Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 5620–5624, August 1988
Genetics

Chromosomal location and structure of the operon encoding peptide-chain-release factor 2 of *Escherichia coli*

(*prfB-herC* operon/nuclease S1 mapping/suppressor of ColE1 replication/translation termination/*Salmonella supK* gene)

KOICHI KAWAKAMI*†, YVONNE H. JÖNSSON‡, GLENN R. BJÖRK‡, HIDEO IKEDA†, AND YOSHIKAZU NAKAMURA*§

Departments of *Tumor Biology and †Molecular Biology, The Institute of Medical Science, The University of Tokyo, P.O. Takanawa, Tokyo 108, Japan; and ‡Department of Microbiology, Umeå University, S-90187 Umeå, Sweden

Communicated by Charles Yanofsky, April 22, 1988 (received for review February 2, 1988)

ABSTRACT The *prfB* gene encodes peptide-chain-release factor 2 of *Escherichia coli*, which catalyzes translation termination at UGA and UAA codons. The gene, identified by sequencing, is located at the 62-min region of the *E. coli* chromosome. The *prfB* gene is followed by an open reading frame encoding a 57,603-Da protein. This downstream open reading frame was identified as *herC*, a gene defined by a suppressor mutation that restores replication of a ColE1 plasmid mutant. RNA blot hybridization and S1 nuclease protection analyses of *in vivo* transcripts showed that *prfB* and *herC* are cotranscribed into a 2800-base transcript in the counterclockwise direction with respect to the *E. coli* genetic map. Thus, we refer to the two genes as the *prfB-herC* operon. Data are presented that suggest that *supK*, a mutation in *Salmonella typhimurium* that suppresses UGA termination, is the structural gene for *Salmonella* release factor 2. Translation control within the *prfB-herC* operon and the relationship of these genes to a tRNA methyltransferase are discussed.

Polypeptide chain termination requires participation of two peptide-chain-release factors that recognize specific termination codons. Release factor 1 (RF1) catalyzes termination at UAA and UAG codons, and release factor 2 (RF2) catalyzes termination at UGA and UAA codons (1).

The RF1 gene has been cloned on the basis of competition between a nonsense suppressor tRNA and a translation release factor. The increased RF1 concentration due to the increased RF1 gene dosage reduces the efficiency of suppression by a glutamine-inserting UAG suppressor, *supE* (2). The gene encoding RF1 has been named *prfA* and is located at 27 min on the *Escherichia coli* genome. Near 27 min on the genetic map, *uar* and *sueB* mutations have also been located. The *uar* mutant is temperature-sensitive for growth, misreads UAA codons, and increases the efficiency of UAA and UAG nonsense suppression (3). The *sueB* mutation enhances the efficiency of UAG nonsense suppression (4, 5). Complementation analysis has disclosed that these mutations occur in the *prfA* gene. The data have been interpreted as direct genetic evidence that RF1 catalyzes translation termination at the UAA and UAG codons in *E. coli* (6).

The RF2 gene has been isolated from the Carbon and Clarke *E. coli* plasmid bank (7) on the basis of RF2 overproduction detected by an anti-RF2 antibody. The RF2 plasmid also reduces the efficiency of tRNA UGA suppressors *in vivo* (7). The deduced amino acid sequences of RF1 and RF2 are similar (8). However, the chromosomal location of the RF2 gene had not been reported.

The present study was initiated in an effort to explain a mutation designated *herC180*, which was isolated as a host suppressor of a replication-deficient ColE1 plasmid (ref. 9 and K.K., unpublished work). This *herC* mutation was mapped at 62 min on the *E. coli* chromosome. Cloning and sequencing analyses with the *herC* region of the chromosome revealed that the RF2 gene is located immediately upstream of *herC* in the same transcriptional unit.¶

MATERIALS AND METHODS

Bacterial, Plasmid, and Phage Strains. Bacterial strains (*E. coli* K-12) were C600 (F^- *thr leu tonA lac thi supE44*) (10) and CSR603 (*uvrA recA phr thr leu pro his thi arg lac gal ara xyl mtl rpsL* $T6^r$) (11). Plasmids pKK945 and pKK941 are derivatives of pACYC184 (12) and carry *herC*; their construction will be described elsewhere. λVIII-prfB contains the 3.0-kilobase (kb) *Eco*RI fragment encoding RF2 in the phage vector λVIII (13). It was isolated from a phage library of the *E. coli* genome by plaque hybridization with the ^{32}P-labeled 381-base-pair (bp) *Eco*RI-*Hin*dIII fragment as a probe. This 3.0-kb *Eco*RI fragment was recloned into pACYC184, giving rise to pKK951. Plasmid pRF2 was kindly provided by Caskey *et al.* (7).

DNA Sequence Analysis. DNA sequence was determined by the dideoxy chain-termination method (14), using a 15-nucleotide-long *lacZ* primer and [α-^{32}P]dCTP.

S1 Nuclease Protection Experiments. RNAs were extracted from cells of the C600 strain as described (15). These RNAs were mixed with ^{32}P-labeled DNA fragments in hybridization buffer [80% (vol/vol) formamide/0.4 M NaCl, 0.04 M Pipes, pH 6.4/1 mM EDTA], incubated at 55°C for 3 hr, and then digested with S1 nuclease at 37°C for 1 hr. Samples were glyoxylated and analyzed by electrophoresis in 1.4% agarose gels. These S1 mapping analyses were conducted essentially as described by Burton *et al.* (16).

