

Sequence Analysis of the Major Outer Membrane Protein Gene of *Chlamydia pneumoniae*

MERCEDES PEREZ MELGOSA, CHO-CHOU KUO, AND LEE ANN CAMPBELL*

Department of Pathobiology, University of Washington, Seattle, Washington 98195

Received 15 January 1991/Accepted 25 March 1991

Compared with the major outer membrane proteins (MOMPs) of the other chlamydial species, the *Chlamydia pneumoniae* MOMP appears to be less antigenically complex, and as determined by immunoblot analysis, it does not appear to be the immunodominant antigen recognized during infection. Nucleotide sequence analysis of the *C. pneumoniae* MOMP gene (*ompA*) revealed that it consisted of a 1,167-base open reading frame with an inferred 39,344-dalton mature protein of 366 amino acids plus a 23-amino-acid leader sequence. A ribosomal-binding site was located in the 5' upstream region, and two stop codons followed by an 11-base dyad forming a stable stem-loop structure were identified. This sequence shares 68 and 71% DNA sequence homology to the *Chlamydia trachomatis* serovar L2 and *Chlamydia psittaci* ovine abortion agent MOMP genes, respectively. Interspecies alignment identified regions, corresponding to the variable domains, which share little sequence similarity with the other chlamydial MOMPs. All seven cysteines conserved in the *C. trachomatis* and *C. psittaci* MOMPs, which are involved in the formation of disulfide cross-linkages, are found in the *C. pneumoniae* MOMP.

Chlamydia pneumoniae has been established as an important cause of acute respiratory diseases in humans (7). Structural studies of this organism have shown that, as in the other *Chlamydia* spp. (12), the major outer membrane protein (MOMP) exists in disulfide-linked protein complexes within the outer membrane complex and functions in maintaining cell wall rigidity (3). Although the MOMP of *C. pneumoniae* plays a similar structural role, antigenic analysis has demonstrated that it has characteristics that differ from those of the other chlamydial MOMPs (3). In contrast to *Chlamydia trachomatis* and *Chlamydia psittaci*, in which the MOMP is the immunodominant antigen recognized during infection, immunoblot analysis of proteins recognized by human sera from *C. pneumoniae* patients has suggested that the MOMP is not the immunodominant antigen recognized during *C. pneumoniae* infection (4).

Monoclonal antibodies (MAbs) against the *C. trachomatis* MOMP have identified serovar-, subspecies-, and species-specific epitopes on the protein (5, 18, 22). The MOMP genes of several *C. trachomatis* and *C. psittaci* strains have been sequenced and were found to be highly conserved except in four variable domains (VDs) (16, 18, 22). Epitope mapping has shown that three of the VDs (VDI, VDII, and VDIV) contain the serovar-, subspecies-, and species-specific antigenic determinants (1, 5, 18, 20). Immunoblot analyses of *C. pneumoniae* with anti-*C. pneumoniae* rabbit sera or human sera from patients with *C. pneumoniae* infection have shown that the recognition of the MOMP is genus reactive (3, 4). To date, no other antigenic reactivities have been identified in the *C. pneumoniae* MOMP. The purpose of our studies was to define the coding sequence of the *C. pneumoniae* MOMP gene and to investigate its relationship to the other *Chlamydia* MOMP genes.

The *C. pneumoniae* isolate AR-39, which had been adapted to grow in HeLa 229 cells, was harvested and purified in a linear gradient of meglumine diatrizoate (Hypaque-76; Winthrop-Breon Laboratories, New York, N.Y.)

(11). The final products usually contained 1.0×10^8 to 5.0×10^8 inclusion-forming units per ml of organisms.

