JOURNAL OF BACTERIOLOGY, Nov. 2000, p. 6239–6242 Vol. 182, No. 21
0021-9193/00/$04.00+0

Identification and Mapping of Sigma-54 Promoters in *Chlamydia trachomatis*

SARAH A. MATHEWS* AND PETER TIMMS

Centre for Molecular Biotechnology, School of Life Sciences, Queensland University of Technology, Brisbane, Queensland, Australia 4001

Received 15 May 2000/Accepted 3 August 2000

The first σ^{54} promoters in *Chlamydia trachomatis* L2 were mapped upstream of hypothetical proteins CT652.1 and CT683. Comparative genomics indicated that these σ^{54} promoters and potential upstream activation binding sites are conserved in orthologous *C. trachomatis* D, *C. trachomatis* mouse pneumonitis strain, and *Chlamydia pneumoniae* (CWL029 and AR39) genes.

Chlamydia is an organism of major medical and veterinary significance; however, its obligate intracellular existence makes genetic investigations a challenge. The unique developmental cycle of *Chlamydia* involves the interconversion between the infectious elementary body and the metabolically active reticulate body (23). Although the key morphological stages of chlamydial development are understood (19, 23) and the developmental expression of over 20 genes has been determined (12, 18), the elements which regulate this developmental gene expression are yet to be elucidated. Our recent investigations (18) identified temporal expression of the three *C. trachomatis* RNA polymerase sigma factors, σ^{66} (major σ^{70} homolog), σ^{28}, and σ^{54}. Several reports have characterized σ^{66} promoters (8, 17, 29), but to date, no promoters have been identified for the developmental-stage-specific sigma factors, σ^{28} and σ^{54}. The present study utilized the complete *Chlamydia trachomatis* genome sequence (27) and a modified fluorescence-based primer extension (PE) assay to identify and map the first σ^{54} promoters for *Chlamydia*.

Transcription initiation from σ^{54} promoters is a multistep process involving the recognition of the promoter by σ^{54}, binding of the core RNA polymerase to the σ^{54} to form a closed complex, and subsequent activation to an open complex following binding by an enhancer binding protein (EBP) (20, 21). In most cases the EBP binds an upstream activator sequence (UAS) located within 200 bp of the promoter (15, 21) and is brought into contact with the σ^{54}-RNA polymerase complex by DNA looping, an event mediated by the integration host factor (IHF) or intrinsic DNA bends (9). In addition to the σ^{54} gene (*rpoN*), recently identified in the *C. trachomatis* genome, genes for the NtrC family EBP (*ntrC*) and IHF (*ihfA*) were also found to be present (27). More detailed analysis of the translated amino acid sequences identified that RpoN (σ^{54}) contains a perfect RpoN box (ARRTVAKYR), which is responsible for recognition of the cognate promoter (30), and the chlamydial NtrC homolog has an exact match to the σ^{54}-binding domain, GAFTGA (7). Furthermore, the chlamydial NtrC has six of seven conserved amino acids of the UAS binding domain, GESGCGK (7) (the underlined amino acid is nonconserved). We previously reported the late-stage-specific expression of *rpoN* (18) and subsequently confirmed that *ntrC* was transcribed by reverse transcription-PCR analysis (data not shown). These observations led us to hypothesize that some chlamydial genes would be regulated by NtrC-activated σ^{54}-mediated transcription initiation. Of the cognate promoters for all eubacterial sigma factors, σ^{54} promoters are the most highly conserved (2) and hence lend themselves to a computational search of the full chlamydial genome. We therefore used the Findpatterns database searching program of the Australian National Genome Information Service to search the *C. trachomatis* D genome for sequences corresponding to the σ^{54} consensus promoter (TGGCAC-N_5-TTGC), allowing up to two mismatches. We identified 427 potential matches and analyzed these for both orientation and proximity to *C. trachomatis* open reading frames (ORFs). Only nine putative σ^{54} promoters were identified within 400 bp upstream of the *C. trachomatis* ORFs, encoding *ychF*, *yebL*, *pkn5*, *cysQ*, and *secA* product homologs and hypothetical proteins CT620, CT652.1, CT683, and CT734. Transcription within each ORF was confirmed by reverse transcription-PCR analysis on *C. trachomatis* RNA (data not shown). While this study was being undertaken Studholme and Buck (28) reported the identification of a σ^{54} promoter upstream of *C. trachomatis* AAC68830 and *Chlamydia pneumoniae* AAD18864 which corresponds to our mapped promoter for CT652.1. Our search strategy identified more candidate promoters and mapped two σ^{54} promoters for *Chlamydia*.

