Identification and Sequence of the *drpA* Gene from *Escherichia coli*

ZONGHAN ZHOU AND MICHAEL SYVANEN*

Department of Medical Microbiology and Immunology, School of Medicine, University of California,
Davis, California 95616

Received 24 April 1989/Accepted 11 October 1989

The *drpA* gene of *Escherichia coli* encodes a factor that is involved in global RNA synthesis. We establish that the *drpA* gene has been successfully cloned and describe the fine-structure map of three *drpA*-(Ts) mutations as well as the complete nucleotide sequence of the *drpA* gene. We identified a major sigma-70 promoter for the *drpA* gene on the bases of (i) its similarity to the consensus sequence and (ii) S1 protection and primer extension mapping data. In addition, the nucleotide sequence revealed a pair of *dnaA* boxes and a factor-independent terminator at the 5' end and 3' end of the gene, respectively. The deduced amino acid sequence of the DrpA protein showed a nucleotide-binding pocket found in some ATPases.

The *drpA* gene encodes a product that is required for the synthesis of RNA and DNA in *Escherichia coli*, though it appears to be needed primarily for RNA synthesis. The gene and its role were revealed by a temperature-sensitive mutation, *drpA1*(Ts), that maps at 5.2 min (7). The phenotype of the *drpA1*(Ts) mutant that indicates its involvement in RNA synthesis is the cessation of uridine incorporation into new RNA when cells growing at 30°C are shifted to 42°C even though the uridine is still taken up by the cells and converted to UTP at 42°C. This cessation of RNA synthesis is quickly followed by the cessation of DNA and protein syntheses (7). The major question posed by these properties is whether *drpA* is directly or indirectly involved in transcription. In an effort to gain a greater understanding of the role of *drpA*, we have cloned a gene from *E. coli* (creating the plasmid pKFL2) that restores growth to the *drpA*(Ts) mutants. All derivatives of pKFL2 that synthesize a 68,000-molecular-weight (MW) protein restore growth at 42°C to the *drpA*(Ts) mutants, whereas insertions in pKFL2 that eliminate this protein do not. In this paper, we establish that pKFL2 carries the *drpA* gene, determine the complete nucleotide sequence of that gene, and identify the coding frame for the DrpA protein.

MATERIALS AND METHODS

Strains. Table 1 lists the bacterial strains used in this study. The 2.8-kilobase (kb) DNA fragment from the *Bam*HI-to-*Hind*III site of plasmid pKFL2 was subcloned into M13 bacteriophage sequencing vectors Mp18 and Mp19 (15) (Fig. 1). Mutations *drpA2*(Ts) and *drpA3*(Ts) were independently isolated by the procedure described for the isolation of *drpA1*(Ts) (7). These two temperature-sensitive alleles were tentatively assigned to the *drpA* gene, since plasmid pKFL2 restored growth at 42°C when transformed into strains with these alleles.

Marker rescue and complementation tests. The M13 derivatives, whose insert sizes are shown in Fig. 1, were spotted onto soft-agar lawns containing 0.1 ml of a saturated culture of *E. coli* SY1164.1, SY1164.106, or SY356.11 (Table 1). The plates were incubated overnight at 30°C to allow the development of a phage spot. About 10⁶ infected bacterial cells from the spot were then plated on a tryptone-yeast agar plate and grown at 42°C. When phage M13 alone was inserted into

the strain carrying *drpA1*(Ts) or *drpA3*(Ts), no viable bacterial colonies resulted. However, when phage M13 was inserted into the strain carrying *drpA2*(Ts), numerous small colonies resulted, reflecting the leakiness of *drpA2*(Ts) at 42°C. When the M13 phage derivatives 18-2 and 19-3 were infected into any of these three strains, confluent growth at 42°C was observed; this was scored as a positive complementation test. With the other phages, either no colonies or a few dozen to a few hundred colonies were seen at 42°C. If no colonies were seen, we concluded that there was no recombination between the DNA fragment and the given *drpA*(Ts) allele. If colonies were seen, we concluded that the *drpA*(Ts) allele and the DNA fragment overlapped. For the leaky *drpA2*(Ts) mutant, the wild-type colonies easily outgrew the mutant colonies and could therefore be distinguished from them.

