Antimutator Mutations in the α Subunit of Escherichia coli DNA Polymerase III: Identification of the Responsible Mutations and Alignment With Other DNA Polymerases

Iwona J. Fijalkowska and Roel M. Schaaper

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Manuscript received December 22, 1992
Accepted for publication April 5, 1993

ABSTRACT

The dnaE gene of Escherichia coli encodes the DNA polymerase (α subunit) of the main replicative enzyme, DNA polymerase III holoenzyme. We have previously identified this gene as the site of a series of seven antimutator mutations that specifically decrease the level of DNA replication errors. Here we report the nucleotide sequence changes in each of the different antimutator dnaE alleles. For each a single, but different, amino acid substitution was found among the 1,160 amino acids of the protein. The observed substitutions are generally nonconservative. All affected residues are located in the central one-third of the protein. Some insight into the function of the regions of polymerase III containing the affected residues was obtained by amino acid alignment with other DNA polymerases. We followed the principles developed in 1990 by M. Delarue et al. who have identified in DNA polymerases from a large number of prokaryotic and eukaryotic sources three highly conserved sequence motifs, which are suggested to contain components of the polymerase active site. We succeeded in finding these three conserved motifs in polymerase III as well. However, none of the amino acid substitutions responsible for the antimutator phenotype occurred at these sites. This and other observations suggest that the effect of these mutations may be exerted indirectly through effects on polymerase conformation and/or DNA/polymerase interactions.

DNA polymerase III holoenzyme is the principal replicative enzyme of Escherichia coli. It consists of an asymmetric dimer, made up of 10 distinct subunits (McHenry 1991; Kornberg and Baker 1992). Polymerization and proofreading activities are contained in a complex (termed core) consisting of three subunits, α, ε and θ. While little is known about the θ subunit, the α and ε subunits are involved in maintaining the high fidelity of DNA replication. The dnaE gene encodes the α subunit responsible for the polymerization step (Welch and McHenry 1982), while the dnaQ gene encodes the ε subunit responsible for the proofreading activity (ε  exonuclease activity) (Scheuermann et al. 1983; Scheuermann and Schöls 1984). The mechanisms that control the fidelity of DNA replication are a matter of considerable interest. It has become clear that these mechanisms are tightly interrelated with the mechanism of DNA polymerization (Goodman 1988; Carroll and Benkovic 1990). These mechanisms are being investigated by a number of approaches, including detailed structural and kinetic studies (Joyce and Steitz 1987; Echols and Goodman 1991). Important tools in these studies are mutant versions of DNA polymerases that affect DNA polymerase activities, including its fidelity (Bebenek et al. 1990; Pollesky et al. 1990; Carroll, Cowart and Benkovic 1991; Eger et al. 1991; Pollesky et al. 1992).

We have isolated seven E. coli antimutator mutants, i.e., strains that display lower mutation rates than the wild-type strain (Fijalkowska, Dunn and Schaaper 1993). By PI transduction and complementation the responsible mutations were shown to reside in the dnaE gene. The mutants were isolated in a mutL strain defective in DNA mismatch repair, in which most mutations can be assumed to result from DNA replication errors. On this basis, it was concluded that these strains represent antimutators with an increased accuracy of DNA replication. Each antimutator decreased the mutation rate in mutL and mutT backgrounds but did not appear to affect error-prone SOS mutagenesis. The specificity of two of the antimutators was further investigated by sequencing spectra of lacI mutants (Schaaper 1993). It was shown that both A·T → G·C and G·C → A·T transitions were reduced, but not transversions or frameshifts.

To understand in more detail the mechanisms by which these antimutators reduce DNA replication errors, we have cloned and sequenced the dnaE gene from each of them. For each antimutator, only one amino acid substitution was found, each different and each located in the middle one-third of the protein. The location of the mutated residues may provide...
alleles dnaE911-917 were termed pIF901-907.

