

Nucleotide Sequences of *dnaE*, the Gene for the Polymerase Subunit of DNA Polymerase III in *Salmonella typhimurium*, and a Variant That Facilitates Growth in the Absence of Another Polymerase Subunit

EDWARD D. LANCY, MIRIAM R. LIFSICS, PATRICIA MUNSON, AND RUSSELL MAURER*

Department of Molecular Biology and Microbiology, Case Western Reserve University, 2119 Abington Road, Cleveland, Ohio 44106

Received 3 May 1989/Accepted 27 June 1989

The *dnaE* gene of *Salmonella typhimurium*, like that of *Escherichia coli*, encodes the α subunit containing the polymerase activity of the principal replicative enzyme, DNA polymerase III. This gene, or one nearby, has been identified as the locus of suppressor mutations that promote growth by cells deleted for *dnaQ*, the gene for the editing subunit of this enzyme complex. Using a combination of nucleotide sequencing and marker rescue experiments, the alteration in one such suppressor was identified as a valine-to-glycine substitution at amino acid 832 of the 1,160-amino-acid α polypeptide. The α polypeptides of *E. coli* and *S. typhimurium* are identical in size and in 97% of their amino acid residues. Their identity includes the valine residue that was changed in the suppressor allele of *S. typhimurium*. We also localized a temperature-sensitive *dnaE* mutation to the 3' half of *dnaE*.

In *Escherichia coli* and *Salmonella typhimurium*, DNA replication is carried out by the multisubunit DNA polymerase III. The intrinsic polymerization activity of this enzyme complex lies in the α subunit, encoded by *dnaE* (8, 20). Other subunits, by associating with α , refine many of its properties, including its rate, processivity, thermostability, and tendency to dimerize (9, 12). In addition, the ϵ subunit encoded by *dnaQ* endows the polymerase complex with an editing capacity (13, 14). These biochemical interactions between α and other polymerase subunits may underlie several genetic interactions that have been described between *dnaE* and genes for other polymerase subunits (5, 6, 10).

In the accompanying paper (7), we have shown that *Salmonella* null mutants of *dnaQ*, although viable, exhibit a severe growth defect that can be remedied by suppressor mutations mapping in or near *dnaE*. Here we present nucleotide sequence data for wild-type *dnaE* and for one of the suppressor variants, *spq-2*. We also present evidence from marker rescue experiments that the single sequence difference in *dnaE* between these two alleles accounts for the suppressor phenotype of *spq-2*. These results provide further evidence for the functional significance of an α - ϵ interaction by showing that the growth-defective phenotype of $DnaQ^-$ can be remedied by an alteration of α .

MATERIALS AND METHODS

Strains. In addition to strains listed in reference 7, we used strains RM1757 and RM1759, derivatives of strain RM10 carrying *dnaE698*(Ts) and *dnaE693*(Ts), respectively (11).

Clones and subclones. Wild-type *dnaE* was subcloned from λ RM310 (11) on a 7.0-kilobase (kb) *SmaI* fragment in the vector pUC8 (19) to give plasmid pFF325. A Southern blot confirmed that *dnaE* is found on a 7.0-kb *SmaI* fragment in

genomic DNA of both wild-type and *spq-2* strains (data not shown). A subclone of the *spq-2* 7.0-kb fragment was made in several steps, beginning with preparation of a library in the vector bacteriophage λ 1059 by using genomic DNA of strain RM746 (7, 11). A *dnaE*-complementing clone, λ RM740, isolated from this library, was the source of the 7.0-kb fragment that was subcloned into pUC8 to give pFF221. Four additional pUC8 plasmid constructs containing incomplete *dnaE* genes were derived from pFF221 and pFF325 by subcloning fragments having one end at a unique *PstI* site corresponding to codons 600 to 601 in *dnaE* and the other end at either a *BamHI* site or an *EcoRI* site derived from the pUC8 multiple-cloning site. The plasmids thus formed were pFF377, carrying the 3' portion of *spq-2*; pFF378, carrying the 3' portion of *dnaE*⁺; pFF381, carrying the 5' portion of *spq-2*; and pFF382, carrying the 5' portion of *dnaE*⁺.

