JOURNAL OF BACTERIOLOGY, Apr. 1988, p. 1775–1782
0021-9193/88/041775-08$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 170, No. 4

Distribution and Diversity of *hsd* Genes in *Escherichia coli* and Other Enteric Bacteria

ANNE S. DANIEL, FRANCES V. FULLER-PACE,† DIANA M. LEGGE, AND NOREEN E. MURRAY*
Department of Molecular Biology, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, United Kingdom

Received 9 November 1987/Accepted 22 January 1988

We screened *Salmonella typhimurium*, *Citrobacter freundii*, *Klebsiella pneumoniae*, *Shigella boydii*, and many isolates of *Escherichia coli* for DNA sequences homologous to those encoding each of two unrelated type I restriction and modification systems (*Eco*K and *Eco*A). Both K- and A-related *hsd* genes were identified, but never both in the same strain. *S. typhimurium* encodes three restriction and modification systems, but its DNA hybridized only to the K-specific probe which we know to identify the *Sty*SB system. No homology to either probe was detected in the majority of *E. coli* strains, but in *C. freundii*, we identified homology to the A-specific probe. We cloned this region of the *C. freundii* genome and showed that it encoded a functional, A-related restriction system whose specificity differs from those of known type I enzymes. Sequences immediately flanking the *hsd* K genes of *E. coli* K-12 and the *hsd* A genes of *E. coli* 15T⁻ were shown to be homologous, indicating similar or even identical positions in their respective chromosomes. *E. coli* C has no known restriction system, and the organization of its chromosome is consistent with deletion of the three *hsd* genes and their neighbor, *mcrB*.

The fundamental attribute of a restriction endonuclease is its ability to recognize and selectively attack foreign DNA. This DNA is usually identified by the absence of a sequence-specific modification (methylation). Although modification is always sequence specific, restriction endonucleases do not necessarily cut within the DNA target sequence identified by methylation. Most enzymes classified as type II do, while others, classified as either type I or III, do not (for reviews, see references 6 and 43).

This paper is concerned with type I restriction systems. These, in every way, are the most complex; they comprise three subunits, have complex cofactor requirements, and cut DNA nonspecifically some thousands of base pairs from the target sequence for modification. The genes encoding these restriction endonucleases are designated *hsd* for host specificity of DNA. The specificity polypeptide, encoded by *hsdS*, associates with that encoded by *hsdM* to form a modification methylase and with those encoded by *hsdM* and *hsdR* to produce a restriction endonuclease.

It is generally accepted that restriction endonucleases function to prevent the acquisition and expression of foreign DNA (2). One consequence would be a degree of protection against infection by bacteriophages, an advantageous feature for the initial establishment of populations of bacteria in new habitats (25). A second consequence could be a reduction in the flow of genetic information between populations of bacteria. This latter suggestion received support from the finding that the classification of many *Salmonella* serotypes, on the basis of standard taxonomical methods, correlates well with their host specificity (*hsd*) systems (11). However, it has been suggested that the ends of DNA fragments resulting from restriction could initiate recombination with homologous chromosomes (16, 31; S. Lederberg, cited in reference 33). Thus, DNA restriction might sometimes serve to stimulate the exchange of genetic information.

* Corresponding author.
† Present address: Research Institute of Scripps Clinic, La Jolla, CA 92037.

Two commonly used laboratory strains, *Escherichia coli* K-12 and *E. coli* B, each have a single, chromosomally encoded, type I system, while *E. coli* 15T⁻ has both a chromosomally encoded type I system, designated A, and a plasmid-encoded type III system, designated P15 (4). Plasmids encoding restriction systems representative of all three types have been isolated from *E. coli* strains, but to date the chromosomally encoded systems are all type I. No *E. coli* strain has been shown to have more than one chromosomally encoded restriction system, but the chromosomal genes of *Salmonella typhimurium* LT2 specify three systems, i.e., LT, SA, and SB (11). It is not known whether more than one of these three systems are type I. A very different system that restricts DNA containing methylated cytosine is encoded, at least in part, by the *mcrB* gene located close to the *hsd* genes in *E. coli* K-12 (34, 36).

Genetic tests dependent on the interaction between the polypeptides of different specificity systems (for example, see reference 10) and molecular experiments based on DNA hybridization and immunological cross-reactivity (28) indicated that the K and B systems of *E. coli* and the SB and SP systems of *S. typhimurium* and *Salmonella potsdam*, respectively, are all related. There is no evidence to indicate that the A system of *E. coli* 15T⁻ is a member of this family, even though it is encoded by three genes that behave as alleles of *hsd* K in P1 transduction experiments (W. Arber, personal communication; N. E. Murray, unpublished observation). On the contrary, the *hsd* A genes are thought to identify a second family of type I systems (19, 41).

This paper reports the search for genes sharing homology with those encoding either the A or K restriction system. DNAs from laboratory and natural isolates of *E. coli* were screened, as well as a small sample of DNAs of other enteric bacteria. Related specificity systems were detected, some of which have been described elsewhere (19, 23, 29), and both the K and A families were represented in more than one genus. Nevertheless, as judged by our failure to detect homologous sequences, most *E. coli* strains lacked either K- or A-related *hsd* genes.

MATERIALS AND METHODS

Bacterial and phage strains. The bacterial strains screened in this survey are shown in Table 1. Lambda phages were propagated on the *E. coli hsd* K and *hsd* A deletion derivatives NM522 (23) and NM555 (19), respectively. The phage vector EMBL3 (17) was used for cloning fragments of the *Cfr*I *hsd* system from *Citrobacter freundii*; λNM1149 and λNM1150 (26) were used for libraries of *E. coli* C and *E. coli* 15T$^-$.

Media and microbial techniques. The media and general methods used were those described by Murray et al. (27).

Enzymes and chemicals. DNA polymerase I and restriction enzymes were purchased from Boehringer Mannheim Biochemicals; deoxycytidine 5'-[α-^{32}P]triphosphate (110 TBq/mmol) was purchased from Amersham International.