Analysis of Proteins Produced by Maxicells. Cultures of CSR603 carrying plasmids were grown in minimal medium containing required amino acids and 25 μg of chloramphenicol per ml. The cells were irradiated with UV light and metabolically labeled with [^{35}S]methionine at 14 μCi/ml (1 μCi = 37 kBq) as described (11). Labeled proteins were precipitated with 10% (wt/vol) trichloroacetic acid, rinsed with acetone, and analyzed by $NaDodSO_4$/PAGE (17). Gels were stained with Coomassie brilliant blue R-250, dried, and exposed to Kodak x-ray film for autoradiography.

Other Procedures. DNA and RNA blot hybridization analyses were conducted as described (18) except that GeneScreen*Plus* Membrane (DuPont) was used and DNA or RNA

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.

Abbreviations: RF2, peptide-chain-release factor 2; ORF, open reading frame; mcmo^{5}U, methyl ester of uridine-5-oxyacetic acid.
§To whom reprint requests should be addressed.
¶The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03795).

was transferred electrophoretically. Experimental procedures and conditions of nick-translation and kinase treatment of DNA were essentially as described (18).

RESULTS

Hybrid Plasmids Carrying *herC*, and the *herC* Nucleotide Sequence. *herC180* was isolated as an *E. coli* mutant that suppresses the defect of ColE1 replication caused by a mutation in primer RNA. This primer RNA mutation, named *cer-114*, is a 1-bp substitution 95 bp upstream of the replication origin. It abolishes initiation of DNA synthesis. Since proper conformation of the primer RNA leads to formation of an RNA·DNA hybrid, a prerequisite to initiation of replication (19, 20), it is suspected that the conformation of the primer RNA is altered by *cer-114*. Therefore, the *herC* gene product may be involved in a reaction that affects RNA stability or conformation. The *herC* gene was located and mapped by P1 phage transduction. It is cotransduced with a Tn*10* transposon inserted in the 62-min locus of the *E. coli* chromosome; the gene order is *thyA–lysA*–Tn*10–herC–serA* (Fig. 1*A*). This Tn*10* was used as a selective marker to clone the *herC* gene. A *Sal* I restriction fragment encoding Tn*10* and *herC* was generated by partial digestion and cloned into a cosmid vector, and several subclone derivatives were constructed. Plasmid pKK945 is one of the derivatives and contains a 1.8-kb *Hin*dIII–*Nru* I fragment encoding an activity that complements the *herC* mutation. Isolation of the *herC* mutation and its characterization with respect to mapping, cloning, and suppression of ColE1 replication will be published elsewhere.

The 1.8-kb bacterial DNA segment in pKK945 was sequenced by the dideoxy chain-termination method according to the strategy shown in Fig. 1*B*. The deduced 1832-bp sequence contains an open reading frame (ORF) that starts with an AUG codon at position 229 and ends with a UAA termination codon at position 1744 (Fig. 2). It encodes a 57,603-Da protein composed of 505 amino acids. Seven base pairs upstream of this ORF, there is a canonical Shine–Dalgarno sequence (GAGG; ref. 21). Codon-preference analysis predicts that the protein encoded by this ORF would be highly expressed in *E. coli* (data not shown).

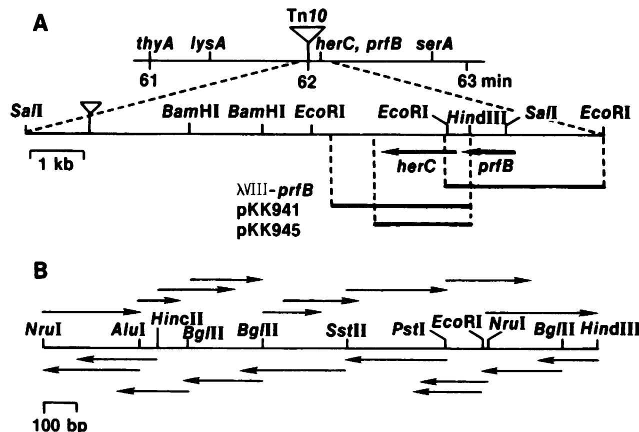

FIG. 1. Chromosomal structure of the *prfB-herC* region (*A*) and strategy of DNA sequencing (*B*). (*A*) Bold bars indicate the bacterial DNA cloned in the plasmid or phage, and bold arrows indicate location and orientation of the genes. Plasmid pKK945 carries the 1.8-kb *Hin*dIII–*Nru* I fragment in the same restriction sites of pACYC184. Plasmid pKK941 was derived from pACYC184 carrying the 3.9-kb *Hin*dIII–*Bam*HI fragment by BAL-31 exonuclease digestion from the *Bam*HI site. The λVIII-prfB phage was isolated from an *E. coli* genomic library and its insert was recloned in pACYC184, yielding pKK951. Triangles show the site of Tn*10* transposon. (*B*) The 1.8-kb DNA carried by pKK945 was sequenced by the indicated strategy. Arrows represent direction and extent of each sequence; only the restriction sites used in sequencing are included.

