The 5S rRNA sequence from *Sphingomonas sp.* SKA58 (GeneID, AAQG01000001) included the T residue at position 1 and the AT residue at positions 118-119.

Table 1. The nucleotide signatures of the 5S rRNA useful in classifying the four clusters of sphingomonads

Strains	Signature at position:											
	1:118	4:115	10:109	15	19:64	21:62	41	52	72:103	81-83:92-94	87	106
Sphingomonas wittichii RW1	A:U	U:G	G:U	A	U:A	U:A	C	C	G:A	AGC:GGU	U	G
Sphingobium chungbukense DJ77	U:A	U:G	A:U	A	U:U	U:A	C	C	G:A	GUG:CCC	C	G
Novosphingobium aromaticivorans F199	U:A	C:G	G:U	A	U:A	C:G	U	C	A:A	AAC:GUU	C	U
Sphingopyxis alaskensis RB2256	U:A	U:G	A:U	G	U:A	A:U	U	C	G:A	GUG:CCC	C	A

The sequence fragments containing the ribosomal RNA transcription units of *Sphingobium chungbukense* DJ77 were found using the NCBI BLAST program and were assembled manually due to problems with the automatic assembly. The rRNA transcription units in which the 16S ribosomal RNA, $tRNA^{Ile}$, $tRNA^{Ala}$, 23S ribosomal RNA, 5S ribosomal RNA, and $tRNA^{fMet}$ genes were sequentially located were highly conserved in the sphingomonads. These conserved rRNA transcription units could also be found in two different genomic sequence contigs of *Sphingobium chungbukense* DJ77, although the complete sequences were not obtained. In these two regions, two complete 5S rRNA sequences were determined and found to be identical. The 5S rRNA gene contains a total of 199 bp (Fig. 1). The 5S rRNA of *Sphingobium chungbukense* DJ77 showed a sequence identity of 96% with *Sphingomonas* sp. SKA58 5S rRNA, 94% with *Sphingopyxis alaskensis* RB2256 5S rRNA, 92% with *Sphingomonas wittichii* RW1 5S rRNA, 88% with *Zymomonas mobilis* ZM4 5S rRNA, 86% with *Novosphingobium aromaticivorans* F199 and *Sphingomonas ursincola* 5S rRNA, and 70% with *E. coli* 5S rRNA. These results demonstrated that *Sphingobium chungbukense* was more closely related with *Sphingopyxis alaskensis* than *Novosphingobium aromaticivorans*.

It has been well established that the secondary structure of the rRNA is generally highly conserved (Gutell *et al.*, 1994). The secondary structure of the 5S rRNA of *Sphingobium chungbukense* was constructed using an *E. coli* model (Fig. 1). Sequence variations were found at 36 positions between the two 5S rRNAs from *Sphingobium chungbukense* DJ77 and *E. coli* K12. The majority of these variations were found in the helix I region, (4:115 pair and contiguous 8-10:110-112 pairs), in the helix II region (contiguous 18-23:60-65 pairs), and in the helix IV region (79:96, 81-82:93-94, and 85-87:88-90 pairs) of *Sphingobium chungbukense* 5S rRNA. Most of the changes had little influence on the secondary structure of the RNA, with the exception of two cases. Interestingly, the D loop of *Sphingobium chungbukense* 5S rRNA was only two nucleotides long, which was one nucleotide shorter than that of *E. coli* 5S rRNA. This is the shortest of all the D loops of bacterial 5S rRNAs that have been reported thus far (Szymanski *et al.*, 2000).

Four 5S rRNA sequences from *Novosphingobium aromaticivorans* F199, *Sphingomonas wittichii* RW1, *Sphingopyxis alaskensis* RB2256, and *Sphingomonas* sp. SKA58, which had previously been reported, were compared with that of *Sphingobium chungbukense* using the CLUSTAL W multiple alignment program. As shown in Fig. 2, there were changes in nine pairs (1:118, 4:115, 10:109, 19:64, 21:62, 72:103, and three consecutive pairs 81-83:92-94), and seven single bases located at the loops. The greatest changes in the *Sphingobium chungbukense* 5S rRNA sequences occurred at nucleotide positions 19 and 64 located in the helix II region. The C19-G64 pair, which was found in *E. coli*, was changed to a U19-U64 pair in the *Sphingobium chungbukense* 5S rRNA (Table 1). Each region in the single 21:62 pair contains the T:A pair in *Sphingobium chungbukense* DJ77 and *Sphingomonas wittichii* RW1, the A:T pair in *Sphingomonas* sp. SKA58 and *Sphingopyxis alaskensis* RB2256, and the C:G pair in *Novosphingobium aromaticivorans*, respectively. Each region in the continuous 81-83:92-94 pairs contains the GTG:CCC pair in *Sphingobium chungbukense* DJ77, *Sphingomonas* sp. SKA58, and *Sphingopyxis alaskensis* RB2256, the AGC:GGT pair in *Sphingomonas wittichii* RW1, and the AAC:GTT pair in *Novosphingobium aromaticivorans* F199, respectively.

