INFECTION AND IMMUNITY, Sept. 1990, p. 2850–2855 Vol. 58, No. 9
0019-9567/90/092850-06$02.00/0
Copyright © 1990, American Society for Microbiology

Multiple Tandem Promoters of the Major Outer Membrane Protein Gene (*omp*1) of *Chlamydia psittaci*

YING YUAN, YOU-XUN ZHANG, D. SCOTT MANNING, AND HARLAN D. CALDWELL*

Laboratory of Intracellular Parasites, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, Hamilton, Montana 59840

Received 9 April 1990/Accepted 24 May 1990

The transcription of *omp*1, the gene encoding the major outer membrane protein, was studied for two strains of *Chlamydia psittaci*, guinea pig inclusion conjunctivitis (GPIC) and mouse pneumonitis (Mn). The transcriptional initiation sites for the *omp*1 of each strain were mapped by S1 nuclease and primer extension analyses. Three different sizes of *omp*1 transcripts were observed for GPIC and four were observed for Mn. The production of these transcripts appeared to be the consequence of multiple tandem promoters. The order in which the *omp*1 RNA transcripts appeared during the growth cycle of the *C. psittaci* strains was found to differ from that of *C. trachomatis*.

Chlamydia, a genus of obligate intracellular procaryotes, is composed of two species, *Chlamydia trachomatis* and *Chlamydia psittaci*. *C. trachomatis* is a pathogen of humans and is one of the primary causes of preventable blindness. It is also a common agent of genital tract infections. *C. psittaci* has a broad nonhuman host range and causes a wide spectrum of animal diseases.

The chlamydiae are distinguished from other intracellular procaryotes by their unique biphasic growth cycle that alternates between the extracellular infectious elementary body (EB) and the intracellular replicative reticulate body (RB). The infectious EB initiates the growth cycle by attaching to the host cell. It is quickly endocytosed and confined within an endocytic vesicle that does not fuse with host lysosomes. Three hours after internalization, the metabolically inactive EB begins its differentiation into the metabolically active RB. By 10 to 12 h postinfection, the mature RB is formed and starts binary fission. At 10 to 20 h after the beginning of cell division, the RBs differentiate back to infectious EBs which are subsequently released from the host cell (11, 12).

Little is known about the genetic regulation of this biphasic life cycle. Stephens et al. (14) have, however, investigated the transcription of the major outer membrane protein (MOMP) gene (*omp*1) during the life cycle of the L2 serotype of *C. trachomatis*. MOMP is believed to serve important functions at both the EB and RB stages of development (1, 2, 9). It contributes to the rigidity of the EB by forming inter- and intramolecular disulfide bonds. In the RB it serves as a porin, facilitating the transport of nutrients. It is probable that the regulation of MOMP expression is a critical event during the chlamydial developmental cycle.

Stephens et al. (14) found that the *omp*1 gene of the *C. trachomatis* L2 serovar had two tandem promoters arranged upstream from a single structural gene. One of these promoters produced an mRNA transcript which was detected 4 h postinfection. The second promoter was responsible for production of a slightly larger transcript which appeared at 12 h postinfection.

Although *omp*1 sequences are largely conserved within a species, there are four domains that vary between serovars or strains. There is about 65% nucleotide sequence homology between members of the *C. psittaci* and *C. trachomatis* species. Compared with *C. trachomatis*, the MOMP variable domains of *C. psittaci* are longer and encode more amino acid variations (17).

In this paper, we describe the identification of multiple tandem promoters for the *omp*1 genes of two *C. psittaci* strains, guinea pig inclusion conjunctivitis (GPIC) and mouse pneumonitis (Mn). These results are contrasted with those obtained for *C. trachomatis* (14).

MATERIALS AND METHODS

Chlamydial strains and RNA extraction. Total RNA was extracted from HeLa cells infected with *C. psittaci* strains GPIC 1 and Mn Cal 10 (4, 7) and the *C. trachomatis* serovar E/Bour 10 as described previously (16). At 24 h postinfection, the cells were lysed in cold 4 M guanidine isothiocyanate lysis buffer and total RNA was extracted with hot acidic phenol. The RNAs were precipitated with 2 volumes of cold ethanol at −20°C, dried, and then suspended in water.

