J. Mol. Biol. (1997) 271, 342–348

JMB

*Kpn*AI, a New Type I Restriction-Modification System in *Klebsiella pneumoniae*

Nan Sook Lee[1], Obed Rutebuka[1], Takeshi Arakawa[1], Thomas A. Bickle[2] and Junichi Ryu[1]*

[1]*Department of Microbiology and Molecular Genetics, School of Medicine, Loma Linda University, Loma Linda CA 92350, USA*

[2]*Department of Microbiology Biozentrum, University of Basel, Klingelberg Strasse 70 CH-4056, Basel, Switzerland*

**Corresponding author*

The *Kpn*AI restriction-modification (R-M) system has been identified in *Klebsiella pneumoniae* strain M5a1. The restriction gene of *Kpn*AI was first cloned into pBR322 using an r^-m^+ M5a1 derivative and phage SBS for screening. Subsequently, an adjacent DNA fragment showing modification activity was cloned into pUC19. A total of 7.2 kb DNA sequencing data revealed three open reading frames, corresponding to *hsdR*, *hsdM* and *hsdS* genes of type I R-M systems. The predicted *hsdR*, *hsdM* and *hsdS*-coded peptides shared 95%, 98% and 44% identity, respectively, with the corresponding peptides of the recently identified *Sty*SBLI system, a prototype of the type ID family. This high homology suggests that *Kpn*AI is also a member of the type ID family. The *Kpn*AI system seems to be the first type I system identified in *Klebsiella* species.

© 1997 Academic Press Limited

*Keywords: Sty*SBLI; restriction endonuclease; modification methylase; type ID; M5a1

Introduction

Restriction-modification (R-M) systems are common in bacteria. They are classified into three different types (I, II and III) based on the subunit composition, DNA cleavage pattern, cofactors involved and recognition sequence symmetry (see recent reviews: Barcus & Murray, 1995; Redaschi & Bickle, 1996). The type I systems were originally found in enteric bacteria. They are the most complex R-M systems and consist of three different subunits, encoded by the *hsdR, hsdM* and *hsdS* genes. Their genetic loci are organized into two transcriptional units, with the genes *hsdM* and *hsdS* forming an operon while *hsdR* is separate (Bickle & Krüger, 1993).

In type I enzymes, all three gene products are required for restriction, whereas *hsdM* and *hsdS* gene products are sufficient for modification methylation. The specificity of the system is determined by the *hsdS* gene. Genetic complementation, DNA homology and antigenic cross-reactivity tests led to the recognition of three distinct families of type I systems: IA, IB and IC (Bickle & Krüger, 1993). The presence of yet another family in *Salmonella* serotypes was suggested from complementation and DNA hybridization studies (Ryu *et al*., 1988). Subsequently, an additional family (ID) was added, based on DNA hybridization studies in both *Salmonella* and *Escherichia coli* (Barcus *et al*., 1995). The nucleotide sequence of the *Sty*SBLI genes confirmed this new family (Titheradge *et al*., 1996).

DNA sequences homologous to type I genes have also been discovered outside the Enterobacteriaceae such as in *Bacillus subtilis* (*Bsu*CI; Xu *et al*., 1995), *Mycoplasma pulmonis* (Dybvig & Yu, 1994), *Spiroplasma citri* (Laigret *et al*., 1996), *Haemophilus influenzae* Rd (Fleischmann *et al*., 1995) and *Methanococcus jannaschii* (Bult *et al*., 1996). These current data suggest that there may be many more type I R-M systems to be discovered.

Several type-II restriction enzymes have been isolated from strains of *Klebsiella pneumoniae* (Roberts & Macelis, 1996; Kravetz *et al*., 1993). To date, however, no clear evidence exists to support the presence of the type I R-M system in *Klebsiella* species.

Two R-M systems have been recognized in the strains of *K. pneumoniae* M5a1 and GM236 (Streicher *et al*., 1974; Satta *et al*., 1978; Bullas *et al*., 1981) and were later designated as *Kpn*AI and

Present address: N. S. Lee, Department of Molecular Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA.

Abbreviations used: Ap^R, ampicillin resistance; NG, nitrosoguanidine; EOP, efficiency of plating; *hsd*, host specificity for DNA; L-broth, Luria-Bertani broth containing 0.1% dextrose; L-agar, L-broth containing 1.5% Difco Bacto agar; SAM, *S*-adenosylmethionine; R-M, restriction and modification.

