JOURNAL OF BACTERIOLOGY, Jan. 2000, p. 518–521 Vol. 182, No. 2
0021-9193/00/$04.00+0

opdA, a *Salmonella enterica* Serovar Typhimurium Gene Encoding a Protease, Is Part of an Operon Regulated by Heat Shock

CHRISTOPHER A. CONLIN† AND CHARLES G. MILLER*

Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Received 14 June 1999/Accepted 21 October 1999

The *opdA* (*prlC*) gene of *Salmonella enterica* serovar Typhimurium and *Escherichia coli* encodes the metalloprotease oligopeptidase A (OpdA). We report that *opdA* is cotranscribed with a downstream open reading frame, *yhiQ*. Transcription of this operon is induced after a temperature shift (30 to 42°C), and this induction depends on the heat shock sigma factor encoded by the *rpoH* (*htpR*) gene.

The *Salmonella enterica* serovar Typhimurium *opdA* gene encodes the metalloprotease oligopeptidase A (OpdA) (17). *opdA* is a homolog of the *Escherichia coli prlC* gene, a site of suppressors of the localization defect conferred by certain signal sequence mutations (2, 4, 14, 15). OpdA is also required for the proteolytic processing of a phage P22 protein in vivo (5), and it can degrade the cleaved *lpp* signal peptide in vitro (10). Comparison of the OpdA amino acid sequence with other protein sequences indicated that it is a member of a subfamily of Zn metalloproteases with representatives in both animals and fungi (3). This family includes the mitochondrial intermediate peptidase, which removes a small peptide from certain proteins that are transported through the mitochondrial membranes (9). In this study, the *opdA* sequence has been extended to include a downstream open reading frame (ORF) of unknown function, *yhiQ*. We show that *opdA* and *yhiQ* form an operon, that transcription of this operon is induced by temperature upshift, and that this induction is dependent on RpoH (σ^{32} or σ^{H}).

Nucleotide sequence of serovar Typhimurium *yhiQ*. The sequence of the *opdA* gene (2) suggested the presence of another ORF beginning downstream of the *opdA* ORF. To characterize this ORF, we determined the nucleotide sequence of an additional 997 bp beyond the end of *opdA* (Fig. 1). This sequence (GenBank accession no. AF137028) contains an ORF encoding a protein of either 253 amino acids (if the GTG at nucleotides [nt] 7 to 9 is used as the translational start) or 221 amino acids (if ATG at nt 103 to 105 is used as the translational start). Given the strong conservation of amino acid sequence between the YhiQ homologs in the region of the ORF between the GTG start codon and the first AUG codon, it seems likely that the protein is translated starting with GTG. There is a possible transcription terminator at nt 34 to 62. Downstream from the *yhiQ* ORF, there are two repetitive extragenic palindromic (REP) sequences (13) in inverted orientation, which have the potential to form a large stem-loop. A potential ribosome binding site precedes both of the possible translation starts, and the codon usage of the ORF is consistent with other expressed *Salmonella* genes (18). In preliminary experiments, a ~28-kDa protein was produced when *yhiQ* was transcribed in vivo from a phage T7 promoter, indicating that *yhiQ* may be translated in vivo.

Comparison of the YhiQ sequence with the sequence databases identified three similar hypothetical proteins of unknown function from *E. coli* (94% identity), *Haemophilus influenzae* (68% identity), and *Neisseria gonorrhoeae* (41% identity). The *E. coli yhiQ* gene is located immediately downstream from *prlC*, the *E. coli* homolog of *opdA*. There are no REPs in the *E. coli* sequence, but there is a potential stem-loop structure (13-bp stem, 4-nt loop) after the *E. coli yhiQ* ORF (12). In *H. influenzae*, the putative homologs of *opdA* and *yhiQ* are unlinked. Clearly, *opdA* and *yhiQ* are not always associated since the genome of *Synechocystis* sp. contains an *opdA* homolog (accession no. D90916) but does not encode a YhiQ-related ORF.

***opdA* and *yhiQ* constitute an operon.** Because there was no obvious promoter preceding *yhiQ*, Northern blot analysis was carried out on RNA isolated from TN4465 (*leuBCD485 opdA10*::Mu*dI1734* containing plasmid pCM138, which carries both *opdA* and *yhiQ*) to determine if *yhiQ* and *opdA* were cotranscribed. As shown in Fig. 2, the *opdA*-specific probe hybridized to two transcripts; one of 2.2 kb and the other of 3 kb. However, the *yhiQ*-specific probe hybridized only to the 3-kb transcript. The sizes of these transcripts are consistent with the lengths of the *opdA* ORF, 2,078 bp, and the combined *opdA* and *yhiQ* ORFs, 2,858 bp. These results show that *opdA* is the first gene in a two-gene operon. The two transcripts may result either from occasional read-through of the potential transcription terminator between *opdA* and *yhiQ* or from the degradation of the longer transcript to yield the shorter one.