This ORF is related to the *herC* gene by the finding that insertion of the *rrnB* T1T2 transcription terminator (22) into the *Taq* I site 37 bp upstream of this ORF (pKK948T) or deletion of the *Bgl* II fragments internal to the ORF (pKK931) abolishes the complementation activity for the *herC180* defect (data not shown). The proteins encoded by pKK945 were analyzed by the maxicell method (11). As shown in Fig. 3, pKK945 directs the synthesis of a 63-kDa polypeptide in addition to the vector (pACYC184)-encoded products. The above insertion and deletion derivatives, pKK948T and pKK931, do not direct the synthesis of the polypeptide (data not shown). In addition, the plasmid carrying the *herC180* mutation produces a protein that migrates slightly slower in $NaDodSO_4$/PAGE than that produced by the $herC^+$ plasmid (data not shown). These observations suggest that the 63-kDa protein is synthesized from this ORF and corresponds to a product of the *herC* gene. Expression of this gene in the pKK945 plasmid seems to be dependent on readthrough transcription from the vector pACYC184, because the 1.8-kb fragment cloned in pKK945 does not encode the promoter, as described below.

Identification of the Gene Encoding RF2. Immediately upstream of the coding region of *herC*, there is another ORF, extending from the *Hin*dIII site (Fig. 1) to UGA at positions 217–219 (Fig. 2). Except for two base differences, this 5′ flanking sequence corresponds to the published sequence of the end of the gene encoding RF2 (*prfB*) (8). One difference is an addition of a guanine residue at position 10 in Fig. 2 and another is substitution of cytosine for thymine at position 228. These differences were confirmed by sequencing the opposite strands. The additional guanine residue at position 10 changes the presumed coding sequence; the new coding frame terminates at a UGA codon at positions 217–219, giving a revised molecular mass of 41,346 Da for the polypeptide. This polypeptide is larger than that deduced previously by 26 amino acids and is more consistent with estimates based on the electrophoretic mobility of purified RF2 in polyacrylamide gels (i.e., 48,000 Da; ref. 7). Further, the codon-preference plot of the entire *prfB* gene is similar to that for the revised COOH terminus of RF2 (data not shown).

Evidence That *prfB* and *herC* Constitute the Same Operon. The *prfB* and *herC* genes are separated only by 9 bp (Fig. 2), implying that they are present in a single transcriptional unit. No canonical promoter sequence is found in the 228-bp sequence preceding *herC*. These observations led us to examine the transcripts of this region directly. First, total RNAs prepared from wild-type *E. coli* cells were examined by blot hybridization analysis. A 381-bp *Hin*dIII–*Eco*RI fragment encoding the 3′ sequence of *prfB* and the 5′ sequence of *herC* was labeled with ^{32}P by nick-translation and used as a hybridization probe. A 2800-base transcript hybridized to the probe (Fig. 4*A*). The size detected seems to be sufficient to encode the *prfB* and *herC* products on the same mRNA.

