JOURNAL OF BACTERIOLOGY, Feb. 1991, p. 1335–1338
0021-9193/91/031335-04$02.00/0

Vol. 173, No. 3

Map Position and Genomic Organization of the *kps* Cluster for Polysialic Acid Synthesis in *Escherichia coli* K1

ERIC R. VIMR

Department of Veterinary Pathobiology, Microbiology Division, University of Illinois, Urbana, Illinois 61801

Received 24 September 1990/Accepted 24 November 1990

The multigenic *kps* cluster in *Escherichia coli* K1 encodes functions for synthesis of a polysialic acid capsule. DNA probes flanking each side of the cluster were hybridized to lambda clones bearing overlapping *E. coli* W3110 genomic fragments. These fragments covered the region between 60 and 70 map units on the chromosome. The results located *kps* to an accretion domain near 64 map units and established the orientation of *kps* cluster genes. Acquisition of *kps* by the *E. coli* genome was apparently the result of an ancestral transpositionlike addition event.

Invasive *Escherichia coli* strains isolated from humans and other animals are often encapsulated (5, 6). The encapsulating structures are usually polysaccharides designated K antigens after the German word for capsule, "Kapsel" (8). Of the more than 70 K antigens that have been described, a few are consistently associated with invasive disease, in which capsules function as antiphagocytic barriers to host immune mechanisms (5, 6). One of these capsules, the K1 antigen, is an unbranched homopolymer of sialic acid and a virulence factor in neonatal meningitis and urinary tract infections (reviewed in reference 13). Identical polysialic acid moieties exist as capsules on *Neisseria meningitidis* group B, *Pasteurella haemolytica* A2, and *Moraxella nonliquefaciens* organisms and on N-linked glycan chains of neural cell adhesion molecule and sodium channel polypeptides in vertebrates (reviewed in reference 13). Since only *E. coli* of a specific K-antigen type contain the genes for that capsule's synthesis (3), the origin of capsule genes encoding identical polysaccharides in otherwise distantly related species is provocative from an evolutionary standpoint.

Several K-antigen gene clusters, including K1, are located near *serA* at a site previously designated *kpsA* (9). Various approaches have helped to identify a common genetic organization for *kps* cluster genes in which a central cassette (region 2) of sugar-specific genes is flanked on each side by genes encoding more general functions (regions 1 and 3) for polysaccharide assembly and perhaps translocation (1, 15). Some of the flanking genes are functionally interchangeable between different K-antigen serotypes (11), suggesting the possibility of homologous recombination between the shared regions 1 and 3 of these serotypes. This mechanism can account for capsule structural diversity and is consistent with the allelism noted previously for *kps* clusters which map near *serA* (9). The capacity to rapidly integrate region 2 genes coding for new capsule structures into different replicons by horizontal (lateral) gene transfer may have selected for the observed cassettelike organization of *kps* clusters. To begin testing these ideas, *kps* was mapped and its overall genetic organization was determined. The results suggest that the cluster for K1 synthesis was inherited as a block, presumably by a transpositionlike event.

Probe derivation and mapping strategy. Partial restriction maps of the *kps* cosmids pSR23 (12) and pSX49 (15) are shown in Fig. 1. As documented previously (15), restriction sites in the 17-kb *kps* clusters of these cosmids were similar. A similar distribution of restriction sites also exists within the *kps* clusters of plasmids pSX48 (15) and pKT274 (3). Since pSX49 was isolated from a hybrid strain constructed by conjugation between a K1 donor and an *E. coli* K-12 recipient (17), it was likely that non-*kps* DNA sequences flanking the cluster would be present in acapsular strains of *E. coli*. Hence, *kps* could be mapped by hybridization by using appropriate nucleic acid probes for flanking sequences.

Plasmids pSR41, pSR26, and pSX49 were digested with *Eco*RI to yield probes A to C, respectively (Fig. 1). Probe D was isolated after *Eco*RI digestion of pSX49 and purification of the 7-kb fragment from agarose with Geneclean (Bio 101, La Jolla, Calif.). Similarly, probes E to G were isolated after appropriate digestion of pSX50 (Fig. 1). A more complete description of each plasmid used in this study is given in Vimr et al. (15).