Nucleotide probes. Three restriction fragments, MNM-1, GPM-1, and DC-9 (17), from cloned MOMP genes of the Mn and GPIC strains were used as probes in Northern (RNA) hybridization and S1 nuclease mapping (Fig. 1). Four 21-base synthetic oligonucleotides, GM1, GM2, GM3, and M4 (Fig. 1), were used in primer extension sequencing and S1 analysis of *omp*1 mRNA.

Northern hybridization. Northern hybridization analysis of RNA collected at different times postinfection was performed by the method of Lehrach et al. (6). Cloned *omp*1 fragments, DC-9 and MNM-1 (Fig. 1), were used as probes. Approximately 40 μg of total mRNA in sample solution (1× MOPS [morpholinepropanesulfonic acid] [pH 7.0], 2.2 M formaldehyde, dye mix, 50% formamide) was denatured at 56°C for 15 min, loaded onto a 2.2 M formaldehyde–1.4% agarose gel (low EEO; Sigma Chemical Co., St. Louis, Mo.), and electrophoresed at 5 V/cm for 4.5 h. After the gel was washed with distilled water, the mRNA was transferred onto GeneScreen-Plus hybridization transfer membrane (NEN Research Products, Wilmington, Del.), prehybridized, and then hybridized at 42°C overnight with MNM-1 or DC-9 fragments randomly labeled with [^{32}P]dCTP (random primed DNA labeling kit; Boehringer Mannheim Biochemicals, Indianapolis, Ind.). The membrane was washed twice in 1×

* Corresponding author.

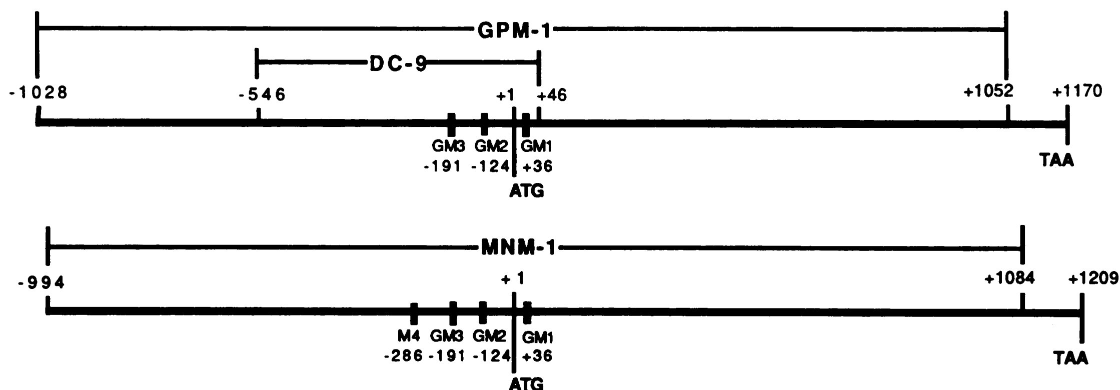

FIG. 1. Locations of DNA fragments and oligonucleotide primers used as probes in Southern blot, S1 nuclease, and primer extension analyses of *C. psittaci* mRNA transcripts. Nucleotide sequence positions are numbered with the adenine residues of the translational initiation codons (ATG) as +1. GPM-1 and DC-9 are cloned fragments of *omp1* of the GPIC strain. MNM-1 is a cloned fragment of *omp1* of the Mn strain. The hybridization sites of the 21-base oligonucleotides used in primer extension assays are indicated by the 3'-terminal nucleotide of each site. The sequences complementary to primers GM-1 (5'-GGCAAACAATAATGCCGATTT), GM-2 (5'-TGAGGAGTGTCTTG CAAGTGA), and GM-3 (5'-CAAACTTTTCTTATTATAGT) are identical for both strains. The primer M-4 is complementary to a sequence (5'-ATGTGTACAAAAATCTGATTAG) that is unique to the Mn strain.