Marker rescue experiments. *dnaE*(Ts) strains RM1757 and RM1759, and derivatives carrying plasmids pFF377, -378, -381, and -382, were grown to saturation at 32°C in LB plus thymine. A 0.1-ml sample of the culture was plated on LB plus thymine agar. After the plate was incubated at 42°C overnight, colonies were counted. In cases where the presence of a plasmid yielded a marked excess of temperature-resistant revertants compared with the number obtained in the control (no plasmid) experiment, individual revertants were purified by restreaking at 32°C. In most cases, these purified revertants were found to have lost their respective plasmids, but the precise point at which plasmid loss occurred was not determined. These revertants were then tested for suppressor phenotype by scoring colony morphology after transduction to *dnaQ*::Tn10 (7).

DNA sequence methodology. DNA sequence was determined by using double-stranded plasmid DNAs (pFF221 and pFF325) as the template. Sequencing was by the dideoxy method, using modified T7 DNA polymerase (Sequenase; U.S. Biochemical Corp.). Initial sequence was obtained by

* Corresponding author.

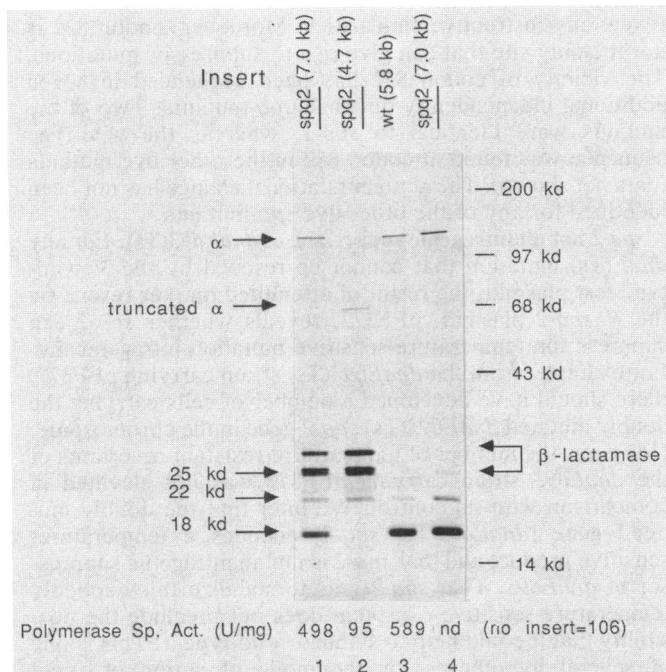


FIG. 2. Expression of gene products under T7 RNA polymerase control from *spq-2* and wild-type DNA fragments and overproduction of DNA polymerase activity. The plasmids used in lanes 1 and 4 were identical except that the entire T7 transcription unit was in the same orientation as the plasmid β -lactamase gene in lane 1 and in the opposite orientation in lane 4. The plasmid for lane 2 differed from that for lane 1 by a deletion of 2.3 kb from the 3' end of the transcription unit. The plasmid for lane 3 differed from that for lane 4 by a deletion of 1.2 kb from the 5' end of the transcription unit. kd, Kilodaltons.

type and *spq-2*. The nucleotide sequence data reported in this paper, as well as flanking sequences not shown, will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession number M26046.

Expression and labeling of proteins expressed under T7 RNA polymerase control. Suitable restriction fragments containing wild-type *dnaE* or *spq-2* were inserted into pT7 plasmids containing a T7 promoter (16). Proteins were expressed under T7 control after thermal induction of a strain in which the gene for T7 RNA polymerase was controlled by a temperature-sensitive λ repressor (16). After the thermal induction, further host RNA synthesis was shut off by the addition of rifampin, to which T7 RNA polymerase is insensitive. The protocols were slightly different in detail in isotopic labeling experiments and overproduction experiments. For labeling, cells were grown at 30°C in LB, washed three times, and suspended in minimal (M9) medium supplemented with thiamine and 18 amino acids (minus cysteine and methionine). After additional growth at 30°C for 1 h, cells were shifted to 42°C for 15 min. Next, rifampin was added at 400 μ g/ml, and incubation was continued at 42°C for 20 min. Cells were then shifted to 37°C for 20 min, and labeling with [³⁵S]methionine took place for 5 min at 30°C. For DNA polymerase overproduction, thermal induction was in LB for 20 min at 42°C; rifampin was added to a final concentration of 100 μ g/ml, and incubation was continued at 42°C for 20 min; the culture was then shifted to 30°C for 1 h before harvesting. DNA polymerase activity in fraction I was determined as described in the accompanying report (7).