PE analysis to determine the transcription start site (TSS) of chlamydial transcripts is difficult for genes expressed at low levels. Since fluorescence detection has greater sensitivity than conventional radioactive methods, we modified recent methodology (1, 24) to perform PE analysis on total RNA isolated from *Chlamydia*-infected cells. *C. trachomatis* L2/434/Bu was propagated in 10^9 HEp2 cells for 30 h before RNA was extracted as previously described (18). Twenty micrograms of total RNA was annealed to 10 pmol of fluorescently end-labeled primer (either 5′6-FAM or 5′TET, synthesized by Pacific Oligos, Lismore, Australia) in a 10-μl volume for 10 min at 75°C in a thermal cycler. The reagents for cDNA synthesis (10 mM dithiothreitol, 1 mM each deoxynucleoside triphosphate, 20 U of RNase inhibitor, and 40 U of Expand reverse transcriptase in buffer [supplied by Rosche]) were added on ice, and the reaction mixture was incubated for 90 min at 42°C. The cDNA was precipitated in 0.3 M sodium acetate with 2.5 volumes of ethanol following RNase incubation using 25 ng of DNase-free RNase (Rosche) for 30 min at 37°C. The PE products were resuspended in 8 μl of 95% (vol/vol) formamide–10 mM EDTA (pH 9.0) and run on a 6 M urea–4.5% polyacrylamide TBE (45 mM Tris-borate, 1 mM EDTA, pH 8.0) gel

* Corresponding author. Mailing address: Centre for Molecular Biotechnology, School of Life Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Queensland, Australia 4001. Phone: 617-3864 5216. Fax: 617-3864 1534. E-mail: s.mathews@qut.edu.au.

