Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9454–9458, December 1990
Biochemistry

Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: A potential transposable element encoding both reverse transcriptase and Dam methylase functions

MEI-YIN HSU, MASAYORI INOUYE, AND SUMIKO INOUYE

Department of Biochemistry, Robert Wood Johnson Medical School at Rutgers, University of Medicine and Dentistry of New Jersey, Piscataway, NJ 08854

Communicated by Howard M. Temin, September 10, 1990 (received for review June 14, 1990)

ABSTRACT **The region (retron-Ec67) required for the biosynthesis of a branched-RNA-linked multicopy single-stranded DNA (msDNA-Ec67) from a clinical isolate of *Escherichia coli* was mapped at a position equivalent to 19 min on the K-12 chromosome. The element containing the retron consisted of a unique 34-kilobase sequence that was flanked by direct repeats of a 26-base-pair sequence found in the K-12 chromosomal DNA. This suggests that the 34-kilobase element was probably integrated into the *E. coli* genome by a mechanism related to transposition or phage integration. In the 34-kilobase sequence an open reading frame of 285 residues was found, which displays 44% sequence identity with the *E. coli* Dam methylase. Interestingly, there are three GATC sequences, the site of Dam methylation, in the promoter region of the gene for reverse transcriptase.**

msDNA-Ec67 is a 67-base multicopy single-stranded DNA from *Escherichia coli* that is linked at its 5′ end to the 2′ OH group of the 15th guanosine residue of a 58-base RNA molecule by a 2′,5′-phosphodiester linkage (1). This peculiar RNA·DNA complex was found in a clinical isolate of *E. coli*, strain Cl-1, and reverse transcriptase (RT) has been shown to be required for msDNA synthesis (for a review, see ref. 2). It has been proposed that a long RNA transcript forms a unique secondary structure, which then serves as primer as well as template for msDNA synthesis by RT (3, 4). The gene for RT has been identified and shown to be closely associated with the region encoding both DNA and RNA molecules of msDNA (1, 5, 6). On the basis of sequence similarities of bacterial RTs to retroviral RTs, their evolutionary homology has been proposed (5). It is possible that the msDNA-synthesizing system may be evolutionarily related to retrotransposons and retroviruses, and Temin (7) proposed designating the system a "retron."

In this report, we first attempted to isolate and sequence DNA fragments* from the chromosome of *E. coli* strain Cl-1 that contain junction sites between the element containing the retron-Ec67 (retron responsible for the synthesis of msDNA-Ec67) and the *E. coli* genome. We found that a 34-kilobase (kb) foreign DNA fragment containing retron-Ec67 was integrated at a site equivalent to 19 min on the *E. coli* K-12 chromosome. This fragment did not hybridize with the K-12 chromosomal DNA and was found to be flanked by direct repeats of a 26-base-pair (bp) sequence found in the K-12 chromosomal DNA. This result raises a possibility that the 34-kb fragment was integrated into the *E. coli* genome by mechanism related to transposition or phage integration. A possibility that the 34-kb element might have been transposed as a retrotransposon will be discussed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Bacterial Strains and Plasmid. pUC9 (8) was used to clone chromosomal DNA and to subclone various DNA fragments. *E. coli* JM83 (8) was used for transformation, and *E. coli* GM33 *dam*$^-$ (9) was used to assay Dam methylase activity. Those *E. coli* cells harboring plasmids were grown in L broth (10) containing ampicillin (50 μg/ml). The λ library of the *E. coli* K-12 genome (11) was obtained from A. Ishihama (National Institute of Genetics, Mishima, Shijuoka-ken, Japan). The λ phage DNA filters were prepared by spotting 0.8 μl of λ phage lysates on a nitrocellulose filter paper according to Kohara *et al.* (11).