The *opdA* operon is a σ^{32}-dependent heat shock operon. A near-consensus σ^{32}-dependent promoter sequence is present upstream from the start of translation of *opdA* (Fig. 3) (2, 4). To determine if this sequence could be the *opdA* promoter, the start of transcription was determined by primer extension analysis. The initial nucleotide in the mRNA was found to be an A, 55 bp 5′ to the start of OpdA translation (Fig. 4). This result is consistent with the identification of the σ^{32} promoter sequence as the functional *opdA* promoter.

To determine if expression of the *opdA* operon was depen-

* Corresponding author. Mailing address: Department of Microbiology, University of Illinois at Urbana-Champaign, B103 CLSL, 601 S. Goodwin, Urbana, IL 61801. Phone: (217) 244-8418. Fax: (217) 244-6697. E-mail: charlesm@uiuc.edu.

† Present address: Department of Biological Sciences, Mankato State University, Mankato, MN 56002-8400.

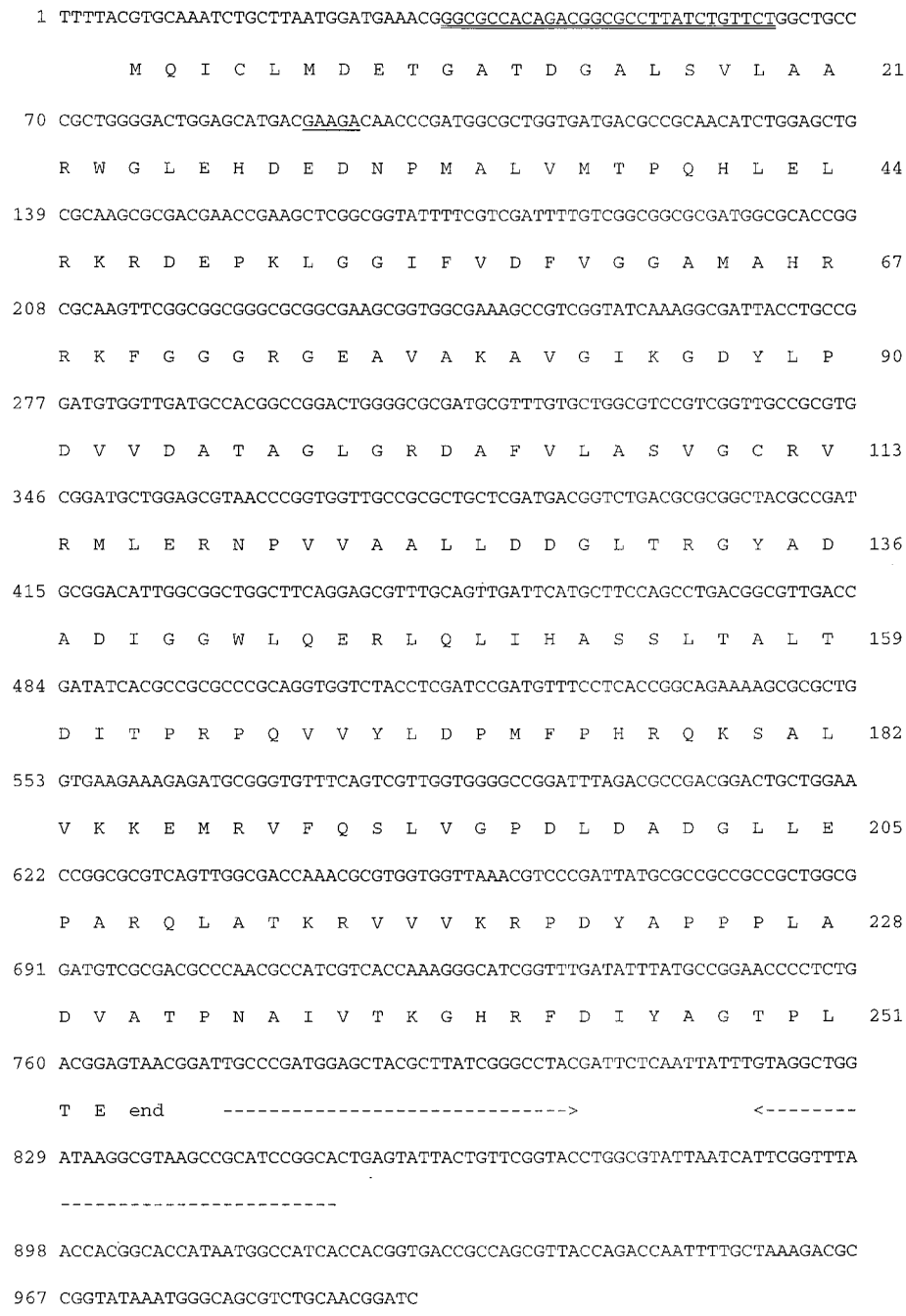

FIG. 1. DNA sequence and translation of serovar Typhimurium *yhiQ*. The sequence shown begins immediately after the end of the *opdA* ORF at nt 2645 in GenBank sequence accession no. M84574. The double underline indicates a potential transcriptional terminator: 6-bp stem, 6-nt loop, and seven T's in the next 11 nt. Either GTG (nt 7 to 9) or ATG (nt 103 to 105) could serve as the translation start for