Map position and orientation of *kps*. Kohara et al. (7) recently constructed a physical map of the entire *E. coli* W3110 chromosome by orienting overlapping lambda clones. To map *kps*, 2- to 2.5-μl samples (10^7 to 10^9 PFU/ml) of each lambda clone, together defining the W3110 genome between 60 and 70 map units (Fig. 2A), were spotted with a lawn of permissive *E. coli* K-12 and incubated overnight at 37°C. The lytically propagated phage were adsorbed onto charged nylon membranes and processed for hybridization by a plaque lift method (2). Probe labeling with [α-^{32}P]dCTP, prehybridization, and hybridization were carried out as described in Steenbergen and Vimr (14). Probe C contained the entire *kps* cluster plus flanking sequences. This probe hybridized to lambda clones 22H4, 3D11, and 1G7 (Fig. 2A and B). To confirm that only non-*kps* flanking sequences were being detected, the set of lambda clones shown in Fig. 2A was hybridized with probes E to G. As expected, if W3110 lacked *kps* sequences, these probes did not hybridize with any Kohara clone. Since probe G did not hybridize, this result also indicated that the most distal region 1 sequences in pSX49 were flanked by little or no non-*kps* DNA (Fig. 1). Taken together, these results indicated that the hybridization signals detected with probe C in Fig. 2B were due to non-*kps* sequences flanking region 3 (Fig. 1). In further support of this conclusion, probe A (Fig. 1) hybridized only to Kohara subclone 3D11 (Fig. 2C). An identical result was obtained with probe D (not shown), as predicted from the physical maps of pSR23 and pSX49 (Fig. 1).

The results shown in Fig. 2A to C were most consistent with *kps* being located near 64 map units on the *E. coli* chromosome. To more accurately map *kps*, probe B (Fig. 1)

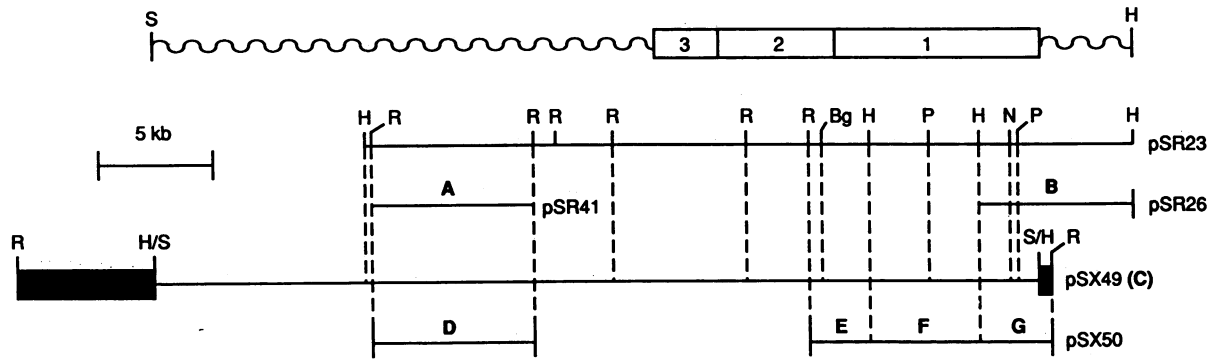

FIG. 1. Cosmid structures and probes used for hybridization mapping of *kps*. Cosmids pSR23 and pSX49 were aligned on the basis of the shared restriction endonuclease sites indicated by vertical dashed lines. The solid box in pSX49 represents cloning vector pHC79. Open, numbered boxes indicate three functional regions of *kps* which are described in the text. Wavy lines designate non-*kps* DNA flanking regions 1 and 3. The seven probes used in this study are designated by bold letters A to G. Restriction sites: Bg, *Bgl*II; H, *Bam*HI; N, *Hin*dIII; P, *Pst*I; R, *Eco*RI; S, *Sau*3A.

was hybridized to a Kohara subset (Fig. 2D) and gave the same result as that obtained with probe A (Fig. 2C). This result demonstrated that one of the two additional Kohara clones (1G7) detected with probe C (Fig. 2B) must represent DNA sequences flanking region 3 that are present on pSX49 but absent from pSR23 (Fig. 1). This conclusion requires the orientation of *kps* shown in Fig. 3. This orientation also was supported by previously locating the nearest chromosomal non-*kps* *Eco*RI site at least 15 kb distal to the region 1 endpoint (14), a result which is consistent with the physical location of *Eco*RI sites in Kohara clones 3D11 and 1G7 (7). The noncontiguous Kohara clone (22H4) detected with probe C (Fig. 2B) may be either the result of a cloning artifact, which presumably would have occurred during construction of pSX49, or truly indicative of a genomic discontinuity between W3110 and the source of pSX49 insert DNA, EV1 (15). The results do not permit distinguishing between these possibilities.