DNA Sequencing and Chromosome Walking. DNA manipulations were performed as described by Maniatis *et al.* (12). To determine DNA sequence, large DNA fragments from chromosomal DNA were digested with various restriction enzymes and subcloned. DNA sequences were determined by the chain-termination method (13) using double-stranded plasmids and synthetic oligonucleotides as primers. To find the junction sites, DNA walking using the pCl-1E DNA as a probe was performed. Two neighboring DNA fragments from restriction enzyme digests of the Cl-1 chromosomal DNA were identified by Southern blot hybridization (14) with use of the left- and righthand end fragments from pCl-1E as probes. DNA walking was continued until DNA fragments from strain Cl-1 hybridized with the K-12 DNA.

RESULTS

Isolation of Junction Fragments. The circular gene map of the entire genome of *E. coli* K-12 (15) and the restriction map of its chromosomal DNA (11) have been established. Furthermore, a λ genomic library encompassing the entire genome of *E. coli* K-12 is also available (11). Therefore, we used *E. coli* K-12 as a reference strain to determine the relative map position of retron-Ec67 on the Cl-1 chromosome and the size of the DNA fragment containing retron-Ec67. Since the original clone of retron-Ec67, pCl-1E containing the 11.6-kb *Eco*RI fragment [E(a)–E(b) in Fig. 1*A*] from strain Cl-1 (1) did not hybridize with the K-12 DNA, DNA walking to the left on the Cl-1 chromosomal DNA was first carried out until a DNA fragment hybridizable with the K-12 DNA was identified. By using the lefthandmost 1-kb *Eco*RI [E(a)]–*Sal* I [S] fragment, a 6-kb *Pst* I fragment [P(a)–P(b)] was cloned from the Cl-1 DNA (see Fig. 1*A*). This *Pst* I fragment was found to hybridize with a 4.3-kb *Eco*RI fragment of the K-12 DNA by Southern blot hybridization (data not shown). The *Pst* I fragment was then used to identify λ clones from the K-12 DNA library (11). Out of 476 λ clones, two λ phage DNAs clearly hybridized with the probe. These two λ phages, 3H12

Abbreviations: RT, reverse transcriptase; ORF, open reading frame; msDNA, multicopy single-stranded DNA.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M55249).

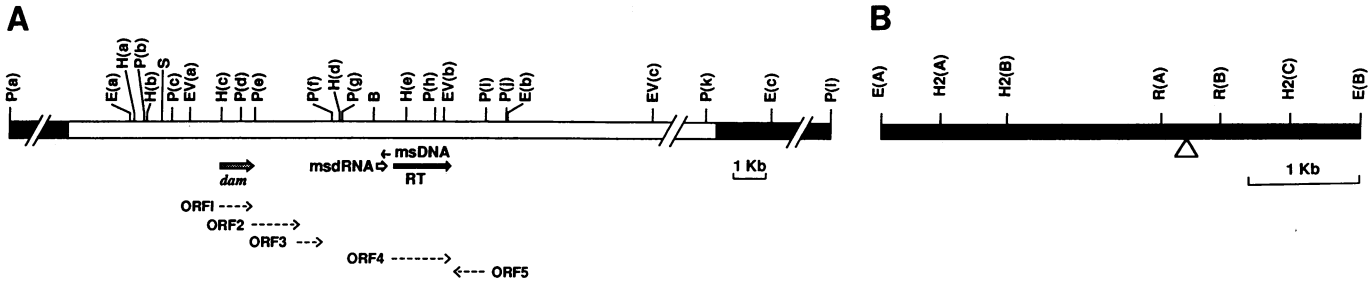

FIG. 1. Restriction map of the 34-kb DNA containing retron-Ec67 (*A*) and the 4.3-kb *Eco*RI fragment from *E. coli* K-12 containing the integration site for retron-Ec67 (*B*). (*A*) The location and the orientation of msDNA and RNA linked to msDNA (msdRNA) are indicated by a small arrow and an open arrow, respectively. ORFs (see Fig. 3) are indicated by dashed arrows. A large solid arrow and a stippled arrow represent the genes for RT and the Dam methylase, respectively, and their orientation. An open bar represents the 34-kb retron Ec67, and solid bars represent DNA sequences that are common to K-12 DNA. A detailed restriction map is shown only between E(a) and E(b). A substantial part (≈15 kb) of the retron is omitted from the map. B, *Bal* I; E, *Eco*RI; EV, *Eco*RV; H, *Hin*dIII; P, *Pst* I; S, *Sal* I. DNA sequence from E(a) to B and junction regions between retron-Ec67 and the *E. coli* chromosomal DNA were determined. DNA sequence from B to E(b) was reported by Lampson *et al.* (1). (*B*) The location of the integration site of retron-Ec67 is shown by an open triangle. The DNA fragment was isolated from a λ phage containing a DNA fragment from 19 min on the *E. coli* K-12 chromosome. E, *Eco*RI; H2, *Hin*cII; R, *Rsa* I. Only two *Rsa* I sites used to determine the integration site are shown.