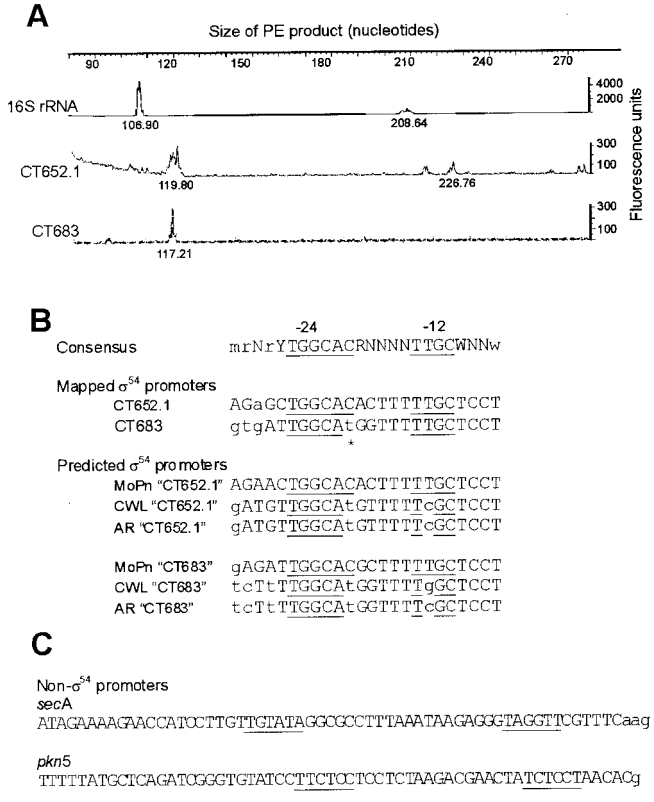

FIG. 1. Analysis of chlamydial promoters. (A) PE analysis of *C. trachomatis* L2 RNA with fluorescent primers against the 16S rRNA (6-FAM labeled), CT652.1 (6-FAM labeled) and CT683 (TET labeled) using GeneScan software. Primer sequences are given in the text, and the sizes of PE products greater than 100 fluorescence units are shown under the traces. Similar traces were obtained in a repeat PE experiment. (B) Consensus σ^{54} promoter (defined in reference 2) in alignment with the mapped *C. trachomatis* L2 σ^{54} promoters for CT652.1 and CT683. Uppercase and lowercase nucleotides in Consensus represent greater than 80% and 60 to 80% conservation to the consensus in σ^{54} mapped promoters, respectively. Uppercase nucleotides in the chlamydial sequences (CT, MoPn, CWL029, and AR39) correspond to homology with the consensus, and lowercase nucleotides represent lack of homology. Underlined nucleotides represent greater than 95% homology to the consensus of all σ^{54} promoters where the TSS has been mapped (2). The CT683 t highlighted with an asterisk is replaced by the consensus C in *C. trachomatis* serovar D. CT, *C. trachomatis* L2; MoPn, *C. trachomatis* MoPn; CWL and AR, *C. pneumoniae* CWL029 and AR39, respectively. (C) Non-σ^{54} promoters (underlined) were predicted on the basis of homology to the σ^{70} consensus (TTGACA-[N_{15-20}]-TATAAT) for *C. trachomatis* L2 *pkn5* and *secA* where TSSs (lowercase) failed to map to the σ^{54} consensus promoter identified in this study.

using an ABI 337 instrument, with GeneScan (Applied Biosystems) analysis using GS 500Rox molecular size markers (PE Biosystems) to determine the size of the 5′6-FAM- or 5′TET-labeled single-stranded DNA.

The 16S rRNA gene was used to establish the PE assay on *C. trachomatis* serovar L2 RNA using primer 16S.PE (5′6-FAM-GAACCAAGATCAAATTCTCAG). We identified two transcripts as seen by fluorescent peaks for PE products at 107 and 209, respectively (Fig. 1A). These correspond to the previously determined TSS for *C. trachomatis* mouse pneumonitis (MoPn) strain, where two promoters were defined (10). PE analysis for the *C. trachomatis pkn5*, *secA*, *cysQ*, *ychF*, CT652.1, and CT863 genes was undertaken using primers pkn5.PE (5′6-FAM-CGAGAAGAGTGCTCATCCACACC), secA.PE (5′6-FAM-TATTCTCTCTTGGGAGGATCCG), cysQ.PE (5′TET-GCATCAGTGACAGCATAGCCTGC), ychF.PE (5′6-FAM-CTACTATTCCACACTCTGTTTGTC), CT652/1.PE (5′6-FAM-CATGGATGTACGCTCTTTCCGAC), and CT683.PE (5′TET-CTGCTTGCTCGTACTCACCACTC), respectively. Definite fluorescent peaks were generated for the *pkn5*, *secA*, CT652.1, and CT683 genes, whereas only nonspecific fluorescent peaks were obtained for *ychF* and *cysQ* PE reactions. The CT652.1 and CT683 PE products were 120 and 117 nucleotides (Fig. 1A), respectively, which map the TSS in the correct position to the predicted σ^{54} promoter and Shine-Dalgarno sequences (Fig. 2). The CT652.1 PE reaction also generated some lower-intensity peaks (with one peak of over 100 fluorescence units at 227 nucleotides) which could be the result of nonspecific priming, since no obvious promoter sequences were identified in the corresponding nontranscribed sequence of CT652.1 (data not shown). The TSSs mapped against the noncoding sequences upstream of *pkn5* and *secA* allowed alternative promoters to be proposed based on homology to the σ^{70}-like consensus promoters (Fig. 1C). The inability to map TSSs for *ychF* and *cysQ* maybe due to either insufficient transcript, failure to induce transcription from these genes under the growth conditions used, or the genes being part of an operon and the promoter not being within 400 bp of the ORF.

Our genomic σ^{54} promoter searching analysis was done against the *C. trachomatis* D genome; however, the TSSs were experimentally determined for *C. trachomatis* L2, and thus the equivalent CT652.1 and CT683 sequences were retrieved from the *C. trachomatis* L2 genome (http://violet.berkeley.edu:4231). When the predicted σ^{54} promoter sequences are compared to the extended consensus proposed by Barrios and colleagues (2), the CT652.1 promoter exactly matches the "uppercase" consensus, and the CT683 promoter is a perfect match with serovar D and has one mismatch with serovar L2 (Fig. 1B). The genome sequences for *C. trachomatis* MoPn (25), *C. pneumoniae* CWL029 (14), and *C. pneumoniae* AR39 (25) were searched for CT652.1 and CT683 homologs, in order to confirm if the chlamydial σ^{54} promoters are conserved between different chlamydial strains and species. Homologs to both CT652.1 and CT683 were found in *C. trachomatis* MoPn (TC0022 and TC0055, respectively), *C. pneumoniae* CWL029 (Cpn0725 and Cpn0693, respectively) and *C. pneumoniae* AR39 (CP0021 and CP0053, respectively). The overall conservation of the predicted σ^{54} promoters (Fig. 1B) indicates that the σ^{54} promoters may be utilized for the equivalent genes in *C. trachomatis* MoPn and *C. pneumoniae* CWL029 and AR39.