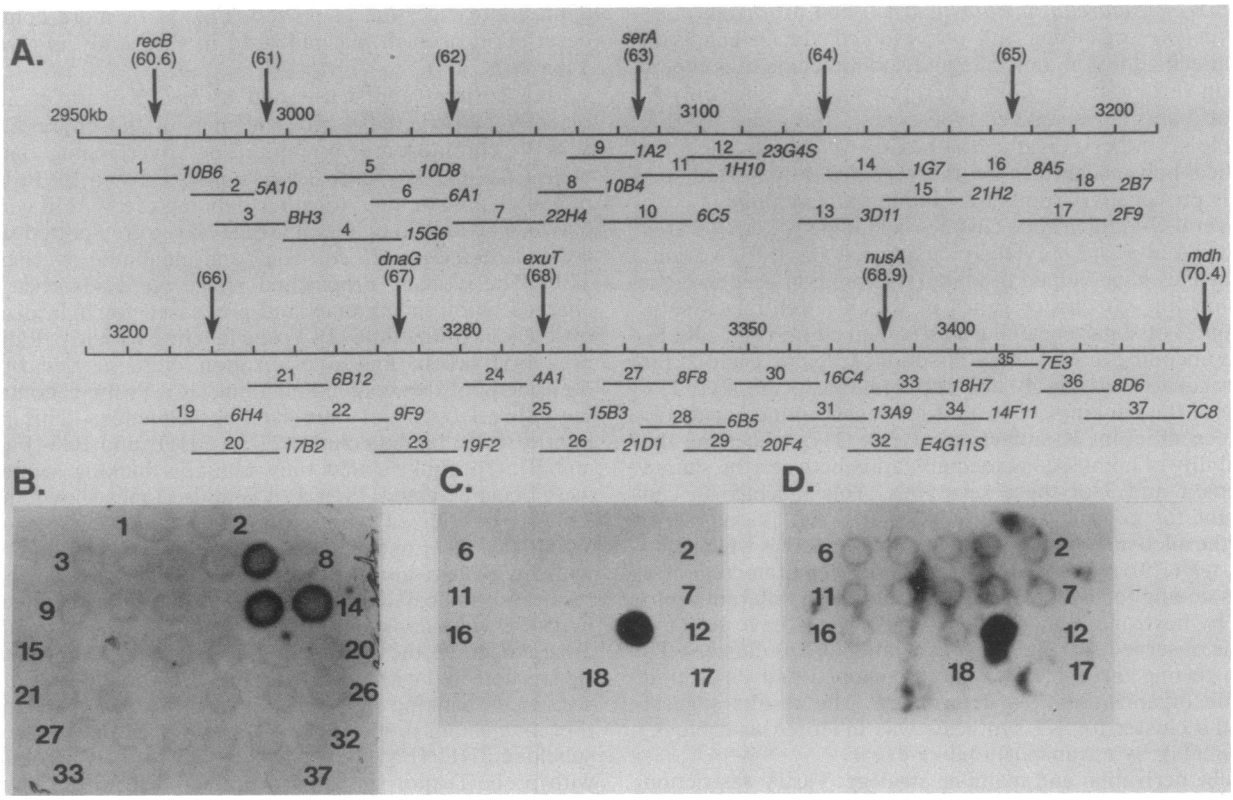

FIG. 2. Physical location of *E. coli* W3110 genomic sequences that flank *kps* in *E. coli* K1 strains. (A) Physical map of *E. coli* W3110 for approximately 10 map units. Distances are in kilobases. Arrows indicate relative positions of genetic markers in *E. coli* K-12 (7). Kohara clones 1 to 37 are identified by the designations to the right of each numbered clone shown below the physical map. (B) Clones 1 to 37 were hybridized with probe C. (C) Clones 2 to 18 were hybridized with probe A. (D) Clones 2 to 18 were hybridized with probe B.