A partial gene bank was constructed by digesting *C. pneumoniae* DNA with *EcoRI*, ligating it into similarly cut lambda gt11, and packaging it into phage particles (9). Using a 1.1-kb *EcoRI* fragment encoding a portion of the *C. trachomatis* serovar L2 MOMP as a probe (15), a 1.5-kb *EcoRI* fragment was isolated and subcloned into the plasmid pTZ18r (Pharmacia, Piscataway, N.J.). The Sanger dideoxy-chain termination method of DNA sequencing (13) was carried out on single-stranded fragments cloned into M13mp18 and M13mp19 by using the Sequenase kit (United States Biochemical Corp., Cleveland, Ohio). Nested deletions were generated by using the Erase-a-Base kit (Promega Biotec, Madison, Wis.). Sequence analyses were performed by using the Pustell Sequence Analysis Program (IBI) and the University of Wisconsin Genetics Computer Group programs. DNA sequence analysis of the 1.5-kb *EcoRI* fragment showed that it contained 0.5 kb of the 3' end of the translated sequence but lacked the 5' terminus of the gene.

The 5'-terminal sequences of the MOMP genes have been shown to be highly conserved in *C. trachomatis* and *C. psittaci* (16, 20, 21). Two different DNA probes that contain this area were used to identify MOMP gene sequences in *C. pneumoniae* by Southern blots. The probes used were a 280-bp *XhoI-EcoRI* fragment (AM-11), which contains the 5' terminus of the *C. trachomatis* serovar A MOMP gene, and GPM-4, which contains the entire coding region of the *C. psittaci* guinea pig inclusion conjunctivitis MOMP (21). Both were graciously provided by Y.-X. Zhang, NIH Rocky Mountain Laboratories, Hamilton, Mont. The 1.5-kb *EcoRI* fragment was recognized by the GPM-4 probe but not by the AM-11 probe, demonstrating that this fragment did not contain the 5' end of the MOMP gene. When AM-11 was used to probe *C. pneumoniae* chromosomal digests, it hybridized to 0.9-kb *HindIII*, 3.5-kb *EcoRI*, 3.8-kb *BamHI*, and 3.7-kb *MspI* fragments. By constructing two partial gene banks of *C. pneumoniae* DNA and using the ^{32}P -labeled AM-11 fragment as a probe, the 0.9-kb *HindIII* and 3.7-kb *EcoRI* fragments were isolated and subsequently sequenced.

* Corresponding author.