Having mapped these first two σ^{54} promoters in *Chlamydia*, we analyzed the sequences upstream of the TSSs and the equivalent sequences in *C. trachomatis* (D and MoPn) and *C. pneumoniae* (CWL029 and AR39) for the presence of NtrC binding sites and other, perhaps *Chlamydia*-specific, elements. Although published NtrC binding sites show limited consensus across eubacteria and often appear in unmatched pairs (the NtrC binds as a dimer), we searched the upstream sequences for patterns resembling the NtrC UAS (5, 6, 16, 26, 31) and found two potential sites in the upstream sequences of both CT652.1 and CT683 (NtrCI and NtrCII in Fig. 2). The sequences had an overall consensus between *C. trachomatis* L2, D, and MoPn, and alternative NtrC binding sites could be identified in the equivalent *C. pneumoniae* upstream sequences.

Our search for common elements between the upstream sequences of both *C. trachomatis* L2 CT652.1 and CT683 revealed three sequence patterns, GAGAA, (A/G)AAAA, and TAAT, located in similar positions within 100 bp upstream of the TSS (Fig. 2). When we searched for intra- and interspecies conservation, we found that (A/G)AAAA is completely conserved and GAGAA was replaced by alternative purine-rich sequences in *C. trachomatis* (D and MoPn) and *C. pneumoniae*

FIG. 2. Analysis of the upstream sequences of *C. trachomatis* CT652.1 and CT683. The upstream sequences from −317 to the start codon (ATG) (boldface) of *C. trachomatis* L2 CT652.1 and CT683 are aligned with respect to the TSS (asterisk). The numbers represent nucleotides upstream of the TSS. The patterns below and above the sequence are represented as follows: underline, predicted Shine-Dalgarno sequence; box, predicted σ^{54} promoter; shaded underbox, predicted NtrC binding sites (NtrCI and NtrCII); overline, predicted IHF binding sites conserved in *C. trachomatis* L2, D, and MoPn; double underline, conservation of purine rich sequence in similar positions between CT652.1 and CT683 in *C. trachomatis* (L2, D, and MoPn) and *C. pneumoniae* (CWL029 and AR39); unidirectional arrows, direct sequence repeats conserved in *C. trachomatis* L2, D, and MoPn; bidirectional arrows, putative NifA binding site conserved in *C. trachomatis* L2, D, and MoPn (TGT-N_{10}-ATA in MoPn). Variations from the *C. trachomatis* L2 NtrCI and NtrCII for *C. trachomatis* D and MoPn are GCAACAATTGCGTTTC (D, CT652.1 NtrCI), CGAACA GCTGCATTTC (MoPn, CT652.1 NtrCI), GCGTAAGGAGATTTC (MoPn, CT652.1 NtrCII), GCTCCTGAACAACTGT (MoPn, CT683 NtrCI), and GTACCGC TGATGTAAT (MoPn, CT683 NtrCI), where the underlined nucleotides represent changes to the sequence (shaded underbox).

(CWL029 and AR39) sequences. The TAAT element is conserved in the same position relative to the predicted σ^{54} promoters across all CT652.1 and CT683 upstream sequences investigated, except that *C. pneumoniae* CT683 has TAGT. Two extra putative regulatory elements for CT652.1 were identified by comparison of different chlamydial strains and species (Fig. 2). First, the UAS, TGT-N10/11-ACA (22), for the NtrC-like EBP, NifA (15), was conserved in the *C. trachomatis* and *C. pneumoniae* sequences. Second, a perfect direct repeat, GC(A/T)AT separated by 9 bp is conserved between the *C. trachomatis* L2, D, and MoPn sequences. The strong conservation of sequence patterns in noncoding regions of the genome, both between genes and between species, supports a role in regulation.