Preparation of DNA. Bacterial DNA was prepared essentially as described by Kaiser and Murray (24), except that dialysis was used instead of precipitation with ethanol. Phage DNA was prepared as described by Wilson et al. (42). Plasmid DNA was purified from cleared bacterial lysates by centrifugation in CsCl-ethidium bromide (13).

Restriction endonuclease digestion and ligation of DNA. Restriction endonuclease digestion and DNA ligation were done as described by Murray et al. (28). Libraries of recombinant phages were recovered by either transfection or in vitro packaging (3). Fragments from cloned genes in recombinant λ phages were subcloned in pBR322, and the recombinant plasmids were recovered by transformation of HB101 (9) or NM522 (23).

Hybridization analysis of bacterial DNAs. Bacterial DNAs were digested with a restriction enzyme, usually *Eco*RI, and after separation through agarose gels, the DNA fragments were transferred to nitrocellulose filters (39). Nick-translated probes were made by using the relevant plasmids as described by Rigby et al (35). Hybridizations were typically carried out overnight at 37°C in a solution containing 50% formamide, 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), 5× Denhardt solution, and 25 μg of sonicated calf thymus DNA per ml. Washes were done at 37°C in 2× SSC–0.1% sodium dodecyl sulfate (twice for 30 min) and subsequently at room temperature in 1× SSC–0.1% sodium dodecyl sulfate (twice for 30 min). Autoradiography was carried out at −70°C by using Cronex X-ray film and intensifying screens.

RESULTS

Screening of bacterial DNAs with K- and A-specific *hsd* probes. Our survey screened DNA from *E. coli*, *C. freundii*, *Klebsiella pneumoniae*, *S. typhimurium*, and some *Shigella* species. A sample of more than 20 *E. coli* strains included standard laboratory strains and isolates from natural sources, most frequently hospital patients. Some of the *E. coli* strains (Table 1) had been characterized on the basis of many protein polymorphisms (12, 38).

The DNA probes were *hsd* genes cloned in pBR322. pRH1 has part of the *hsd* region of *E. coli* K-12, and pFFP20 (and pFFP32) has a segment of the *hsd* region of the A specificity determinants. The *hsd*-specific segment in pRH1 and pFFP32 is an entire *hsdM* gene flanked by part of *hsdR* and part of *hsdS*; the segment in pFFP20 is only part of *hsdM* and part of *hsdS* (Fig. 1). These probes will be referred to as

TABLE 1. Bacterial strains used to provide DNA

Strain	Relevant features	Source[a]	Reference
C600	*E. coli hsd* K		1
WA2899	*E. coli hsd* A in K-12		4
15T$^-$	*E. coli hsd* A; P15	W.A.	
B	*E. coli* B *hsd* B	R.S.H.	40
C-1a	*E. coli* C		5
W	*E. coli* W		22
A58	*E. coli* API	K.K.	15
A101	*E. coli* API	K.K.	15
E166[b]	*E. coli* API	K.K. and K.C.	
629[b]	*E. coli* API	K.K. and K.C.	
653[b]	*E. coli* API	K.K. and K.C.	
E147B[b]	*E. coli* API	K.K. and K.C.	
E171[b]	*E. coli* API	K.K. and K.C.	
E136[b]	*E. coli* API	K.K. and K.C.	
E163[b]	*E. coli* API	K.K. and K.C.	
BLXA[c]	*E. coli*; 19 of 20 like B	B.R.L.	38
BLD4[c]	*E. coli*; 20 of 20 like K	B.R.L.	
BRL-ET2[c]	*E. coli*; 15 of 15 like K	B.R.L.	12
BRL-ET13[c]	*E. coli*; 14 of 15 like K	B.R.L.	12
RM66A[c]	*E. coli*; 19 of 20 like K	B.R.L.	38
RM74A[c]	*E. coli*; 19 of 20 like K	B.R.L.	38
RM01A[c]	*E. coli*; 6 of 15 like K	B.R.L.	
4247	*S. typhimurium* LT2; *hsd* SA, LT, SB		11
NCTC 9750	*C. freundii*	NCTC	
M5al	*K. pneumoniae*	C.K.	14
KK47	*Shigella boydii*	K.K. and K.C.	

[a] W.A., Werner Arber; R.S.H., Richard Hayward; K.K., Kim Kaiser; K.C., Keith Cartwright; B.R.L., Bruce Levin; C.K., Christina Kennedy; NCTC, National Collection of Type Cultures, London.

[b] These strains were isolated by K. Cartwright in the Western General Hospital, Edinburgh, United Kingdom, and were classified by him as *E. coli* by using the Analytical Profile Index 20E system (API). The DNAs were donated by K. Kaiser.

[c] These *E. coli* strains were sent by B. R. Levin, and the electrophorectic mobilities of at least 15 of their proteins had been compared with those of *E. coli* K-12 and *E. coli* B; these data are summarized as a relevant feature: e.g., 19 of 20 like B indicates that the electrophoretic mobilities of 19 of 20 proteins analyzed were identical to those of *E. coli* B.

K-specific or A-specific. Bacterial DNAs were digested with a restriction enzyme and, after separation through agarose gels, the DNA fragments were transferred to nitrocellulose (39). Duplicate filters were hybridized to radiolabeled vector DNA (pBR322) and subsequently to either the K- or A-specific probe. Additional bands detected by the second hybridization indicated homology with the respective *hsd* sequence. One such experiment is illustrated in Fig. 2, and the results of many are summarized in Table 2. In some cases, DNA fragments hybridizing to pBR322 sequences themselves precluded unambiguous conclusions, since additional fragments of identical mobility would not be resolved. We chose to exclude bacterial DNAs that gave complex patterns with vector DNA rather than use probes of labeled *hsd* DNA separated from the vector plasmid, since traces of pBR322 sequences in the probe would detect homologous sequences that could be misinterpreted as *hsd* specific.

