

## recF in *Actinobacillus pleuropneumoniae*

Barbara M.Loynds, Paul R.Langford and J.Simon Kroll\*

Molecular Infectious Diseases Group, Department of Paediatrics, Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK

Submitted December 13, 1991

EMBL accession no. X63626

Part of a gene resembling *recF* was identified in the course of studies of genes responsive to oxidative stress in the respiratory tract pathogen *Actinobacillus pleuropneumoniae* serotype III, the cause of severe and economically-important swine pleuropneumonia. RecF from *Escherichia coli* is a single-stranded DNA binding protein (1) essential for the efficient repair of damaged DNA and for efficient induction of the SOS DNA repair genes. As it is therefore potentially important in the response to oxidative stress in *A. pleuropneumoniae*, further DNA was sequenced on both strands using synthetic oligonucleotide primers to reveal a 1179 bp open reading frame. On translation this DNA would encode a 360 amino acid protein with a sequence 56% identical to RecF from *E.coli* (a highly significant match,  $Z = 65.9$ , (2)). Our identification of this *A. pleuropneumoniae* gene as *recF* is substantiated by further extended similarity of its translated sequence to the *E. coli* protein in terms of biochemically conservative amino acid substitutions, and by the presence in the translated sequence of conserved domains previously recognised in RecF and other members of the recently-described superfamily of UvrA-related NTP-binding proteins (3) (Fig. 1). Domains I and III, corresponding to the N-terminal (A) and C-terminal (B) nucleotide-binding domains of Walker *et al.* (4), have a particularly close match to the *E. coli* RecF sequence. Gorbalenya and Koonin (3) have suggested that further domains II and IV flanking nucleotide-binding domain III are also conserved among DNA repair/recombination proteins. Domain II is found in the *A. pleuropneumoniae* RecF sequence, but the short domain IV ending in a supposedly conserved histidine (ISAEH in *E. coli* RecF) is not. A review of published sequences of RecF from other organisms identified a similar mismatch in RecF from the enteric organism *Proteus mirabilis*, in which the pentapeptide ITSGQ is found in the corresponding position to our ITKDKQ (5). It seems likely therefore that the stringency of conserved sequence in domain IV of UvrA-related NTP-binding proteins is less than previously suggested.

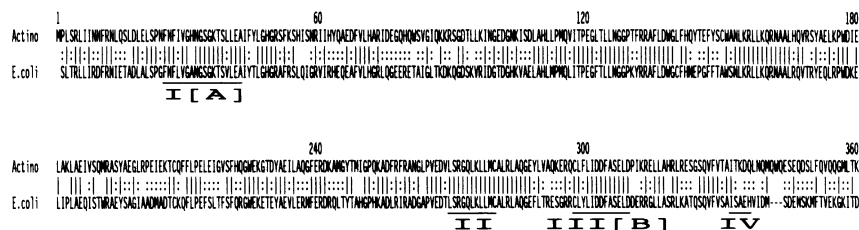
All bacterial *recF* genes previously described have been found within a cluster of genes encoding products essential for DNA replication. The gene order in *E. coli*, *dnaA-dnaN-recF-gyrB*, is also found in *Proteus mirabilis*, *Bacillus subtilis* and *Pseudomonas putida* (5, 6, 7, 8), and it has been suggested that this gene array is conserved across species (8). This is refuted by our findings in *A. pleuropneumoniae*, for while a gene highly similar to *dnaN* precedes *recF* as expected, *gyrB* is not found 3' to *recF*. Instead, *recF* is followed immediately by a run of adenylate residues, a stem-loop structure and a run of thymidylate residues, features characteristic of a rho-independent terminator on which two genes on opposite strands converge. The converging gene in question is *sodC*, a novel bacterial [Cu,Zn]-superoxide dismutase (9, 10) to be described in detail elsewhere.

### ACKNOWLEDGEMENTS

This work was supported by grants to JSK from the Wellcome Trust and the British Lung Foundation.

### REFERENCES

- Griffin, T.J., IVth and Kolodner, R.D. (1990) *J. Bacteriol.* **172**, 6291–6299.
- Lipman, D.J. and Pearson, W.R. (1985) *Science* **227**, 1435–1441.
- Gorbalenya, A.E. and Koonin, E.V. (1990) *J. Mol. Biol.* **213**, 583–591.
- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, M.J. (1982) *EMBO J.* **1**, 945–951.
- Skovgaard, O. (1990) *Gene* **93**, 27–34.
- Blonar, M.A., Sandler, S.J., Armengod, M.-E., Ream, L.W. and Clark, A.J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4622–4626.
- Ogasawara, N., Moriya, S., von Meyenburg, K., Hansen, F.G., and Yoshikawa, H. (1985) *EMBO J.* **4**, 3345–3350.
- Fujita, M.Q., Yoshikawa, H. and Ogasawara, N. (1989) *Mol. Gen. Genet.* **215**, 381–387.
- Kroll, J.S., Langford, P.R. and Loynds, B.M. (1991) *J. Bacteriol.* **173**, 7449–7457.
- Langford, P.R., Loynds, B.M. and Kroll, J.S. (1992) *J. Gen. Microbiol.* **138**, in press.



Alignment of the deduced amino acid sequence of RecF from *A. pleuropneumoniae* (360 amino acids) with that of the *E. coli* protein. Dashes between the sequences indicate identical matches and double dots, biochemically-conservative substitutions. Conserved domains I to IV (3) and nucleotide binding domains A and B (4) are indicated in the *E. coli* sequence.

\* To whom correspondence should be addressed