

Nucleotide sequence of the *deoC* gene of *Mycoplasma pneumoniae*Steve Loechel<sup>1</sup>, Julia M. Inamine<sup>1</sup> and Ping-chuan Hu<sup>1,2</sup>Departments of <sup>1</sup>Pediatrics and <sup>2</sup>Microbiology and Immunology, University of North Carolina, Chapel Hill, NC 27599-7220, USA

Submitted November 23, 1988

Accession no. X13544

DNA sequence analysis of a 2.3 kilobase *Hind* III fragment from *Mycoplasma pneumoniae* revealed an open reading frame coding for a protein of 224 amino acid residues. It has 32% identity with the *E. coli* protein deoxyribose-phosphate aldolase (EC 4.1.2.4) (1). This enzyme, coded for by the *deoC* gene, is part of the pathway for catabolism of deoxyribonucleosides. To show that the mycoplasma gene does in fact code for deoxyribose-phosphate aldolase, we inserted it in the *E. coli* expression vector pKK223-3 (2) and tested for complementation of *E. coli* strain S0063, which is *deoC*<sup>-</sup> (3). Cells transformed with the recombinant plasmid were able to utilize thymidine as a carbon source, confirming the identity of the gene. Lys-167 of the *E. coli* enzyme has been tentatively identified as being involved at the active site (1), and this residue is conserved in the *M. pneumoniae* enzyme. The entire 2.3 kilobase *Hind* III fragment was sequenced by making nested deletions with the Erase-a-base kit from Promega. Dideoxy sequencing was done on both strands.

```

          30          60          90          120
ATGAAATAGAAATACAACCGCATTATTGACAGTACACTCTTAAAGCTGATACCTTACCCACGAAATGATGCTTTATGTGCTGATGCTCATAAGTACCAGTTTTACGCAGTTTGTGTC
M K L E Y N R I I D S T L L K A D T L P H E I D A L C A D A H K Y Q P Y A V C V

          150          180          210          240
AATCCTAGCTATGTGGGGTATGCCAAAATATCCTTAAAGGGTACGGGGGTGAAACTGTGCACTGTGCTGGTTTTCCCTTAGGCAAAACAACACAGCCGAGAAGGTGTATGAACTAAG
N P S Y V R Y A K N I L K G T G V K L C T V V G F P L G Q T T Q R Q K V Y E T K

          270          300          330          360
ATTGCCATTAAGGAAGGTGCGGATGAAATGACATGTTAATGAACATTGCCGAGTTCAAAAAGCGCTGTGCCTGTGCTCATTAGTGAATAAGGGCAGTCAAAAAGGTTTTGTGTAAGCGA
I A I K E G A D E I D M V M N I A E F K K R C A C V I S E I R A V K K V C G K R

          390          420          450          480
ACGCTCAAAGTCATTATTGAAACGGCTTTGTTAAACGAGGACGAAATGGTGATGCCGTTAATGTCTGTATTGATGGTAAACGCTGACTTTGTCAAACCTCTACCGGGTTTTCCATCGCT
T L K V F I E T A L L N Q D E I R D A V N V C I D G N A D P V K T S T G F S M R

          510          540          570          600
GGTGCATCATTGGAAAGATATCACAAATATGCGCGAAGCTAGTGGTAATCTCATTAAAGATCAAAGCATCGGGTGGGGTCAAACAGCGCAGCAATTCCTTGACTTCTTTAATGCTGGAGTT
G A S L E D I T I M R E A S G N L I K I K A S G G V Q T A Q Q F L D F F N A G V

          630          660
AGTCGGATTGGCACTTCCAATTGCGGTCAAGATTATGGAAGAATTGCACAACTAGAAAAGCCATGAACATCGTTAA
S R I G T S N A V K I M E E L H K L E S H E H R *

```

**ACKNOWLEDGEMENTS**

*E. coli* strain S0063 was kindly provided by Poul Valentin-Hansen. This work was supported by NIH grant HL-19171 and by Cooperative Agreement CR807392 from EPA.

**REFERENCES**

- 1) Valentin-Hansen, P. *et al.* (1982) *Eur. J. Biochem.* 125, 561-566.
- 2) Brosius, J. and Holy, A. (1984) *Proc. Natl. Acad. Sci. USA.* 81, 6929-6933.
- 3) Valentin-Hansen, P. *et al.* (1982) *EMBO J.* 1, 317-322.