0022–2836/97/330342–07 $25.00/0/mb971202 © 1997 Academic Press Limited

*Kpn*BI, respectively (Valinluck *et al*., 1989). The restriction gene for *Kpn*BI has been cloned and sequenced (Valinluck *et al*., 1995). Amino acid homology and mutant studies suggest that *Kpn*BI is a member of either the type I or type III systems. From our previous studies (Valinluck *et al*., 1989), we suggested that *Kpn*AI is also either a type I or type III system, since both r^-m^+ and r^-m^- mutants were obtained in approximately equal numbers when r^- strains were selected.

We have further studied the *Kpn*AI system, and report here the cloning and sequencing of the complete *Kpn*AI locus and discuss evidence that this system is a new member of the type ID R-M system.

Results

Presence of an R-M system in M5a1

To show the presence of an R-M system in M5a1, a classical R-M test was performed using phage SBS. An r^-m^- derivative, 5022, was used as control. A typical result is shown in Table 1. As a comparison, Table 1 includes the EOP results of strain GM236, carrying the *Kpn*BI system, and its r^-m^- derivative (GM238). The result clearly indicates that strain M5a1 has an R-M system with different specificity from GM236. The R-M system in M5a1 was designated *Kpn*AI (Valinluck *et al*., 1989).

Cloning of the restriction and modification genes of the *Kpn*AI system

The restriction (R) gene of the *Kpn*AI system was cloned in pBR322 using a genetic complementation strategy. Four different R clones were obtained and designated pNL1 to pNL4. Modification tests showed that none of these clones contained the modification genes. The R gene was then located within a 3.6 kb region common to all four R clones (Figure 1). The shortest clone (pNLA2) was used for further subcloning and sequencing.

Since all previously described restriction and modification genes are clustered (Wilson, 1988), we assumed that the *Kpn*AI modification gene(s) is located close to the restriction gene. Hence, a chromosomal restriction map was first constructed using the 3.2 kb *Hin*dIII fragment of pNLA2 as a probe (Figure 1). Using the same probe, we cloned adjacent DNA fragments. A 6.7 kb *Pst*I fragment cloned in pUC19 was found to contain the *Kpn*AI modification activity. This fragment is located upstream of the R gene and partially overlapping with the R gene (Figure 1). A total of seven clones (pJR31 to pJR37) were obtained. The restriction pattern obtained from *Kpn*I digestion of these seven clones shows that they were all inserted in the same orientation in pUC19. Later analysis suggested that all the genes cloned are transcribed in the direction opposite to the *lac*Z promoter. Thus the overexpression of the *Kpn*AI modification genes in the high copy vector may be lethal and could be subcloned in only one orientation, because the same *Pst*I fragment was able to be subcloned in both orientations in the low copy number plasmid pACYC177. Subsequent studies were done using a representative plasmid pJR31. Several subclones of pJR31 were derived and used for DNA sequencing and complementation tests (Figure 1).

Complementation tests between two clones in *E. coli*

Plasmids containing either R or MS (M + S) genes of *Kpn*AI were transferred to *E. coli* DH5α and semi-quantitative R-M tests were performed using phage lambda. To solve the plasmid incompatibility problem, the MS genes were first subcloned into pACYC177 and the resultant plasmid was designated pJR41 (Figure 1). In this experiment, we noticed that these two plasmids had to be transferred sequentially into DH5α: first pJR41 (MS genes) and second pNLA2 (R gene). No transformants were obtained when cells were transformed in the reverse order or when two plasmids were transferred simultaneously. This result suggests that modification methylation is required prior to the establishment of those two plasmids. When lambda phages were used to infect the strain DH5α harboring the two plasmids (pJR41 and pNLA2), they were restricted to EOP values of 10^{-5} to 10^{-6}. This strong restriction suggests that lambda DNA has multiple restriction sites for the *Kpn*AI enzyme. Once lambda phages were

Table 1. The efficiency of plating of SBS phages propagated in four different *K. pneumoniae* strains and plated on the same four strains