and UTH2 were obtained from C. McHenry. All constructs were made by inserting a BamHI site of plasmid pHSG576, a low copy number plasmid, into DNA polymerase I family conferring resistance to chloramphenicol (Takeshita et al. 1987). The ligation products were electroporated into strain E486, but srl360::Tn10, recA56, as The E. coli strains and plasmids used in this study are described in Table 1. Strains E486 and UTH2 were obtained from C. McHenry. All constructions were by P1 transduction using P1virA. NR10109 was constructed by first linking dnaE486 in strain E486 with strain NR9800 as a result of the mapping experiments described previously (Fijalkowska, Dunn and Schaaper 1993). This configuration was first obtained in strain NR9907) and from the wild-type strain JW353 by the method described in Silhavy, Berman and Enquist (1984), and digested with restriction enzyme BamHI. Fragments of ~6.4 kb were isolated from agarose gels and ligated into the BamHI site of plasmid pHSG576, a low copy number plasmid conferring resistance to chloramphenicol (Takeshita et al. 1987). The ligation products were electroporated into strain UTH2 (recA56, dnaE486) and chloramphenicol-resistant transformants selected at 42°C. Plasmids isolated from the transformants were analyzed by agarose gel electrophoresis after digestion with BamHI to check for an insert of the expected size. The presence of the dnaE486 gene was confirmed by a second round of transformation into strain dnaE486. The plasmid containing the wild-type dnaE gene was termed pIF900, and those carrying the antimutator alleles dnaE911-917 were termed pIF901-907.

Papillation experiments: Tests for levels of mutagenesis using the papillation assay were done as before (Fijalkowska, Dunn, and Schaaper 1993). This assay scores the reversion of a galK2 mutation to Gal+ in growing colonies: the number of red papillae within a colorless colony is a measure of the mutation frequency from Gal− to Gal+ (Fijalkowska, Dunn and Schaaper 1993). To test for an antimutator effect of the cloned dnaE alleles in trans, 20 transformants for each plasmid in strain NR10109 (mulT, dnaE486) were toothpicked onto MacConkeyGal plates containing kanamycin and chloramphenicol. The number of papillae was counted after 48 h at 42°C.

DNA sequencing: Double-stranded plasmid DNA was isolated using a kit from Qiagen Inc. (Chatsworth, California). Sequencing of the dnaE genes was on double-stranded DNA, using 2 µg of plasmid DNA per reaction, denatured by the method of Kraft et al. (1988) and sequenced by the dideoxy chain-termination method of Sanger, Nicklen and Coulson (1977) using Sequenase version 2.0 (U.S. Biochemical Corp., Cleveland, Ohio). Nineteen 32P-end-labeled primers were used, spaced approximately 200 bases apart and chosen based on the known sequence of the dnaE gene (Tomasiwicz and McHenry 1987).

Polymerase alignment: A search for the preserved amino-acid motifs A, B and C as described by Delarue et al. (1990) was performed manually. Motif B (K...·...YG) was found at residues 674-682 (KPVLEFTYG) (see Figure 2). This motif could be assigned confidently since no other sequence containing the K and YG residues separated by six or seven residues is present in polymerase III. Motif A was then located by searching the region between 30 and 15 D residues of the polymerase α or polymerase I family as described by Delarue et al. (1990). Of the 15 D residues in this stretch, D524 appeared the most plausible since it contained the highly conserved residues Y at position +2 and H at position +12 (relative to D, see Figure 2). The nearby residue D542, although lacking the above residues, may be an alternative possibility based on some homology to the polymerase I family (S at position +2, E at position +5). Motif C was located by inspecting the 200 amino acids to the right of motif B for either a DXD or DE sequence. A DMD sequence was found at position 790, a DE sequence at 802, either of which could represent motif C. Because of greater homology to the polymerase α family in case of motifs A and B, the DMD sequence was preferred. The

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pHSG576</td>
<td>dnaE from JW353</td>
<td>This work</td>
</tr>
<tr>
<td>pIF900</td>
<td>dnaE from NR9907</td>
<td>This work</td>
</tr>
<tr>
<td>pIF901-907</td>
<td>dnaE from NR9901-9907</td>
<td>This work</td>
</tr>
</tbody>
</table>

**MATERIALS AND METHODS**

**Strains and plasmids:** The E. coli strains and plasmids used in this study are described in Table 1. Strains E486 and UTH2 were obtained from C. McHenry. All constructions were by P1 transduction using P1virA. NR10109 was constructed by first linking dnaE486 in strain E486 with strain NR9800 as a result of the mapping experiments described previously (Fijalkowska, Dunn and Schaaper 1993). This configuration was first obtained in strain NR9800 as a result of the mapping experiments described previously (Fijalkowska, Dunn and Schaaper 1993).