DNA sequencing. The DNA fragments were sequenced by the dideoxy-chain termination reaction with T7 DNA polymerase provided by Sequenase. The dideoxy-terminated DNA fragments were separated on an 80-cm polyacrylamide gel. Nucleotide and deduced amino acid sequences were analyzed with Genetics Computer Group software provided by the University of Wisconsin.

S1 protection assay. Total RNA was extracted from cells carrying plasmid pKFL2 as described elsewhere (11). The RNA was hybridized to the DNA probe described in Results, and the hybrid was digested with the S1 single-stranded endonuclease as described previously (1). The digestion temperature was 45°C, and the samples were removed at different times between 5 and 80 min.

Primer extension assay. Total RNA from cells carrying pKFL2 (11) was annealed to the oligonucleotide primer and reacted with reverse transcriptase as described elsewhere (2).

RESULTS

Fine-structure map of *drpA*(Ts) mutations. Strains with mutations *drpA1*(Ts), *drpA2*(Ts), and *drpA3*(Ts) do not grow at 42°C. However, growth is restored by the presence of plasmid pKFL2, which is pBR322 carrying a 2.8-kb insert from wild-type *E. coli*. pKFL2 restores the growth of *drpA1*(Ts) because it carries a gene whose product is a 68,000-MW protein. This is inferred because pKFL2 carrying Tn5 insertions that prevent synthesis of the 68,000-MW protein fails to restore growth of the *drpA1*(Ts) mutant at 42°C (7). The most likely possibility is that pKFL2 carries a

* Corresponding author.

TABLE 1. Strain list

Strain, plasmid, or phage	Genotype	Reference
Bacterial strain		
JM101	$\Delta(lac\ pro)\ supE\ thi/F'\ traD36\ lacI^a$	15
SY1164.1	JM101 <i>drpA</i> (Ts) <i>zae::Tn10</i>	This study
SY1164.106	JM101 <i>drpA</i> (Ts) <i>zae::Tn10</i>	This study
SY356.110	<i>ara</i> $\Delta(lac-pro)XIII\ argE(Am)\ nalA\ rif\ thi\ thyA\ strA\ drpA3(Ts)/F'\ lacZ::Tn5\ pro^+$	7
Plasmid		
pKFL2	pBR322:: <i>drpA</i> ⁺	7
Phage		
Mp18	M13::cloning vector with polylinker	15
Mp19	M13::cloning vector with polylinker	15

wild-type copy of the *drpA* gene and that the *drpA*(Ts) mutations are recessive and complemented by the plasmid. The alternative explanation is that pKFL2 carries some other gene that, when carried on a high-copy-number plasmid, bypasses or suppresses the *drpA*(Ts) defect. To establish that pKFL2 does in fact carry the *drpA* gene, we demonstrate here that the inserted DNA sequences in pKFL2 can recombine with the chromosomal *drpA*(Ts) mutations and restore the wild-type phenotype. Indeed, this permits us to determine the map order of the three *drpA*(Ts) mutations.

The initial criterion for determining that the three mutations are in *drpA* is that the mutants fail to grow at 42°C but

their growth at 42°C is restored when they carry pKFL2. In preparation for sequencing the insert in pKFL2, a series of subclones was made in the M13 phage vectors Mp18 and Mp19. The physical maps of the insert in pKFL2 and the various M13 subclones are shown in Fig. 1A and B, respectively. The subclones are of two types. When an M13 subclone of the first type was infected into a *drpA*(Ts) mutant, the infected cells grew at 42°C; when the second type of subclone was inserted, *drpA*(Ts) cells did not grow at 42°C. The first type of subclone (18-2 and 19-3) contains the entire DNA sequence that encodes the 68,000-MW protein known to restore growth of *drpA*(Ts) at 42°C; the second type contains only fragments of that region. However, from some populations of the second type of infected cells we recovered cells that grew at 42°C; these appeared at a frequency of 1×10^{-2} to 5×10^{-4} per cell. In other words, the fragment on the M13 subclone recombined with the chromosomal mutation, thereby rescuing the wild-type phenotype.