YhiQ. A potential ribosome binding site for the ATG start site is indicated by the single underline. A potential ribosome binding site for the GTG start is present in the *opdA* coding region not shown in Fig. 1. Dashed lines identify two REP sequences in inverted orientation.

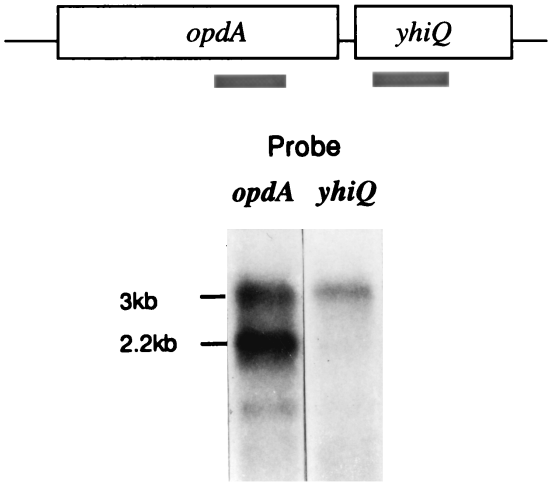

FIG. 2. Northern blot analysis of *opdA* RNA. Total RNA was isolated from TN4465 (1), treated with RNase-free DNase, fractionated by agarose-formaldehyde gel electrophoresis, and vacuum blotted to a nylon membrane (11). Half the membrane was hybridized to a radiolabeled 1.36-kb *opdA*-specific DNA fragment, and the other half was hybridized to a 0.55-kb *yhiQ*-specific fragment. These DNA probes were generated by PCR amplification and labeled with [α-^{32}P]dATP. Hybridization was conducted at 60°C in 20 mM sodium phosphate–5× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate)–7% sodium dodecyl sulfate–2× Denhardt's solution–100 μg of denatured salmon sperm DNA per ml for 18 h (11). The membrane was washed extensively, concluding with a final wash in 0.1× SSC–1% sodium dodecyl sulfate at 60°C. Bands were visualized by autoradiography. The positions of the 3-kb message containing both *opdA* and *yhiQ* and of a 2.2-kb message containing only *opdA* are indicated. A third, weaker band may represent a partially degraded fragment of the *opdA* message.