SSC (0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate for 15 min at room temperature and then subjected to autoradiography. As a positive control, total RNA from cells infected with *C. trachomatis* E serovar (E/Bour) was extracted at different times postinfection and analyzed by Northern blot. A random labeled 40-base oligonucleotide from a conserved region (+80 to +120) of *C. trachomatis omp1* was used as the probe.

S1 nuclease mapping. The 5' ends of *omp1* transcripts were analyzed by S1 nuclease protection assays (13, 15). The oligonucleotide primer GM-1 (100 ng, [^{32}P]dATP labeled) and plasmids (20 μg each) containing either MNM-1 or GPM-1 gene fragments were denatured and annealed at 40°C for 15 min in 0.1 M Tris (pH 8.0) containing 100 mM $MgCl_2$ (TM buffer). The primers were extended at 37°C for 30 min in a reaction solution with 400 μM deoxynucleoside triphosphate mix and 10 U of the Klenow fragment of *Escherichia coli* DNA polymerase I. The extension products were cleaved by restriction enzymes, *Xba*I for MNM-1 and *Hin*dIII for GPM-1. The two radiolabeled DNA fragments were isolated by 1.2% alkaline agarose gel electrophoresis and used as probes in the S1 nuclease analysis. Approximately 2×10^4 to 4×10^4 Cerenkov cpm of probe were mixed with 80 μg of RNA in 0.3 M sodium acetate and precipitated by 70% ethanol. The mixture was added to 30 μl of S1 hybridization buffer containing 80% formamide, 40 mM PIPES (piperazine-*N*,*N*'-bis[2-ethanesulfonic acid]) (pH 6.4), 0.4 M NaCl, and 1 mM EDTA, heated to 90°C for 5 min, and then incubated at 42°C overnight. After hybridization, the mixtures were iced, and 300 μl of cold solution containing 0.28 M NaCl, 0.05 M sodium acetate (pH 4.6), 4.5 mM $ZnSO_4$, and 300 U of nuclease S1 (225,000 U/ml; Pharmacia LKB Corp., Piscataway, N.J.) was added. After 5 min on ice, the mixtures were incubated at 25°C for 1 h. The digestion was stopped by extracting with phenol-chloroform. The protected fragments were precipitated by ethanol and analyzed by electrophoresis on 6% denaturing polyacrylamide-urea sequencing gels.

Primer extension assay. The methods for primer extension sequencing have been described elsewhere (16). Four primers homologous to sequences located in *omp1* (Fig. 1) were used to identify the 5'-terminal nucleotide sequences of the mRNA transcripts of the GPIC or Mn strain. These 21-base oligonucleotide primers were 5'-end labeled with [α-^{32}P]dATP (3,000 Ci/mmol; NEN Research Products) by T4 polynucleotide kinase (10,000 U/ml; New England BioLabs, Beverly, Mass.). The 30 to 70 μg of total RNA was hybridized to the primer in annealing buffer by allowing the heated mixture to cool down slowly to 45°C, and then the primer extensions were performed in a reaction solution catalyzed by 0.6 to 0.8 U of avian myeloblastosis virus reverse transcriptase (20,000 U/ml; Pharmacia LKB Corp.) per μl. The samples were analyzed on 8% polyacrylamide gels.

RESULTS AND DISCUSSION

Multiple mRNA transcripts. In both the GPIC and Mn strains, Northern analysis of mRNA at 24 h postinfection showed two major transcripts with molecular weights of approximately 1,350 (T1) and 1,500 (T2) (Fig. 2A). There also appeared to be specific RNA species migrating above the 1.6-kilobase band for both GPIC and Mn. S1 nuclease mapping and primer extension (Fig. 2B) confirmed an additional higher-molecular-weight transcript in GPIC (T3) and two more in Mn (T3 and T4). The T3 and T4 transcripts were probably not clearly resolved by Northern blot analysis because they were less abundant and migrated close to the T2 transcripts on agarose gels.

Southern analysis of genomic DNAs digested with a variety of restriction enzymes and probed with fragments of cloned *omp1* indicated that there was a single gene copy (data not shown). The multiple transcripts of *omp1* must then result from a single gene with tandem transcriptional initiation sites.