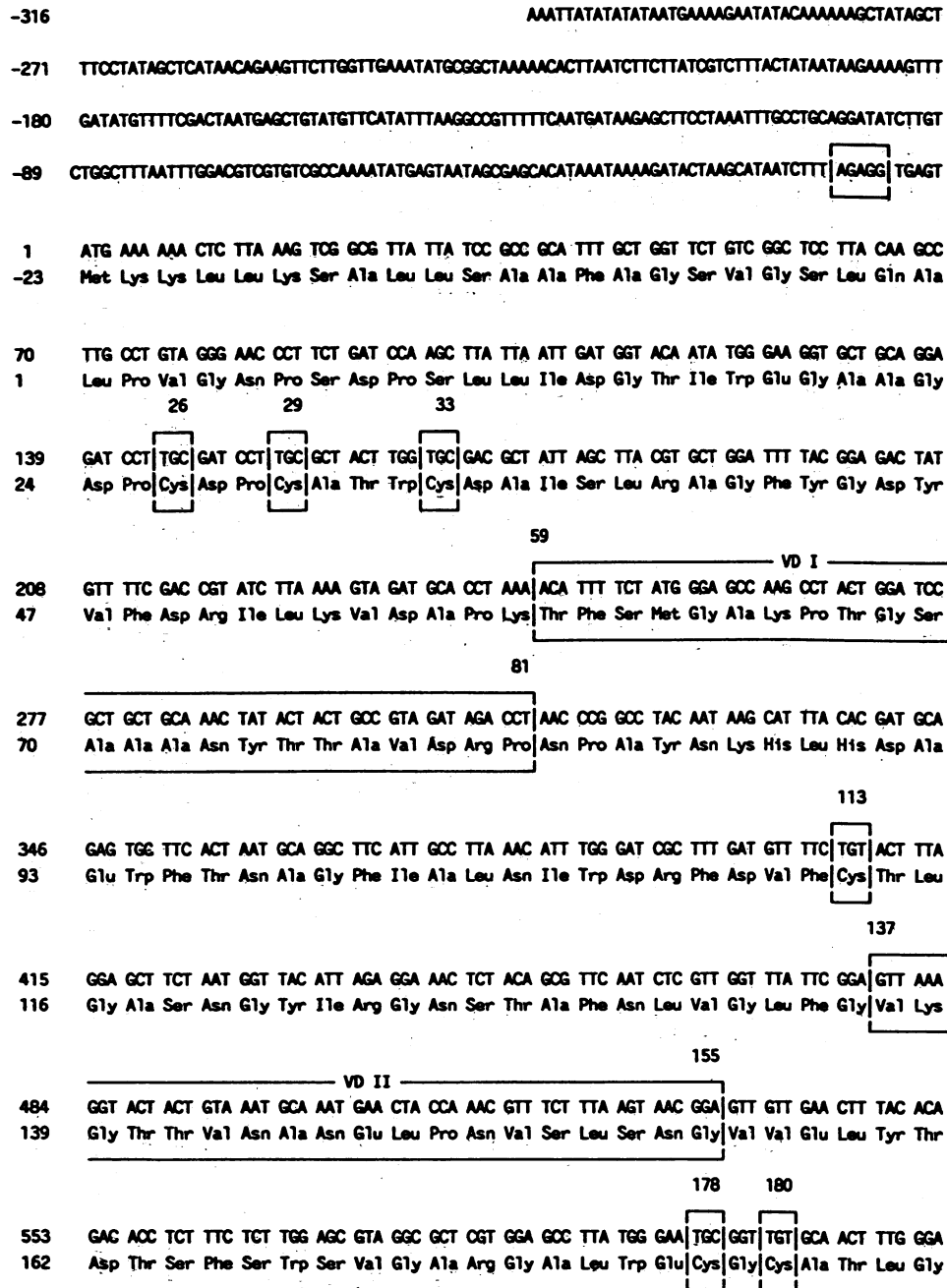


FIG. 1. Nucleotide sequence and inferred amino acid sequence of the gene encoding the *C. pneumoniae* AR-39 MOMP. The gene contains an ATG start codon followed by a 1,167-nucleotide open reading frame and ends with two in-frame TAA and TAG stop codons. The 1 by the gene sequence marks the transitional initiation codon ATG, and the 1 by the peptide sequence marks the N terminus of the mature MOMP. The VDs and the conserved cysteines are boxed. A putative Shine-Dalgarno sequence is underlined. The 11-base transcription terminator dyad is indicated by inverted arrows from nucleotide 1202 to 1229.

The complete sequence of the *C. pneumoniae* MOMP structural gene and its predicted amino acid (AA) sequence is shown in Fig. 1. This gene is analogous to the *C. trachomatis* and *C. psittaci omp1*. Based on the recent recommendations by Yuan et al. (19), the *C. pneumoniae* MOMP gene has been designated *ompA*.

The *C. pneumoniae* MOMP gene open reading frame consists of 1,167 bp which encode a 389-AA pre-MOMP sequence. By comparison with known MOMP sequences, in

which the mature N terminus is a leucine preceded by a 22-AA leader sequence (16, 20, 21), we deduced that the *C. pneumoniae* MOMP leader sequence is composed of 23 AA. The leader peptide has the conserved basic amino terminus and the hydrophobic core found in the other chlamydia MOMPs, while the rest of the sequence shows significant heterogeneity. The first six residues of the mature MOMP sequences of *C. trachomatis* (serovars A, B, C, and L2), *C. psittaci*, guinea pig inclusion conjunctivitis, meningopneu-