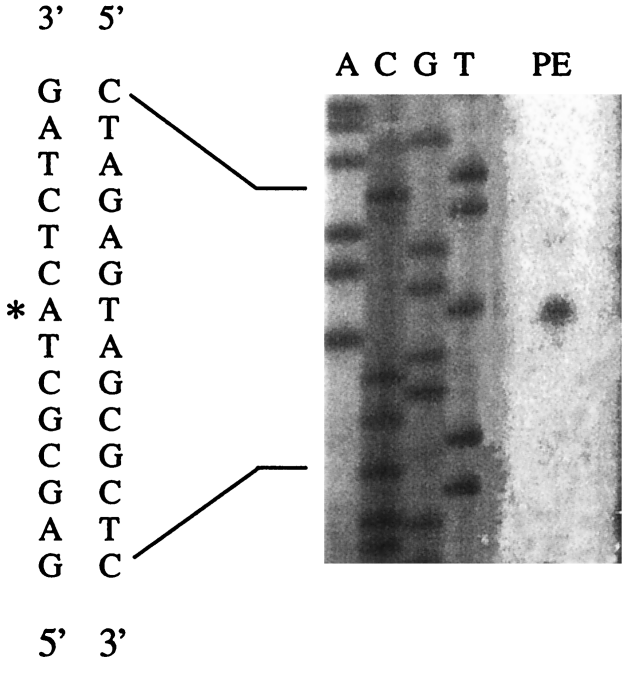

FIG. 4. Identification of the transcriptional start site for OpdA. Strain TN4465 containing pCM138 was grown overnight at 30°C, diluted 1:100, and incubated at 30°C to an optical density at 600 nm of 0.4. The culture was then shifted to 42°C for 10 min, and total RNA was isolated by using Qiagen's RNeasy Mini kit. A 25-nt primer complementary to nt 603 to 627 of GenBank sequence accession no. M84574 was used for extension, and the product is shown in lane PE. Lanes A, C, G, and T contain sequence reactions generated with the same primer. Radioactivity was detected by autoradiography and, in addition, by the use of a Molecular Dynamics Storm 860 PhosphorImager. An asterisk indicates the transcription start in the DNA sequence shown on the left.

dent on σ^{32}, Northern blot analysis was done with RNA isolated from strains carrying an *opdA* plasmid and lacking functional RpoH. Plasmid pCM138, containing the entire *opdA* operon, was transformed into *E. coli* strains with three different *rpoH* alleles: *rpoH* wild type (SC122), *rpoH156* (K165) (6), and Δ*rpoH30*::Kan (CAG9333) (19). Cells were grown at 30°C and shifted to 42°C, and total RNA was isolated from samples taken at various times after the shift. The RNA was fractionated by agarose-formaldehyde gel electrophoresis and transferred to a nylon membrane as described above. The membrane was probed with a radiolabeled *opdA* probe. As shown in Fig. 5, *opdA* transcription increases after a temperature upshift, and this transient induction requires the *rpoH* gene product, σ^{32}. Thus, *opdA* and *yhiQ* are indeed previously unidentified heat shock genes.

The role of OpdA and YhiQ in the cell's response to stress is unknown. Strain TN3101, which contains a polar *opdA*::Mu*dI1734* insertion, showed no defect in growth rate at 30, 37, or 42°C. There is also no obvious correlation between the predicted molecular weight of either OpdA or YhiQ and any of the identified heat shock-induced proteins in the *E. coli* protein index of VanBogelen et al. (16). YhiQ appears to be unnecessary for OpdA function, since plasmid pCM258, which carries *opdA* but not *yhiQ*, was able to complement all known defects conferred by an insertion mutation in *opdA*.

Several other protease genes in *E. coli* are known to be part of the heat shock regulon, and it seems likely that the degradation of irremediably misfolded proteins is a major component of the heat shock response (7). Only one natural macromolecular substrate for OpdA has been identified (phage P22 gp7), and this protein is not degraded but rather is specifically processed by OpdA. It is not clear whether OpdA participates directly in the degradation of misfolded proteins or whether it plays a more specialized role in the heat shock response. Perhaps the identification of cellular substrates for OpdA will clarify its functional role. The function of YhiQ remains completely mysterious.

	-35		-10	+1
σ^{32} promoter sequence:	CTTGAAA	12-16bp	CCCCAT-T	
opdA sequence:	ATTGAAATTCCCCTTACTATCCCCCATTCTAGATCTCAT			

FIG. 3. Comparison of *opdA* promoter region and consensus σ^{32} promoter. The consensus sequence is from reference 8. The double underline indicates the 5′ end of the *opdA* mRNA as determined by primer extension analysis and corresponds to nt 548 in GenBank sequence accession no. M84574.

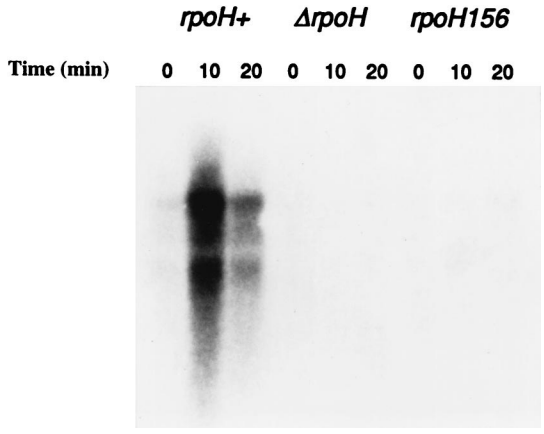

FIG. 5. RpoH-dependent heat shock induction of *opdA* transcription. *E. coli* strains containing pCM138 and one of three different *rpoH* alleles (*rpoH* wild type, *rpoH156*, or Δ*rpoH30*::Kan) were grown at 30°C and shifted to 42°C, and samples were taken at the indicated times after the shift. Total RNA was isolated, separated by agarose-formaldehyde gel electrophoresis, and transferred to a nylon membrane. The membrane was probed with a radiolabeled *opdA* probe.