The observation that there are three or four *omp1* gene transcripts produced by the GPIC and Mn strains of *C. psittaci* is in contrast to a report that there are only two *omp1* transcripts in the L2 serovar of *C. trachomatis* (14). We have conducted Northern and primer extension analyses on all 15 serovars of *C. trachomatis* and found that only two *omp1* transcripts were produced by all members of this species (unpublished results).

Transcription patterns during the growth cycle. In Northern blot analysis of the mRNA obtained at different times postinfection (Fig. 3), T2 was the first transcript detected for both Mn and GPIC. It appeared as early as 4 h postinfection

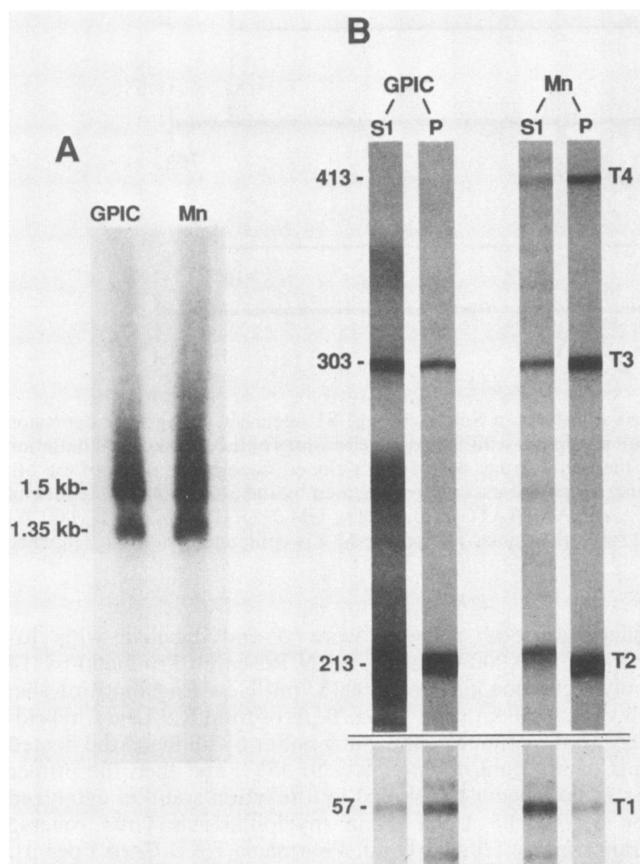

FIG. 2. (A) Identification of *omp*1 mRNA transcripts in *C. psittaci* strains GPIC and Mn. Total RNA was extracted from HeLa 229 cells 24 h postinfection. The DNA fragment DC-9 was randomly labeled with ^{32}P and used as a probe in Northern analysis. kb, Kilobases. (B) Primer extension (P) and S1 nuclease (S1) analyses of *omp*1 transcripts of GPIC and Mn strains were compared. The oligonucleotide GM-1 was used for primer extensions analysis. Probes for S1 nuclease analysis were synthesized by using DNA fragments DC-9 and MNM-1 as templates and GM-1 as a primer. Fragments of similar sizes (indicated in bases) were obtained by each method. Four transcripts were identified and designated T1 through T4, with T1 having the lowest molecular weight.