Of 20 isolates of *E. coli*, other than *E. coli* K-12, the DNA of 4 isolates, including *E. coli* B, hybridized to *hsd* K (pRH1), and the DNA of 2 isolates, excluding *E. coli* 15T$^-$, hybridized to *hsd* A (pFFP20 or pFFP32). None hybridized to both, as expected from the allelic behavior of the *hsd* A and K genes in transduction experiments.

The chromosome of *S. typhimurium* 4247 encodes the SA, SB, and LT specificity systems (11); the SB system is known to be related to the K system (10, 28). The *hsd* K probe identified a single band, while the A-specific probe hybridized to none, implying that A is not related to either SA or LT. The DNA of *C. freundii* includes DNA that hybridized to the A-specific probe; the DNA of *K. pneumoniae* was negative with both probes. The *Shigella* DNAs were not easy to analyze because these hospital isolates had extensive homology with pBR322. We did not detect additional bands with either pRH1 or pFFP20 but were cautious in accepting these negative results and included *Shigella boydii* only in Table 2.

Interspecific hybridization with the K probe was weaker than intraspecific hybridization, and subsequent comparisons of the sequences of part of the *hsd* region of *E. coli* K-12 and *S. typhimurium* have indicated about one difference per 10 nucleotides (20; J. Kelleher, personal communication). Intraspecific hybridization was consistently weaker for sequences detected by the A probe than for those detected by *hsd* K, at least in part because of a difference in the length of the *hsd* sequences in the plasmid probes (Fig. 1). The faintness of the A-specific bands (Fig. 2) emphasizes the inherent problems in the interpretation of negative results.

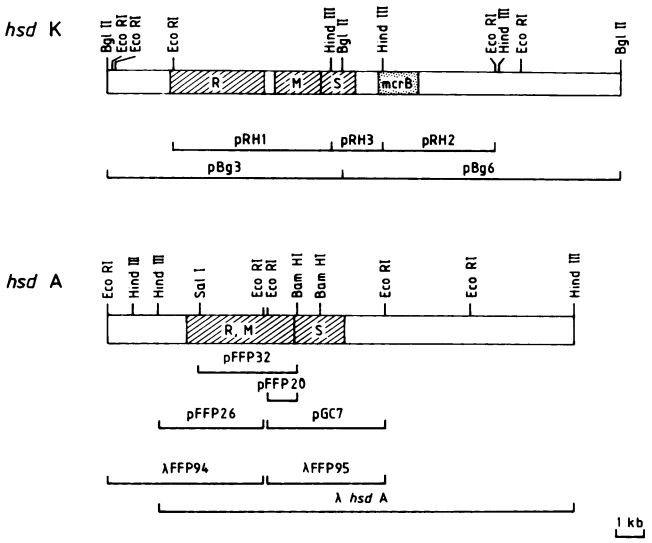

FIG. 1. Maps of the *hsd* regions of *E. coli* K-12 and *E. coli* WA2899 (a derivative of *E. coli* K-12 with *hsd* A genes). Shaded areas identify known genes. The extent of the region carried in plasmids and phage is indicated. Note that λFFP94 and λFFP95 were cloned directly from *E. coli* 15T$^-$ and not *E. coli* WA2899.

TABLE 2. Hybridization to probes for restriction genes

Strain[a]	Hybridization to the following probe[b]			
	Vector	*hsd* K	*hsd* A	*mcrB*
C600	−	+	−	+
15T$^-$	−	−	+	−
B	−	+	−	+
C-la	−	−	−	−
A58	−	−	+	−
A101	−	−	+	−
E166	−	+	−	−
BLXA	−	+	−	+
RM01A	−	−	−	−
RM74A	−	−	−	−
629	−	+	−	+
653	−	−	−	−
RM66A	−	−	−	−
S. typhimurium (4247)	−	+	−	−
C. freundii (NCTC 9750)	−	−	+	NT
K. pneumoniae (M5al)	−	−	−	NT
W	+	−	−	−
E147B	+	−	−	NT
E171	+	−	−	−
BRL-ET2	(++)	−	−	−
BRL-ET13	(+)	−	−	−
BLD4	+	−	−	−
Shigella boydii	+(+)	−	−	NT
E136	+(+++)	−	−	−
E163	+(+++)	−	−	−

[a] Strain descriptions are given in Table 1; strains are *E. coli* unless designated otherwise. The strains in this table are subdivided by the results obtained when the vector alone was used as the probe.

[b] Symbols: +, hybridization; −, no hybridization. For the vector probe, each + sign indicates a band identified by a vector sequence; parentheses indicate weak hybridization. NT, Not tested.

Diversity of specificities. The homologous sequences indicated by the *hsd* probes (Table 2), other than that identified in strain BLXA, were cloned in λ vectors, and their specificities were confirmed or identified (Table 3). Strain BLXA encoded a B-specific modification system, which we presume is encoded by DNA sharing homology with the K-specific probe. The B specificity was inferred from the following experiment. BLXA will not serve as an indicator for phage λ; nevertheless, lysogens of a λ *bla* phage were selected as ampicillin-resistant colonies. These lysogens produced only low yields of phage, but the resulting phage had replaced their K-specific modification with that of B (data not shown).

The five *E. coli* strains with K-like *hsd* genes identify three specificity systems (23; this paper), and the three with A-like systems identify two specificities (19). The cloning and analysis of the *hsd* genes of *C. freundii* are reported below.