Bacterial strains	R-M phenotype	SBS phage propagated on strain: 5022 (SBS.0)[a]	M5a1 (SBS.*Kpn*AI)	GM238 (SBS.0)[a]	GM236 (SBS.KpnBI)
5022	$r^-_{KpnAI}\ m^-_{KpnAI}$	1.0	1.0	1.0	1.0
M5a1	$r^+_{KpnAI}\ m^+_{KpnAI}$	7×10^{-3}	0.9	5×10^{-3}	1×10^{-2}
GM238	$r^-_{KpnBI}\ m^-_{KpnBI}$	1.0	1.1	1.1	1.3
GM236	$r^+_{KpnBI}\ m^+_{KpnBI}$	4×10^{-3}	1×10^{-3}	4×10^{-3}	1.2

EOP values relative to strain 5022 are shown.
[a] Zero stands for no modification.

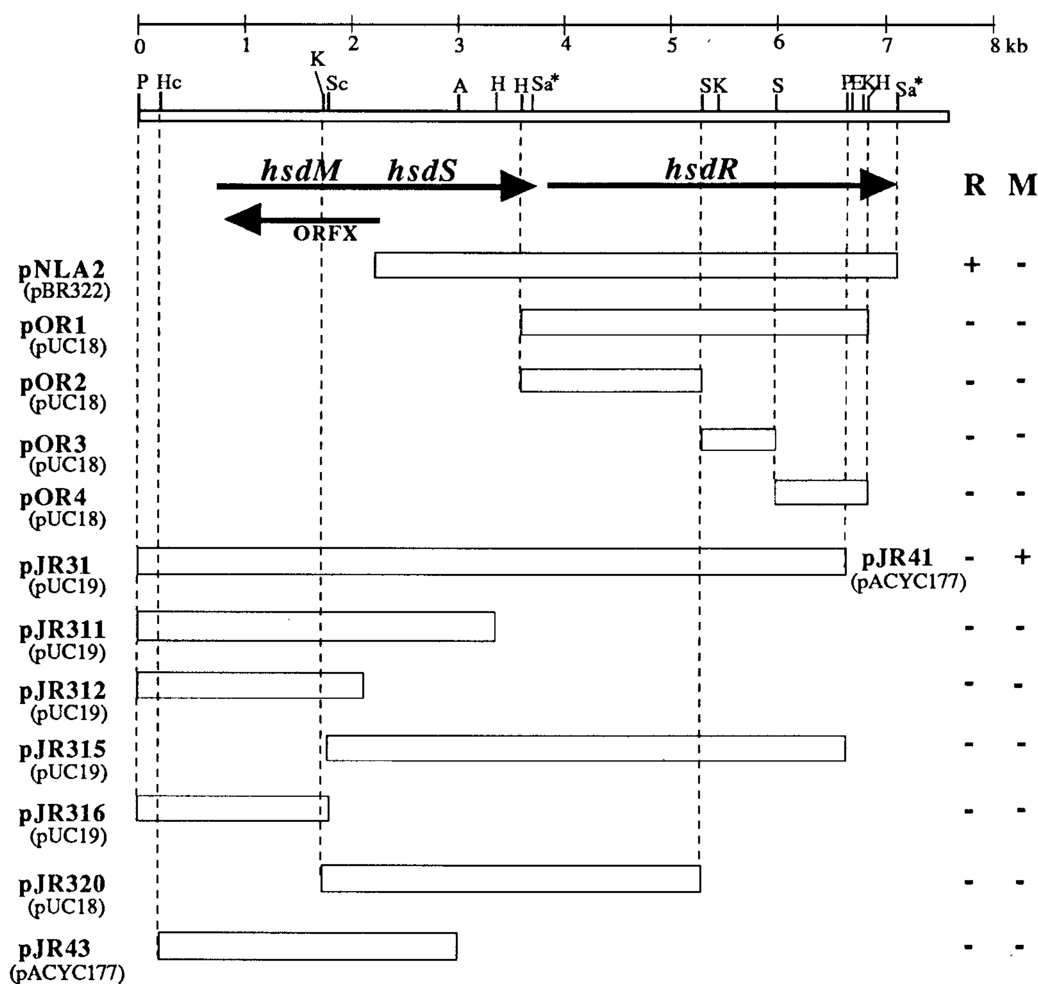

Figure 1. A chromosomal restriction map of the *Kpn*AI R-M system of M5a1. Two original clones, pNLA2 and pJR31, and their subclones, are shown. Parental plasmids are given in parenthesis. Arrows indicate the structural genes and their direction of transcription. R-M phenotypes of bacteria carrying each clone are also shown. Restriction enzyme abbreviations: A, *Apa*L1; E, *Eco*RI; H, *Hin*dIII; Hc, *Hin*cII; K, *Kpn*I; P, *Pst*I; S, *Sal*I; Sa*, *Sau*3AI partial digests; Sc, *Sac*I.

propagated in the cell, their DNA was modified completely (EOP 1.0) as expected.