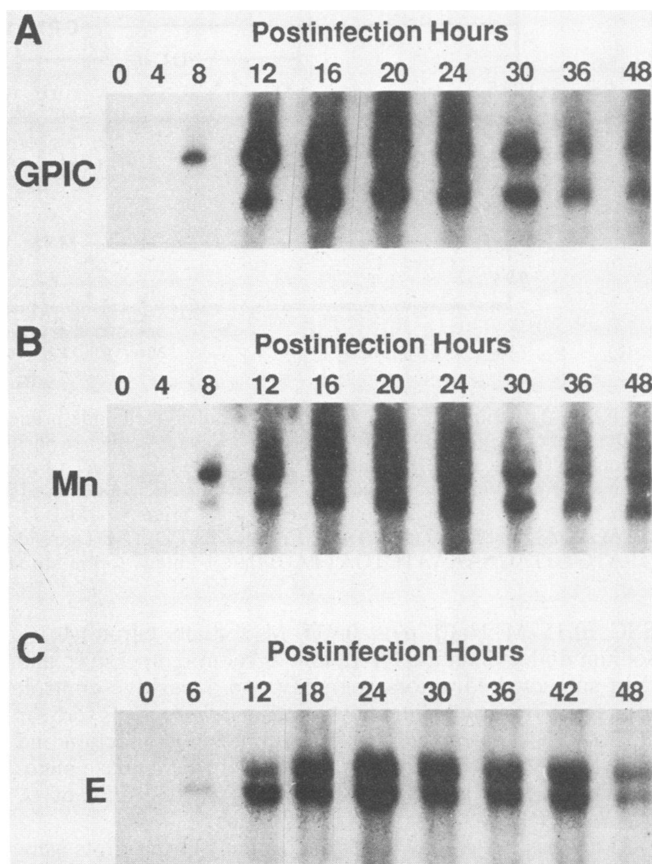

FIG. 3. Northern blot analysis of total RNA extracted from chlamydia-infected HeLa 229 cells at different times postinfection. (A) RNA transcripts from the GPIC strain were hybridized with the DNA fragment DC-9. (B) RNA transcripts for the Mn strain were also detected with DC-9. (C) RNA transcripts from *C. trachomatis* serovar E were detected by hybridization with a 40-base oligonucleotide which recognizes a conserved region (+81 to +120) of *omp*1 in all *C. trachomatis* serovars.

in Mn and 8 h in GPIC. The lower-molecular-weight transcript, T1, was observed 4 h after T2. The signals for the T1 and T2 transcripts increased until 24 to 30 h postinfection. They then began to decrease but were still detectable until 48 h, when one growth cycle was completed. These results contrast with those for *C. trachomatis* serovars E (Fig. 3C) and L2 (14) in which the lower-molecular-weight transcript T1 is the first to appear.

To determine when transcripts T3 and T4 were produced, primer extension analysis was conducted on RNA obtained at different times postinfection with Mn (Fig. 4). The T3 transcripts were produced at the same time as T1. The T4 transcript appeared the latest and was only a minor component of the *omp*1 mRNA.

Determination of the 5′ end(s) of *omp*1 transcripts. The transcriptional initiation sites of the *omp*1 of GPIC and Mn were identified by using oligonucleotide primer extension sequencing (results for Mn shown in Fig. 5). The locations of the initiation sites were −21, −177, −267, and −377 in Mn and −21, −177, and −267 in GPIC, with the adenine residue of the translational initiation codon, ATG, as +1 in each case. The locations of putative promoter sequences are shown in Fig. 6A. The comparisons of the predicted promoter regions show that the nucleotide sequences of Mn and GPIC are identical within the −10 and −35 regions of the putative promoters P1, P2, and P3, except that there is one nucleotide difference in the −35 region of the P3s. There were two nucleotide differences in the −10 region and one in the −35 region of the Mn P4 and the corresponding sequences of GPIC. These differences may account for the lack of P4 activity in GPIC. The P2 and P3 putative promoters contain TATGTT and TATTAG in −10 regions and TTTACT and TGTACA in −35 regions (Fig. 6B). These sequences are similar to the consensus sequences of canonical *E. coli* promoters (−10 region, TATAAT; −35 region, TTGACA) (5). The −10 region of the Mn P4 (TAATGA) has less similarity to the consensus sequence, and this may be why P4 is the weakest among four promoters.