A-related restriction system (*Cfr*I) encoded by *C. freundii*. A library of the *C. freundii* genome was made by recovering *Sau*3A partial digestion products in a λ vector. Plaques hybridizing to the A-specific probe were detected, and six independent clones were selected for analysis. Phage from

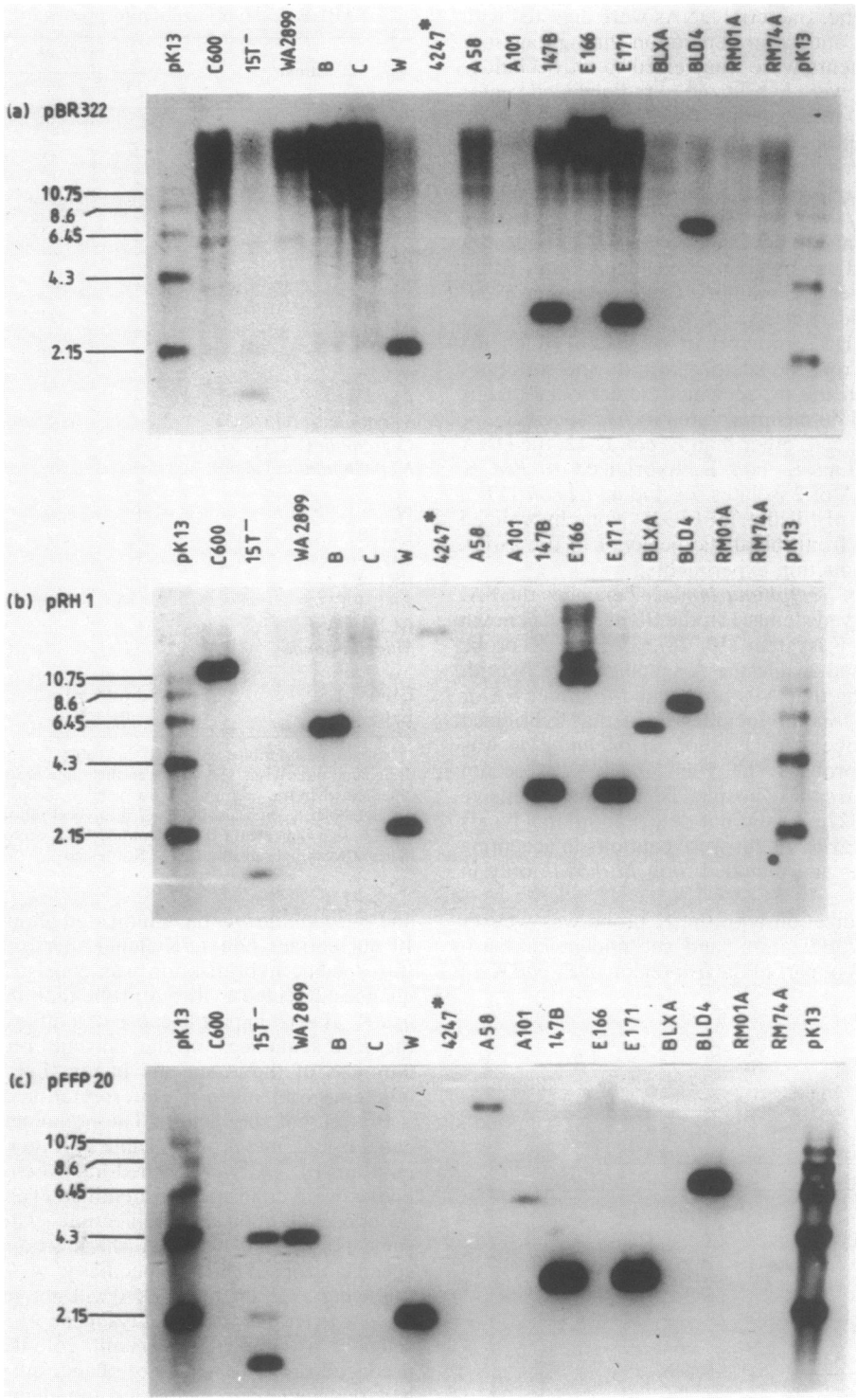

FIG. 2. An example of the screening of natural isolates of enteric bacteria for the presence of K- and A-related *hsd* genes. The bacterial DNAs were digested with *Eco*RI and, after electrophoresis through an agarose gel, transferred to a nitrocellulose filter before hybridization. (a) One of two filters probed with pBR322 to detect background bands due to vector DNA; (b) one of the duplicate filters shown in panel a subsequently probed with pRH1 (K specific); (c) the other filter shown in panel a subsequently probed within pFFP20 (A specific). pK13 (8) was used as a marker; marker sizes are given in kilobases. Strains and plasmids are indicated across the top.*, *S. typhimurium* 4247, which encodes SA, SB, and LT systems.

each of these were screened for an active *hsdS* gene by a simple complementation test previously devised to confirm the relatedness of the A and E specificity systems (19). In this test, the active *hsdR* and *hsdM* genes of the A system are provided by a plasmid in a restriction- and modification-deficient ($r^- m^-$) strain used as an indicator for the λ *hsd* phages. If the S polypeptide of the *Cfr*I system, encoded by the λ *hsd Cfr* phage, can associate with the resident R and M

TABLE 3. Specificity of cloned *hsd* genes[a]

Strain	Family	Specificity	Reference
E. coli K-12 C600	K (prototype)	K	7
E. coli B	K	B	23
E. coli E166	K	D	23
E. coli 629	K	B	23
S. typhimurium LT2	K	SB	18
E. coli 15T$^-$	A (prototype)	A	19
E. coli A58	A	E	19
E. coli A101	A	A	19

[a] Note that the K-like *hsd* genes of BLXA have not been cloned, but the data in the text indicate that they confer B specificity.

polypeptides of A to make an active restriction endonuclease, the unmodified host chromosome would be susceptible to attack, and, consequently, the efficiency of plating (EOP) of a λ *hsdS*$^+$ phage would be low.

Four of the six phages plated with an EOP of 10^{-3} on the r$^-$ m$^-$ strain carrying the *hsdR*$^+$*M*$^+$ plasmid (Table 4) and are presumed to include an active *hsdS* gene from an A-related system. The remaining two, while including part of the *hsd* region, are assumed to be deleted for part of the *hsdS* gene.