DNA sequencing and prediction of the peptide sequences

The DNA sequence of a 7.2 kb region containing the *Kpn*AI locus was determined using pNLA2, pJR31 and their subclones as templates and the DNA sequence was stored in GenBank (accession no. U93843). This region covers both the restriction and modification genes as well as a 0.8 kb upstream region of the M gene (Figure 1) The DNA sequence revealed three ORFs that we designated as *hsdR*, *hsdM* and *hsdS*. We speculated that these three genes are transcribed in two separate units in the same direction (Figure 1). Another ORF, similar in size to the *hsdM* gene, but oriented in the opposite direction (ORFX in Figure 1) was also found. On the basis of the presence of a putative ribosome binding site as well as potential promoter sequences, we speculated that the longest ORF with 3,306 bp is the coding region for the R polypeptide (Figure 1). The seven helicase motifs common to all the type I R peptides (Gorbalenya & Koonin, 1991) were also found.

The DNA sequence upstream of the R ORF revealed two ORFs of 1614 bp and 1314 bp. These genes were designated as *hsdM* and *hsdS*, respectively (Figure 1). The predicted amino acid of the M peptide contains a potential SAM binding domain (ILANPPF) unique to N^6-adenine DNA methylases.

The *hsdM* ORF was preceded by a putative promoter and a ribosome binding site. However, as seen in other type I methylase genes, no obvious promoter sequence was found for the *hsdS* gene. The *hsdM* and *hsdS* coding sequences overlapped by five nucleotides. This feature is also common to other type I systems and allows the transcriptional coupling of *hsdM* and *hsdS* genes (Loenen *et al.*, 1987). Immediately after the stop codon (TAA) of the *hsdS* ORF, a sequence with dyad symmetry followed by an A + T rich sequence (a potential rho-independent termination) was found.

Confirmation of *hsdM* and *hsdS* genes

It is known that both HsdM and HsdS subunit are essential components of the type I modification methylase and that neither subunit alone shows any modification activity. To test whether the *Kpn*AI system requires both *hsdM* and *hsdS* genes for modification, plasmids containing either one of those genes were constructed from pJR31 for *in vivo* complementation experiments. First, the 3 kb *Apa*LI-*Pst*I fragment containing the *hsdM* gene was subcloned into pACYC177 to construct pJR43 (Figure 1). Similarly, a 4 kb *Kpn*I-*Sal*I fragment containing *hsdS* gene was subcloned into pUC18 to create pJR320 (Figure 1). The *hsdS* gene was intentionally placed downstream of the *lacZ* promoter to provide an exogenous promoter. Neither of those clones showed any modification activity by themselves. However, when these two plasmids were transferred to DH5α for the modification test, the full modification activity (EOP of lambda, 1.0) was restored. This experiment confirmed the necessity of the two genes (*hsdM* and *hsdS*) for modification activity.

The presence of a *hsdS* gene promoter was investigated by subcloning the 5.8 kb *Sac*I -*Pst*I fragment containing *hsdS* gene into pUC19 to construct pJR315 (Figure 1). In this construct, no exogenous promoter was provided for the *hsdS* gene. When this plasmid was transferred to DH5α harboring pJR43, modification activity was not restored. This result supports the idea that the *hsdS* gene lacks its own promoter.