The 5′ end of this mRNA was determined by S1 nuclease protection experiments. Since it was expected that the 5′ end would be located upstream of the *Hin*dIII site, a 3.0-kb *Eco*RI fragment carrying *prfB* and its 5′ flanking region was cloned from the wild-type *E. coli* chromosome into a phage vector by plaque hybridization to the 381-bp *Eco*RI–*Hin*dIII fragment (λVIII-prfB; see Fig. 1). Once isolated, this *Eco*RI fragment was labeled with ^{32}P at its 5′ ends by bacteriophage T4 polynucleotide kinase, hybridized to total *E. coli* RNA, and analyzed by agarose gel electrophoresis after S1 treatment. A 1300-base-long region was protected from S1 digestion (Fig. 4*B*). This 1300-base fragment was shortened to 1100 bases by cleavage of the probe DNA with *Sal* I prior to hybridization. There is only one *Sal* I site within the bacterial DNA, and it

```
                    *
   1 AAG CTT TAT GAA CTG GAG ATG CAG AAG AAA AAT GCC GAG AAA CAG GCG ATG GAA GAT AAC AAA TCC GAC ATC GGC TGG GGC AGC CAG ATT CGT TCT TAT GTC
     Lys Leu Tyr Glu Leu Glu Met Gln Lys Lys Asn Ala Glu Lys Gln Ala Met Glu Asp Asn Lys Ser Asp Ile Gly Trp Gly Ser Gln Ile Arg Ser Tyr Val
                                                  ••  •
 103 CTT GAT GAC TCC CGC ATT AAA GAT CTG CGC ACC GGG GTA GAA ACC CGC AAC ACG CAG GCC GTG CTG GAC GGC AGC CTG GAT CAA TTT ATC GAA GCA AGT TTG
     Leu Asp Asp Ser Arg Ile Lys Asp Leu Arg Thr Gly Val Glu Thr Arg Asn Thr Gln Ala Val Leu Asp Gly Ser Leu Asp Gln Phe Ile Glu Ala Ser Leu
                                *                   =                        ••  •
 205 AAA GCA GGG TTA TGA GGA ACC AAC ATG TCT GAA CAA CAC GCA CAG GGC GCT GAC GCG GTA GTC GAT CTT AAC AAT GAA CTG AAA ACG CGT CGT GAG AAG CTG
     Lys Ala Gly Leu End             Met Ser Glu Gln His Ala Gln Gly Ala Asp Ala Val Val Asp Leu Asn Asn Glu Leu Lys Thr Arg Arg Glu Lys Leu
 307 GCG AAC CTG CGC GAG CAG GGG ATT ACC TTC CCG AAC GAT TTC CGT CGC GAT CAT ACC TCT GAC CAA TTG CAC GCA GAA TTC GAC GGC AAA GAG AAC GAA GAA
     Ala Asn Leu Arg Glu Gln Gly Ile Ala Phe Pro Asn Asp Phe Arg Arg Asp His Thr Ser Asp Gln Leu His Ala Glu Phe Asp Gly Lys Glu Asn Glu Glu
 409 CTG GAA GCG CTG AAC ATC GAA GTC GCC GTT GCT GGC CGC ATG ATG ACC CGT CGT ATT ATG GGT AAA GCG TCT TTC GTT ACC CTG CAG GAC GTT GGC GGT CGC
     Leu Glu Ala Leu Asn Ile Glu Val Ala Val Ala Gly Arg Met Met Thr Arg Arg Ile Met Gly Lys Ala Ser Phe Val Thr Leu Gln Asp Val Gly Gly Arg
 511 ATT CAG CTG TAC GTT GCC CGT GAC GAT CTC CCG GAA GGC GTT TAT AAC GAG CAG TTC AAA AAA TGG GAC CTC GGC GAC ATC CTC GGC GCG AAA GGT AAG CTG
     Ile Gln Leu Tyr Val Ala Arg Asp Asp Leu Pro Glu Gly Val Tyr Asn Glu Gln Phe Lys Lys Trp Asp Leu Gly Asp Ile Leu Gly Ala Lys Gly Lys Leu
 613 TTC AAA ACC AAA ACC GGC GAA CTG TCT ATC CAC TGC ACC GAG TTG CGT CTG CTG ACC AAA GCA CTG CGT CCG CTG CCG GAT AAA TTC CAC GGC TTG CAG GAT
     Phe Lys Thr Lys Thr Gly Glu Leu Ser Ile His Cys Thr Glu Leu Arg Leu Leu Thr Lys Ala Leu Arg Pro Leu Pro Asp Lys Phe His Gly Leu Gln Asp
 715 CAG GAA GCG CGC TAT CGT CAG CGT TAT CTC GAT CTC ATC TCC AAC GAT GAA TCC CGC AAC ACC TTT AAA GTG CGC TCG CAG ATC CTC TCT GGT ATT CGC CAG
     Gln Glu Ala Arg Tyr Arg Gln Arg Tyr Leu Asp Leu Ile Ser Asn Asp Glu Ser Arg Asn Thr Phe Lys Val Arg Ser Gln Ile Leu Ser Gly Ile Arg Gln
 817 TTC ATG GTG AAC CGC GGC TTT ATG GAA GTT GAA ACG CCG ATG ATG CAG GTG ATC CCT GGC GGT GCC GCT GCG CGT CCG TTT ATC ACC CAC CAT AAC GCG CTG
     Phe Met Val Asn Arg Gly Phe Met Glu Val Glu Thr Pro Met Met Gln Val Ile Pro Gly Gly Ala Ala Ala Arg Pro Phe Ile Thr His His Asn Ala Leu
 919 GAT CTC GAC ATG TAC CTG CGT ATC GCG CCG GAA CTG TAC CTC AAG CGT CTG GTG GTT GGT GGC TTC GAG CGT GTA TTC GAA ATC AAC CGT AAC TTC CGT AAC
     Asp Leu Asp Met Tyr Leu Arg Ile Ala Pro Glu Leu Tyr Leu Lys Arg Leu Val Val Gly Gly Phe Glu Arg Val Phe Glu Ile Asn Arg Asn Phe Arg Asn
1021 GAA GGT ATT TCC GTA CGT CAT AAC CCA GAG TTC ACC ATG ATG GAA CTC TAC ATG GCT TAC GCA GAT TAC AAA GAT CTG ATC GAG CTG ACC GAA TCG CTG TTC
     Glu Gly Ile Ser Val Arg His Asn Pro Glu Phe Thr Met Met Glu Leu Tyr Met Ala Tyr Ala Asp Tyr Lys Asp Leu Ile Glu Leu Thr Glu Ser Leu Phe
1123 CGT ACT CTG GCA CAG GAT ATT CTC GGT AAG ACG GAA GTG ACC TAC GGC GAC GTG ACG CTG GAC TTC GGT AAA CCG TTC GAA AAA CTG ACC ATG CGT GAA GCG
     Arg Thr Leu Ala Gln Asp Ile Leu Gly Lys Thr Glu Val Thr Tyr Gly Asp Val Thr Leu Asp Phe Gly Lys Pro Phe Glu Lys Leu Thr Met Arg Glu Ala
1225 ATC AAG AAA TAT CGC CCG GAA ACC GAC ATG GCG GAT CTG GAC AAC TTC GAC TCT GCG AAA GCA ATT GCT GAA TCT ATC GGC ATC CAC GTT GAG AAG AGC TGG
     Ile Lys Lys Tyr Arg Pro Glu Thr Asp Met Ala Asp Leu Asp Asn Phe Asp Ser Ala Lys Ala Ile Ala Glu Ser Ile Gly Ile His Val Glu Lys Ser Trp
1327 GGT CTG GGC CGT ATC GTT ACC GAG ATC TTC GAA GAA GTG GCA GAA GCA CAT CTG ATT CAG CCG ACC TTC ATT ACT GAA TAT CCG GCA GAA GTT TCT CCG CTG
     Gly Leu Gly Arg Ile Val Thr Glu Ile Phe Glu Glu Val Ala Glu Ala His Leu Ile Gln Pro Thr Phe Ile Thr Glu Tyr Pro Ala Glu Val Ser Pro Leu
1429 GCG CGT CGT AAC GAC GTT AAC CCG GAA ATC ACA GAC CGC TTT GAG TTC TTC ATT GGT GGT CGT GAA ATC GGT AAC GGC TTT AGC GAG CTG AAT GAC GCG GAA
     Ala Arg Arg Asn Asp Val Asn Pro Glu Ile Thr Asp Arg Phe Glu Phe Phe Ile Gly Gly Arg Glu Ile Gly Asn Gly Phe Ser Glu Leu Asn Asp Ala Glu
1531 GAT CAG GCG CAA CGC TTC CTG GAT CAG GTT GCC GCG AAA GAC GCA GGT GAC GAC GAA GCG ATG TTC TAC GAT GAA GAT TAC GTC ACC GCA CTG GAA CAT GGC
     Asp Gln Ala Gln Arg Phe Leu Asp Gln Val Ala Ala Lys Asp Ala Gly Asp Asp Glu Ala Met Phe Tyr Asp Glu Asp Tyr Val Thr Ala Leu Glu His Gly
1633 TTA CCG CCG ACA GCA GGT CTG GGA ATT GGT ATC GAC CGT ATG GTA ATG CTG TTC ACC AAC AGC CAT ACC ATC CGC GAC GTT ATT CTG TTC CCG GCG ATG CGT
     Leu Pro Pro Thr Ala Gly Leu Gly Ile Gly Ile Asp Arg Met Val Met Leu Phe Thr Asn Ser His Thr Ile Arg Asp Val Ile Leu Phe Pro Ala Met Arg
1735 CCG GTA AAA TAA GCATTACGTTATGCTCACAACCCCGGCAAATGTCGGGGTTTTTTATTTAAGCTGGGTAAATGGAGATAATCGTTTTCTGGCTTCG 1832
     Pro Val Lys End
```