Specificity of the *hsdS* gene of *C. freundii* different from the specificities of A and E. At least one of the λ *hsd Cfr* phages encodes a functional restriction and modification system. When this phage (λ *hsd Cfr*I 6) is integrated into the chromosome of r$^-$ m$^-$ *E. coli* (NM555), it endows the host with an r$^+$ m$^+$ phenotype. (Since the λ vector is *imm*$^\lambda$ *c*IΔ, an *imm*^{434} *c*I$^+$ derivative was made and introduced by a λ *imm*^{21} *c*I$^+$ helper phage.) Three of the six λ *hsd Cfr* phages plated with an EOP of 1 on the (heteroimmune) dilysogen, while three plated with a low EOP indicative of restriction (Table 5). The three phages which plated with an EOP of 1 were all previously identified as *hsdS*$^+$ (Table 4) and presumably also include a functional *hsdM* gene, so that they modify their genomes and are resistant to attack by the *Cfr*I system. The other λ *hsd* phages, presumably, do not include the entire coding sequence for the methylase and, like λ *vir*, plated with an EOP of 1 only if previously propagated in the dilysogen endowed with the *Cfr*I specificity system.

*Cfr*I is a new specificity system since phage λ modified with either the A or E specificity remained sensitive to it (Table 5). Similarly, λ modified by the *Cfr*I methylase was restricted by both the A and E systems. The *Cfr*I specificity has also been checked against other type I systems; λ phage modified by the *Cfr*I specificity were restricted by all the known type I systems of *E. coli* and by the SA, SB, SP, SQ, and SJ *Salmonella* systems (data not shown, but for identification of the systems, see references 11 and 21); and the *Cfr*I system restricted λ modified with these same *E. coli* and *Salmonella* specificities.

TABLE 4. Test for A-related specificity (*hsdS*) gene

Infecting phage	EOP on NM522 (pFFP21)[a]	Genotype of phage
λ *hsd Eco*E Δ 5[b]	1	*hsdS*
λ *hsd Eco*E Δ 6[b]	10^{-3}	*hsdS*$^+$
λ *hsd Cfr*I 1	1	*hsdS*
λ *hsd Cfr*I 2	10^{-3}	*hsdS*$^+$
λ *hsd Cfr*I 3	1	*hsdS*
λ *hsd Cfr*I 4	10^{-3}	*hsdS*$^+$
λ *hsd Cfr*I 5	10^{-3}	*hsdS*$^+$
λ *hsd Cfr*I 6	10^{-3}	*hsdS*$^+$

[a] pFFP21 has the *hsdR* and *hsdM* genes of the A system in pBR322. The EOP on NM522 (pBR322) was 1 for all strains.
[b] See reference 19.

TABLE 5. *C. freundii* A-related system of different specificity from that of *Eco*A or *Eco*E

Phage	EOP on NM555 (λ *hsd Cfr*I 6)[a]
λ *hsd Cfr*I 1 (*S*)	10^{-3}
λ *hsd Cfr*I 2 (*S*$^+$)	10^{-2}
λ *hsd Cfr*I 3 (*S*)	10^{-3}
λ *hsd Cfr*I 4 (*S*$^+$)	1
λ *hsd Cfr*I 5 (*S*$^+$)	1
λ *hsd Cfr*I 6 (*S*$^+$)	1
λv.0[b]	10^{-3}
λv.A[b]	10^{-3}
λv.E[b]	10^{-3}
λv.*Cfr*I[b]	1

[a] The EOP on NM555 was 1 for all phages.
[b] λv.A, λv.E, and λv.*Cfr*I indicate λ *vir* modified against the relevant restriction system. λv.0 is unmodified.

A preliminary restriction map (Fig. 3) of the *hsd* region of *C. freundii* was based on the analysis of the DNA of the six λ *hsd* phages. The approximate locations of the *hsd* genes were deduced from the phenotypes of these phages.

Organization of the chromosome of *E. coli* C in the *hsd* region. The genome of *E. coli* C lacks homology with a probe for *hsd*K and *mcrB* (37), but flanking probes from the *hsd* region of *E. coli* K-12 (pBg3 and pBg6; Fig. 1) each identify a fragment of 5.1 kilobases (kb) in a total *Eco*RI digest of *E. coli* C DNA (data not shown).

*Eco*RI fragments of the *E. coli* C genome were cloned in λ, and plaques were screened with radiolabeled probes for the flanking regions of the *hsd* genes of *E. coli* K-12 (pBg3 and pBg6). The same plaques hybridized to both flanking probes. DNA was purified from two phages, and each included a 5.1-kb *Eco*RI fragment that hybridized to sequences upstream (pBg3) and downstream (pBg6) of the *hsd* K genes. The genome of *E. coli* C may be deleted for a segment of DNA including the *hsd* and *mcrB* genes; alternatively, it may be a descendant of ancestral bacteria that never included this region.

Screening of bacterial DNAs with a probe for the *mcrB* gene. The finding that *E. coli* C differs from *E. coli* K-12 by the absence of a segment of DNA, including both the *hsd* and *mcrB* genes, raises the possibility that this entire region was acquired concomitantly from a mobile element. This notion might anticipate a correlation between the presence of an *mcrB* gene and K-like *hsd* genes in the various bacterial DNAs. The segment of the *E. coli* K-12 chromosome in pRH2 (Fig. 1) includes most, though not all, of the *mcrB* gene and some downstream sequence (36). We used this plasmid as a screen for *mcrB* in our bacterial DNAs and detected strong signals only in the three *E. coli* strains that confer B specificity, in addition to the K-12 control (see the last column of Table 2). Although the only evidence for *mcrB* was in those strains encoding a K-family *hsd* system, homology was not detected in all such strains, notably E166 (*Eco*D) and *S. typhimurium*.