**Cloning antimutator mutations:** Chromosomal DNA was isolated from the seven antimutator strains (NR9901-9907) and from the wild-type strain JW353 by the method described in Silhavy, Berman and Enquist (1984), and digested with restriction enzyme BamHI. Fragments of ~6.4 kb were isolated from agarose gels and ligated into the BamHI site of plasmid pHSG576, a low copy number plasmid conferring resistance to chloramphenicol (Takeshita et al. 1987). The ligation products were electroporated into strain UTH2 (recA56, dnaE486) and chloramphenicol-resistant transformants selected at 42°C. Plasmids isolated from the transformants were analyzed by agarose gel electrophoresis after digestion with BamHI to check for an insert of the expected size. The presence of the dnaE486 gene was confirmed by a second round of transformation into strain dnaE486. The plasmid containing the wild-type dnaE gene was termed pIF900, and those carrying the antimutator alleles dnaE911-917 were termed pIF901-907.

**Papillation experiments:** Tests for levels of mutagenesis

**TABLE 1**

**Bacterial strains and plasmids**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E486</td>
<td>thi, leu, thy, lac, rpsL, met, tonA, supE, dnaE486</td>
<td>Welch and McHenry (1982)</td>
</tr>
<tr>
<td>UTH2</td>
<td>as The E486, but srl360::Tn10, recA56</td>
<td>Welch and McHenry (1982)</td>
</tr>
<tr>
<td>JW353</td>
<td>thr-1, leuB6, zar-502::Tn10, thy6, met-89, thi-1, desC1, lacY1, rpsL67, tonA21, X−, supE44</td>
<td>B. Bachmann</td>
</tr>
<tr>
<td>NR9355</td>
<td>ara-9, fhuA1, ts-3, supE44, galE2, X−, hisG4(Oc), rfbD19, thr-3(Oc), rpsL8 or rpsL9, malA1(λ), metE64, thi-1</td>
<td>Fijalkowska, Dunn and Schaaper (1993)</td>
</tr>
<tr>
<td>NR9606</td>
<td>ara-9, fhuA1, ts-3, supE44, galE2, X−, hisG4(Oc), rfbD19, thr-3(Oc), rpsL8</td>
<td>NR9606, but dnaE486, zar-502::Tn10</td>
</tr>
<tr>
<td>NR9901-9907</td>
<td>ara-9, fhuA1, ts-3, supE44, galE2, X−, hisG4(Oc), rfbD19, thr-3(Oc), rpsL8</td>
<td>NR9901, but mutL::Tn5, lacZ118(Oc)</td>
</tr>
<tr>
<td>NR9900</td>
<td>ara-9, fhuA1, ts-3, supE44, galE2, X−, hisG4(Oc), rfbD19, thr-3(Oc), rpsL8</td>
<td>NR9900, but dnaE911-917, zar-502::Tn10, zar::Tn10d-Cam</td>
</tr>
<tr>
<td>Plasmids</td>
<td>pSC101 replicon, pUC9-polylinker, Cam*</td>
<td>Takishita et al. (1987)</td>
</tr>
<tr>
<td>pIF900</td>
<td>pHSG576 [dnaE from JW353]</td>
<td>This work</td>
</tr>
<tr>
<td>pIF901-907</td>
<td>pHSG576 [dnaE911-917 from NR9901-9907]</td>
<td>This work</td>
</tr>
</tbody>
</table>
greater overall homology to the polymerase α than the polymerase I family is further supported by the presence of six (rather than seven) residues between the K and YG residues in motif B, as well as the general intermotif distances (see Figure 2).

RESULTS AND DISCUSSION

Cloning of the antimutator genes: To determine the nature and location of the previously isolated antimutator mutations, which had been mapped to the dnaE gene (FiJALKOWSKA, DUNN and SCHAAPER 1993), we cloned the dnaE gene from the seven strains carrying the antimutator mutations and from the wild-type strain as a control. Previous studies describing the cloning and DNA sequence of the wild-type dnaE gene (WELCH and MCHENRY 1982; TOMASIEWICZ and MCHENRY 1987) revealed that it is about 3,480 nucleotides long and resides on a 6.4-kb BamHI chromosomal fragment. We therefore isolated BamHI DNA chromosomal fragments of this size, cloned them into the BamHI site of the low copy number plasmid pHS576 (TAKESHITA et al. 1987), and selected for restored temperature resistance of the dnaE486(Ts) strain. In this manner, we obtained eight plasmids pIF900, carrying the wild-type allele and pIF901–pIF907 carrying the respective antimutator alleles dnaE911–dnaE917.