FIG. 2. Nucleotide and predicted amino acid sequence of *herC* and the COOH-terminal coding region of *prfB*. Putative ribosomal binding site is underlined, and putative ρ-independent termination signal is indicated by arrows followed by broken line. Two bases that do not agree with the sequence of Craigen *et al.* (8) are shown by asterisks. The 3′ end of the DNA segment cloned in pRF2 is marked by double line at position 244 (8). Possible UAG termination codons, one of which (positions 140–142) was presented previously (8), are marked by dots.

is located 1.1 kb from the *Eco*RI site internal to *herC*. These results indicate that the 2800-base transcript starts just upstream of the initiation codon of RF2 and extends to at least the internal *Eco*RI site labeled by kinase. The location of this start site is consistent with that predicted (8).

The 3′ end of the transcript that extends beyond the *herC* gene was similarly determined by S1 mapping. Two probes were used, plasmid pKK941 cleaved with *Hin*dIII and the 4.2-kb *Eco*RI fragment isolated from pKK941, and each was labeled at the 3′ ends by the Klenow fragment of DNA polymerase I. Only one end of the labeled probes contains DNA encoding *prfB* (*Hin*dIII) or *herC* (*Eco*RI). The other end of each probe encodes vector DNA. Thus, only one end of each fragment can be protected by bacterial RNA. A transcript was detected that extends about 1800 bases from the *Hin*dIII site or about 1400 bases from the *Eco*RI site (Fig. 4*C*). When the DNA was cut with *Hin*cII before S1 mapping, the 1800-base fragment from *Hin*dIII digestion was shortened to 1450 bases, a length identical to that of the *Hin*dIII–*Hin*cII probe. These data indicate that the transcript extending from the *Hin*dIII site terminates about 350 bases beyond the *Hin*cII site. A potential ρ-independent termination signal located immediately downstream of the stop codon of *herC* may determine the 3′ end of this transcript (Fig. 2). These results led us to conclude that the *prfB* and *herC* genes are cotranscribed and are in a single operon in *E. coli*.

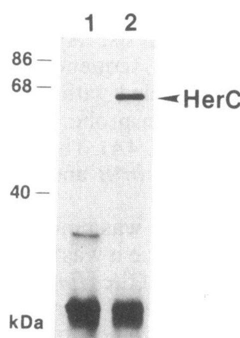

FIG. 3. Autoradiograph of the *herC* protein synthesized in maxicells. [^{35}S]Methionine-labeled proteins produced in maxicells carrying pACYC184 (lane 1) or pKK945 (lane 2) were analyzed by electrophoresis in a 10–20% linear gradient polyacrylamide gel. A protein at the 63-kDa position in lane 2 seems to be a product of the *herC* gene. A 35-kDa protein seen in lane 1 corresponds to the tetracycline-resistance (*tet*^{r}) gene product, which is not synthesized by pKK945 because of insertion of the *herC* fragment into the *tet*^{r} gene. Sizes of marker proteins run in parallel are represented at left.

Is *supK* the Structural Gene for *Salmonella* RF2? In *Salmonella typhimurium*, a gene, *supK*, has been characterized whose mutations mediate suppression of termination at UGA. Mutations in *supK* are recessive and map at 62 min (23)—i.e., in the same region as *prfB* in *E. coli*. To test whether or not the *supK* gene of *S. typhimurium* encodes RF2, various plasmids carrying the *E. coli prfB* gene were introduced into *supK* strains and into another UGA suppressor strain, *supU*. The *supU* strains harbor dominant UGA suppressors and are likely to be mutated in ribosomal protein S4 or S5 (ref. 24; G.R.B., unpublished data). As shown in Table 1, the episome KLF16, which carries a large segment of *E. coli* DNA from the 62-min region of the genome, complements both the *supK* and *supU* mutations, restoring UGA termination in the *his* gene. The plasmid pKK951, which carries only 3 kb of *E. coli* DNA including *prfB*, also complements, thereby suggesting that complementation involves the expression of *prfB* and not of another gene located near 62 min. Complementation of the recessive *supK* strains presumably is due to the overproduction of *E. coli* RF2, which competes with the mutant *supK* gene product for