We thank Tina Knox for performing the primer extension experiment and John Cronan and Carol Gross for providing strains.

This work was supported by a grant (AI10333) from the National Institute for Allergy and Infectious Diseases.

REFERENCES

1. **Aiba, H., S. Adhya, and B. de Crombrugghe.** 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:**11905–11910.
2. **Conlin, C. A., and C. G. Miller.** 1992. Cloning and nucleotide sequence of *opdA*, the gene encoding oligopeptidase A in *Salmonella typhimurium*. J. Bacteriol. **174:**1631–1640.
3. **Conlin, C. A., and C. G. Miller.** 1995. Dipeptidyl carboxypeptidase and oligopeptidase A from *Escherichia coli* and *Salmonella typhimurium*. Methods Enzymol. **248:**567–579.
4. **Conlin, C. A., N. J. Trun, T. J. Silhavy, and C. G. Miller.** 1992. *Escherichia coli prlC* encodes an endopeptidase and is homologous to the *Salmonella typhimurium opdA* gene. J. Bacteriol. **174:**5881–5887.
5. **Conlin, C. A., E. R. Vimr, and C. G. Miller.** 1992. Oligopeptidase A is required for normal phage P22 development. J. Bacteriol. **174:**5869–5880.
6. **Cooper, S., and T. Ruettinger.** 1975. A temperature sensitive nonsense mutation affecting the synthesis of a major protein of *Escherichia coli* K12. Mol. Gen. Genet. **139:**167–176.
7. **Gottesman, S., S. Wickner, and M. R. Maurizi.** 1997. Protein quality control: triage by chaperones and proteases. Genes Dev. **11:**815–823.
8. **Gross, C. A.** 1996. Function and regulation of the heat shock proteins, p. 1382–1399. *In* F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd ed. ASM Press, Washington, D.C.
9. **Isaya, G., and F. Kalousek.** 1995. Mitochondrial intermediate peptidase. Methods Enzymol. **248:**556–567.
10. **Novak, P., and I. K. Dev.** 1988. Degradation of a signal peptide by protease IV and oligopeptidase A. J. Bacteriol. **170:**5067–5075.
11. **Sambrook, J., E. F. Fritsch, and T. T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. **Sofia, H. J., V. Burland, D. L. Daniels, G. R. Plunkett, and F. R. Blattner.** 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. Nucleic Acids Res. **22:**2576–2586.
13. **Stern, M. J., G. F. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37:**1015–1026.
14. **Trun, N. J., and T. J. Silhavy.** 1987. Characterization and in vivo cloning of *prlC*, a suppressor of signal sequence mutations in *Escherichia coli* K12. Genetics **116:**513–521.
15. **Trun, N. J., and T. J. Silhavy.** 1989. PrlC, a suppressor of signal sequence mutations in *Escherichia coli*, can direct the insertion of the signal sequence into the membrane. J. Mol. Biol. **205:**665–676.
16. **VanBogelen, R. A., P. Sankar, R. L. Clark, J. A. Bogan, and F. C. Neidhardt.** 1992. The gene-protein database of *Escherichia coli*: edition 5. Electrophoresis **13:**1014–1054.
17. **Vimr, E. R., L. Green, and C. G. Miller.** 1983. Oligopeptidase-deficient mutants of *Salmonella typhimurium*. J. Bacteriol. **153:**1259–1265.
18. **Wada, K., Y. Wada, H. Doi, F. Ishibashi, T. Gojobori, and T. Ikemura.** 1991. Codon usage tabulated from the GenBank genetic sequence data. Nucleic Acids Res. **19**(Suppl.)**:**1981–1986.
19. **Zhou, Y. N., N. Kusukawa, J. W. Erickson, C. A. Gross, and T. Yura.** 1988. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor σ^{32}. J. Bacteriol. **170:**3640–3649.