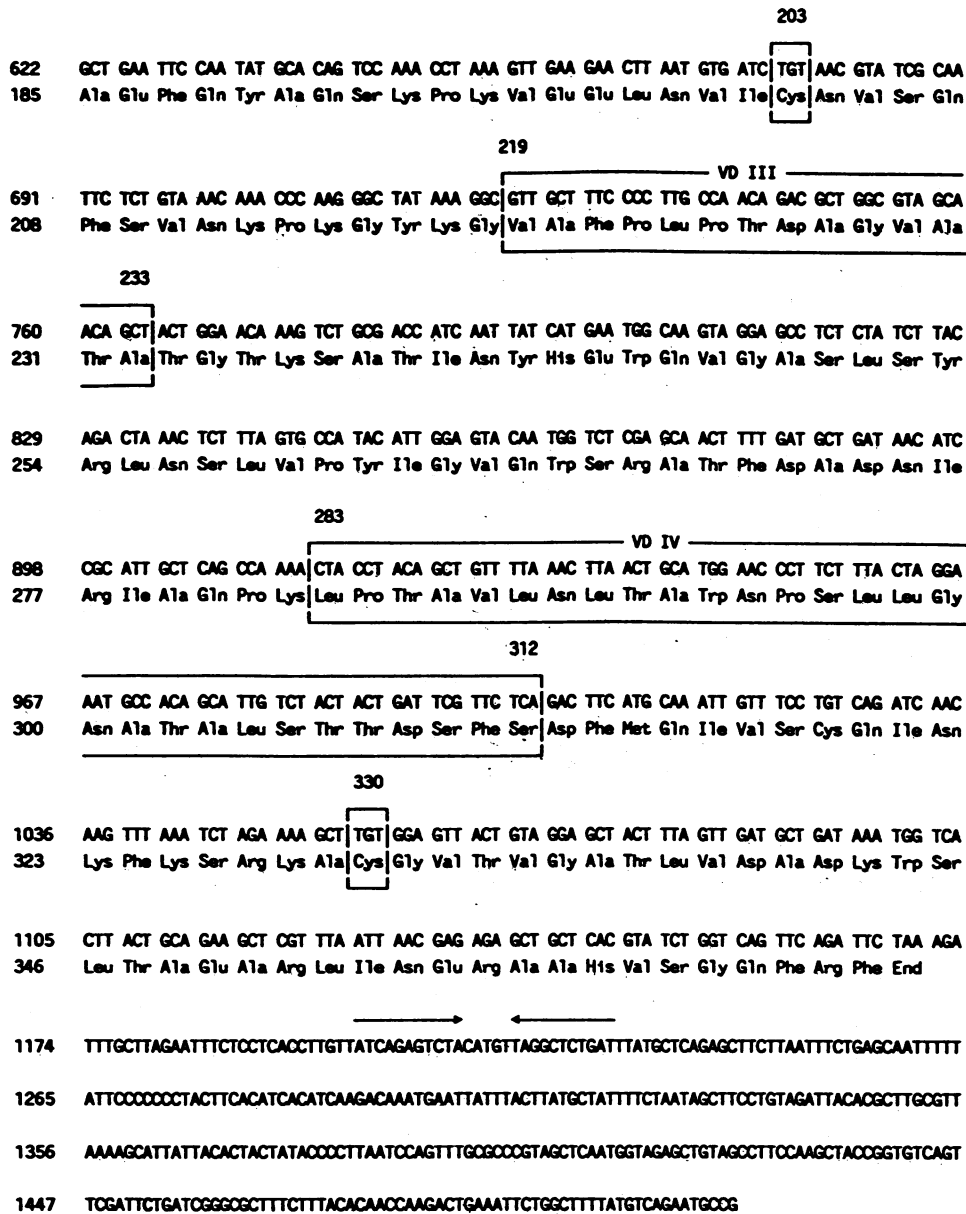


FIG. 1—Continued

monitis (22), ovine abortion (OA), and *C. pneumoniae* are identical. The mature *C. pneumoniae* MOMP contains 366 AA with a predicted molecular weight of 39,344, which is consistent with the observed mobility of the *C. pneumoniae* MOMP on sodium dodecyl sulfate-polyacrylamide gels (3). Comparison of the inferred AA sequence of the *C. pneumoniae* MOMP with those of the *C. trachomatis* L2 MOMP and the *C. psittaci* OA MOMP demonstrated 73 and 80% similarity, respectively (Fig. 2). The *C. pneumoniae* MOMP contains 33 basic and 31 acidic AA residues and has a weakly acidic estimated pI, as for the MOMPs of *C. psittaci* meningopneumitis and guinea pig inclusion conjunctivitis (21). In the *C. trachomatis* MOMPs the number of acidic residues is significantly higher, 39 to 42, resulting in a lower pI (2, 21).