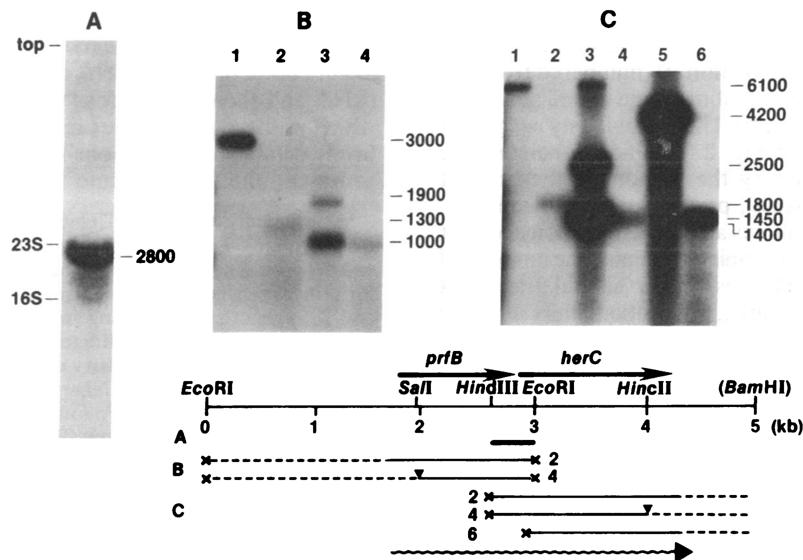

FIG. 4. RNA blot hybridization (*A*) and S1 nuclease mapping of 5′ (*B*) and 3′ (*C*) ends of the *prfB-herC* transcript. Probes and transcripts detected are illustrated below the restriction map; solid lines indicate DNA fragments protected from S1 digestion, and broken lines indicate S1 nuclease-sensitive portions of the fragments. Numbers at the end of the solid lines correspond to lane numbers in *B* and *C*. Wavy arrow indicates the RNA transcript. Sizes of DNA fragments were determined by using the ^{32}P-labeled *Eco*RI–*Hin*dIII fragments of λ phage DNA as markers. (*A*) RNA blot hybridization. RNAs isolated from C600 cells were subjected to electrophoresis in 1% agarose gel and transferred to the GeneScreen*Plus* membrane. The 381-bp *Hin*dIII–*Eco*RI fragment labeled with ^{32}P by nick-translation was the hybridization probe. A 2800-base RNA was detected. (*B*) S1 mapping of the 5′ end. The 3.0-kb *Eco*RI fragment was isolated from λVIII-prfB DNA, labeled with ^{32}P at the 5′ ends by phage T4 polynucleotide kinase, and used as an S1 protection probe. Lane 1, ^{32}P-labeled DNA; lane 2, protected fragment after S1 treatment; lane 3, ^{32}P-probe DNA recut with *Sal* I; lane 4, protected fragment of the *Sal* I-cut DNA after S1 treatment. (*C*) S1 mapping of the 3′ end. Lanes 1, 3, and 5 represent ^{32}P-labeled DNAs used for hybridization and lanes 2, 4, and 6 represent protected fragments after S1 treatment. ^{32}P-labeled DNAs were as follows. Lanes 1 and 2, pKK941 DNA cleaved with *Hin*dIII and labeled at the 3′ ends by Klenow enzyme; lanes 3 and 4, the same DNA as used in lanes 1 and 2 except that it was cut with *Hin*cII; lanes 5 and 6, 4.2-kb *Eco*RI fragment isolated from pKK941 and labeled at the 3′ ends (lanes 5 and 6). Only one end of each labeled probe (lanes 1, 2, 5, and 6), derived from the bacterial insert, can be protected by bacterial RNA.

binding to the ribosomal complex. Reversal of the dominant *supU* suppression could be due to enhanced RF2 binding to mutant ribosomes, again due to overproduction. On the basis of the above results, we cannot rule out that complementation is caused by a gene other than *prfB* carried by pKK951. However, the failure to complement both *supK* and *supU* strains with either of two point mutations of KLF16, *sup*^{0}-*1190* (ICR-191-induced) and *sup*^{0}-*1189* (spontaneous) (23), means that the complementation involves a single gene. The simplest interpretation of these results is that complementation is due to *prfB*. This strongly suggests that *supK* encodes RF2 in *S. typhimurium*.

Table 1. Complementation/competition of *supK*- and *supU*-mediated UGA suppression by various plasmids

	Suppressor activity (His$^+$ phenotype)					
UGA suppressor	No plasmid	KLF16	*sup*^{0}-*1189*	*sup*^{0}-*1190*	pKK951	pRF2
supK584	+	ND	ND	ND	−	+
supK599	+	−	+	+	−	+
supK1292	+	−	+	+	−	+
supK1293	+	−	+	+	−	+
supU1283	+	−	+	+	−	(+)
supU1285	+	−	+	+	−	(+)

All strains used were derivatives of *S. typhimurium* and carry the indicated suppressor mutations and a *his* UGA mutation. Episomes (wild-type KLF16 or KLF16 carrying *sup*^{0}-*1189* or *sup*^{0}-*1190* mutation) were transferred into strains also carrying *serA790* mutation by selecting for Ser$^+$. In each case where complementation occurred, the original His phenotype was regained upon loss of the episome. The His phenotype was scored after 2 days of incubation at 37°C. (+), Weak inhibition of suppression. ND, not done. The episomal mutation *sup*^{0}-*1190* was induced by ICR-191 (a frameshift-inducing mutagen); *sup*^{0}-*1189* arose spontaneously (23). *supK* mutations are recessive UGA suppressors and map at 62 min of chromosome (23). *supU1283* and *supU1285* are dominant UGA suppressors and map in the ribosome cluster (24). Strains carrying *supU1286*, another allele of *supU*, have an altered ribosomal protein S4 (G.R.B., unpublished data). Thus *supU* is likely to code for ribosomal protein S4 or S5, since mutated forms of these two ribosomal proteins are known to suppress nonsense codons (25, 26).