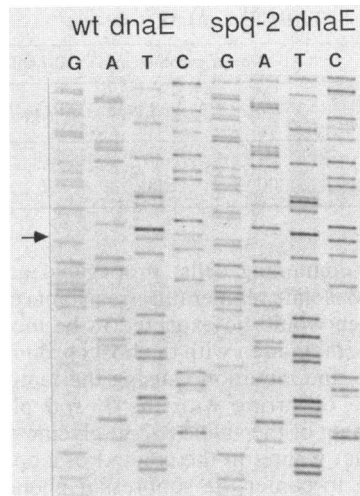


FIG. 3. Portion of a DNA sequence gel showing the difference between wild type (wt) and *spq-2* (arrow).

RESULTS

The *S. typhimurium dnaE* (wild type) open reading frame (ORF) and some of its flanking sequence are shown in Fig. 1. The 3,480-base ORF encodes an 1,160-residue protein of calculated molecular weight 130,133. That this ORF is in fact *dnaE* was supported by several pieces of evidence, including (i) the striking similarity at the amino acid level between this sequence and the *E. coli* α peptide sequence deduced from the *E. coli dnaE* gene sequence (18); (ii) the ability of this sequence, when present on a λ phage or a plasmid, to complement authentic *dnaE*(Ts) mutations of *E. coli* or *S. typhimurium* (11; unpublished data); (iii) synthesis of a protein of the predicted size when a fragment containing this region in the appropriate orientation was transcribed from a phage T7 promoter (Fig. 2, lanes 1, 3, and 4); (iv) overproduction of DNA polymerase activity dependent on T7 RNA polymerase in vivo from such constructs (Fig. 2); and (v) loss of genetic complementation of mutants, overproduction of polymerase activity, and synthesis of the presumptive α protein when the reading frame was truncated (Fig. 2, lane 2).

The same region was sequenced by using *spq-2* DNA as the template. The only change observed in the entire *dnaE* reading frame was found at codon 832. In the *spq-2* gene, the sequence encoded glycine in place of the wild-type valine (Fig. 3). There were no differences found in any flanking sequences.

Marker rescue of *spq-2* and localization of *dnaE*(Ts) mutations. Without sequencing much more flanking DNA on the 5' side of *dnaE*, we could not be sure that the valine 832-to-glycine change found in *spq-2 dnaE* was the only difference between the *spq-2* and wild-type clones. Instead, we carried out a marker rescue experiment to show that the identified change was sufficient to produce the suppressor phenotype. We prepared four plasmids, which we shall refer to as 5' wild type, 5' *spq*, 3' wild type, and 3' *spq*. Each plasmid contained approximately half of *dnaE* (the dividing point being at a *Pst*I site corresponding to codons 600 to 601), as well as flanking sequences on the same side, amounting to 506 bases on the 3' side and about 3,000 bases on the 5' side. These plasmids were introduced into two different *dnaE*(Ts) mutants of *S. typhimurium*, where, as expected, they did not complement the temperature sensi-

TABLE 1. Marker rescue

<i>dnaE</i> (Ts) allele	No. of colonies at 42°C/0.1 ml				
	No plasmid	3' <i>spq-2</i>	3' wild type	5' <i>spq-2</i>	5' wild type
<i>dnaE693</i>	7	23	17	2	26
<i>dnaE698</i>	47	300	200	6	3

tivity. Plasmid-containing cells, propagated at 32°C, were plated at 42°C to isolate temperature-resistant revertants, the number of which would be expected to be increased if the plasmid, by recombination with the host chromosome, could replace (rescue) the mutation causing the temperature sensitivity. Rescue occurring with the 3' *spq* plasmid might result in placement of the valine 832-to-glycine mutation into the chromosome, where, in the context of a complete *dnaE* gene, its ability to confer the suppressor phenotype on the cells could be assessed as described in the accompanying report (7).

dnaE698 was efficiently rescued by the 3' wild-type and by the 3' *spq* plasmids but not by the 5' plasmids (Table 1). Moreover, all of 16 temperature-resistant revertants recovered from the experiment with the 3' *spq* plasmid and none of 16 recovered from the experiment with the 3' wild type exhibited the suppressor phenotype. It is important to note that complete sequence of the 3' plasmids was known, including the flanking regions for both wild type and *spq-2*; since these plasmids differed only by the valine-to-glycine change, the recovery of suppressor cells must be attributed to this difference. This experiment incidentally localized *dnaE698* to the 3' portion of *dnaE*, and most probably to a region close to the *spq-2* mutation, since the sites of *dnaE698* and *spq-2* were not separated in any of 16 recombinants. Another *dnaE* allele, *dnaE693*, was not efficiently rescued by either 3' or 5' plasmids. This result may be explained if the mutation in *dnaE693* was located close to the *Pst*I site in *dnaE*, as this positioning would tend to diminish the frequency of recombination with the plasmid that ought to give marker rescue in principle.

