

Nucleotide sequence of the major outer membrane protein gene from *Chlamydia trachomatis* serovar H

Paul T.Hamilton* and Douglas P.Malinowski

Cell Biology Department, Becton Dickinson Research Center, Research Triangle Park, NC 27709, USA

Submitted September 13, 1989

EMBL accession no. X16007

The gene encoding the major outer membrane protein (MOMP) from *Chlamydia trachomatis*, serovar H was isolated from a genomic library using the MOMP gene from serovar L2 as the probe (1). The recombinant insert was subcloned into a plasmid vector (Bluescript), the gene localized by Southern blotting and the nucleotide sequence was determined by the dideoxy chain termination method.

The H MOMP gene is 1193 bp long and encodes a polypeptide of 397 amino acids including a 22 amino acid leader sequence. This corresponds to a mature H MOMP of 40,672 daltons, in agreement with the molecular mass predicted by SDS-PAGE (2). The H and L2 MOMP genes are 74% homologous at the nucleotide level and are 84% homologous in their protein sequences. The differences in the two sequences are clustered in four regions (435-483, 655-687, 906-952 and 1083-1194).

```

CTCGAGCATTGAACGACAGTTCCTCGATTAAGGCTGCTTTTACTTGCAAGACATTCCTCAGGCCATTAATTC
TACAGGACATCTTTGCTGGCTTAACTAGGACGACGTGCCGCCAGAAAAAGATAGCGAGCACAAAGAGAGCT-144
AATTATACAATCTTAGAGGTAAGAATGAAAAAATCTTGAATCGGTATTAGTATTGCCGCTTTGAGTCTT
M K K L L K S V L V F A A L S S
GCTTCCCTCTGCAAGCTCGCTGTGGGAACTCCTGCTGAACCAAGCCTTATGATCGACGGAAATCTGTGG-288
A S S L Q A L P V G N P A E P S L M I D G I L W
GAAGTTTGGCGGAGATCTTCGATCCTTGCACCTTGTGTGACCGTATACAGTATCGGTGTGGTTC
E G F G C D P C D P C A T W C D A I S M R V G Y
TAGCGAGACTTGTGTTTCGACCTGTGTTGAAAACCTGATGGAATAAGAATTCAGATGGGAGCGCGCT-432
Y G D F V F D R V L K T D V N K E F Q M G A A P
ACTACCAACGATGCAGCAGACTTACAAAACGATCCAAAAACAATGTGTGCTGCCAAATCCGCTTATGGC
T T N D A A D L Q N D P K T N V A R P N P A Y G
AAACACATGCAAGATGCTGAAATGTTTACGAACGCTGCTTACATGGCATTAAATATCTGGGATCTTTTGT-576
K H M Q D A E M F T N A A Y M A L N I W D R F D
GTATTTGTACATTTGGGAGCACTACCCGTTATTTAAAAGGAACTCCGCTTCTCAACTTAGTTGGATTA
V F C T L G A T T G Y L K G N S A S F N L V G L
TTCGGAACAAAAAATACTTCTGATTTTAAATACAGCGAAGCTTGTCTCAACTGCTTTGAATCGAGCT-720
F G T K T K S S D F N T A K L V P N I A L N R A
GTGGTTGAGCTTTATACAGACACTACCTTTGCTTGGAGCGTAGGTCTGCTGAGCTCTGGGAATGGGG
V V E L Y T D T T F A W S V G A R A A L W E C G
TGTGCAACGTTAGGAGCTTCTTCCAATGCTCAATCAAACCTAAAGTAGAAGATTAAATGTTCTTTGT-864
C A T L G A S F Q Y A Q S K P K V E E L N V L C
AATGCATCCGAATTTACTATTAAAGCCGAAAGGATATGTTGGGCGGAATTTCCACTTGTATATACCGCA
N A S E F T I N K P K G Y V G A E F P L D I T A
GGAACAGAAGCTGCGCAGGGACTAAGGATGCCTCTATTGACTACCATGAGTGGAAGCAAGTTTAGCCCTT-1008
G T E A A T G T K D A S I D Y H F W Q A S L A L
TCTTACAGACTAAATATGTTCACTCTTACATTTGAGTAAATGGTCTAGAGTAAGTTTAGTATGCCACAGC
S Y R L N M F T P Y I G V K W S R V S F D A D T
ATCCGATCGCTCAGCTAAATTTGGCTGAAGCAATCTTGGATGTCACACTACTTAAACCCGACCTCGCTGGT-1152
I P I A Q P K L A E A I L D V T L N P T I A G
AAAGAACTGTGGTCCGTTCCGGAAGCATAACGACTGGCTGATACAATGCAATCGTTTCTGGAGTTG
K G T V V A S S D N D L A D T M S Q L
ACAAGATGAAATCTAGAAAATCTGCGGTATTGACGTAGGAACGACTATTGTAGATCGAGCAAAATACGCA-1296
N K M K S R K S C G I A V G T I V D A D K Y A
GTTACAGTTGAGACTCGCTTGTGATGAGAGAGCAGCTCACGTAATGCACATTCGCTTCTAATAATG
V T V E T R L I D E R A A H V N A Q F R F -
TATAATATATAGAACTTTTCTTAGTAACTTATCTTTGTAATAACACTATCCGATTTCTGGGCTCGG-1440
CTTCGGTCCGGTCCAGTTTTTTATAAACTTGTGTTGCTGACTTCTCTATCTTTTGAACGCAAAAT
AAAATTTCTCTAAAAGAAGATTGCA-1537
    
```

*To whom correspondence should be addressed

REFERENCES (1) Stephens, R.S., R. Sanchez-Pescador, E.A. Wagar, C. Inouye and M.S. Urdea (1987) *J. Bacteriol.* 169, 3879-3885. (2) Newhall, W.J., B. Batteiger and R.B. Jones. (1982) *Infect. Immun.* 38, 1181-1189.