We also tested the plasmid pRF2, which had been characterized previously and was presumed to carry the entire RF2 gene (7). pRF2 fails to complement *supK* strains and only weakly affects *supU*-mediated suppression (Table 1). Further, pRF2 inhibits the Hirsh UGA suppressor (27) by <14%, whereas pKK951 inhibits it by >90% (data not shown). These results indicate that pKK951 encodes a biologically active RF2, whereas pRF2 encodes an RF2 with much lower activity. The difference in behavior of the two plasmids may be due to the fact that pRF2 lacks DNA near the 3′-terminal region of *prfB* and therefore may express an altered RF2 (see *Discussion*).

DISCUSSION

The present work establishes that the *prfB* gene, encoding RF2, is located at 62 min on the *E. coli* chromosome. The *prfB* gene is followed immediately by a gene named *herC*. A mutation in *herC* was first isolated as a host mutation that suppresses a defect in ColE1 plasmid replication. Mapping, cloning, and sequencing of the *herC* region of the chromosome showed that the region upstream from the 5′ end of *herC* coincides with the 3′ sequence of *prfB*. By RNA blot hybridization and S1 nuclease protection experiments, a 2800-base transcript encoding *prfB* and *herC* was identified. The data indicate that these two genes constitute an operon that we call the *prfB-herC* operon.

The sequence of *prfB* contains two base differences from the sequence reported previously (7). One of these is an additional guanine residue at position 10 (numbered from the *Hin*dIII site), which shifts the reading frame to end at a UGA at positions 217–219 and not at the UAG previously suggested (7) and shown by the dots in Fig. 2. This additional guanine residue has been confirmed by the authors of the original sequence report (C. T. Caskey, personal communication).

The RF2 coding region contains a UGA termination codon at amino acid position 26 from the amino terminus; the amino acid sequence of the protein indicates that a natural frameshift occurs at the UGA codon during translation, thus allowing complete translation of RF2 (8). This provides a natural mechanism of autotranslational control of RF2 expression, since RF2 normally catalyzes termination at UGA codons; i.e., in limiting RF2 conditions, frameshifting would be favored over termination (28). In the revised RF2 sequence, the termination codon of the second, COOH-terminal reading frame is also UGA. Translation termination at this UGA codon would generate an RF2 protein of 41,346 Da. If frameshifting also were to occur at this second UGA codon, then RF2 translation would extend an additional 15 codons and terminate at the UAG codon located in *herC*, generating an RF2 of 43,083 Da. pKK951, but not pRF2, is capable of generating the putative longer RF2 protein because the extended RF2 coding sequence in pRF2 ends at the position corresponding to 244, six codons upstream from the UAG codon (see Fig. 2). If only the extended product is fully active, RF2 encoded by pRF2 may be less active, since its COOH terminus would be truncated. This may explain why a low activity of RF2 was observed in complementation of *supK*-mediated UGA suppression by pRF2 (Table 1). A frameshift at the end of the RF2 coding region might also interfere with initiation and translation of *herC*. The molecular mass of RF2 is 47–50 kDa as measured by $NaDodSO_4$/PAGE (29, 30). Since these values are much higher than the two calculated molecular masses, they are not helpful in distinguishing which RF2 form is expressed in cells. The actual COOH-terminal sequence of RF2 remains to be determined.

In *E. coli*, mutations in *prfB* have not been reported. In *S. typhimurium*, a recessive UGA suppressor, *supK*, has been isolated (23, 31, 32) that maps in the same location in the genome as does *prfB* in *E. coli*. We report complementation experiments that strongly indicate that *E. coli* RF2 is responsible for reversing mutant phenotypes of *supK*. In the *supK* strains, reduced levels of tRNA-($mcmo^5U$) methyltransferase ($mcmo^5U$ is the methyl ester of uridine-5-oxyacetic acid) were observed and might cause UGA suppression by failure to modify tRNA (33–36). However, *E. coli* mutants with even greater defects in tRNA methylation, such as the *aroA*, *-B*, *-C*, *-D*, and *-E* mutants (37), do not affect UGA suppression (unpublished observation). Further, some *supK* strains have a normal level of $mcmo^5U$ in their tRNA (G.R.B., unpublished data). These results are not consistent with the explanation that a tRNA lacking $mcmo^5U$ is the suppressing agent in the *supK* strains. However, they are consistent with our suggestion that *supK* is the structural gene for the *Salmonella* RF2.

