JOURNAL OF BACTERIOLOGY, Mar. 1992, p. 1709–1710
0021-9193/92/051709-02$02.00/0

Vol. 174, No. 5

Alignment of Genes from the 9-Minute Region (*araJ* to *tsx*) of the *Escherichia coli* K-12 Linkage Map to the Physical Map

JAN M. LUCHT, WINFRIED BOOS, AND ERHARD BREMER*

Department of Biology, University of Konstanz, P.O. Box 5560, D-7750 Konstanz, Germany

A number of genes with diverse functions have been genetically mapped to the 9-min region of the *Escherichia coli* K-12 linkage map (1), and their DNA sequences have recently been established by several laboratories (Table 1). We report here the alignment of the genetic and physical maps of *E. coli* from this region for a 22.5-kb DNA segment (1, 7). With the exception of an approximately 2.5-kb segment, the complete DNA sequence of this region is now known (Fig. 1).

From a compilation of *E. coli* DNA sequences by Kröger et al. (8; electronic update, release 8), a contiguous DNA segment of 9,668 bp comprising the *araJ*, *sbcC*, and *phoBR* genes and an open reading frame (*orf45*) are known (Fig. 1 and Table 1). The restriction sites predicted from the DNA sequence of this segment closely match the pattern of sites established by Kohara et al. (7) for the DNA region of the *E. coli* physical map between coordinates 421.6 kb and 431.3 kb (Fig. 1). This contiguous stretch of DNA corresponds to the *E. coli* genetic map from 8.93 to 9.14 min (Table 1 and Fig. 1).

We have assembled from published DNA sequences (3, 5, 8, 12, 13) a contiguous DNA segment of 10,364 bp which extends from the *malZ* gene to *tsx* (Table 1 and Fig. 1). The DNA sequence established by Tapio et al. (13) for *malZ* overlaps 376 bp of the 3,545-bp DNA segment sequenced by Reuter et al. (12). This overlapping sequence carries the 3′ end of *malZ* and part of a convergently transcribed open reading frame (*orf14*) (Fig. 1). The stop codons for the *malZ* genes and for *orf14* are separated by only 4 bp (12, 13). Divergently transcribed from *orf14* are the *queA*, *tgt*, *orf12*, *secDF*, and *orf6* genes (Fig. 1). This cluster of genes with known functions (Table 1) and the two open reading frames might constitute a operon (5, 12), although individual transcriptional initiation sites for several of these genes have been detected (12). Polypeptides which might correspond to the gene products of *orf14*, *orf12*, and *orf6* have been found by maxicell analysis and by in vitro translation of cloned DNA fragments (5, 12). Hence, these open reading frames most likely constitute real genes.

The *secD* and *tsx* genes are known to be genetically and physically tightly linked (1, 4). The DNA sequence reported by Gardel et al. (5) for the *secD*, *secF*, and *orf6* genes (see Fig. 1) ends with a *Stu*I restriction site. Such a site is also present at the 3′ end of the *tsx* sequence established by Bremer et al. (3). We sequenced across the *Stu*I site present behind the *tsx* gene and found that both sequences are

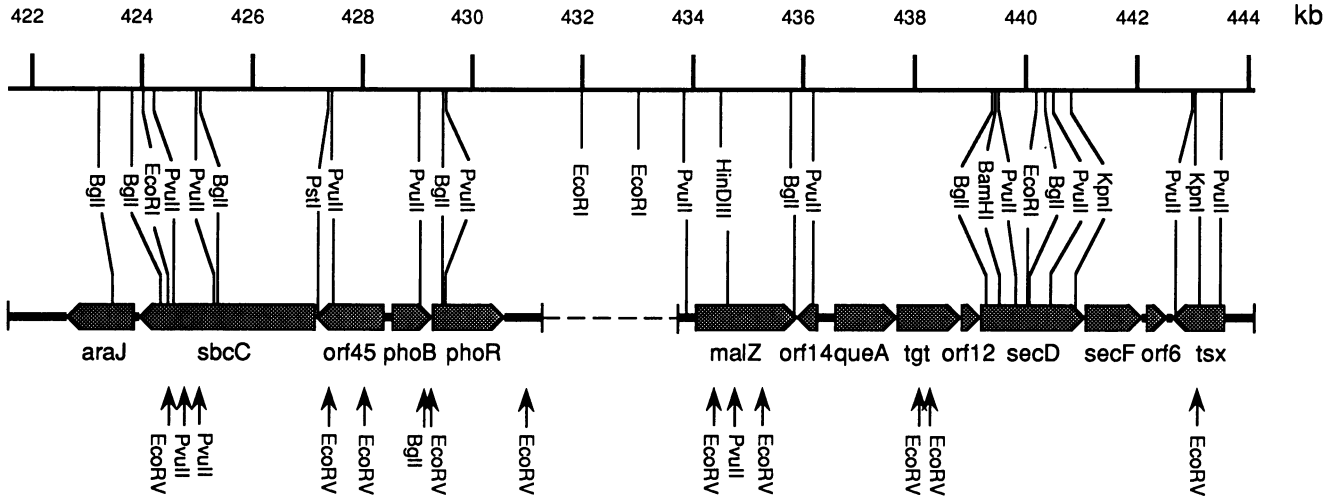

FIG. 1. Physical and genetic map of the 9-min region (422 to 444 kb) of the *E. coli* chromosome. A portion of the *E. coli* chromosomal restriction map (7), with map coordinates in kilobase pairs and positions of restriction sites, is shown in the upper line. The lower line, drawn to the same scale, shows the aligned genetic map of this region. Segments of sequenced DNA are represented by thick horizontal lines; a thin dashed line represents an unsequenced DNA region. The positions and directions of transcription of genes with known functions and of several open reading frames (*orf*) are indicated by pointed grey boxes. Positions of restriction sites predicted from the assembled DNA sequences and which are also present in the Kohara (7) restriction map are indicated between the upper and lower line. Below the genetic map, several sites not present in the Kohara map are indicated by arrows. The *Eco*RV restriction map from this segment of the *E. coli* chromosome is known to be incomplete (7).