To confirm that we had cloned the expected dnaE alleles, the plasmids were transformed into strain NR10109 (mutL, dnaE486). The resulting strains were analyzed using the papillation assay on MacConkeyGal plates at 42°. At this temperature, the only functional α subunit is provided by the antimutator gene residing on the plasmid, and this method allows one to observe the antimutator effect in trans. Compared to the strain carrying the wild-type dnaE gene, a 3–10-fold antimutator effect was produced by the dnaE antimutator alleles (Figure 1).

DNA sequence analysis: The seven cloned dnaE antimutator genes as well as the wild-type gene were sequenced in their entirety, starting from the putative dnaE promoter region (TOMASIEWICZ and MCHENRY 1987) located some 200 bases upstream of the coding sequence. The locations of the mutations are shown in Table 2. For each antimutator, one amino acid substitution was observed for dnaE911 (nucleotide 2236, CTT → TGT) and dnaE914 (nucleotide 2601, CGC → CGT). The mutated base is underlined. The numbering system is as in TOMASIEWICZ and MCHENRY (1987). Additional (but silent) base-pair changes were observed for dnaE911 (nucleotide 2236, CTT → TGT) and dnaE914 (nucleotide 2601, CGC → CGT).

The mutated base is underlined. The numbering system is as in TOMASIEWICZ and MCHENRY (1987). Additional (but silent) base-pair changes were observed for dnaE911 (nucleotide 2236, CTT → TGT) and dnaE914 (nucleotide 2601, CGC → CGT).
responsible for the observed effect.

Lancy et al. (1989) determined the nucleotide sequence of the dnaE gene from the closely related bacterium Salmonella typhimurium and compared the resulting amino acid sequence to that of E. coli. Among the 1,160 residues, 30 amino acid substitution differences were noted, none of which resided in the central third of the gene. These and our data suggest that the central third of the protein is the region most important for DNA polymerase function, including its accuracy.

**Amino acid alignment of DNA polymerase III with other DNA polymerases:** Since very little is known about the functional regions of DNA polymerase III, a comparison was undertaken to other DNA polymerases. Polymerases from a large number of sources have been aligned using amino acid homologies to reveal regions of possible functional significance (Jung et al. 1987; Wang, Wong and Korn 1989; Delarue et al. 1990; Blanco et al. 1991; Ito and Braithwaite 1991; Blanco, Bernad and Salas 1992). Several families of related polymerases have been described, of which the most prominent are the eukaryotic polymerase α family and the E. coli polymerase I family (Jung et al. 1987; Delarue et al. 1990).

E. coli DNA polymerase III has generally not been included in such analyses, presumably because it displays little or no homology to the other enzymes, except its close relative S. typhimurium (Lancy et al. 1989). Limited homology has also been reported to DNA polymerase III from Bacillus subtilis (Hammond et al. 1991; Ito and Braithwaite 1991) and a separate family for the three enzymes has been suggested (Ito and Braithwaite 1991).

We have used the alignments proposed by Delarue et al. (1990) since these authors, by emphasizing the most critically preserved residues, were able to bridge the gap between the eukaryotic polymerase α and the E. coli polymerase I family, and were also able to include even more distantly related enzymes such as DNA polymerase β, DNA-dependent RNA polymerases, RNA-dependent DNA polymerases, and RNA dependent RNA polymerases. Three highly conserved sequence motifs were deduced, termed A, B and C, which occurred in the same linear order in each primary sequence and were proposed to constitute a universal polymerase “core” (see Figure 2). The relevance of the motifs is substantiated by their location in the crystal structure obtained for E. coli DNA polymerase I (Ollis et al. 1985; Joyce and Steitz 1987; Delarue et al. 1990) and site-specific mutagenesis experiments altering the critical residues (Polesky et al. 1990, 1992). On the assumption that the three “core” motifs would also be present in polymerase III, we inspected its amino acid sequence for their presence (see MATERIALS AND METHODS). Indeed, the three motifs could be found, as shown in Figure 2.

The crystal structure derived for E. coli DNA polymerase I has allowed a tentative assignment of functional domains (Ollis et al. 1985; Joyce and Steitz 1987). The salient feature is the presence of an extended cleft, which is presumed to hold the DNA in its proper orientation for the chain-elongation reaction. Motifs A, B and C are close together on the inner surface of the cleft and constitute essential components of the catalytic site (Delarue et al. 1990). The regions outside or in between the motifs correspond roughly to the “finger, thumb and palm” areas of the polymerase (Kohlstaedt et al. 1992) that define the polymerase cleft and that may be expected to be in contact with the DNA. In Figure 3, we have placed the seven antimutator mutations on the linear dnaE sequence relative to the three “core” motifs. Four substitutions are located on the amino-terminal side of motif A, one between motifs A and B, and the remaining two between motifs B and C. Thus, none of the mutations occurs directly at the sites most directly implicated in catalysis (at least by extrapolation from DNA polymerase I). These findings suggest that the effects of the antimutator mutations may not be exerted by a direct interference with the catalysis of the reaction but, instead, indirectly (see below).