Comparison of the *C. trachomatis* L2, *C. psittaci* OA, and *C. pneumoniae* AR-39 MOMP AA sequences (Fig. 2) indicate that, as observed for other chlamydial MOMPs, the

proteins are interspersed with four VDs (VDI to VDIV) and that all insertions and deletions occurred in these VDs (16, 20, 21). The seven cysteines observed in the MOMPs of the other chlamydial species (21) are in exactly the same positions in the *C. pneumoniae* MOMP. Three of them are located in the N-terminus region before VDI, one is located between VDI and VDII, two are located between VDII and VDIII, and the last one is located in the C-terminus region after the VDIV domain (Fig. 2). One cysteine (in position 203) which is conserved in all known *C. trachomatis* MOMP but not in the *C. psittaci* MOMP appears to be conserved in the *C. pneumoniae* MOMP (Fig. 2). The cysteine residues allow the formation of intra- and interdisulfide bonds (12, 21). These play an important role in the maintenance of the structural integrity of chlamydia, in the formation of diffusion channels, and in the regulation of chlamydial differentiation (8, 12).

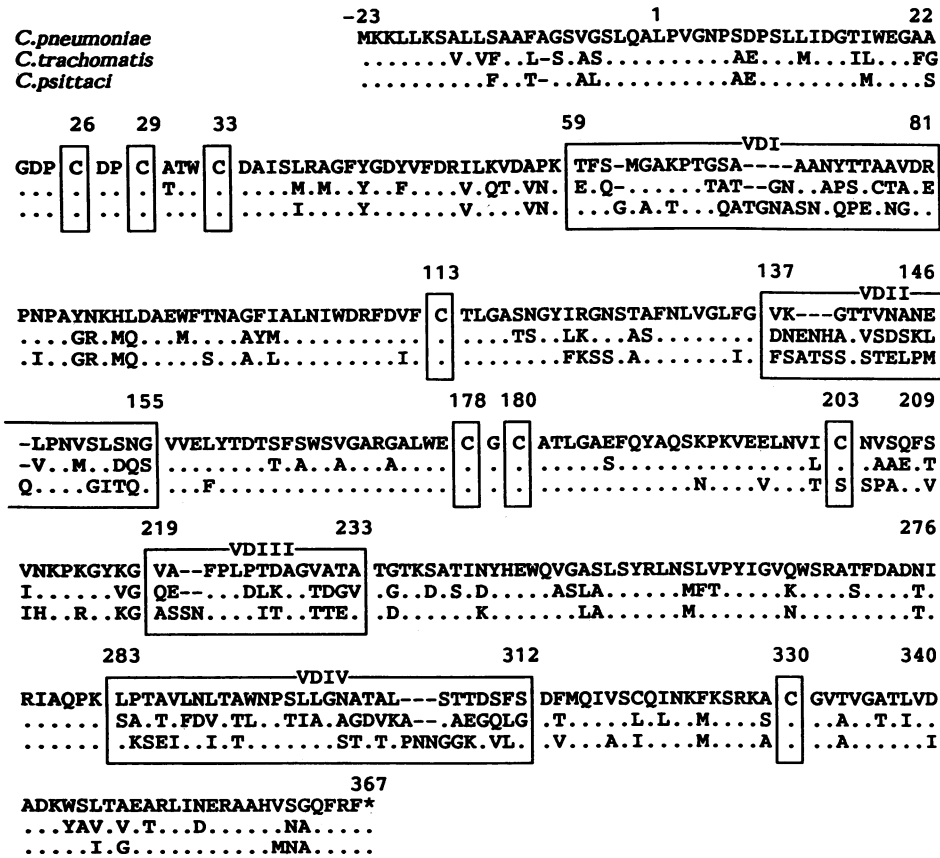


FIG. 2. AA comparison of *C. pneumoniae* AR-39 MOMP with *C. trachomatis* L2 and *C. psittaci* OA MOMPs. A dot indicates the same AA as in the *C. pneumoniae* MOMP and a dash represents a gap in the sequence. The four VDs and the conserved cysteines are boxed. Numbers above the sequences denote AA residues of the *C. pneumoniae* MOMP.