DISCUSSION

The main result presented here is the determination of the *S. typhimurium dnaE* sequence for both the wild-type gene and a mutant allele that promotes growth of DnaQ⁻ strains. We argue elsewhere (7) that the *spq-2* allele of *dnaE* encodes an active α subunit of DNA polymerase III. This argument is reinforced by the data presented here showing overproduction of DNA polymerase activity in vivo from an *spq-2* expression plasmid (Fig. 2). We infer, therefore, that the change of valine 832 to glycine alters some property or properties of α in a way that mimics the effect of the normal α - ϵ interaction. Although valine 832 could be a residue that normally interacts with ϵ , our data do not require this to be so. Studies of the wild-type and *spq-2* α proteins, now in progress, have so far failed to identify any property of α that is altered in the mutant.

The sequence change in *spq-2 dnaE* and its location within *dnaE* offer little help in understanding the mode of action of the *spq-2* suppressor. In general, *dnaE* shows little similarity to other proteins of related function (18), and codon 832 does not lie within any regions that have previously been suggested as significantly similar to other polymerases. A data base search for protein sequences similar to the 201-amino-acid *dnaE* subsequence surrounding codon 832 failed to

reveal any instructive similarities. Moreover, codon 832 is not the only site that can give rise to suppressor mutations. The vicinity of codon 832 has been sequenced in seven additional independently obtained *spq* mutants. Two of the mutants were identical to *spq-2*, whereas the wild-type sequence was found at codon 832 in the other five mutants (data not shown). The actual mutational change has not been identified for any of the other five *spq* mutants.

***spq-2* not an intragenic suppressor of *dnaE693*(Ts).** For any *dnaE*(Ts) mutation that cannot be rescued by the 3' wild-type test plasmid, the result of attempted marker rescue by the 3' *spq-2* plasmid, pFF377, reveals whether *spq-2* can suppress the temperature-sensitive mutation intragenically. For example, from the *dnaE693*(Ts) strain carrying pFF377, there should have developed a number of cells carrying the doubly mutated *dnaE693*(Ts) *spq-2* gene in the chromosome. However, the number of temperature-resistant revertants of the *dnaE693* strain carrying pFF377 was not elevated in comparison with its control. We infer that the doubly mutated gene [*dnaE693*(Ts) *spq-2*] encodes a temperature-sensitive product and that *spq-2* is not an intragenic suppressor of *dnaE693*. That *spq-2* fails to stabilize this explicitly temperature-sensitive α protein does not preclude the possibility that it could help to stabilize wild-type α . This is one of several hypotheses for the mode of action of *spq-2* discussed elsewhere (7) and is a particularly attractive one because it would form a coherent picture with data showing that ϵ (whose function we suppose *spq-2* mimics) stabilizes wild-type α (8, 17).

General comparison of *dnaE* between *E. coli* and *S. typhimurium*. Like other DNA replication genes whose sequences are known from both *E. coli* and *S. typhimurium* (*dnaA*, *dnaB*, *dnaG*, and fragments of *dnaN* and *dnaQ* [4, 7, 15, 21]), the *dnaE* genes and even more so their encoded protein sequences are extensively similar in the two species. The predicted proteins are identical in length, nearly identical in charge, and identical in 97% of amino acid residues. Among the 36 amino acid differences, many are conservative substitutions, including glutamate for aspartate, lysine for arginine, isoleucine for valine or leucine, and threonine for serine. The distribution of substitutions is not random, since changes are concentrated toward the amino-terminal portion of the protein and are correspondingly scarce in the middle of the protein. For example, there are 23 substitutions within the first 350 residues but only 3 substitutions within the next 350 residues. On average, given the total of 36 differences between the two proteins, about 11 changes would be expected over an interval of 350 residues. The uneven distribution of substitutions may indicate the presence of important functional domains in the most strongly conserved regions, but the small total number of substitutions and their predominantly conservative character make it difficult to put much confidence in this interpretation.

Although the two DnaE proteins are identical in length, they are not perfectly aligned. To maintain the alignment shown in Fig. 1, it was necessary to place a single codon gap in the *E. coli* sequence following codon 159 and to match this with a single codon gap in the *S. typhimurium* sequence following codon 180. Given this adjustment, the intervening sequences encode identical amino acids in 18 of 20 positions. Apparently, α can tolerate such variation in the precise disposition of these amino acid residues, whose role in polymerase activity is unknown.