and 5F4, have been shown to contain DNA fragments from 19 min on the K-12 chromosome (11), indicating that retron-Ec67 is integrated into the Cl-1 genome at the position equivalent to 19 min on the K-12 chromosome. From the λ DNA identified above, a 4.3-kb *Eco*RI fragment was cloned, which hybridized with the 6-kb *Pst* I [P(a)–P(b); see Fig. 1*A*] fragment of the Cl-1 DNA.

This *Eco*RI fragment is expected to contain the retron-Ec67 integration site and was mapped as shown in Fig. 1*B*. It was found that the 0.55-kb E(A)–H2(A), the 0.60-kb H2(A)–H2(B), and the 2.55-kb H2(B)–H2(C) fragments were able to hybridize with the 6-kb P(a)–P(b) fragment (Fig. 1*A*) from the Cl-1 DNA, while the 0.6-kb H2(C)–E(B) did not (data not shown). By cloning the 2.55-kb *Hin*cII [H2(B)–H2(C)] fragment, further mapping of this fragment with *Rsa* I was carried out, and it was found that the righthand end 0.6-kb R(B)–H2(C) fragment could not hybridize with the 6-kb P(a)–P(b) fragment, but the 0.55-kb R(A)–R(B) fragment did (data not shown). DNA sequencing at the junction regions revealed that the 0.55-kb R(A)–R(B) fragment was directly connected to the 0.6-kb R(B)–H2(C) fragment as shown in Fig. 1*B*. These results indicate that the integration site of retron-Ec67 was within the R(A)–R(B) fragment of the K-12 DNA.

To isolate a DNA fragment that contains the junction between the righthand end of retron-Ec67 and the host chromosomal DNA, the 1.9-kb EV(b)–E(b) fragment (Fig. 1*A*) was used as a probe. As a result a 6.6-kb *Eco*RV fragment [EV(b)–EV(c) in Fig. 1*A*] was cloned. This *Eco*RV fragment hybridized to the 11.6-kb *Eco*RI fragment from E(a) to E(b) as well as to another 22-kb *Eco*RI fragment. We were unsuccessful in cloning this 22-kb *Eco*RI fragment from the Cl-1 chromosomal DNA. Therefore, to isolate a junction fragment from the Cl-1 DNA containing the righthand junction region, the 0.6-kb R(B)–H2(C) fragment from the K-12 DNA (Fig. 1*B*) was used as probe. Subsequently, a 2.1-kb *Pst* I–*Eco*RI fragment [P(k)–E(c) in Fig. 1*A*] was cloned. This 2.1-kb *Pst* I–*Eco*RI fragment hybridized to the 22-kb *Eco*RI fragment in Southern blot hybridization as did the 6.6-kb *Eco*RV fragment (data not shown). Since the 6.6-kb *Eco*RV fragment contained only one *Eco*RI site, it was concluded that the 22-kb *Eco*RI fragment was the adjacent fragment [E(b)–E(c)] to the 11.6-kb *Eco*RI fragment [E(a)–E(b)]. Thus the restriction map around the retron-Ec67 integration site within the 39-kb P(a)–E(c) region was determined as shown in Fig. 1*A*.