Since IHF binding sites are often found between the UAS and σ^{54} promoter, we examined the sequence between the σ^{54} promoter and putative UAS for IHF binding sites by searching for the consensus WATCAA-N_4-WTR (13, 32). We identified two putative IHF binding sites in the upstream sequences of CT652.1 and CT683 (Fig. 2). Without a transformation system for *Chlamydia*, it is difficult to obtain functional data to support the involvement of the above-mentioned sequence patterns in the two-component σ^{54} regulatory system.

These first two chlamydial σ^{54} promoters have been mapped upstream of hypothetical proteins. The ORF for CT652.1 shows no homology to any other known proteins by BLAST analysis, and the CT683 ORF contains the ubiquitous tetratricopeptide repeat module involved in protein-protein interaction (3). The chlamydial CT683 has been classified as both an O-linked GlcNAc transferase (27) and a type III secretion chaperone (25), proteins which play a direct role in signal transduction (4, 11). Since CT652.1 and CT683 are probably regulated by σ^{54}-mediated transcription initiation, we propose that they might be required for reticulate body-to-elementary body conversion, since our earlier investigations identified *rpoN* expression during mid- to late-stage-specific chlamydial development (18). Further analysis to determine the temporal expression and function of CT652.1 and CT683 will be required to elucidate their role in chlamydial development and pathogenesis. It is hard to imagine that *Chlamydia* has *rpoN* and *ntrC* genes for regulating only two genes; thus, different search strategies are required to identify other σ^{54} promoters in *Chlamydia*. Since σ^{66}-regulated promoters often show limited homology to the major σ factor consensus (8, 12, 17, 29), it is possible that a unique class of σ^{54} promoters exist in *Chlamydia*.

This work was supported by National Health and Medical Research Council grant 981383, and S.A.M. is the recipient of an Australian Research Council postdoctoral fellowship.

We are grateful to Stephen Myers for technical assistance and to Kelly Ewen (Australian Genome Research Facility, Melbourne, Australia) for the GeneScan analysis.

REFERENCES

1. **Altermann, E., R. Jurgen, K. Klein, and B. Henrich.** 1999. Synthesis and automated detection of fluorescently labeled primer extension products. BioTechniques **26:**96–101.
2. **Barrios, H., B. Valderrama, and E. Morett.** 1999. Compilation and analysis of σ^{54}-dependent promoter sequences. Nucleic Acids Res. **27:**4305–4313.
3. **Blatch, G. L., and M. Lassle.** 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays **21:**932–939.
4. **Brummel, J. H., O. Steele-Mortimer, and B. B. Finlay.** 1999. Bacterial invasion: force feeding by Salmonella. Curr. Biol. **9:**R277–R280.
5. **Cheema, A. K., N. R. Choudhury, and H. K. Das.** 1999. A- and T-tract-mediated intrinisic curvature in native DNA between the binding site of the upstream activator NtrC and the nifLA promoter of *Klebsiella pneumoniae* facilitates transcription. J. Bacteriol. **181:**5296–5302.
6. **Cullen, P. J., W. C. Bowman, D.-F. Hartnett, S. C. Reilly, and R. G. Kranz.** 1998. Translational activation by an NtrC enhancer-binding protein. J. Mol. Biol. **278:**903–914.
7. **Dischert, W., P. M. Vignais, and A. Colbeau.** 1999. The synthesis of *Rhodobacter capsulatus* HupSL hydrogenase is regulated by the two-component HupT/HupR system. Mol. Microbiol. **34:**995–1006.
8. **Douglas, A. L., and T. P. Hatch.** 1996. Mutagenesis of the P2 promoter of the major outer membrane protein gene of *Chlamydia trachomatis*. J. Bacteriol. **178:**5573–5578.
9. **Dworkin, J., G. Jovanovic, and P. Model.** 1997. Role of upstream activation sequences and integration host factor in transcriptional activation by the constitutively active prokaryotic enhancer-binding protein PspF. J. Mol. Biol. **273:**377–388.
10. **Engel, J. N., and D. Ganem.** 1987. Chlamydial rRNA operons: gene organization and identification of putative tandem promoters. J. Bacteriol. **169:** 5678–5685.
11. **Hart, G. W.** 1997. Dynamic O-linked glycosylation of nuclear cytoskeletal