K and A *hsd* genes possibly allelic. In P1 transduction experiments using either *serB*$^+$ or *dnaC*$^+$ as a selectable marker, no transductant with both A and K specificities has been isolated (W. Arber, personal communication; N. E. Murray, unpublished observation), yet the genes are sufficiently dissimilar that no cross-hybridization has been de-

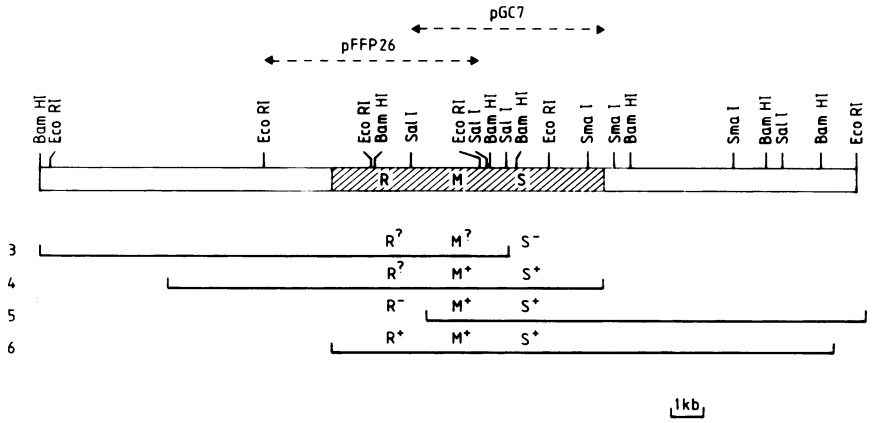

FIG. 3. Map of the *hsd* region of *C. freundii*. The orientation of the three genes was determined by hybridization to probes (Fig. 1) pFFP26 (*hsdRM* A) and pGC7 (*hsdMS* A); the fragments which hybridized to each are shown (⟷). The regions carried in four λ *hsd Cfr* phages (numbered 3 to 6 in the figure) are shown (⌊⌋); the phenotypes of these phages indicate the positions of *hsdR*, *hsdM*, and *hsdS* (▨).

tected (28). We used our cloned fragments of the *hsd* genes to compare their relative locations in the *E. coli* K-12 and *E. coli* 15T⁻ chromosomes by checking for hybridization between flanking sequences.

Two *Eco*RI fragments from the *hsd* A region of strain 15T⁻, each of which includes a short (2- to 3-kb) segment of flanking sequence, were cloned in a λ vector (Fig. 1). One, a 5.4-kb *Eco*RI fragment in λFFP94, included the beginning of *hsdM*, all of *hsdR*, and some upstream sequence; the other, a 4.1-kb *Eco*RI fragment in λFFP95, included the distal end of *hsdM*, all of *hsdS*, and a short segment of downstream sequence. Various digests of these two DNAs and λ *hsd* A, which covers the whole *hsd* region (Fig. 1), were assayed for hybridization to probes derived from the *hsd* region of *E. coli* K-12. Five plasmids (pRH1, pRH2, pRH3, pBg3, and pBg6; Fig. 1) which contain segments of the *hsd* K genes and their flanking sequences were used as the probes. Figure 4 illustrates the information gained from these hybridization experiments. The K-specific *hsd* probe (pRH1), as expected, detected no homology within the *hsd* A DNA. Plasmid pBg3, which includes sequences upstream of the *hsdR* gene of *E. coli* K-12, hybridized to the 5.4-kb *Eco*RI fragment present in λFFP94 and more particularly to the 3.2-kb *Eco*RI-*Sal*I fragment containing mainly DNA upstream of *hsdR* A. It did not hybridize to the 4.1-kb *Eco*RI fragment in λFFP95. Plasmid pBg6, which includes sequence downstream of the *hsdS* gene, hybridized to the 4.1-kb *Eco*RI fragment in λFFP95 and more particularly to the 2.3-kb *Bam*HI-*Eco*RI segment extending downstream of *hsdS* A. Plasmid pBg6 did not hybridize to the 5.4-kb *Eco*RI fragment in λFFP94. The insert in pBg6 is 10 kb, so shorter probe sequences were used to refine the analysis. Plasmid pRH3, a 1.8-kb segment that extends only 900 base pairs downstream of *hsdS*, shared homology with the 4.1-kb *Eco*RI fragment of λFFP95. Our hybridization experiments were consistent with identical positions for the *hsd* K and A genes.

DISCUSSION

Our screening of bacterial DNAs with probes from within both the *hsd* K and *hsd* A genes detected homologous sequences within functional *hsd* genes, some of which endowed new specificities. A more remarkable observation was the frequent absence of homology to the *hsd* K and *hsd* A probes. *E. coli* C is one such strain, and in this case, a 5.1-kb *Eco*RI fragment spanned all or most of what in the *E. coli* K-12 genome is an 18-kb segment. This apparent deletion in *E. coli* C removes both the *hsd* and *mcrB* genes. We do not have a definitive analysis for any other *E. coli* strain whose genomic DNA lacks homology with both the *hsd* K and A probes, but preliminary experiments using sequences that flank the *hsd* K genes as probes suggest considerable variability.

The K- and A-like *hsd* systems were assigned to different families on the basis of a number of criteria. This division relied in part on lack of homology, as judged by failure to detect hybridization between DNA fragments, as well as on dissimilarity at the polypeptide level implied from both complementation tests and the absence of cross-reactivity

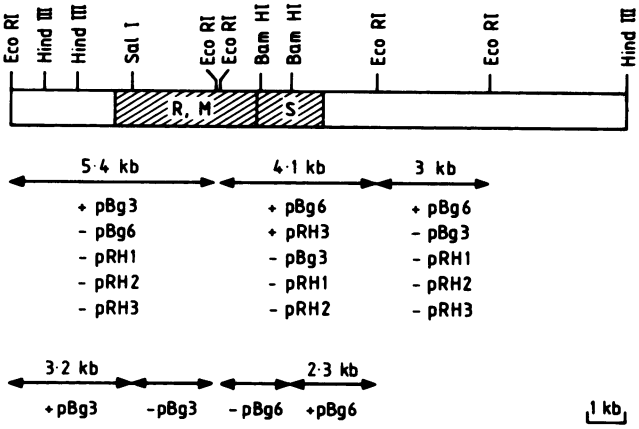

FIG. 4. Positions of the *hsd* A genes in the chromosome compared with the position of *hsd* K. The restriction map was deduced from λ *hsd* A (derived from strain WA2899). The *Eco*RI fragments from *E. coli* 15T⁻ cloned in λ agree with this map, and λFFP94 extends it leftwards. The arrows (⟷) indicate the relevant fragments of 15T⁻ or λ *hsd* A; underneath are listed the plasmids used as probes. Symbols: +, hybridization to the relevant fragment; −, no hybridization detected. The region of the *E. coli* K-12 genome carried in each plasmid is diagrammed in Fig. 1. In analyses of the digests of genomic DNA from *E. coli* 15T⁻ (data not shown), neither the 5.4-kb nor the 4.1-kb *Eco*RI fragment was identified with any of the probes shown above. The 3-kb fragment hybridized to pBg6, but no homology was detected with pRH2, a plasmid including most of the *mcrB* gene.