From the results described above, we conclude that the 10.8-kb *Pst* I fragment from the Cl-1 DNA [P(k)–P(1)] contained the junction site between the righthand end of retron-Ec67 and the host chromosomal DNA. To determine the junction site, the *Pst* I fragment was digested with *Eco*RI, which generated two fragments, the 2.1-kb P(k)–E(c) fragment and the remaining 8.7-kb fragment (see Fig. 1*A*). The P(k)–E(c) fragment was found to hybridize to the 0.6-kb R(B)–H2(C) fragment from the K-12 DNA (Fig. 1*B*), indicating that the junction site exists in the P(k)–E(c) fragment. It was found that the P(k)–E(c) fragment hybridized to the same λ clones that hybridized to the P(a)–P(b) probe (see Fig. 1*A*). This indicates that the host chromosomal sequences at the left- and the righthand junctions are located in a close proximity in the K-12 chromosomal DNA.

DNA Sequences of the Junction Sites. To determine the junction sequences at the left and right ends of retron-Ec67, we first determined the DNA sequence of the 0.55-kb R(A)–R(B) fragment from the K-12 chromosomal DNA (Fig. 1*B*), which is considered to contain the retron-Ec67 integration site. Subsequently, the junction sequences from the Cl-1 DNA were searched using the DNA sequence of the R(A)–R(B) fragment determined above as a reference. At the lefthand end, ≈2.3 kb leftward from the P(b) site (Fig. 1*A*), the sequence found in the K-12 fragment appeared as shown in Fig. 2. For the righthand end, the K-12 sequence also appeared ≈0.3 kb rightward from the P(k) site (Fig. 1*A*) (Fig. 2). The same 26-bp sequence found in the K-12 DNA (boxed in Fig. 2) is repeated at both ends (Cl-1L and Cl-1R) of the

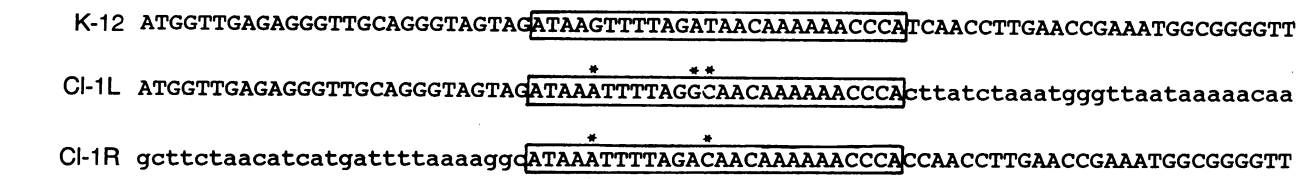

FIG. 2. DNA sequence of the integration site of retron-Ec67 in K-12 DNA and regions from strain Cl-1. DNA sequences from K-12 or Cl-1 are shown with uppercase letters and the retron-Ec67 are shown with lowercase letters. Cl-1L and Cl-1R represent left and right junction regions in Fig. 1*A*, respectively. Twenty-six bases shared among the three DNA sequences are boxed. Asterisks represent base substitutions between the sequences.

FIG. 3. DNA sequence of the 8.2-kb E(a)–H(e) fragment from retron-Ec67. The amino acid sequences of ORF1, -2, and -3 and the sequence of the amino-terminal region of ORF4 are shown by single-letter code. The DNA sequence from residue 7422 to residue 8221 has been determined (1). The restriction enzyme sites used to detect Dam methylase activity are shown. P, *Pst* I site at P(g); H, *Hin*dIII site at H(e) in Fig. 1*A*. The three *Cla* I sites between P(g) and H(e) are also indicated by C. Three GATC sequences upstream of *msr* are boxed. The sequences of RNA linked to msDNA and msDNA are boxed, and their orientations are indicated by open arrows. The branched G residue at position 7684 is circled. The inverted repeat sequences required for biosynthesis of msDNA-Ec67 are shown by arrows with a1 and a2 (1).

retron DNA, except that there were three and two mismatches between K-12 and Cl-1L and between K-12 and Cl-1R, respectively (indicated by asterisks). It should be noted that there is only one mismatch between Cl-1L and Cl-1R. The DNA sequence of the 292-bp sequence upstream of Cl-1L was determined and was found to have only four single-base mismatches when compared with the corresponding region of the K-12 DNA. Similarly there were two mismatches in the 220-bp sequence downstream of Cl-1R. On the contrary, there was little homology between the downstream sequence of Cl-1L and the corresponding region of the K-12 DNA and between the upstream sequence of Cl-1R and the corresponding sequence of the K-12 DNA. These results clearly demonstrate that a 34-kb foreign DNA fragment

(retron Ec67) was integrated into the Cl-1 genome and is flanked by a duplicated 26-bp host sequence.