The overall interspecies conservation of the *C. pneumoniae* MOMP gene with the *C. trachomatis* and *C. psittaci* MOMP genes is 68 and 71%, respectively. This degree of similarity is comparable to that observed between the *C. trachomatis* and *C. psittaci* MOMP genes (21). In the *C. pneumoniae* MOMP, the translation initiator ATG is preceded by a Shine-Dalgarno-like ribosomal-binding site, as has been shown for other MOMP genes (Fig. 1) (16, 21). While putative promoter and terminator regulatory regions have been identified in other chlamydial genes (14), no chlamydial consensus promoter sequences have been identified. Within the *C. trachomatis* and *C. psittaci* MOMP genes, multiple tandem promoters have been mapped (17, 19). The 5' flanking area of the *C. pneumoniae* MOMP gene shows 75.5 and 86.7% similarity with the corresponding regions of the *C. trachomatis* L2 and *C. psittaci* OA MOMP genes. The terminator sequences reported for other chlamydial genes demonstrate typical rho-independent stem-loop structures (6, 16). In contrast, the region downstream of the translational stop signals in the *C. pneumoniae* MOMP gene revealed an 11-base dyad which forms a stem-loop structure followed by a second stem-loop structure, but there is no string of thymidine residues characteristic of the rho-independent terminator (Fig. 1). Transcription of the 11-base dyad would result in the folding of the mRNA into a stem-loop structure with a calculated least free energy of -14.0 kcal. If the transcription of the second stem-loop structure is included in this calculation, a lower calculated

least free energy of -27.4 kcal would result. The terminator sequence for the *C. pneumoniae* 75-kDa gene also contains two stem-loop structures and lacks a poly(T) tail (10).

In the case of *C. trachomatis* and *C. psittaci*, the MOMP is a complex antigen that is immunodominant. Molecular analysis of the *C. pneumoniae* MOMP gene has shown that it has significant similarities in structure and sequence to MOMPs of the other chlamydial species, except in the VDs, where they share little sequence similarity. For *C. trachomatis*, the AA sequence variation within the VDs provides the molecular basis for serologic classification of *C. trachomatis*. To date, all monoclonal antibodies that we have prepared which recognize the *C. pneumoniae* MOMP are genus specific. Thus, we have been unable to identify any serological domains on the *C. pneumoniae* MOMP. Further sequence analysis and comparison of the MOMPs of other *C. pneumoniae* isolates will aid in elucidating the relationship of the VDs to the antigenic nature of the *C. pneumoniae* MOMP.

This study was supported by Public Health Service grant AI 21885.

We thank Lynne C. Kikuta for technical assistance and David T. Grayston for assistance in computer analysis of the data.

REFERENCES

1. Baehr, W., Y.-X. Zhang, T. Joseph, H. Su, F. E. Nano, K. D. E. Everett, and H. D. Caldwell. 1988. Mapping antigenic domains