The marker rescue test allowed us to map each of the *drpA*(Ts) mutations against each M13 subclone. The results of this mapping are given in Fig. 1C. Those subclones that recombine with the corresponding *drpA*(Ts) mutations can be identified as those with a region that lies above a region indicated by a stippled box for each indicated mutation (Fig. 1). For *drpA*1(Ts) and *drpA*3(Ts), the results of this mapping were unambiguous: all subclones that gave wild-type recombinants lie above a region indicated by a stippled box. For *drpA*2(Ts), on the other hand, the subclones P2, R1, and R4 failed to give wild-type recombinants; however, on the basis of the pattern of recombination with the other subclones, we would expect either R1 and R4 or P2 to yield recombinants. The one feature common to the three subclones is that one

FIG. 1. Physical map of the *drpA* gene. (A) The 2.8-kb fragment with some relevant restriction sites. B, *Bam*HI; H, *Hind*III; R, *Eco*RI; Pv, *Pvu*II; Ps, *Pst*I. (B) The 2.8-kb fragment shown in panel A was subcloned into the phage M13 sequencing vectors Mp18 and Mp19 (15). The extents of the different subclones and their isolation numbers are shown. Phages with the 18-2 and 19-3 inserts complement the mutants carrying *drpA*1(Ts), *drpA*2(Ts), and *drpA*3(Ts), while the others do not. The arrows on the lines are the regions for which clearly resolvable sequences that provide the sequence shown in Fig. 2 were obtained. The position for the primer is indicated (●). (C) Some of the phage vectors recombined with the *drpA*(Ts) mutations to yield wild-type recombinants and are shown above the corresponding stippled boxes.

end of each subclone is the same *EcoRI* site. One explanation for this result is that the *drpA2*(Ts) mutation lies too close to the *EcoRI* site to permit recombination when the DNA fragment ends at that site. Alternatively, R1 and R2 might fail to rescue *drpA2*(Ts) because they are too short.

These mapping data allow us to conclude that the *drpA* gene is in fact carried on pKFL2, and we can describe the ability of pKFL2 to restore growth to the *drpA*(Ts) mutations as complementation. Thus, the three *drpA*(Ts) mutations are recessive. This also establishes that the temperature-sensitive *drpA* mutants fail to grow at 42°C because they are missing an essential function at that temperature.

Complete nucleotide sequence of the *drpA* gene. The DNA fragments diagrammed in Fig. 1 were sequenced by the dideoxy-chain termination reaction. The complete nucleotide sequence of the 2.8-kb fragment was determined, and within this sequence the *drpA* gene can be located with reasonable certainty. There is one long open reading frame (ORF) that extends from base pair (bp) 645 to 2192. Three Tn5 insertions in pKFL2 that eliminate *drpA* activity and eliminate the 68,000-MW protein lie within this long ORF, and two mini-Mu-*lacZ* insertions that produce active β -galactosidase-protein fusions (M. Syvanen, unpublished data) are inside and correctly oriented with this long ORF. One of the three Tn5 insertions elaborates a polypeptide that is 19% smaller than the wild-type DrpA protein (unpublished data); this insertion lies in the ORF near its C-terminal end.

The set of 516 codons from the long ORF displays the rare codon bias typical for *E. coli*. For example, none of the 29 isoleucines use ATA, none of the 27 prolines use CCC, only 2 of 34 arginines use CGG or CGA, and 35 of 41 leucines use CTG. The long ORF encodes a protein with a net charge of -11, consistent with the acid isoelectric point of the DrpA protein identified in maxicells (C. Georgopoulos, personal communication). The only inconsistency with our earlier data on the DrpA protein is that this long ORF codes a 57,300-MW protein, whereas the measured MW of maxicell proteins as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis is 68,000. Although not common, such a discrepancy is not unprecedented.

The deduced amino acid sequence of the DrpA protein was examined for sequence homology to other proteins and for the presence of a variety of sequence motifs. We found no significant homology between the DrpA protein and any of the protein sequences maintained by the National Biomedical Research Foundation (available through GenBank; updated version, September 1988), nor was there any homology between the nucleotide sequence of *drpA* and the GenBank sequence data base. We did find the consensus of a nucleotide-binding pocket; this is an e1r---aG-G--G-V sequence beginning at position 1617 (Fig. 2). The lowercase letters in this sequence are from the *drpA* sequences that are not part of the consensus sequence but are encountered in some protein kinases (4). It seems probable that DrpA protein binds ATP and possible that it is an ATPase. In addition, the DrpA protein could be a phosphotransferase, since the VIDR sequence that begins a sequence of seven residues on the C-terminal side of the nucleotide-binding pocket is encountered on some phosphotransferases.