Dam Methylase Linked to Retron-Ec67. The DNA sequence of 7.4 kb from E(a) to B (see Fig. 1*A*) was determined. The DNA sequence 4.2 kb downstream of site B, which contains *msr*, *msd*, and the RT gene, has been determined. Fig. 3 shows the 8221-kb DNA sequence from residue 1 [E(a)] to residue 8221 [H(e)] of which the sequence from residue 7422 (B site) to 8221 was determined (1). In the region from E(a) to E(b), there are five large open reading frames (ORFs; >200 amino acid residues) including the ORF for RT. All of the ORFs are in the same orientation except for ORF5, with the first (ORF1) from residue 2821 to residue 3675, the second (ORF2) from residue 3734 to residue 5134, the third (ORF3) from residue 5017 to 5625, the fourth (ORF4; RT-coding ORF) from residue 7839 to residue 9596 (1), and the fifth (ORF5) from residue 10,667 to residue 9690 (data not shown; available from GenBank) (see Fig. 1*A*). Although the amino acid sequences of ORF2, ORF3, and ORF5 show no significant similarities to any known protein sequences, ORF1 shows significantly high sequence similarity to *E. coli* DNA adenine methylase (Dam methylase; ref. 16), as shown in Fig. 4. Out of 285 residues, there are 125 identical residues (44% identity) between the two Dam methylases.

To examine if ORF1 is able to produce a functional Dam methylase, we transformed Dam$^-$ *E. coli* strain GM33 with pCl-1E harboring the 11.6-kb E(a)–E(b) fragment, which contains all the ORFs described above (see Fig. 1*A*). After overnight culture, the plasmid DNA was isolated and digested with *Pst* I and *Hin*dIII. Subsequently, the 1.8-kb P(g)–H(e) fragment (Fig. 1*A*) was purified and examined by digestion with *Cla* I. As a control, strain GM33 was also transformed with pCl-1EP5 containing the 5-kb P(g)–E(b) fragment, which does not contain ORF1 (see Fig. 1*A*). The 1.8-kb P(g)–H(e) fragment was also isolated from PCl-1EP5 and digested with *Cla* I. If a guanosine residue preceded a *Cla* I site (ATCGAT), a Dam methylase recognition site (GATC) is created. All three *Cla* I sites within the 1.8-kb fragment do form the Dam methylation sites. Therefore, in the presence of Dam methylase, all the *Cla* I sites are methylated to become resistant to *Cla* I digestion. As shown in Fig. 5, lane 6, the fragment from pCl-1EP5 was completely digested by *Cla* I, yielding four bands at 0.73, 0.69, 0.34, and 0.12 kb, as expected. On the contrary, the 1.8-kb fragment from pCl-1E was hardly digested with *Cla* I (lane 4). It should be noted that a few faint bands under the undigested band are observed. The band densities did not change after prolonged digestion or digestion with more enzyme, indicating that the *Cla* I sites in the fragment were partially unmethylated. The Dam methylase of retron-Ec67 may be less active than the *E. coli* chromosomal Dam methylase.