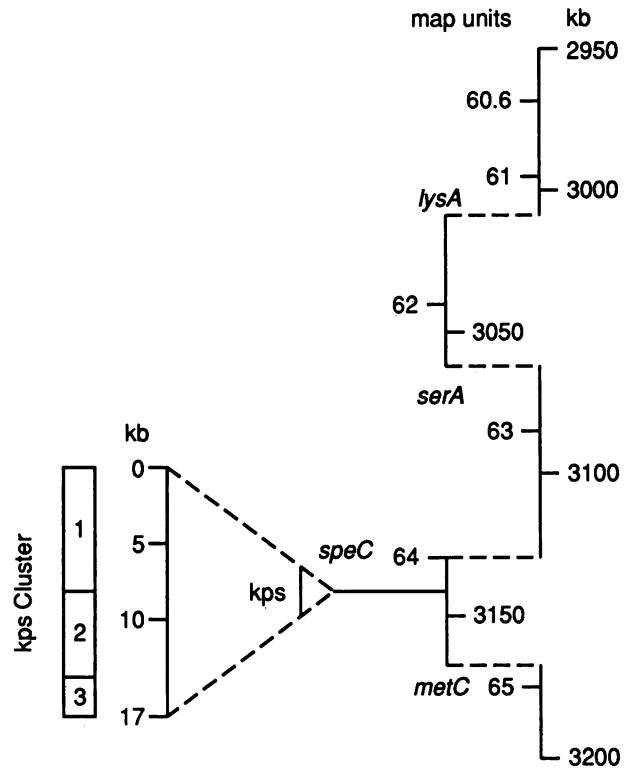

FIG. 3. Orientation of *kps* cluster genes. The physical (kilobases) and genetic (map units) maps of *E. coli* in the vicinity of *kps* are indicated. Horizonal dashed lines indicate accretion domains as defined in the text.

Implications of *kps* acquisition for *E. coli* genome evolution. The map of *E. coli* K-12 summarized in Riley and Krawiec (10) identified two accretion domains near *serA* at 62 and 64 map units (Fig. 3). The *kpsA* locus was placed arbitrarily nearest *serA* at map unit 62 (10). Results presented here more accurately locate *kps* to the accretion domain nearest map unit 64. This conclusion suggests that *E. coli* K-12 and W3110 contain the same accretion domain but that *kps* addition must have occurred after establishment of this domain, since only encapsulated strains contain *kps* sequences. Although the nucleotide sequences flanking *kps* were similar enough to permit detection by cross hybridization (Fig. 2), some sequence divergence may exist. For example, we have been unable to transfer *kps* to *E. coli* K-12 by P1 transduction, even though *kps* represents at most 20% of P1 DNA carrying capacity, and transductions between K-12–K1 hybrid strains are readily carried out (15, 16). These results indicate that *kps* flanking sequences may be sufficiently mismatched so that productive homologous recombination does not occur. This conclusion is consistent with an independent origin of *kps* and the accretion domain in which it is located. In this context, it will be interesting to test whether other *kps* clusters for capsules that are structurally unrelated to polysialic acid, but which share regions 1 and 3 (1, 11), map near or in the site for K1. The capsule gene cluster allelism noted previously is consistent with either possibility (9).

Results presented in Fig. 2C and D show that DNA sequences flanking both *kps* ends are present on the same Kohara clone, 3D11. This conclusion indicates that only *kps* sequences were added to the *E. coli* genome, presumably by a transposition event into the accretion domain shown in Fig. 3. If this mechanism is correct, then *kps* sequences at each cluster endpoint may retain nucleotides reminiscent of an ancestral recombination event. This prediction may be tested directly by sequencing both *kps* ends, one of which is now located at the region 1 endpoint in pSX49 (Fig. 1).