* Corresponding author.

TABLE 1. Physical and genetic locations of genes in the *araJ*-to-*tsx* region of the *E. coli* K-12 chromosome

Gene	Physical map location (kb)[a]	Genetic map location (min)[b]	Gene function	Reference
araJ	422.7–423.9	8.96–8.98	Arabinose-inducible gene	6
sbcC	424.0–427.2	8.98–9.05	Genetic recombination	11
orf45	427.2–428.4	9.05–9.07	Unknown	11
phoB	428.6–429.3	9.08–9.10	Regulatory gene	9
phoR	429.3–430.6	9.10–9.12	Regulatory gene	10
malZ	434.1–435.9	9.20–9.24	Maltodextrin glucosidase	13
orf14	435.9–436.3	9.24	Unknown	12
queA	436.6–437.7	9.25–9.27	Queuine biosynthesis	12
tgt	437.7–438.9	9.27–9.30	tRNA guanine transglycosylase	12
orf12	438.9–439.2	9.30	Unknown	12
secD	439.2–441.1	9.30–9.35	Protein secretion	5
secF	441.1–442.1	9.35–9.37	Protein secretion	5
orf6	442.2–442.5	9.37–9.38	Unknown	5
tsx	427.7–443.6	9.38–9.40	Nucleoside-specific channel	3

[a] The physical map locations refer to the Kohara map (7). The values given may need to be slightly adjusted in the future when more DNA sequences are aligned with the physical map.

[b] Positions in minutes are derived from the physical map position in kilobase pairs by dividing by 47.2 (8).

contiguous (2; Fig. 1). By aligning the *E. coli* physical map (7) with the restriction sites predicted from the DNA sequences of the continuous 10,364-bp segment from *malZ* to *tsx*, we were able to position this DNA segment between 433.7 and 444.1 kb on the Kohara map. This region corresponds to 9.18 to 9.41 min on the genetic map (1) (Fig. 1). The *araJ*-to-*phoR* segment and the *malZ*-to-*tsx* segment are separated by a unsequenced DNA segment of approximately 2.5 kb (Fig. 1). Restriction maps of this region have been reported (11, 13, 14), and these data are in agreement with the Kohara map.

Financial support for this work was provided by grants from the Deutsche Forschungsgemeinschaft through SFB-156 and by the Fond der Chemischen Industrie. J.M.L. gratefully acknowledges the receipt of a fellowship through the Graduiertenförderung des Landes Baden-Württemberg.

We thank V. Koogle for her help in preparing the manuscript.

REFERENCES

1. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54:**130–197.
2. **Bremer, E., and A. Middendorf.** Unpublished results.
3. **Bremer, E., A. Middendorf, J. Martinussen, and P. Valentin-Hansen.** 1990. Analysis of the *tsx* gene, which encodes a nucleoside-specific channel-forming protein (Tsx) in the outer membrane of *Escherichia coli*. Gene **96:**59–65.
4. **Gardel, C., S. Benson, J. Hunt, S. Michaelis, and J. Beckwith.** 1987. *secD*, a new gene involved in protein export in *Escherichia coli*. J. Bacteriol. **169:**1286–1290.
5. **Gardel, C., K. Johnson, A. Jacq, and J. Beckwith.** 1990. The *secD* locus of *E. coli* codes for two membrane proteins required for protein export. EMBO J. **9:**3209–3216.
6. **Hendrickson, W., C. Stoner, and R. Schleif.** 1990. Characterization of *Escherichia coli araFGH* and *araJ* promoters. J. Mol. Biol. **215:**497–510.
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. **Kröger, M., R. Wahl, and P. Rice.** 1991. Compilation of DNA sequences of *Escherichia coli* (update 1991). Nucleic Acids Res. **19:**2023–2043.
9. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulon of *Escherichia coli* K-12. J. Mol. Biol. **190:**37–44.
10. **Makino, K., H. Shinagawa, M. Amemura, and A. Nakata.** 1986. Nucleotide sequence of the *phoR* gene, a regulatory gene for the phosphate regulon of *Escherichia coli*. J. Mol. Biol. **192:**549–556.
11. **Naom, I. S., S. J. Morton, D. R. F. Leach, and R. G. Lloyd.** 1989. Molecular organization of *sbcC*, a gene that affects genetic recombination and the viability of DNA palindromes in *Escherichia coli* K-12. Nucleic Acids Res. **20:**8033–8045.
12. **Reuter, K., R. Slany, F. Ullrich, and H. Kersten.** 1991. Structure and organization of *Escherichia coli* genes involved in biosynthesis of the deazaguanine derivative queuine, a nutrient factor for eukaryotes. J. Bacteriol. **173:**2256–2264.
13. **Tapio, S., F. Yeh, H. A. Shuman, and W. Boos.** 1991. The *malZ* gene of *Escherichia coli*, a member of the maltose regulon, encodes a maltodextrin glucosidase. J. Biol. Chem. **266:**19450–19458.
14. **Tommassen, J., P. de Geus, B. Lugtenberg, J. Hacket, and P. Reeves.** 1982. Regulation of the *pho* regulon of *Escherichia coli* K-12. Cloning of the regulatory genes *phoB* and *phoR* and identification of their gene products. J. Mol. Biol. **157:**265–274.