Nucleotide sequence of the *Rickettsia prowazekii* *greA* homolog

G.Lynn Marks and David O.Wood^{1*}

Departments of Medicine and ¹Microbiology and Immunology, Laboratory of Molecular Biology, University of South Alabama College of Medicine, Mobile, AL 36688, USA

Submitted May 26, 1992

EMBL accession no. Z12122

Rickettsia prowazekii is an obligate intracellular parasitic bacterium and the etiologic agent of epidemic (louse-borne) typhus. Rickettsial pathogenicity is based on the ability of these organisms to invade eukaryotic cells, multiply intracellularly unbounded by a phagosomal membrane, and escape by lysis from the cell in order to repeat the cycle. Characterization of the genes involved in rickettsial pathogenicity and their transcriptional regulation is ongoing. A focal point of the transcriptional studies is the characterization of the DNA dependent RNA polymerase of *R.prowazekii* (1, 2). RNA polymerase holoenzyme of *Escherichia coli* and *R.prowazekii* is composed of core subunits (α_2 , β , β') and a sigma factor which are produced by *rpoA*, *rpoB*, *rpoC*, and *rpoD* respectively (1, 3, 4). Mutations in *rpoB* of *E.coli* are classically associated with resistance to rifampicin. A single nucleotide transition mutation in *E.coli rpoB* (*rpoB*_{ts8}, SP209) results in pleiotropic effects (5) that include conditional lethality, competence in antitermination by the λ N gene product with a defective *nusA* allele, and proficiency in termination with a defective *rho* allele.

The *greA* and *greC* genes of *E.coli* are able to suppress the temperature conditional lethality imposed by the *rpoB*_{ts8} mutation (6) by an undefined mechanism. We have isolated and characterized the *R.prowazekii greA* homolog. DNA sequence analysis of a 6.4 kilobase *EcoRI* fragment of *R.prowazekii* chromosomal DNA contained in pBluescript SKII vector was performed by the dideoxynucleotide chain termination method of Sanger. Overlapping templates of both strands confirmed the presence of a 486 base pair open reading frame encoding a protein of 162 amino acids. The predicted molecular weight (18,230 daltons) compares closely to *E.coli* (17,640). The putative -10 and -35 promoter region of the *R.prowazekii greA* gene is essentially that of the *E.coli* consensus sequence. The -35 region (TTGACA) is identical to the *E.coli* consensus and 17 bp downstream (an optimal distance in *E.coli*) lies the -10 region (TATACT) which differs in one base from the *E.coli* consensus (TATAAT). However, an *E.coli*-type ribosomal binding site is not apparent which may explain the lack of expression (data not shown) in *E.coli*.

The predicted amino acid sequence encoded by the *R.prowazekii greA* gene was analyzed via a TFASTA (7) search of GenBank. The search revealed the *greA* gene of *E.coli* as the best match with an optimization score of 381. No other sequence within the data base produced an optimization score greater than 83. A GAP (7) comparison of the *R.prowazekii* and *E.coli* peptide sequences produced 50.3% amino acid identity and 66.9% amino

acid similarity. Conservation of amino acid residues was particularly evident from *R.prowazekii* position 34 to 64 (Figure 1) with 83% amino acid identity in this region. Despite this degree of amino acid conservation, the *R.prowazekii greA* gene was unable to suppress the *E.coli rpoB*_{ts8} temperature sensitive mutation (*E.coli* mutant and control were kindly provided by Dr Asis Das, Department of Microbiology, University of Connecticut Health Center, Farmington, CT). The lack of complementation may be due to an inability of the rickettsial protein to interact with the *E.coli* components, poor expression of the rickettsial gene in *E.coli*, or rapid degradation of the rickettsial gene product in the *E.coli* background. The significance of the highly conserved region and the role of the *E.coli* and *R.prowazekii* GreA as transcription regulation factors remain to be determined.

ACKNOWLEDGEMENTS

We are grateful to Betty Williams and Marie Solomon for expert technical assistance and to Dr Herbert Winkler for guidance and helpful discussions. Computer analyses were performed on the University of South Alabama VAX. This work was supported by Public Health Service grants AI00922 (Physician Scientist-GLM) and AI20384 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

1. Ding, H.-F. and Winkler, H.H. (1990) *J. Bacteriol.* **172**, 5624-5630.
2. Marks, G.L., Winkler, H.H. and Wood, D.O. (1992) *Gene* in press.
3. Burgess, R.R., Travers, A.A., Dunn, J.J. and Bautz, E.K.F. (1969) *Nature* **221**, 43-46.
4. Burgess, R.R. (1969) *J. Biol. Chem.* **244**, 6160-6167.
5. Sparkowski, J. and Das, A. (1992) *Genetics* **130**, 411-428.
6. Sparkowski, J. and Das, A. (1990) *Nucleic Acids Res.* **18**, 6443.
7. Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395.

```

1 MQAIPMTLRGAELREELDFLKSVRRPEIIAAIAEAAREHGDLENAEYH 49
  * * * * *
1 MNTKFPITAKGPKLEHEHLKHLKHVERKKISEDIAEAAREHGDLENAEYE 50
  * * * * *
50 AAREQQGFCEGRIDIEAKLSNAQVIDVTGPNHNGRVIFGATVTVLNLDS 99
  * * * * *
51 AAREKQAFIEGRILEDMTARAEIIDICKLSGDN . IKFGATVTLIDDDT 99
  * * * * *
100 DEEQTVRVGDDDEADFKQNLISVNSPIARGLIGKEEDVVVKTGGEVE 149
  * * * * *
100 EEEVTYIIIVGEYADITKKRVSIASPIAKALIGKSVGDFEVVTPKGLKS 149
  * * * * *
150 FEVIKVEYL 158 E. coli
  * * *
150 YEVVTVEYKELDL 162 R. prowazekii

```

Figure 1. GAP comparison of the predicted amino acid sequence of *E.coli* and *R.prowazekii* GreA. * denotes an identical amino acid residue at the position.

* To whom correspondence should be addressed