proteins. Annu. Rev. Biochem. **66:**315–335.
12. **Hatch, T. P.** 1999. Developmental biology, p. 29–67. *In* R. S. Stephens (ed.), Chlamydia: intracellular biology, pathogenesis, and immunity. ASM Press, Washington, D.C.
13. **Hoover, T. R., E. Santero, S. Porter, and S. Kustu.** 1990. The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. Cell **63:**11–22.
14. **Kalman, S., W. Mitchell, R. Marathe, C. Lammel, J. Fan, R. W. Hyman, L. Olinger, J. Grimwood, R. W. Davis, and R. S. Stephens.** 1999. Comparative genomes of *C. pneumoniae* and *C. trachomatis*. Nat. Genet. **21:**385–389.
15. **Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of sigma 54 (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53:**367–376.
16. **Lu, C.-D., and A. T. Abdelal.** 1999. Role of ArgR in activation of the *ast* operon, encoding enzymes of the arginine succinyltransferase pathway in *Salmonella typhimurium*. J. Bacteriol. **181:**1934–1938.
17. **Mathews, S. A., and K. S. Sriprakash.** 1994. The RNA polymerase of *Chlamydia trachomatis* has a flexible sequence requirement at the −10 and −35 boxes of its promoters. J. Bacteriol. **176:**3765–3789.
18. **Mathews, S. A., K. M. Volp, and P. Timms.** 1999. Development of a quantitative gene expression assay for *Chlamydia trachomatis* identified temporal expression of σ factors. FEBS Lett. **458:**354–358.
19. **McClarty, G.** 1994. Chlamydiae and the biochemistry of intracellular parasitism. Trends Microbiol. **2:**157–164.
20. **Merrick, M. J.** 1993. In a class of its own—the RNA polymerase sigma factor σ^{54} (σ^{N}). Mol. Microbiol. **10:**903–909.
21. **Morett, E., and L. Segovia.** 1993. The sigma-54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. J. Bacteriol. **175:**6067–6074.
22. **Morett, E., and M. Buck.** 1989. In vivo studies on the interaction of the RNA polymerase-σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti* nifH promoters. J. Mol. Biol. **210:**65–77.
23. **Moulder, J. W.** 1991. Interaction of chlamydiae and host cells in vivo. Microbiol. Rev. **55:**143–190.
24. **Peltonen, T., and P. Mantsala.** 1999. Isolation and characterization of a purC(orf)QLF operon from *Lactobacillus lactis* MG1614. Mol. Gen. Genet. **261:**31–41.
25. **Read, T. D., R. C. Brunham, C. Shen, S. R. Gill, J. F. Heidelberg, O. White, E. K. Hickey, J. Peterson, T. Utterback, K. Berry, S. Bass, K. Linher, J. Weidman, H. Khouri, B. Craven, C. Bowman, R. Dodson, M. Gwinn, W. Nelson, R. DeBoy, J. Kolonay, G. McClarty, S. L. Salzberg, J. Eisen, and C. M. Fraser.** 2000. Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Res. **28:**1397–406.
26. **Reitzer, L. J., and B. Magasanik.** 1986. Transcription of glnA in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45:**785–792.
27. **Stephens, R. S., S. Kalman, C. Lammel, J. Fan, R. Marathe, L. Aravind, W. Mitchell, L. Olinger, R. L. Tatusov, Q. Zhao, E. V. Koonin, and R. W. Davis.** 1998. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science **282:**754–759.
28. **Studholme, D. J., and M. Buck.** 2000. New roles of σ^{N} in small genomes. Microbiology **146:**4–5.
29. **Tan, M., T. Gaal, R. L. Gourse, and J. N. Engel.** 1998. Mutational analysis of the *Chlamydia trachomatis* rRNA P1 promoter defines four regions important for transcription in vitro. J. Bacteriol. **180:**2359–2366.
30. **Taylor, M., R. Butler, S. Chambers, M. Casimiro, F. Badii, and M. Merrick.** 1996. The RpoN-box motif of the RNA polymerase sigma factor σ^{N} plays a role in promoter recognition. Mol. Microbiol. **22:**1045–1054.
31. **Wassem, R., E. M. de Souza, M. G. Yates, F. de O. Pedrosa, and M. Buck.** 2000. Two roles for integration host factor at an enhancer dependent nifA promoter. Mol. Microbiol. **35:**756–764.
32. **Weiner, L., J. L. Brissette, N. Ramani, and P. Model.** 1995. Analysis of the proteins and cis-acting elements regulating the stress-induced phage shock protein operon. Nucleic Acids Res. **23:**2030–2036.