- expressed by *Chlamydia trachomatis* major outer membrane protein genes. Proc. Natl. Acad. Sci. USA **85**:4000-4004.
2. Bavoil, P., A. Ohlin, and J. Shachter. 1984. Role of disulfide bonding in outer membrane structure and permeability in *Chlamydia trachomatis*. Infect. Immun. **44**:479-485.
 3. Campbell, L. A., C.-C. Kuo, and J. T. Grayston. 1990. Structural and antigenic analysis of *Chlamydia pneumoniae*. Infect. Immun. **58**:93-97.
 4. Campbell, L. A., C.-C. Kuo, S.-P. Wang, and J. T. Grayston. 1990. Serological response to *Chlamydia pneumoniae* infection. J. Clin Microbiol. **28**:1261-1264.
 5. Conlan, J. W., I. N. Clarke, and M. E. Ward. 1988. Epitope mapping with solid-phase peptides: identification of type-, sub-species-, species-, and genus-specific antibodies binding domains on the major outer membrane protein of *Chlamydia trachomatis*. Mol. Microbiol. **2**:673-679.
 6. Danilition, S., I. W. Maclean, R. Peeling, S. Winston, R. C. Brunham. 1990. The 75 kilodalton protein of *Chlamydia trachomatis*: a member of the heat shock protein 70 family? Infect. Immun. **58**:189-196.
 7. Grayston, J. T., L. A. Campbell, C.-C. Kuo, C. H. Mordhorst, P. Saiku, D. H. Thorn, and S.-P. Wang. 1990. A new respiratory tract pathogen: *Chlamydia pneumoniae*, strain TWAR. J. Infect. Dis. **161**:618-625.
 8. Hackstadt, T., W. T. Todd, and H. D. Caldwell. 1985. Disulfide-mediated interactions of the chlamydial major outer membrane protein: role in differentiation of chlamydiae? Infect. Immun. **161**:25-31.
 9. Huynh, T. V., R. A. Young, and R. W. Davis. 1985. Constructing and screening DNA libraries in lambda gt10 and lambda gt11, p. 49-78. In D. M. Glover (ed.), DNA cloning, vol. I. IRL Press, Washington, D.C.
 10. Kornak, J. M., C.-C. Kuo, and L. A. Campbell. 1991. Sequence analysis of the gene encoding the *Chlamydia pneumoniae* dnaK protein homolog. Infect. Immun. **59**:721-725.
 11. Kuo, C.-C., H.-H. Chen, S.-P. Wang, and J. T. Grayston. 1986. Identification of a new group of *Chlamydia psittaci* strains called TWAR. Infect. Immun. **24**:1034-1037.
 12. Newhall, W. J., V. 1988. Macromolecular and antigenic composition of chlamydiae, p. 47-70. In A. L. Barron (ed.), Microbiology of *Chlamydia*. CRL Press, Inc., Boca Raton, Fla.
 13. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.
 14. Sardinia, L. M., J. N. Engel, and D. Ganem. 1989. Chlamydial gene encoding a 70-kilodalton antigen in *Escherichia coli*: analysis of expression signals and identification of the gene product. J. Bacteriol. **171**:335-341.
 15. Stephens, R. S., C.-C. Kuo, G. Newport, and N. Agabian. 1985. Molecular cloning and expression of *Chlamydia trachomatis* major outer membrane protein antigens in *Escherichia coli*. Infect. Immun. **47**:713-718.
 16. Stephens, R. S., R. Sanchez-Pescador, E. A. Wagar, C. Inouye, and M. S. Urdea. 1987. Diversity of *Chlamydia trachomatis* major outer membrane protein genes. J. Bacteriol. **169**:3879-3885.
 17. Stephens, R. S., E. A. Wagar, and U. Edman. 1988. Developmental regulation of tandem promoters for the major outer membrane protein gene of *Chlamydia trachomatis*. J. Bacteriol. **170**:744-750.
 18. Stephens, R. S., E. A. Wagar, and G. K. Schoolnik. 1988. High-resolution mapping of serovar-specific and common antigenic determinants of the major outer membrane protein of *Chlamydia trachomatis*. J. Exp. Med. **167**:817-831.
 19. Yuan, Y., Y.-X. Zhang, D. S. Manning, and H. D. Caldwell. 1990. Multiple tandem promoters of the major outer membrane protein gene (*omp1*) of *Chlamydia psittaci*. Infect. Immun. **58**:2850-2855.
 20. Yuan, Y., Y.-X. Zhang, N. G. Watkins, and H. D. Caldwell. 1989. Nucleotide and deduced amino acid sequences for the four variable domains of the major outer membrane proteins of the 15 *Chlamydia trachomatis* serovars. Infect. Immun. **57**:1040-1049.
 21. Zhang, Y.-X., S. G. Morrison, H. D. Caldwell, and W. Baehr. 1989. Cloning and sequence analysis of the major outer membrane protein genes of the outer membrane proteins of two *Chlamydia psittaci* strains. Infect. Immun. **57**:1621-1625.
 22. Zhang, Y.-X., S. J. Stewart, and H. D. Caldwell. 1989. Protective monoclonal antibodies to *Chlamydia trachomatis*. Serovar- and serogroup-specific major outer membrane protein determinants. Infect. Immun. **57**:636-638.