Homology search

Each of the predicted amino acid sequences of *Kpn*AI genes was compared to the GenBank data base using the BLAST program. PC/Gene and GCG software programs were also used for more detailed analysis. To our surprise, both the predicted HsdR and HsdM peptide sequences from the *Kpn*AI locus showed an extremely high degree of identity with the predicted HsdR (95%) and HsdM (98%) peptide sequences of the recently identified *Sty*SBLI system (Titheradge *et al.*, 1996). On the contrary, the HsdS peptide showed limited identity (44%) between the two enzymes. Figure 2 shows the result of comparison of the HsdS peptides of *Kpn*AI and *Sty*SBLI using DOTPLOT. It is obvious that two clusters of highly homologous regions exist at the central region and the C-terminal region. The central homologous region consists of 76 amino acids (155 to 230), whereas the C-terminal homologous region consists of a shorter 32 amino acids domain (408 to 439). The former region has 95% identity (4% similarity), whereas the latter region has 88% identity (0% similarity) between the two systems. Similarly, the N-terminal (proximal) variable region and the distal variable region have 40% identity (14% similarity) and 14% identity (20% similarity), respectively. This pattern of homology with an alternative constant and a variable region within the *hsdS* gene is a characteristic of the members of the type I family. The two variable regions comprise the recognition domains of the type I enzymes (Gann *et al.*, 1987).

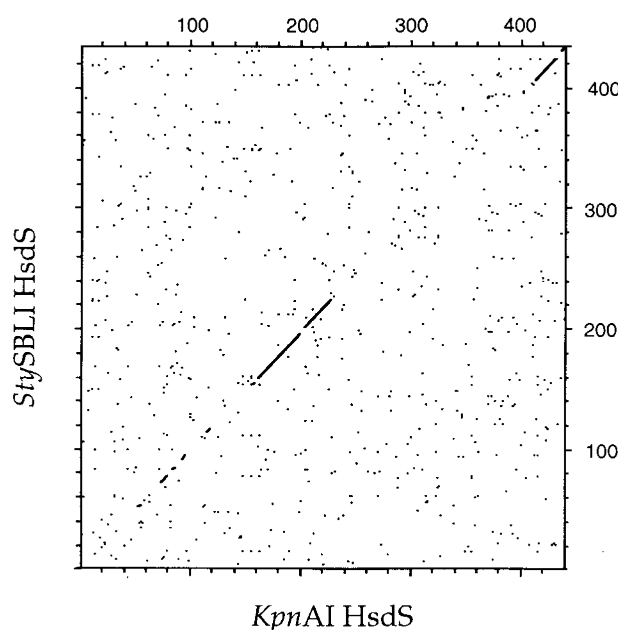

Figure 2. A comparison of the amino acid sequences of the *Sty*SBLI and *Kpn*AI HsdS proteins. The COMPARE program (GCG) was used with a word of size 2 and the results are displayed with DOTPLOT.

Further sequence analysis revealed that the predicted HsdR peptide of the *Kpn*AI enzyme also has the best homology (26% identity) with another *hsdR*-like sequence (HI 1285) of *H. influenzae* Rd (Fleishmann *et al.*, 1995). Similarly, the HsdM peptide sequence is highly homologous (49% identity) to another HsdM-like sequence (HI1287) of *H. influenza* Rd (Fleishmann *et al.*, 1995).

Discussion

Here we describe the cloning of the *Kpn*AI restriction and modification genes. Both restriction and modification genes were cloned individually in high copy vector plasmids. This two-step cloning strategy has an advantage over the cloning of the entire genes at one time, especially when the resultant clone is harmful to the cell. Our results show that the cloned plasmids (pNLA2 for *hsdR*, pJR41 for *hsdM* and *hsdS*) are unable to be transferred simultaneously into the recipient cells. Therefore, it is likely that the entire *Kpn*AI genes cannot be cloned in one vector especially when unmodified host bacteria are used.

Three distinct ORFs, which correspond to *hsdR*, *hsdM* and *hsdS* genes of type I R-M systems, have been identified. The complementation study clearly shows that two genes (*hsdM* and *hsdS*) are involved in the modification activity. This eliminates the possibility that the *Kpn*AI is a type III

system, since a single gene (*mod*) is sufficient for the modification activity in type III systems. Although another ORF (designated ORFX) was identified (Figure 1), it is not clear from the DNA sequence whether this gene is transcribed or whether it influences the *Kpn*AI R-M function. A similar ORF was reported from the *Sty*SBLI system (RF6; Titheradge *et al*., 1996).