between polyclonal antisera (19, 28, 32). The classification of the *hsd* A and K genes into different families contrasts with their apparent allelism, as detected by P1 transduction. For this reason, we compared their positions in the *E. coli* chromosome by molecular tests.

Cloned segments of the *hsd* regions of the K-12 and 15T$^-$ chromosomes were checked for cross-hybridization (Fig. 4). Since none was detected between the *hsd*-specific sequences (28; this paper), we anticipate that hybridization will be confined to flanking sequences. The preliminary nucleotide sequence of the *hsdS* A gene (Gill Cowan, personal communication) confirms that this gene shares no homology with *hsdS* K and allows us to conclude that the homology between the *Hin*dIII fragment in pRH3 and the *Eco*RI fragment in λFFP95 is confined to the 0.9-kb sequence downstream of *hsdS* K and the 1.6-kb segment downstream of *hsdS* A. The 3.2-kb *Eco*RI-*Sal*I fragment immediately upstream of *hsdR* A shares homology with pBg3, which includes approximately 2 kb of DNA upstream of *hsdR* K. Our experiments indicate that the *hsd* genes are in the same orientation and that the immediate short flanking sequences share homology. The data are consistent with an identical location for the two families of *hsd* genes. If the nucleotide sequences reinforce this suggestion, two alternative possibilities must be considered: either two clusters of analogous but unrelated genes have an identical location or the identity of location reflects a common origin but the genes have diversified so extensively that the tests applied fail to detect relatedness.

Our present analysis of type I restriction systems did not detect any simple correlation between the specificity of the restriction system and other enzyme polymorphisms. One strain classified as resembling *E. coli* B on the basis of enzyme polymorphisms (B. R. Levin, personal communication) did have a B specificity system, but others classified as K-like had no specificity system, as judged by our homology-dependent screening.

We already know that the *hsdS* genes of the K family have diverged so much that only short sections of these genes retain homology (23). Two large segments share no similarity at the DNA level; nevertheless, there is evidence of relatedness if the encoded amino acid sequences are compared (21). Similar divergence may have occurred within the *hsdR* and *hsdM* genes to create apparent alternative families of which K and A are but two representatives. Complementation tests involving various *Salmonella* systems suggest that a division into two families may be an oversimplification (J. Ryu and L. R. Bullas, manuscript in preparation).

The absence of both K- and A-like *hsd* genes in many *E. coli* strains is difficult to reconcile with their presence in *S. typhimurium* and *C. freundii*, respectively. Related *hsd* genes in *E. coli* and *S. typhimurium* or *E. coli* and *C. freundii* may indicate horizontal transfer subsequent to the separation of ancestral lines of different genera, perhaps via generalized transduction. Cross-hybridization tests, however, do indicate that the *hsd* genes of *E. coli* K-12 are much more similar to those of *E. coli* B than to those of *Salmonella* species (28). Only a 226-base-pair segment of the *hsdM* gene has been sequenced from *E. coli* and *Salmonella* species; within this segment, the proportion of silent sites that differ when *E. coli* K-12 and *E. coli* B are compared is only 8%, while figures of 26% and 30% result from interspecific comparisons. The latter figures, nevertheless, fall short of the 58% that represents an average value for comparisons of *E. coli* and *Salmonella* species (30). The nucleotide sequences of K- and A-like *hsd* genes and appropriate searches for alternative systems are necessary for an understanding of the relationships between *hsd* systems. Whether most strains have alternative *hsd* genes or lack them completely is basic to any consideration of the relevance of type I restriction enzymes.

ACKNOWLEDGMENTS

We thank Jill Gough for invaluable help and advice in the early stages of this work, and Len Bullas for making available unpublished work. We are also grateful to Kim Kaiser for donation of bacterial DNAs, Gill Cowan, Alexander Gann, and Bruce Levin for critical discussion, Jeanette Ferguson for preparation of the manuscript, and Annie Wilson for drawing the figures.

The work was supported by the Medical Research Council.