Identification of the DrpA promoter. Preceding the DrpA protein beginning at bp 551 is the sequence --TTGACA(N)₁₅TAAACT--, which has the appearance of an active sigma-70 promoter. This sequence scores a high 54% by the algorithm of Mulligan et al. (14), which is designed to test the significance of a putative promoter sequence. To show that this sequence is in fact used to initiate transcrip-

tion, we extracted mRNA from cells carrying pKFL2 and mapped the 5' end of the DrpA message by the S1 nuclease protection (1) and primer extension (2) techniques.

In the S1 procedure, we prepared a DNA probe that extended from the *PstI* site at position 299 to the *PvuI* site at position 842, where the 5' phosphate at position 842 was labeled with ³²P. If the sequence identified above is the promoter, then the resulting transcript should begin very close to position 583 and therefore its resulting mRNA should protect a 232-nucleotide region of the probe from S1 digestion. When we performed this digestion, we found that the protected probe was 220 nucleotides long, 12 bases shorter than expected. It seems possible that the probe is 12 bases shorter than expected because the S1 nuclease digested some of the double-stranded RNA-DNA probe from the unlabeled end. This could occur because, if the position near 583 is the first base of the message, then the 5' end of the resulting RNA would have 12 A's or U's in the first 14 nucleotides, and thus a low-melting-temperature duplex would be formed. The notion that the S1 nuclease digests some double-stranded ends and produces an artifactually shorter measure of the actual message is supported by the experiment whose results are shown in Fig. 3. In this experiment, we varied the digestion time of the RNA-DNA probe hybrid from 5 to 80 min; the 80-min time is from the standard protocol. As can be seen with shorter times of digestion, the protected fragment is longer and approaches a size consistent with the A or C at position 583 or 584, respectively. In fact, in the ladder of bands in the 5-min lane (Fig. 3), a very faint band corresponding to this A or C can be seen.

To resolve the uncertainty of the S1 mapping, we used a primer extension assay (2) to measure the 5' end of the DrpA message. A ³²P-labeled 21-nucleotide primer (with the sequence CGGTCGGCAGCCAGGTATATA) from nucleotides 780 to 760 was annealed to cellular RNA, and DNA fragments complementary to the 5' end of the message were synthesized with reverse transcriptase. The resulting fragments were resolved on a DNA sequencing gel (Fig. 4, lane 1). The size standard for this gel was the dideoxy sequencing ladder produced from the primer and an appropriate template. As can be seen, the largest band from this experiment corresponds to a 5' end that begins with the C at residue 584. This supports the conclusion from the S1 protection assay that the sequence between positions 550 and 580 is a major promoter for the *drpA* gene. The multiple bands below position 584 are probably due to homogeneity in the mRNA preparation and to prematurely terminated reverse transcriptase reactions. For example, the strong bands ending at positions 613 and 616 may be the results of an RNA secondary structure, since position 616 represents the base of a G-C-rich 8-bp stem-loop structure that could form in a longer RNA molecule. The results illustrated in Fig. 3 and 4 point to position 584 as representing the 5' end of the major RNA for the *drpA* gene. It is possible that a shorter, minor message is synthesized from this region; if so, it is not detected by both of the techniques whose results are shown in Fig. 3 and 4. We therefore conclude that the DrpA message begins at the C at position 584 and that the -35 region of the promoter begins with the sequence TTGACA at position 551.

***dnaA* boxes found 5' to DrpA promoter.** On the basis of an examination of sequences on the 5' side of the DrpA promoter, two regions are possibly *dnaA* boxes. Sequence TTATCCGCA at position 363 and sequence TTATCCGCG at position 383 are similar to the consensus *dnaA* box of