Despite the sequence variations, all the sphingomonad 5S rRNA sequences reported thus far could be folded into a general secondary structure. The two-base-long D loops were also found in all of the sphingomonad strains. The most unique structural feature of the 5S rRNA from *Sphingobium chungbukense* was the U19-U64 non-canonical pair (Nagaswamy *et al.*, 2000) located in the helix II region, whereas the U19-A64 pair was found among all of the other sphingomonads.

Acknowledgements

This work was supported by a research grant from Chungbuk National University in 2005.

References

Ewing, B. and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8, 186-194.

Ewing, B., L. Hillier, M.C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. accuracy assessment. *Genome Res.* 8, 175-185.

Felsenstein, J. 1993. PHYLIP (phylogeny inference package), version 3.5c. University of Washington.

Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8, 195-202.

Gutell, R.R., N. Larsen, and C.R. Woese. 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative prospective. *Microbiol. Rev.* 58, 10-26.

Kim, C.K., J.W. Kim, Y.C. Kim, and T.I. Mheen. 1986. Isolation of aromatic hydrocarbon-degrading bacteria and genetic characterization of their plasmid genes. *Kor. J. Microbiol.* 24, 67-72.

Kim, S.J., J. Chun, K.S. Bae, and Y.C. Kim. 2000. Polyphasic assignment of an aromatic-degrading *Pseudomonas* sp., strain

DJ77, in the genus *Sphingomonas* as *Sphingomonas chungbukensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 50, 1641-1647.

Lee, J.S., S.J. Jin, and H.S. Kang. 2001. Molecular organization of the ribosomal RNA transcription unit and the phylogenetic study of *Zymomonas mobilis* ZM4. *Mol. Cells* 11, 68-74.

Lee, K.Y., H.R. Kwon, W.H. Lee, and Y.C. Kim. 2005. Nucleotide sequence and secondary structure of 16S rRNA from *Sphingomonas chungbukensis* DJ77. *Kor. J. Microbiol.* 41, 125-129.

Nagaswamy, U, N. Voss, Z. Zhang, and G.E. Fox. 2000. Database of non-canonical base pairs found in known RNA structures. *Nucleic Acids Res.* 28, 375-376.

Pal, R., S. Bala, M. Dadhwal, M. Kumar, G. Dhingra, O. Prakash, S.R. Prabagaran, S. Shivaji, J. Cullum, C. Holliger, and R. Lal. 2005. Hexachlorocyclohexane-degrading bacterial strains *Sphingomonas paucimobilis* B90A, UT26 and Sp+, having similar lin genes, represent three distinct species, *Sphingobium indicum* sp. nov., *Sphingobium japonicum* sp. nov. and *Sphingobium francense* sp. nov., and reclassification of [*Sphingomonas*] *chungbukensis* as *Sphingobium chungbukense* comb. nov. *Int. J. Syst. Evol. Microbiol.* 55, 1965-1972.

Szymanski, M., M.Z. Barciszewska, J. Barciszewski, and V.A. Erdmann. 2000. 5S ribosomal RNA database Y2K. *Nucleic Acids Res.* 28, 166-167.

Takeuchi, M., K. Hamana, and A. Hiraishi. 2001. Proposal of the genus *Sphingomonas sensu stricto* and three new genera, *Sphingobium*, *Novosphingobium* and *Sphingopyxis*, on the basis of phylogenetic and chemotaxonomic analyses. *Int. J. Syst. Evol. Microbiol.* 51, 1405-1417.

Thompson, H., M.J. Saunders, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.