

An anticodon nuclease gene inserted into a *hsd* region encoding a type I DNA restriction system

P.Linder, R.Doelz, M.Gubler and T.A.Bickle
Departments of Microbiology and Biocomputing, Biozentrum, Basel University, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Submitted October 22, 1990 EMBL accession no. X52284

The DNA sequence of the *E. coli prr* locus has recently been published (1, EMBL data base accession number X52284). The locus encodes four polypeptides, PrrA–PrrD. PrrC codes an enzyme that cleaves $tRNA^{Lys}$ within the anticodon following T4 phage infection; the other three proteins are thought to be involved in masking this activity in non-infected cells.

We have noticed that the *prr* locus has a great deal of homology with the *hsdM*, *S* and *R* genes encoding type I restriction and modification enzymes, *Eco*R124 and *Eco*R124/3 (2, accession number X13145). We note that the *prr* locus was earlier shown to be associated with a DNA restriction system (3). The homology is shown as a dot plot in the Figure. The *prr* sequence (1) begins in the middle of an *hsdM* gene (the Met codon indicated as the start of *prrA* in (1) is internal to this reading frame), continues with an entire *hsdS* gene and ends with a truncated *hsdR* gene. The *prrC* gene is inserted into the 120 base pairs that normally separate *hsdS* and *hsdR*. There are only three amino acid sequence differences between HsdM and the 145 amino acids of PrrA translated from the beginning of the sequence. A PrrB and HsdS comparison gives results typical for HsdS proteins that recognise different DNA sequences (eg. ref.4): good homology at the ends and in the middle of the proteins, no homology in the two DNA recognising domains. Finally PrrD and HsdR are 76% identical, a figure which increases to 84% if the nucleotide at position 3713 of the *prr* sequence is deleted, an operation that results in identity between the two protein sequences beyond the deletion.

All previously identified members of the *Eco*R124 family of restriction systems are coded on conjugative plasmids. The *prr* system is chromosomal; however, strains that are *prr* negative lack 27 kb of DNA including the *prr* region (1). It is thus likely that *prr* was originally on a plasmid and was transferred to the chromosome by either cointegration or by a transposition event. It is not clear whether the ancestral *Eco*R124 family *hsd* region contained a *prrC* gene which was subsequently lost in the other members, or whether *prrC* is a recent addition.

The *prrC* anticodon nuclease can be considered to be a host defence system because its only known role is to abort infection by phage T4 mutated in either the RNA ligase or polynucleotide kinase genes (3). Indeed, the only known physiological function of the ligase and kinase is to repair the lesion caused by the anticodon nuclease. The close association of *prrC* with *hsd* genes encoding a classical DNA restriction system is reminiscent of the 'immigration control region' of the *E. coli* K12 chromosome where genes for three different restriction systems are clustered within 14 kb (5).

ACKNOWLEDGEMENT

This project was supported by grants from the Swiss National Science Foundation.

REFERENCES

1. Levitz,R., Chapman,D., Amitsur,M., Green,R., Snyder,L. and Kaufmann,G. (1990) *EMBO J.* **9**, 1383–1389.
2. Price,P., Lingner,J., Bickle,T.A., Firman,K. and Glover,S.W. (1989) *J. Mol. Biol.* **205**, 115–125.
3. Abdul-Jabbar,M. and Snyder,L. (1984) *J. Virol.* **51**, 522–529.
4. Gann,A.F.F., Campbell,A.J.B., Collins,J.F., Coulson,A.F.W. and Murray,N.E. (1987) *Molec. Microbiol.* **1**, 13–22.
5. Dila,D., Sutherland,E., Moran,L., Slatko,B. and Raleigh,E.A. (1990) *J. Bacteriol.* **172**, 4888–4900.
6. Maizel,J.V. and Leak,R.P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7665–7669.
7. Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucl. Acids Res.* **12**, 387–395.

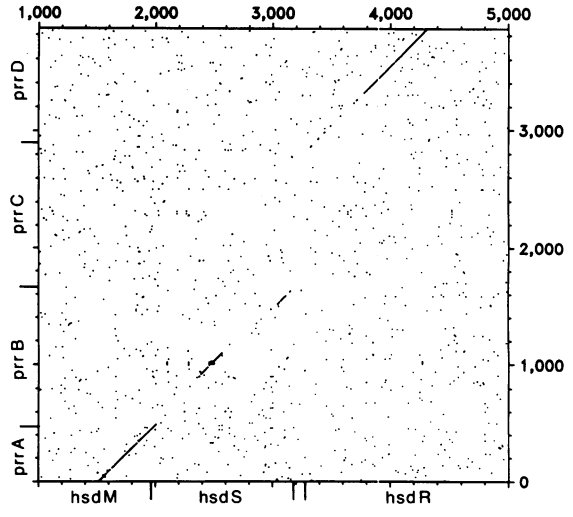

Figure 1. A comparison of the *prr* and *Eco*R124/3 *hsd* DNA sequences using the dot plot algorithm (6) as implemented in the GCG program package (7). A dot has been printed for every stretch of 7 identical nucleotides.