There are a total of nine additional ORFs of >100 amino acid residues in the regions where no ORFs are assigned in Fig. 3, ORFa (bases 616–939; 108 amino acid residues), ORFb (bases 975–1481; 169 residues), ORFc (bases 1656–1994; 113 residues), ORFd (bases 2522–2821; 100 residues), ORFe (bases 5625–6086; 154 residues), ORFf (bases 6369–6671; 101 residues), ORFg (bases 7203–6889; 105 residues), ORFh (bases 1110–799; 104 residues), and ORFi (bases 625–32; 198 residues). There is one overlapping ORF (ORFj) with ORF1 (bases 3164–3493; 110 residues). Interestingly, when these ORFs were analyzed by the Genetics Computer Group Sequence Program (26) and the TFastA Program (27), it was found that ORFi has 30% identity to the *c*I gene product and ORFb 70% to the *cp76* gene product of bacteriophage 186 (28, 29). In addition, these two genes are situated in a similar manner as found in the DNA sequence shown in Fig. 3. This finding raises an intriguing possibility that the 33-kb foreign DNA fragment might be a prophage related to phage 186, a member of the temperate P2 family of *E. coli* phages.

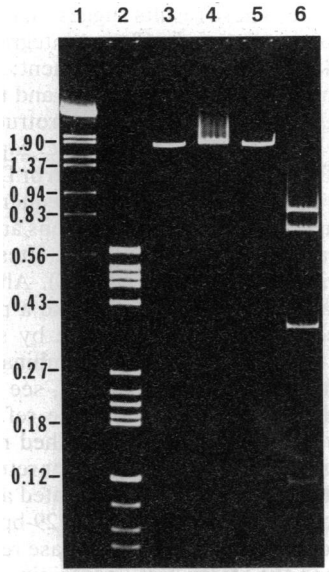

FIG. 5. *Cla* I digestion patterns of the 1.8-kb P(g)–H(e) fragment isolated from pCl-1E and pCl-1EP5. Plasmids were isolated from GM33 (dam$^-$) (9) and digested by *Pst* I and *Hin*dIII. The 1.8-kb P(g)–H(e) fragments were isolated by polyacrylamide gel electrophoresis and subjected to *Cla* I digestion followed by polyacrylamide gel electrophoresis. The bands were visualized by staining with ethidium bromide. Lanes: 1, 0.3 μg of λ DNA digested with *Hin*dIII and *Eco*RI; 2, pBR322 digested with *Hae*III as molecular weight markers; 3 and 4, 1.8-kb fragment isolated from pCl-1E; 5 and 6, the 1.8-kb fragment isolated from pCl-1EP5; 3 and 5, without *Cla* I digestion; 4 and 6, with *Cla* I digestion. Numbers to the left are sizes of DNA in kb.

DISCUSSION

From the hybridization analysis between the *E. coli* K-12 DNA and strain Cl-1 DNA, the junction sites between the DNA fragment containing retron-Ec67 and the host chromosomal DNA were identified. The size of this fragment was determined to be 34 kb and to have little homology to K-12 DNA. The 34-kb fragment was flanked by 26-bp direct repeats, the sequence of which was found at 19 min on the

```
Ec   Dam   MKKNRAFLKWAGGKTPLLDDIKRHLPKGECLVEPFVGAGSVFLNTDFSRYILADINSDLISLYNIVKMRTDEYVQAAR   78
Ec67 Dam   MSTI---LKWAGNKTAIMSELKKHLPAGPRLVEPFAGSCAVMMATDYPSYLVADINPDLINLYKKIAADCEAFISRAR   75
           •      ●●●●● ●● oo oo●●●●● ● ●●●●● ●  ● o ●●o  ●oo●●●●● ●●● ●● o   o oo  ●●

Ec   Dam   ELVFPETNCAEVYYQFREEFNKSQDP--FRRAVLFLYLNRYGYNGLCRYNLRGEFNVPFGRYKKPYFPEAELYHFAEK  154
Ec67 Dam   VL-FEIANREVAYYNIRQEFNYSTEITDFMKAVYFLYLNRHGYRGLCRYNKSGHFNIPYGNYKNPYFPEKEIRKFAEK  152
            • •  •    ●●o ● ●●● ● o   ● o●● ●●●●●● ●● ●●●●●●  ● ●●o●o● ●● ●●●●● ●o  ●●●●

Ec   Dam   AQNAFFYCESYADSMARADDASVVYCDPPYAPLSATANFTAYHTNSFTLEQQAHLAEIAEGLV-ERHIPVLISNHDTM  231
Ec67 Dam   AQRATFICASFDETLAMLQVGDVVYCDPPYDG---T--FSGYHTDGFTEDDQYHLASVLEYRSSEGH-PVIVSNSDTS  224
           ●● ● ● ● ●o oooo    o ●●●●●●●●      ●  ●oo●●●  ●● o ● ●●● o ●    ● ● ●●oo●● ●●

Ec   Dam   LTREWYQRAKLHVVKVRRSISSNGGTRKKVDELLALYKPGVVSPAKK                                  278
Ec67 Dam   LIRSLYRNFTHHYIKAKRSIGVSAGESKSATEIIAVSGARCWVGFDPSRGVDSSAVYEVRV                    285
           ● ● ●     ● o● o●●●   o●  ●   ●ooo●o
```