Amino acid homology scores are useful to categorize the members of the type I restriction enzymes. Comparison of the amino acid homology of all the reported M and R proteins of type I R-M systems clearly leads us to the expected three families (IA to IC). That is, intra-family homology of R and M peptides is at the 80 to 90% level, whereas inter-family homology is only at the 20 to 30% level (Sharp *et al*., 1992). In our present study, high DNA and peptide sequence homology (more than 97%) was observed between the R and M peptides of *Kpn*AI and the recently identified *Sty*SBLI. This high homology strongly supports the possibility that the *Kpn*AI is a type I R-M system and belongs to the newly proposed ID family (Titheradge *et al*., 1996). Thus *Kpn*AI is the first R-M system clearly identified as type I in *Klebsiella* species. Including our finding, type ID system have now been identified in three species of enteric bacteria, *E. coli* (*Eco*90I), *Salmonella enterica serovar blegdam* (*Sty*SBLI) (Titheradge *et al*., 1996) and *K. pneumoniae* (*Kpn*AI). It would be interesting to know how widely spread the type ID homologous genes are both inside and outside the Enterobacteriaceae.

Contrary to the extremely high homology of the R and M subunits in the *Kpn*AI and *Sty*SBLI systems, the differences between their *hsdS* genes are obvious, suggesting that they have their own distinct recognition DNA sequences. Typical type I enzymes recognize two-part, asymmetric DNA sequences that consist of three nucleotides at the 5′ end and four at the 3′ end separated by a nonspecific spacer of six to eight nucleotides. Each of the two variable regions in the HsdS peptide recognizes the corresponding target sequence. Amino acid homology in the variable regions is usually only 20 to 30% even within the same family. The type IA *Eco*KI and *Sty*SPI enzymes, however, share homologous (92%) 5′ variable regions and recognize the same 5′ domain sequence (AAC). Their 3′ variable regions share 38% homology and the *Sty*SPI enzyme recognizes GTRC, which is a degenerate sequence of GTGC recognized by the *Eco*KI enzyme. Since the homology in the 5′ variable regions between HsdS peptides of the *Kpn*AI and *Sty*SBLI systems is 54%, it is possible that both enzymes recognize the same or similar 5′ target sequence.

The DNA sequence presented here also reveals that three *hsd* genes are organized in two transcriptional units, *hsdM* and *hsdS* as an operon and *hsdR* as a single gene. In type I R-M systems, all three genes are transcribed in the same direction. However, the gene order is characteristic for each family. Both the type IA and IB families share the same gene order (*hsdR*, *hsdM* and *hsdS*), whereas type IC and ID possess another gene order (*hsdM*, *hsdS* and *hsdR*). The *hsd* genes in the latter group have higher homology with the recently identified type I DNA sequences of *H. influenzae* (Fleishmann *et al*., 1995) and *M. jannaschii* (Bult *et al*., 1996).

Type IC genes are mainly found on plasmids, whereas type ID genes are chromosomally linked (Titheradge *et al*., 1996; Bullas *et al*., 1980; our unpublished data). The *hsd* genes of type IA, IB and ID mapped so far are alleles of the same locus that lies at the 98 minute region on the chromosome and are linked to *serB*. Although a detailed chromosome map is not available for *K. pneumoniae*, our preliminary mapping data suggest that *Kpn*AI is linked to the chromosome but not located in the *serB* region at 98 minutes (unpublished). Thus, it is possible that *Kpn*AI occupies a unique position on the chromosome among the type ID group.

Materials and Methods

Bacteria and phages

K. pneumoniae strain M5a1 and r^-_{KpnAI} m^-_{KpnAI} derivative 5022 (Streicher *et al*., 1974) were from C. Kennedy *via* L. Bullas. Other *K. pneumoniae* strains 236 and r^-_{KpnBI} m^-_{KpnBI} derivative 238 have been described (Valinluck *et al*., 1995). An r^-m^+ derivative M5a1R was isolated after NG treatment. *E. coli* DH5α (Woodcock *et al*., 1989) and XL1-Blue (Bullock *et al*., 1987) were also used for cloning and the complementation test. L-broth and L-agar were used for growing bacterial cultures. Appropriate amounts of antibiotics were added to the medium when necessary.

Restriction-modification tests

Semiquantitative R-M tests were done as described by Bullas *et al*. (1980). Phage SBS was used for the *K. pneumoniae* strain and lambda was used for the *E. coli* strain.