Like the P1 of *C. trachomatis omp*1 (14), the P1s of both GPIC and Mn do not appear to have any known consensus promoter sequences even though a substantial portion of the *omp*1 transcripts map to the region. This suggests that some novel form of regulation may be involved in the production of transcripts from these sites. There is a high degree of homology between the sequences of *C. psittaci* and *C.*

FIG. 4. To determine the order in which each of the *omp*1 transcripts appears, total RNA from HeLa cells infected with the Mn strain was extracted at different times postinfection and analyzed by primer extension. The oligonucleotide GM-1 was used as the primer. The four transcripts are labeled T1 through T4.

trachomatis serovars for approximately the first 200 nucleotides 5′ of the *omp*1 open reading frame (Fig. 6). The significance of these conserved sequences is unclear at this time. They may, however, contain recognition sites for alternative sigma factors or other regulatory proteins for the putative P1 promoters. It is interesting that the first appearance of the smallest transcripts (T1s) occurs at different times during the developmental cycles of the two species. In *C. trachomatis* production of the smallest transcript begins soon after infection and continues throughout the growth cycle; thus, it has been described as being constitutively regulated (14). The production of the larger transcript begins at a time which correlates with the onset of binary fission and thus has been proposed to be under developmental regulation (14). In *C. psittaci*, T2 is produced earliest and T1, the smallest transcript, is not produced until a time which correlates with the beginning of binary fission. The largest transcripts, T3 and T4, also appear at this time even though their promoter regions resemble those of T2. Thus, the hypothesis that the promoter for the smallest transcript, P1, is constitutively regulated while the promoters containing recognizable consensus sequences are developmentally regulated appears not to hold for *C. psittaci*.

A 9-base-pair inverted repeat separated by 13 base pairs occurs at the 5′ end of the conserved sequence in the *C. psittaci* and *C. trachomatis omp*1 promoters. The −35 region of P2 in the *C. psittaci* strains lies between the inverted repeats. A perfect −10 consensus sequence, TATAAT, is also present in this region, but since it is located 26 bases from the site of initiation of transcription and does not have a corresponding −35 region, it is unclear whether this sequence is recognized as a promoter. The significance of the conserved inverted repeat structure is unknown, although it has been observed (14) that a similar structure is found in a *Bacillus subtilis* promoter which is

FIG. 5. Identification of the 5′ ends of Mn strain *omp*1 transcripts by primer extension mRNA sequencing. Total RNA was extracted from HeLa 229 cells 24 h after infection. Primers GM-1, GM-2, GM-3, and M-4 were used to identify the initiation sites of *omp*1 transcripts T1 to T4, respectively. Dots in the sequences indicate residues at the 5′ end which could not be determined with certainty. Lanes are labeled with the nucleotide complimentary to the dideoxynucleotide used for each primer extension reaction.