These conclusions further suggest that the polysialic acid biosynthetic functions in species unrelated to *E. coli* may have a common origin, in which case interspecies transfer might be expected (4). As appreciation of horizontal gene transfer as a mechanism for acquisition of novel biochemical traits increases, it may become apparent that the 15% or so of the *E. coli* genome that has apparently evolved by genomic accretion events (10) is a minimal estimate. Thus, while horizontal transfer may not be sufficient to explain all of microbial diversity, it may represent a widespread mechanism for niche expansion that potentially includes transfer of virulence traits such as K1. It follows that the evolution of pathogenic potential may derive largely from the acquisition of virulence determinants provided by well-established mechanisms for horizontal gene transfer. Determining the origins of laterally inherited DNA is likely to provide new insight into genome evolution.

R. P. Silver kindly provided several plasmids used in this study. T. Vanden Boom and J. E. Cronan very generously supplied the lambda bank.

This research was supported by Public Health Services grant AI 23039.

REFERENCES

1. **Boulnois, G. J., and K. Jann.** 1989. Bacterial polysaccharide capsule synthesis, export and evolution of structural diversity. Mol. Microbiol. **3**:1819–1823.
2. **Dillon, J. R., G. S. Bezanson, and K. Yeung.** 1985. Basic techniques, p. 1–126. *In* J. R. Dillon, A. Nasim, and E. R. Nestman (ed.), Recombinant DNA methodology. John Wiley & Sons, Inc., New York.
3. **Echarti, C., B. Hirschel, G. J. Boulnois, J. M. Varley, F. Waldvogel, and K. N. Timmis.** 1983. Cloning and analysis of the K1 capsule biosynthesis genes of *Escherichia coli*: lack of homology with *Neisseria meningitidis* group B DNA sequences. Infect. Immun. **41**:54–60.
4. **Falkow, S.** 1990. The "Zen" of bacterial pathogenicity, p. 1–9. *In* B. H. Iglewski and V. L. Clark (ed.), The bacteria: a treatise on structure and function, vol. XI. Molecular basis of bacterial pathogenesis. Academic Press, Inc., San Diego.
5. **Jann, K., and B. Jann.** 1982. The K antigens of *Escherichia coli*. Prog. Allergy **33**:53–79.
6. **Jann, K., and B. Jann.** 1987. Polysaccharide antigens of *Escherichia coli*. Rev. Infect. Dis. **9**:(Suppl):S517–S526.
7. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
8. **Ørskov, I., F. Ørskov, B. Jann, and K. Jann.** 1977. Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriol. Rev. **41**:667–710.
9. **Ørskov, I., V. Sharma, and F. Ørskov.** 1976. Genetic mapping of the K1 and K4 antigens of K (L) antigens and K antigens of O8:K27(A), O8:K8(L) and O9:K57(B). Acta Pathol. Microbiol. Scand. Sect. B **84**:125–131.
10. **Riley, M., and S. Krawiec.** 1987. Genome organization, p. 967–981. *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.
11. **Roberts, I. S., R. Mountford, R. Hodge, K. B. Jann, and G.**

Boulnois. 1988. Common organization of gene clusters for production of different capsular polysaccharides (K antigens) in *Escherichia coli*. J. Bacteriol. **170**:1305–1310.

12. **Silver, R. P., W. F. Vann, and W. Aaronson.** 1984. Genetic and molecular analyses of *Escherichia coli* K1 antigen genes. J. Bacteriol. **157**:568–575.
13. **Silver, R. P., and E. R. Vimr.** 1990. Polysialic acid capsule of *Escherichia coli* K1, p. 39–60. *In* B. H. Iglewski and V. L. Clark (ed.), The bacteria: a treatise on structure and function, vol. XI. Molecular basis of bacterial pathogenesis. Academic Press, Inc., San Diego.
14. **Steenbergen, S. M., and E. R. Vimr.** 1990. Mechanism of polysialic acid chain elongation in *Escherichia coli* K1. Mol. Microbiol. **4**:603–611.
15. **Vimr, E. R., W. Aaronson, and R. P. Silver.** 1989. Genetic analysis of chromosomal mutations in the polysialic acid gene cluster of *Escherichia coli* K1. J. Bacteriol. **172**:1106–1117.
16. **Vimr, E. R., and S. Steenbergen.** Unpublished data.
17. **Vimr, E. R., and F. A. Troy.** 1985. Regulation of sialic acid metabolism in *Escherichia coli*: role of *N*-acylneuraminate pyruvate-lyase. J. Bacteriol. **164**:854–860.