Cloning

To clone the *hsdR* gene of *Kpn*AI, chromosomal DNA of M5a1 was isolated using a G NOME® DNA isolation kit (BIO 101). The DNA was partially digested with *Sau*3AI and ligated into *Bam*HI-digested pBR322. The ligated mixture was electroporated into an r^- m^+ mutant (M5a1R). Cells were plated on L-agar containing 1 mg Ap/ml, and incubated overnight at 30°C. To screen the potential R clones, the strategy used for the *Kpn*BI system (Valinluck *et al*., 1995) was also used. Thus, the Ap^R transformants were replica plated onto L-plates with a lawn of SBS.0 phages (10^4 phages/plate) spread on the surface. It was expected that bacteria carrying a plasmid containing the $hsdR_{KpnAI}$ gene would restrict the SBS.0 phage and thus survive on the plate, whereas the remaining cells would be lysed.

For cloning the modification genes, a restriction map was constructed using a 3.2 kb *Hin*dIII fragment as a probe. A 6.7 kb *Pst*I fragment was identified as a candidate for chromosomal walking. Thus the chromosomal DNA of M5a1 was digested with *Pst*I and the 6.7 kb

region was taken out of the agarose, cleaned with Geneclean® (BIO 101) and ligated with *Pst*I digested (and dephosphorylated) pUC19 and transferred to DH5α cells. Plasmids of Ap^R transformants were extracted and hybridized to the 3.2 Kb *Hin*dIII *hsdR* probe. The probe was labeled with digoxigenin-dUTP and the hybrids were detected by enzyme immunoassay (Boehringer Mannheim).

DNA sequencing and sequence analysis

DNA sequencing was performed by the dideoxynucleotide termination method (Sanger *et al*., 1977) using both the manual and automated sequencing system. In the manual sequencing, the Sequenase® (United State Biochemical) kit was used with [α-^{35} S]dATP. For the automated sequencing, an Applied Biosystems automated sequencer (ABI 373A) was used with the Dye Terminator Cycle Sequencing method. Both strands of the cloned genes and various subclones shown in Figure 1 were used to determine the sequences. The sequences of DNA fragments were compiled by the ABI Sequencher® and PC/Gene (Intelligenetics) software.

PC/Gene software and BLAST of the National Center for Biotechnology Information were used for sequence similarity searches in GenBank and SWISS-PROT database. Protein sequences were aligned using either the PALIGN or the CLUSTAL program. For the similarity between the *hsdS* genes, COMPARE and DOTPLOT (Genetics Computer Group, Wisconsin) were also used.

Acknowledgments

We are grateful to Dr Leonard Bullas for phage SBS and *Klebsiella* strains, Dr Garry Larson and Thanhtam Huynh for DNA sequencing, and Dr Adam Bailis for useful comments on the manuscript. Work in Basel was supported by a grant from the Swiss National Foundation.

References

Barcus, V. A. & Murray, N. E. (1995). Barriers to recombination: restriction. In *Population Genetics of Bacteria* (Baumberg, S., Young, J. P. W., Saunders, S. R. & Wellington, E. M. H., eds), pp. 31–58, Cambridge University Press, Cambridge, UK.

Barcus, V. A., Titheradge, A. J. B. & Murray, N. E. (1995). The diversity of alleles at the *hsd* locus in natural populations of *Escherichia coli*. *Genetics*, **140**, 1187–1197.

Bickle, T. A. & Krüger, D. H. (1993). Biology of DNA restriction. *Microbiol. Rev.* **57**, 434–450.

Bullas, L. R., Colson, C. & Neufeld, B. (1980). Deoxyribonucleic acid restriction and modifcation systems in *Salmonella*: chromosomally located systems in different serotypes. *J. Bacteriol.* **141**, 275–291.

Bullas, L. R., Ryu, J. & Merriam, L. (1981). The DNA restriction and modification systems of *Klebsiella pneumoniae*. In *Abstracts of the Annual Meeting of the American Society for Microbiology*, pp. 136, Am. Soc. Microbiol, Washington, DC.

Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *BioTechniques*, **5**, 376–379.

Bult, C. J., *et al* (1996). Complete genome sequence of the Methanogen Archaeon, *Methanococcus jannaschii*. *Science*, **273**, 1058–1073.