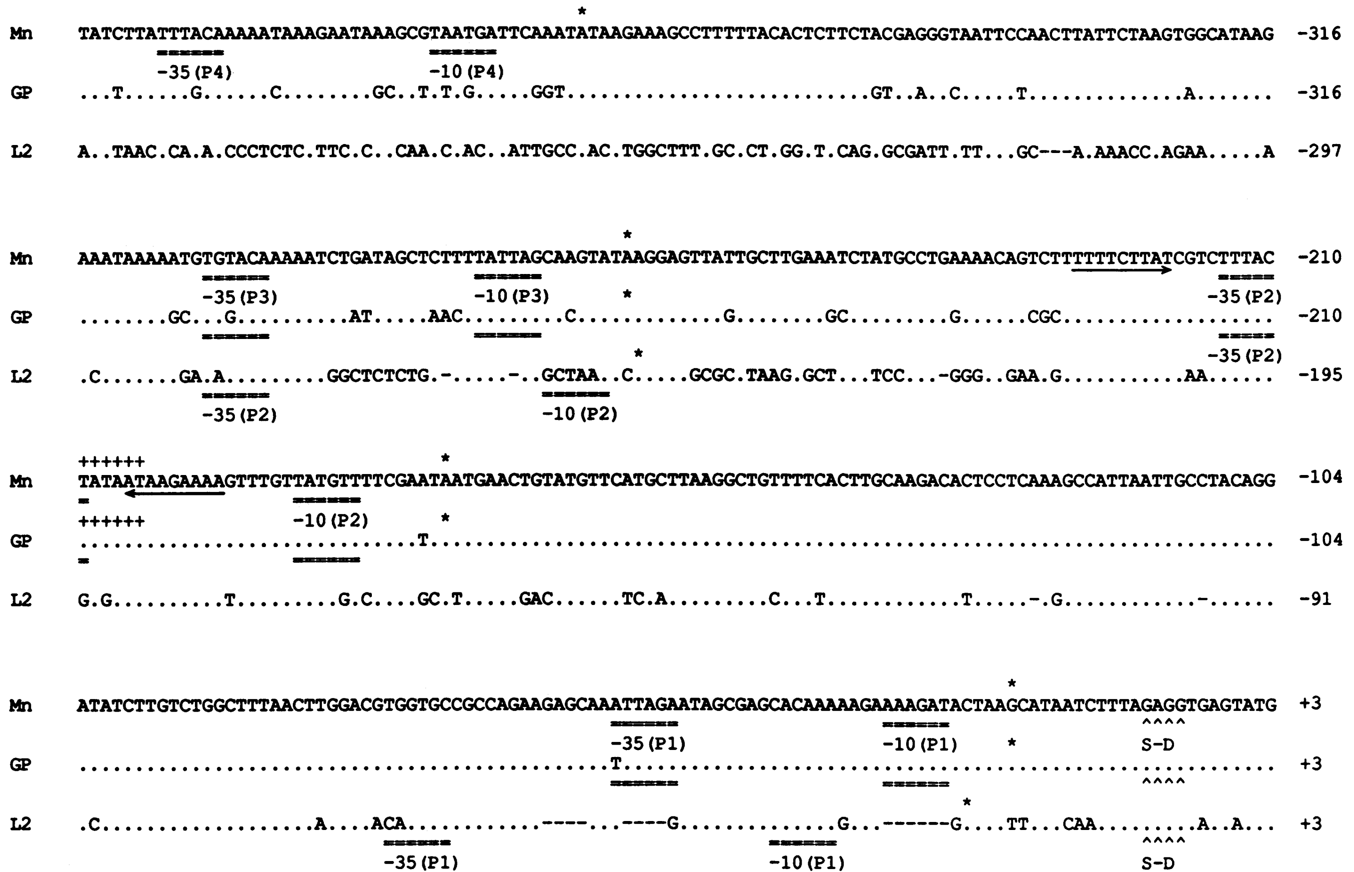

FIG. 6. Sequence comparison of the promoter regions of *C. psittaci* strains GPIC and Mn and the *C. trachomatis* serovar L2. The sequence of the Mn strain is shown as a prototype. For GPIC and L2, identical bases are indicated by dots and deletions are shown by dashes. The predicted promoters for transcripts T1 to T4 are underlined and labeled P1 to P4, respectively. The sites of initiation for each transcript are indicated by a star. Plus signs designate the location of a canonical *E. coli* −10 sequence. Putative Shine-Dalgarno (S-D) sites are marked. Arrows denote a conserved inverted repeat.

regulated by a shift in sigma factors associated with sporulation (10).

While the P2 of the *C. psittaci* strains apparently resides in the region of the inverted repeats, it does not appear that a promoter exists in the corresponding region of the L2 serovar of *C. trachomatis*. The P2 of *C. trachomatis* does, however, reside in a region with some homology to the P3 of the *C. psittaci* strains (Fig. 6).

The adenine-plus-thymidine (A+T)-rich nature of the promoter regions of *C. trachomatis* is also observed for *C. psittaci*. There is approximately 70% A+T among the 500 nucleotides upstream from the methionine codon of GPIC and Mn *ompl* open reading frames. Promoters of several highly expressed genes have been found to be A+T rich (8), and it has been suggested that the A+T richness of the *ompl* promoter regions may contribute to RNA polymerase binding (14).