**Possible antimutator mechanisms:** We have previously mentioned three different (although not mutually exclusive) ways by which mutations in the DNA polymerase could reduce error rates (Fijalkowska, Dunn and Schaaper 1993). First, the antimutator mutations may improve base selection. Second, altered interaction or coordination between the α (polymerase) and ε (editing) subunits could improve discrimination via the proofreading step. This is possible since the exonuclease acts in coordination with the polymerase (Maki and Kornberg 1987; Studwell and O’Donnell 1990; Reems, Grief and McHenry 1991) and does not, by itself, discriminate between correctly and incorrectly paired termini (Brenowitz et al. 1991). Relevant to this type of mechanism may be the dnaE173 mutator mutation which resides in the polymerase subunit, but whose mutator phenotype reportedly results from a proofreading deficiency (Maki, Mo and Sekiguchi 1991). Third, the mutant DNA polymerases may simply be catalytically less competent, having an increased probability of stalling at certain points in the DNA replication cycle (Kuchta, Benkovic and Benkovic 1988). Such stalling (or slow down) would provide increased time for the exonuclease to act, leading to higher fidelity. This type of mechanism has been proposed for the T4 antimutator polymerase CB120, which is defective in strand displacement (in addition to a presumed increase in accuracy at the insertion step) (Gillin and
are from DNAE antimutators were found at higher frequency than proficient ones (mutators). On the other hand, it is

described above, is, in the simplest view, most easily reconciled with the last model. Obviously, catalytically impaired mutants may be obtained throughout the enzyme. Furthermore, the location of the mutations relative to three core motifs indicates that they may be part of the polymerase molecule that forms the cleft. This suggests that they may affect protein-DNA interactions of which many occur throughout the cleft (POLESKY et al. 1992). Alternatively, or in addition, they may affect DNA polymerase conformation(s). The mechanism would also be most compatible with the observation (FJALKOWSKA, DUNN and SCHAAWER 1993) that during the isolation of the new DNAl derivatives, DNAl antimutators were found at higher frequency than DNAl mutators. Obviously, within the framework of a “sailing” model, catalytically impaired mutants (antimutators) would be more readily obtained than more proficient ones (mutators). On the other hand, it is

less clear how this mechanism can account for the defined specificity displayed by some of the antimutators (SCHAAWER 1993). For example, one antimutator reduced transition errors, but increased transversion and frameshift mutations. Thus, multiple mechanisms may play a role. It is our intention to further address the question of the antimutator mechanisms by analyzing the properties of purified antimutator α subunits in vitro.

We thank K. BERENEX, J. DRAKE, J.-Y. MO and A. OLLER of NIEHS and Y. PAVLOV, University of Sankt Petersburg for helpful comments on this manuscript.

LITERATURE CITED


Figure 2.—Alignment of E. coli DNA polymerase III (middle) with various other DNA polymerases in the region of motifs A, B and C, as defined by DELAIRE et al. (1990). The motifs for DNA polymerase III were derived as described in MATERIALS AND METHODS. The group of polymerases at the top represents DNA polymerases of the human polymerase I type. Amino acid residues that are most strictly conserved are in bold and other highly conserved residues are underlined. The amino acid sequences were taken from the SWISS-PROT data base and aligned as described in DELAIRE et al. (1990). The polymerase α type sequences are from human DNA polymerase α, Autographica calijornia nuclear polyhedrosis virus, bacteriophage T4, Epstein-Barr virus, cytomegalovirus, vaccinia virus, varicella-Zoster virus, and E. coli DNA polymerase II. The polymerase I type sequences are from E. coli, S. pneumoniae, Thermus aquaticus and bacteriophages SPO2, T5 and T7. The numbers indicate the number of amino acids that precede or reside between the presented sequences.

Figure 3.—Location of amino acid substitutions in DNAl antimutator mutants relative to the three “core” motifs (A, B and C) as defined by DELAIRE et al. (1990). The number below the motifs indicates the position of the invariant aspartate (D), lysine (K), and aspartate (D) in these motifs, respectively (see Figure 2).

Nossal 1976a, b).