FIG. 4. Amino acid sequence alignment of the putative retron Ec-67 Dam methylase with *E. coli* dam methylase. Amino acid sequences are aligned using IBI Pustell sequence analysis programs. ● and ○, Identical and functionally homologous amino acids, respectively.

K-12 chromosome. These results suggest that the 34-kb DNA fragment containing retron-Ec67 was integrated into the *E. coli* genome like a transposon-like element; a staggered cut was made at the 26-bp target sequence, and the retron-Ec67 containing DNA was then joined at the protruding ends of the chromosomal DNA. Subsequently, the gap at the target sequence was filled in by a repair reaction or DNA replication resulting in the duplication of the 26-bp sequence at both ends of the DNA fragment. Various transposons are considered to be integrated at their individual target sites, as described above (for reviews, see refs. 17 and 18). Alternatively, the 34-kb DNA fragment was integrated like a prophage at the 26-bp sequence on the *E. coli* genome by site-specific recombination. Such site-specific recombination has been shown for bacteriophages (for a review, see ref. 19) as well as for some plasmids (for example, see ref. 20). We have found (J. Sun, M.I., and S.I., unpublished results) another clinical *E. coli* strain that contains another retron. In this case the retron-containing fragment is integrated at 82 min on the *E. coli* chromosome and is flanked by 29-bp direct repeats that did not show any homology to 26-base repeats in Fig. 3.

How and when the retron-Ec67 containing DNA fragment was integrated into the genome of strain Cl-1 remains unanswered. However, comparative analysis of various enzymes of strain Cl-1 with other well-established *E. coli* strains [*E. coli* Reference Collections (ECOR); ref. 21] revealed that strain Cl-1 is closely related to ECOR35 strain among other 72 strains (22). Importantly, we found that ECOR35 also contained msDNA, which is highly similar to msDNA-Ec67. Furthermore, the chromosomal DNA isolated from ECOR35 contained a DNA fragment that hybridized with the RT gene from retron-Ec67. In the phylogenetic tree established for the ECOR collection, ECOR35 and ECOR36 are most related. Surprisingly, ECOR36 neither contained msDNA nor contained the RT gene. These results suggest that retron-Ec67 was recently acquired by the genome of strain Cl-1.

A most intriguing question is whether the RT function is required for transposition of the retron-Ec67 containing fragment. If the DNA fragment is a retrotransposon, the entire 34-kb DNA sequence had to be first transcribed. The resulting RNA transcript then had to be converted to cDNA by RT. The single-stranded cDNA was also converted to double-strand DNA by RT and is then assumed to have integrated into an *E. coli* genome. Although we have demonstrated that RT from retron-Ec67 is able to synthesize cDNA as well as double-stranded DNA *in vitro* by using synthetic primers and templates (4), it remains to be seen if retron-Ec67 is able to synthesize cDNA *in vivo*. It is interesting to point out that a complementary sequence to the 3′ end of msDNA-Ec67, 5′---TccTGCG-3′, is found at the junctions between the 34-kb DNA fragment and the righthand direct repeat sequence, Cl-1R (see Fig. 2). In this sequence, 5′-GGCAtaA-3′, the first three bases (GGC) are from retron Ec67 and the remaining four bases are from the direct repeat (AtaA), where lowercase letters represent mismatches. It is possible that the 3′ end of msDNA may serve as a primer for DNA synthesis from the very end of retron Ec67, which may play a role in transposition.