In *E. coli*, *dnaE* is embedded in a proposed operon containing at least seven genes, including *lpxA* and *lpxB* (1-3, 18). The order of the genes and the sizes of their protein

<i>S. t.</i> protein	AlaLeu GlyLeuValSerTER	(<i>dnaE</i>)	MetSerGlu
<i>S. t.</i> DNA	CTCT-GGGACTTGTGCTGATCTTGTATCGAGA-----AATCTGAAGATGCTCGAA		
<i>E. c.</i> DNA	CACTTGGGACTTGGCTCTGATCTTGTGTCGAGATTAAGTAAACCGGAATCTGAAGATGCTCGAA		
<i>E. c.</i> protein	AlaLeuGlyThrCysValLeuIleLeuValSerArgLeuSerLysProGluSerGluAspValTER	(<i>dnaE</i>)	MetSerGlu

FIG. 4. Nucleotide sequence and translation termination of the ORF upstream of *dnaE* in *E. coli* and *S. typhimurium*. The reading frame for the *S. typhimurium* upstream sequence was chosen on the basis of strong homology at the protein level to ORF₂₃ of *E. coli*. The portion of the sequence shown in the figure (as well as the next 115 nucleotides on the 5' side) was determined on both strands, and the two deletions were visualized clearly on both strands.

products (subscripts, expressed in kilodaltons) are 5'-ORF₇-ORF₁₇-*lpxA*₂₈-*lpxB*₄₂-ORF₂₃-*dnaE*₁₃₀-ORF₃₇. The 7.0-kb fragment of *Salmonella* DNA used in our studies gave rise to four labeled translation products (including the presumptive *dnaE* product) when it was transcribed in vivo by T7 RNA polymerase in one direction (Fig. 2, lane 4) and no products when transcribed in the other direction (data not shown). By similar analysis of smaller fragments related to the 7.0-kb fragment by deletion at one end or the other (Fig. 2, lanes 2 and 3), it was possible to arrive at an unambiguous order for the ORFs in *S. typhimurium*: 5'-ORF₂₅-ORF_{22.5}-*dnaE*₁₃₀-ORF_{18.5}. We know from the sequence determination that the 18.5-kilodalton translation product resulted from truncation of a longer ORF at the *Sma*I site. We did not detect any labeled product similar in size to the *E. coli lpxB* protein, but even in the experiments of Crowell et al. (2), this protein was poorly labeled and ran as a diffuse band. The 7.0-kb fragment is sufficiently large to encode a protein of 42 kilodaltons in addition to the proteins we detected. If such an ORF were to be located between ORF₂₅ and ORF_{22.5}, then the organization of genes and the sizes of gene products would be similar in *E. coli* and *S. typhimurium*. This is our working model for this portion of the chromosome.

Another feature of *dnaE* in *E. coli* in an overlap of five bases between the end of the upstream reading frame (ORF₂₃) and the *dnaE* reading frame. In *S. typhimurium*, this overlap is disrupted by two deletions. These deletions are highlighted in Fig. 4, which shows the sequence through this region and the translation products for *E. coli* and *S. typhimurium*. The more important deletion is that of a single nucleotide 40 bases upstream of the *dnaE* initiation codon. This deletion brings a termination codon into frame after five additional amino acid residues. The other deletion, of 13 nucleotides, does not affect the reading frame since it is in the newly created intergenic region. The net result is that *dnaE* is separated from the upstream ORF by 26 nucleotides in *S. typhimurium*. This difference may be significant for the expression of *dnaE*, since it has been hypothesized that the overlapping reading frames of the genes of this operon in *E. coli* lead to translational coupling of their expression (1, 3, 18). In *S. typhimurium*, such coupling could not occur. Instead, the level of translational activity for *dnaE* message would depend on the exposure and affinity of a ribosome-binding site and might be greater or diminished compared with that of the translationally coupled *E. coli* gene.

ACKNOWLEDGMENTS

We thank Stan Tabor for sending us the T7 expression system.

This work was supported by Public Health Service grants AI19942 and GM39419 from the National Institutes of Health and by a grant from the G. Harold and Leila Y. Mathers Charitable Foundation.