It appears that in both *C. psittaci* and *C. trachomatis* species the transcription of *ompl* is required early in the growth cycle and multiple promoters in some way provide for supplemental transcription as binary fission commences. The mechanisms involved in developmental regulation of these transcriptional events, however, remain poorly understood. The role of the smallest transcripts (T1s) of both species is particularly intriguing, since the putative promoter regions contain no recognizable consensus promoter sequences. The *C. trachomatis* rRNA genes also do not have recognizable promoter consensus sequences (3). In this case two tandem transcripts are produced coordinately beginning as early as 9 h postinfection. Comparisons of the putative P1 promoter regions of the *ompl* with the rRNA promoters do not reveal any obvious regions of homology. As more chlamydial promoters are studied, regulatory sequences may be identified for those promoters which do not appear to have conventional consensus sequences. Since *ompl* genes are poorly expressed in *E. coli* (14) and since transformation of chlamydia is not yet possible, site-directed mutagenesis cannot be used to identify what sequences are critical to transcription in chlamydial promoters. More direct investigation of the sequences involved in the regulation of these chlamydial genes may then require the design of new experimental approaches. Expression of cloned chlamydial sigma factors and other regulatory proteins in *E. coli* may allow for analysis of chlamydial promoters in an *E. coli* background. Alternatively, ways will have to be found to reintroduce cloned genes into chlamydia.

As a closing comment, we would like to suggest to our colleagues that the designation for the gene encoding MOMP be changed from *ompl* to *ompA*. This designation would be in conformity with general genetic nomenclature in which capital, italicized letters are used for cistron designations.

ACKNOWLEDGMENTS

We thank Sandra Morrison for synthesizing oligonucleotides, Gary Hettrick for photography, Susan Smaus for secretarial assistance, Paul Policastro for suggesting the Northern blot protocol, and Patricia Rosa and Robert Belland for constructive review of the manuscript.

LITERATURE CITED

1. **Bavoil, P., A. Ohlin, and J. Schachter.** 1984. Role of disulfide bonding in outer membrane structure and permeability in *Chlamydia trachomatis*. Infect. Immun. **44:**479–485.
2. **Caldwell, H. D., J. Kromhout, and J. Schachter.** 1981. Purification and partial characterization of the major outer membrane protein of *Chlamydia trachomatis*. Infect. Immun. **31:**1161–1176.
3. **Engel, J. N., and D. Ganem.** 1987. Chlamydial rRNA operons: gene organization and identification of putative tandem promoters. J. Bacteriol. **169:**5678–5685.
4. **Francis, T., Jr., and T. O. Magill.** 1938. An unidentified virus producing acute meningitis and pneumonia. J. Exp. Med. **68:** 147–160.
5. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
6. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determinations by gel electrophoresis and denaturing conditions, a critical reexamination. Biochemistry **16:**4743–4751.
7. **Murray, E. S.** 1964. Guinea pig inclusion conjunctivitis virus. I. Isolation and identification as a member of the psittacosis-lymphogranuloma-trachoma group. J. Infect. Dis. **114:**1–12.
8. **Nakamura, K., and M. Inouye.** 1979. DNA sequence of the gene for the outer membrane lipoprotein of *E. coli*: an extremely AT-rich promoter. Cell **18:**1109–1117.
9. **Newhall, W. J., V., and R. B. Jones.** 1983. Disulfide-linked oligomers of the major outer membrane protein of chlamydiae. J. Bacteriol. **154:**998–1001.
10. **Rather, P. N., R. E. Hay, G. L. Ray, W. G. Haldenwang, and C. P. Moran, Jr.** 1986. Nucleotide sequences that define promoters that are used by *Bacillus subtilis* Sigma-29 RNA polymerase. J. Mol. Biol. **192:**557–565.
11. **Schachter, J.** 1988. The intracellular life of Chlamydia. Curr. Top. Microbiol. Immunol. **138:**109–139.
12. **Schachter, J., and H. D. Caldwell.** 1980. Chlamydiae. Annu. Rev. Microbiol. **34:**285–309.
13. **Sharp, P. A., A. J. Berk, and S. M. Berget.** 1980. Transcription maps of adenovirus. Methods Enzymol. **65:**750–768.
14. **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.
15. **Weinrich, S. L., E. G. Niles, and D. E. Hruby.** 1985. Transcriptional and translational analysis of the vaccinia virus late gene L65. J. Virol. **55:**450–457.
16. **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.
17. **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 two *Chlamydia psittaci* strains. Infect. Immun. **57:**1621–1625.