If one accepts that *supK* encodes *Salmonella* RF2, one can speculate on different possibilities to explain the deficiency of tRNA-($mcmo^5U$) methyltransferase in several *supK* strains (33). One possibility is that the *herC* gene encodes the methyltransferase or a protein that stabilizes the methyltransferase *in vitro*. In fact, there is evidence for such a stabilizing factor from studies on the purification of the enzyme (34). The *herC* gene product could be at lower levels in *supK* strains because translation of RF2 might extend beyond the UGA codon located just upstream of the *herC* gene, causing translational interference with *herC* expression. A second possibility is that synthesis of the methyltransferase is controlled directly by RF2, possibly by action at a crucial UGA codon. In this case the structural gene for the enzyme must be located outside the *prfB-herC* operon. If *herC* affects the methyltransferase, the *herC180* mutant might affect the modification of primer RNA of the ColE1 plasmid as well as tRNA and thereby suppress the *cer-114* mutant. To decide which of these possibilities is correct, it will be necessary to firmly establish the relationship between the *supK* and *herC* genes and the tRNA-($mcmo^5U$) methyltransferase.

We sincerely thank H. Uchida for much valuable advice and discussion, C. T. Caskey for communicating the resequencing data and for providing the pRF2 plasmid, D. Court and J. W. B. Hershey for critical reading of the manuscript and helpful advice, C. Nolan and T. Sugimoto for the typing and editing of the manuscript, and K. Egawa for continuous encouragement. This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan (to Y.N. and H.I.), Uehara Memorial Foundation, Japan (to Y.N.), the Nissan Science Foundation, Japan (to H.I.), the Swedish National Science Foundation (BU-2930 to G.R.B.), the Swedish Cancer Society (project 680, to G.R.B.), and the Swedish Board for Technical Development (to G.R.B.).

1. Scolnick, E., Tompkins, R., Caskey, T. & Nirenberg, M. (1968) *Proc. Natl. Acad. Sci. USA* **61**, 768–774.
2. Weiss, R. B., Murphy, J. P. & Gallant, J. A. (1984) *J. Bacteriol.* **158**, 362–364.
3. Ryden, S. M. & Isaksson, L. A. (1984) *Mol. Gen. Genet.* **193**, 38–45.
4. Oeschger, M. P., Oeschger, N. S., Wiprud, G. T. & Woods, S. L. (1980) *Mol. Gen. Genet.* **177**, 545–552.
5. Oeschger, M. P. & Wiprud, G. T. (1980) *Mol. Gen. Genet.* **178**, 293–299.
6. Ryden, M., Murphy, J., Martin, R., Isaksson, L. & Gallant, J. (1986) *J. Bacteriol.* **168**, 1066–1069.
7. Caskey, C. T., Forrester, W. C., Tate, W. & Ward, C. D. (1984) *J. Bacteriol.* **158**, 365–368.
8. Craigen, W. J., Cook, R. G., Tate, W. P. & Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3616–3620.
9. Naito, S., Kitani, T., Ogawa, T., Okazaki, T. & Uchida, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 550–554.
10. Appleyard, R. K. (1954) *Genetics* **39**, 429–439.
11. Sancar, A., Hack, A. M. & Rupp, W. D. (1979) *J. Bacteriol.* **137**, 692–693.
12. Chang, A. C. Y. & Cohen, S. N. (1978) *J. Bacteriol* **134**, 1141–1156.
13. Murray, N. E. & Murray, K. (1974) *Nature (London)* **251**, 476–481.
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
15. Aiba, H., Adhya, S. & Crombrugghe, B. (1981) *J. Biol. Chem.* **256**, 11905–11910.
16. Burton, Z. F., Gross, C. A., Watanabe, K. K. & Burgess, R. R. (1983) *Cell* **32**, 335–349.
17. Nakamura, Y. & Uchida, H. (1983) *Mol. Gen. Genet.* **190**, 196–203.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
19. Masukata, H. & Tomizawa, J. (1984) *Cell* **36**, 513–522.
20. Masukata, H. & Tomizawa, J. (1986) *Cell* **44**, 125–136.
21. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
22. Amann, E., Brosius, J. & Ptashne, M. (1983) *Gene* **25**, 167–178.
23. Reeves, R. H. & Roth, J. R. (1971) *J. Mol. Biol.* **56**, 523–533.
24. Johnston, H. M. & Roth, J. R. (1980) *J. Bacteriol.* **144**, 300–306.
25. Piepersberg, W., Böck, A. & Wittmann, H. G. (1975) *Mol. Gen. Genet.* **140**, 91–100.
26. Gorini, L. (1971) *Nature (London) New Biol.* **234**, 261–264.
27. Hirsh, D. (1970) *Nature (London)* **228**, 57.
28. Craigen, W. J. & Caskey, C. T. (1986) *Nature (London)* **322**, 273–275.
29. Klein, H. A. & Capecchi, M. R. (1971) *J. Biol. Chem.* **246**, 1055–1061.
30. Ratliff, J. C. & Caskey, C. T. (1977) *Arch. Biochem. Biophys.* **181**, 671–677.
31. Riddle, D. L. & Roth, J. R. (1972) *J. Mol. Biol.* **66**, 483–493.
32. Wu, T. T. (1966) *Genetics* **54**, 405–410.
33. Reeves, R. H. & Roth, J. R. (1975) *J. Bacteriol.* **124**, 332–340.
34. Pope, W. T. & Reeves, R. H. (1978) *J. Bacteriol.* **136**, 191–200.
35. Pope, W. T., Brown, A. & Reeves, R. H. (1978) *Nucleic Acids Res.* **5**, 1041–1057.
36. Björk, G. R., Ericson, J. U., Gustafsson, C. E. D., Hagervall, T. G., Jönsson, Y. H. & Wikström, P. M. (1987) *Annu. Rev. Biochem.* **56**, 263–287.
37. Björk, G. R. (1980) *J. Mol. Biol.* **140**, 391–410.