Dybvig, K. & Yu, H. (1994). Regulation of a restriction and modification system via DNA inversion in *Mycoplasm pulmonis*. *Mol. Microbiol.* **12**, 547–560.

Fleischmann, R. D., *et al* (1995). Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, **269**, 496–512.

Gann, A. A. F., Campbell, A. J. B., Collins, J. F., Coulson, A. F. W. & Murray, N. E. (1987). Reassortment of DNA recognition domains and the evolution of new specificities. *Mol. Microbiol.* **1**, 13–22.

Gorbalenya, A. E. & Koonin, E. V. (1991). Endonuclease (R) subunits of type I and III restriction-modification enzymes contain a helicase-like domain. *FEBS Letters*, **291**, 277–281.

Kravetz, A. N., Zakharova, M. V., Beljetzkaja, I. V., Pertzev, A. V., Spivak, O. I. & Solonin, A. S. (1993). Two novel restriction endonucleases from *Klebsiella pneumoniae*. *Nucl. Acids Res.* **21**, 1501.

Laigret, F., Gaurivaud, P. & Bove, J. (1996). The unique organization of the *rpoB* region of *Spiroplasma citri*: a restriction and modification system gene is adjacent to *rpoB*. *Gene*, **171**, 95–98.

Loenen, W. A. M, Daniel, A. S., Braymer, H. D. & Murray, N. E. (1987). Organization and sequence of the *hsd* genes of *Escherichia coli* K-12. *J. Mol. Biol.* **198**, 159–170.

Redaschi, N. & Bickle, T. A. (1996). DNA restriction and modification systems. In Escherichia coli *and* Salmonella typhimurium. *Cellular and Molecular Biology* (Neidhardte, F. C., ed.), pp. 773–781, American Society for Microbiology, Washington, DC.

Roberts, R. J. & Macelis, D. (1996). Rebase-restriction enzymes and methylases. *Nucl. Acids Res.* **24**, 223–235.

Ryu, J., Rajadas, P. T. & Bullas, L. R. (1988). Complementation and hybridization evidence for additional family of type I DNA restriction and modification genes in *Salmonella* serotypes. *J. Bacteriol.* **170**, 5785–5788.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.

Satta, G., Debbia, E., Pruzzo, C. & Calegari, L. (1978). The peculiar behavior of coliphage p1*vir* mutants on restricting hosts. *Microbios*, **22**, 93–102.

Sharp, P. M., Kelleher, J. E., Daniel, A. S., Cowan, G. M. & Murray, N. E. (1992). Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc. Natl Acad. Sci. USA*, **89**, 9836–9840.

Streicher, S. L., Shanmugam, S. L., Ausubel, F., Morandi, C. & Goldberg, R. B. (1974). Regulation of nitrogen fixation in *Klebsiella pneumoniae*: evidence for a role of glutamine synthetase as a regulator of nitrogenase synthesis. *J. Bacteriol.* **120**, 815–821.

Titheradge, A. J. B., Ternent, D. & Murray, N. E. (1996). A third family of alleic *hsd* genes in *Salmonella enterica*: Sequence comparisons with related proteins identify conserved regions implicated in restriction of DNA. *Mol. Microbiol.* **22**, 437–447.

Valinluck, B., Reynold, M. & Ryu, J. (1989). Two new restriction-modification systems, *Kpn* AI and *Kpn* BI, in *Klebsiella pneumoniae*. In *Abstracts of the Annual Meeting of the American Society for Microbiology*, pp. 180, Ame. Soc. Microbiol, Washington, DC.

Valinluck, B., Lee, N. S. & Ryu, J. (1995). A new restriction-modification system, *Kpn*BI, recognized in *Klebsiella pneumoniae*. *Gene,* **167**, 59–62.
Wilson, G. G. (1988). Cloned restriction-modification systems: a review. *Gene,* **74**, 281–289.
Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989). Quantitative evalution of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucl. Acids Res.* **17**, 3469–3478.
Xu, G., Willert, J., Kapfer, W. & Trautner, T. A. (1995). *Bsu*CI, a type-I restriction and modification system in *Bacillus subtilis*. *Gene,* **157**, 59.

Edited by K. Yamamoto

(Received 17 March 1997; received in revised form 9 June 1997; accepted 10 June 1997)