LITERATURE CITED

1. **Appleyard, R. K.** 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. Genetics **29**:440–452.
2. **Arber, W.** 1965. Host-controlled modification of bacteriophage. Annu. Rev. Microbiol. **19**:365–378.
3. **Arber, W., L. Enquist, B. Hohn, N. E. Murray, and K. Murray.** 1983. Experimental methods for use with lambda, p. 433–466. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
4. **Arber, W., and D. Wauters-Willems.** 1970. Host specificity of DNA produced by *Escherichia coli*. XII. The two restriction and modification systems of strain 15T$^-$. Mol. Gen. Genet. **108**: 203–217.
5. **Bertani, L. E.** 1968. Abortive induction of bacteriophage *P2*. Virology **36**:87–103.
6. **Bickle, T. A.** 1982. The ATP dependent restriction endonucleases, p. 85–108. *In* S. M. Linn and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **Borck, K., J. D. Beggs, W. J. Brammar, A. S. Hopkins, and N. E. Murray.** 1976. The construction *in vitro* of transducing derivatives of phage lambda. Mol. Gen. Genet. **146**:199–207.
8. **Bouché, J. P., J. P. Gélugne, J. Louarn, J. M. Louarn, and K. Kaiser.** 1982. Relationships between the physical and genetic maps of a 470 × 10^3 base-pair region around the terminus of *Escherichia coli* K12 DNA replication. J. Mol. Biol. **154**:21–32.
9. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
10. **Bullas, L. R., and C. Colson.** 1975. DNA restriction and modification systems in *Salmonella*. III. SP, a *Salmonella potsdam* system allelic to the SB system in *Salmonella typhimurium*. Mol. Gen. Genet. **139**:177–188.
11. **Bullas, L. R., C. Colson, and B. Neufield.** 1980. Deoxyribonucleic acid restriction and modification systems in *Salmonella*: chromosomally located systems of different serotypes. J. Bacteriol. **141**:275–292.
12. **Caugant, D. A., B. R. Levin, and R. K. Selander.** 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. Genetics **98**:467–490.
13. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA **62**:1159–1166.
14. **Dixon, R., C. Kennedy, A. Kondorosi, V. Krishnapillai, and M. Merrick.** 1977. Complementation analysis of *Klebsiella pneumoniae* mutants defective in nitrogen fixation. Mol. Gen. Genet. **157**:189–198.
15. **Duguid, J. P., I. W. Smith, G. Dempster, and P. N. Edmunds.** 1955. Non-flagellar filamentous appendages ("fimbrae") and haemagglutinating activity in Bacterium coli. J. Pathol. Bacteriol. **70**:335–348.
16. **Endlich, B., and S. Linn.** 1985. The DNA restriction endonuclease of *Escherichia coli* B. II. Further studies of the structure of DNA intermediates and products. J. Biol. Chem. **260**:5729–

5738.
17. **Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
18. **Fuller-Pace, F. V., L. R. Bullas, H. Delius, and N. E. Murray.** 1984. Genetic recombination can generate altered restriction specificity. Proc. Natl. Acad. Sci. USA **81**:6095–6099.
19. **Fuller-Pace, F. V., G. M. Cowan, and N. E. Murray.** 1985. *Eco*A and *Eco*E: alternatives to the *Eco*K family of type I restriction and modification systems of *Escherichia coli*. J. Mol. Biol. **186**: 65–75.
20. **Fuller-Pace, F. V., and N. E. Murray.** 1986. Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. Proc. Natl. Acad. Sci. USA **83**:9368–9372.
21. **Gann, A. A. F., A. J. B. Campbell, J. F. Collins, A. F. W. Coulson, and N. E. Murray.** 1987. Reassortment of DNA recognition domains and the evolution of new specificities. Mol. Microbiol. **1**:13–22.
22. **Glover, S. W., and G. Kerszman.** 1967. The properties of a temperate bacteriophage Wϕ isolated from *Escherichia coli* strain W. Genet. Res. **9**:135–139.
23. **Gough, J. A., and N. E. Murray.** 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J. Mol. Biol. **166**:1–19.
24. **Kaiser, K., and N. E. Murray.** 1979. Physical characterisation of the "Rac prophage" in *Escherichia coli* K12. Mol. Gen. Genet. **175**:159–174.
25. **Levin, B. R.** 1986. Restriction-modification immunity and the maintenance of genetic diversity in bacterial populations, p. 669–688. *In* S. Karlin and E. Nero (ed.), Evolutionary processes and theory. Academic Press, Inc., New York.
26. **Murray, N. E.** 1983. Phage lambda and molecular cloning, p. 395–432. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Murray, N. E., W. J. Brammar, and K. Murray.** 1977. Lambdoid phages that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **150**:53–61.
28. **Murray, N. E., J. A. Gough, B. Suri, and T. A. Bickle.** 1982. Structural homologies among type I restriction modification systems. EMBO J. **1**:535–539.
29. **Nagaraja, V., M. Stieger, C. Nager, S. M. Hadi, and T. A. Bickle.** 1985. The nucleotide sequence recognised by the *Escherichia coli* D type I restriction and modification enzyme. Nucleic Acids Res. **13**:389–399.
30. **Ochman, H., and A. C. Wilson.** 1987. Evolutionary history of enteric bacteria, p. 1649–1654. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
31. **Price, C., and T. A. Bickle.** 1986. A possible role for DNA restriction in bacterial evolution. Microbiol. Sci. **3**:296–299.
32. **Price, C., T. Pripfl, and T. A. Bickle.** 1987. *Eco*R124 and *Eco*R124/3: the first members of a new family of type I restriction and modification systems. Eur. J. Biochem. **167**:111–115.
33. **Radding, C. M.** 1973. Molecular mechanisms in genetic recombination. Annu. Rev. Genet. **7**:87–111.
34. **Raleigh, E. A., and G. Wilson.** 1986. *Escherichia coli* K12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA **83**:9070–9074.
35. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**: 237–251.
36. **Ross, T. K., and H. D. Braymer.** 1987. Localization of a genetic region involved in McrB restriction by *Escherichia coli* K-12. J. Bacteriol. **169**:1757–1759.
37. **Sain, B., and N. E. Murray.** 1980. The *hsd* (host specificity) genes of *E. coli* K12. Mol. Gen. Genet. **180**:35–46.
38. **Selander, R. K., and B. R. Levin.** 1980. Genetic diversity and structure in *Escherichia coli* populations. Science **210**:545–547.
39. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
40. **Studier, F. W.** 1969. The genetics and physiology of bacteriophage T7. Virology **39**:562–574.
41. **Suri, B., and T. A. Bickle.** 1985. *Eco*A: the first member of a new family of type I restriction and modification systems. J. Mol. Biol. **186**:77–85.
42. **Wilson, G. G., V. I. Tanyashin, and N. E. Murray.** 1977. Molecular cloning of fragments of bacteriophage T4 DNA. Mol. Gen. Genet. **156**:203–214.
43. **Yuan, R.** 1981. Structure and mechanism of multifunctional restriction endonucleases. Annu. Rev. Biochem. **50**:285–315.