The present finding of the *dam* gene in the element containing retron-Ec67 raises the intriguing possibility that the retron Dam methylase may play a regulatory role in the functions of the retron. Dam methylases have been shown to play positive and negative roles in expression of several important genes in *E. coli*, which contain GATC sequences in their promoters (for reviews, see refs. 23 and 24). Furthermore, transposition of Tn*10* has been shown to be regulated by Dam methylase (25). It is interesting to note that there are three GATC sequences clustered within the 30-bp sequence (residue 7486 to residue 7515 in Fig. 3; the GATC sequences are boxed) 154 bp upstream of the *msr* gene. Methylation of these sequences may be involved in the regulation of the transcription of the *msr–msd*–RT region.

We thank Drs. Monica Roth and Bert Lampson for helpful discussions and critical reading of this manuscript. We also thank Dr. A. Ishihama for the λ library of the *E. coli* K-12 genome. This work was supported by a grant from the U.S. Public Health Service (GM26843) and a grant from Takara Shuzo Company, Ltd.

1. Lampson, B. C., Sun, J., Hsu, M.-Y., Vallejo-Ramirez, J., Inouye, S. & Inouye, M. (1989) *Science* **243**, 1033–1038.
2. Varmus, H. E. (1989) *Cell* **56**, 721–724.
3. Dhundale, A., Lampson, B., Furuichi, T., Inouye, M. & Inouye, S. (1987) *Cell* **51**, 1105–1112.
4. Lampson, B. C., Viswanathan, M., Inouye, M. & Inouye, S. (1990) *J. Biol. Chem.* **265**, 8490–8496.
5. Inouye, S., Hsu, M.-Y., Eagle, S. & Inouye, M. (1988) *Cell* **56**, 709–717.
6. Lim, D. & Maas, W. K. (1989) *Cell* **56**, 891–904.
7. Temin, H. M. (1989) *Nature (London)* **339**, 254–255.
8. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
9. Marinus, M. G. & Morris, N. R. (1974) *J. Mol. Biol.* **85**, 309–322.
10. Miller, H. J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
11. Kohara, Y., Akiyama, K. & Isono, K. (1987) *Cell* **50**, 495–508.
12. Maniatis, T., Fritsch, E. F. & Sambrook. J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
13. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
15. Bachman, B. J. (1972) *Bacteriol. Rev.* **36**, 525–557.
16. Brooks, J. E., Blumenthal, R. M. & Gingeras, T. R. (1983) *Nucleic Acids Res.* **11**, 837–851.
17. Grindley, N. D. F. & Reed, R. R. (1985) *Annu. Rev. Biochem.* **54**, 863–896.
18. Mizuuchi, K. & Craigie, R. (1986) *Annu. Rev. Genet.* **20**, 385–429.
19. Craig, N. (1988) *Annu. Rev. Genet.* **22**, 77–105.
20. Brown, D. P., Idler, K. B. & Katz, L. (1990) *J. Bacteriol.* **172**, 1877–1888.
21. Selander, R. K., Caugant, D. A. & Whittam, T. S. (1987) in *Escherichia coli and Salmonella Typhimurium*, ed. Neidhart, F. C. (Am. Soc. Microbiol., Washington), p. 1625.
22. Herzer, P. T., Inouye, S., Inouye, M. & Whittam, T. S. (1990) *J. Bacteriol.*, in press.
23. Barras, F. & Marinus, M. G. (1989) *Trends Genet.* **5**, 139–143.
24. Messer, W. & Noyer-Weidner, M. (1988) *Cell* **54**, 735–737.
25. Roberts, D., Hoopers, B. C., McClure, W. & Kleckner, N. (1985) *Cell* **43**, 117–130.
26. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
27. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
28. Kalionis, B., Dodd, I. B. & Egan, J. B. (1986) *J. Mol. Biol.* **191**, 199–209.
29. Richardson, H., Puspus, A. & Egan, J. B. (1989) *J. Mol. Biol.* **206**, 251–255.