LITERATURE CITED

- Coleman, J., and C. R. H. Raetz. 1988. First committed step of lipid A biosynthesis in *Escherichia coli*: sequence of the *lpxA* gene. *J. Bacteriol.* **170**:1268-1274.
- Crowell, D. N., M. S. Anderson, and C. R. H. Raetz. 1986. Molecular cloning of the genes for lipid A disaccharide synthase and UDP-*N*-acetylglucosamine acyltransferase in *Escherichia coli*. *J. Bacteriol.* **168**:152-159.
- Crowell, D. N., W. S. Reznikoff, and C. R. H. Raetz. 1987. Nucleotide sequence of the *Escherichia coli* gene for lipid A disaccharide synthase. *J. Bacteriol.* **169**:5727-5734.
- Erickson, B. D., Z. F. Burton, K. K. Watanabe, and R. R. Burgess. 1985. Nucleotide sequence of the *rpsU-dnaG-rpoD* operon from *Salmonella typhimurium* and a comparison of this sequence with the homologous operon of *Escherichia coli*. *Gene* **40**:67-78.
- Horiuchi, T., H. Maki, and M. Sekiguchi. 1981. Conditional lethality of *Escherichia coli* strains carrying *dnaE* and *dnaQ* mutations. *Mol. Gen. Genet.* **181**:24-28.
- Kuwabara, N., and H. Uchida. 1981. Functional cooperation of the *dnaE* and *dnaN* gene products in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78**:5764-5767.
- Lancy, E. D., M. R. Lifshits, D. G. Kehres, and R. Maurer. 1989. Isolation and characterization of mutants with deletions in *dnaQ*, the gene for the editing subunit of DNA polymerase III in *Salmonella typhimurium*. *J. Bacteriol.* **171**:5572-5580.
- Maki, H., and A. Kornberg. 1985. The polymerase subunit of DNA polymerase III of *Escherichia coli*. II. Purification of the alpha subunit, devoid of nuclease activities. *J. Biol. Chem.* **260**:12987-12992.
- Maki, H., and A. Kornberg. 1987. Proofreading by DNA polymerase III of *Escherichia coli* depends on cooperative interaction of the polymerase and exonuclease subunits. *Proc. Natl. Acad. Sci. USA* **84**:4389-4392.
- Maurer, R., B. C. Osmond, and D. Botstein. 1984. Genetic analysis of DNA replication in bacteria: *dnaB* mutations that suppress *dnaC* mutations and *dnaQ* mutations that suppress *dnaE* mutations in *Salmonella typhimurium*. *Genetics* **108**:25-38.
- Maurer, R., B. C. Osmond, E. Shekhtman, A. Wong, and D. Botstein. 1984. Functional interchangeability of DNA replication genes in *Salmonella typhimurium* and *Escherichia coli* demonstrated by a general complementation procedure. *Genetics* **108**:1-23.
- McHenry, C. S. 1988. DNA polymerase III holoenzyme of *Escherichia coli*. *Annu. Rev. Biochem.* **57**:519-550.
- Scheuermann, R., T. Schuman, P. M. J. Burgers, C. Lu, and H. Echols. 1983. Identification of the ϵ -subunit of *Escherichia coli* DNA polymerase III holoenzyme as the *dnaQ* gene product: a fidelity subunit for DNA replication. *Proc. Natl. Acad. Sci. USA* **80**:7085-7089.
- Scheuermann, R. H., and H. Echols. 1984. A separate editing exonuclease for DNA replication: the ϵ subunit of *Escherichia coli* DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. USA* **81**:7747-7751.
- Skovgaard, O., and F. Hansen. 1987. Comparison of *dnaA* nucleotide sequences of *Escherichia coli*, *Salmonella typhimurium*, and *Serratia marcescens*. *J. Bacteriol.* **169**:3976-3981.
- Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074-1078.
- Takano, K., Y. Nakabeppu, H. Maki, T. Horiuchi, and M. Sekiguchi. 1986. Structure and function of *dnaQ* and *mutD*

- mutators of *Escherichia coli*. *Mol. Gen. Genet.* **205**:9-13.
18. **Tomasiewicz, H. G., and C. S. McHenry.** 1987. Sequence analysis of the *Escherichia coli dnaE* gene. *J. Bacteriol.* **169**:5735-5744.
 19. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**:259-268.
 20. **Welch, M., and C. McHenry.** 1982. Cloning and identification of the product of the *dnaE* gene of *Escherichia coli*. *J. Bacteriol.* **152**:351-356.
 21. **Wong, A., L. Kean, and R. Maurer.** 1988. Sequence of the *dnaB* gene of *Salmonella typhimurium*. *J. Bacteriol.* **170**:2668-2675.