1 GGATCCTTTTCTGCTTTTCATCAAAAGCATGGAAGCGGCTGGCGTCCAGTGCGTCCACCGCGCTCGCGGTAACCGAAGTGGGGTTTCGCAACA
101 CGCTCTACTTTCCGCCAATTAACCAATTGGCATGTCGCTGGTAGAGTGAAAGAGGTGTTTGGCATAAAGACAGCGTATTCTGAAGCTCGGCAGTCTGG
201 ATCTGGTCGATGGTACGCCGTAGTGGATATCAAAACGATATCCCTTTGCCGAATCGCTTCCGATGCCAGTGCCAGCTATCGAAAGTGGCGCAGCT
301 GCAGAGATGGCGGTGAGTTTACCAGCAGAGGTGAAAGCAGCTTTTGACGCTGGAGAAGCGTTATCCGAGTTAACGCTGTTTATCCGCGAAGTACTGG
401 CGCAGGACCCGCGCCGGCTATCGTAAAGTGAGGAAACGGGCAAAACCTATGCCGTCTGGCTACATGATTTTAACTGTCGGCGGTACCGACGCA
501 GGTTTTGAAGTCTTTGCGCTGGAACCGGTTAAATTCACGCCCTTCTTTTGGACATTTCTTTTGGCACTGGTAACTAAATCACTTTTTTTTGTCCAGG
601 CTCGCCCTGAGGCTGTTCTACCTTCAACTGGAACCGTAACAACTGCGTACTAGCAATACCTGCTCTCCACTCTCAAGGAGACACCTGCCGACGCCGA
701 GGTGATCAGCCATCAGCTGATGCTGCGCGCGGGATGATCCGCAAGCTGGCTCCGGTTATATACCTGGCTGCCGACCGCGTGGCTTCTGAAAAA
801 GTCGAAACATCGTGGTGAAGAGTGAACACCGCGGTGCGATCGAGGTGTCGATCGCGGTGTCAGCCAGCGGATTTGTGGCAAGAGAGTGGTGGT
901 GGAACAGTACGCTCGGAACTGCTGCTTTTGTGACCGTGGCGAGCGTCCGTTCTACTCGGCCCACTCATGAAGAAGTTATCACTGACCTGATTG
1001 TAACGAGCTTAGCTTCAAAACAGTGGCGCTGAACCTTCTATCAGATCCAGACCAAGTTCCGCGACGAAGTGGCTGCGGTTTCCGGCTCATGGCTTC
1101 CGCGAATTCCTGATGAAGATGCTTCTTCCATCTTCTCAGGAATCCCTCAGGAAACCTACGATGCAATGATGCGGCTACAGCAAAATCTTCA
1201 CGCGATGGGGTGGATTTCGCGCGCTACAAGCGACACCGGTTCTATCGCGCGGACGCTCTCAGCAATTCAGGTGCTGGCGCAGAGGTGAAGACGA
1301 TGTGGTCTTCTCGACACCTCTGACTATGACGAGCACTGAACTGGCAGAAGCTATCGCGCGGAAAGAACCGCGCTGCTGCTACCGAGGAATGACG
1401 CTGGTGTATACGCGCAACGCGAAACCATCGGAACTGGTGAACAGTTCAATCTGCGGATTGAGAAACGGTTAAGACTCTGCTGGTTAAAGCGGTTG
1501 AAGCGACGAGCTTCCGCGAGTGGCTGCTGGTGGCGGTGATCAGGCTGAACGAAGTTAAAGCAAAACCTGCCGAGGTGCAACCGCGCTGAC
1601 TTTCGACCGCAAGAAAGAAATTCGTCGCGGTGTTAAAGCGGTCGCGGTTCTACGGTCCGGAATACGCGATTCCGGTGGTATGACCTACCGTT
1701 GCGCGATGAGTGATTTCGCTGCTGGTGAATCATGATGTTAACTACTTCTGGCATCACTGGGATCGGATGTCGCTACCCGGAAGTTGACAGATA
1801 TCCGTAACGTGGTGGTGGCGATCAAGCCGGATGGCGAGGTAGGCTGCTGATCAACGTGGTATCGAAGTTGGTCACATCTCCAGCTGGGTACCAA
1901 GTACTCGAAGCACTGAAAGCCTCCGTACAGGGTGAAGATGGCGTAACCAATCTGACGATGGTGTGCTACGGTATCGGGTAAACGCTGTGGTAGT
2001 GCGCGATTGAGCAGAACTACGACGAACGAGGCTATGCGCTGACGCTATCGCGCGGTTCCAGGTGGCGGATCTGCGCATGACATGCACAACTCT
2101 TCCGCTACAGAGCTTCTGAGAACTGTACAGCACTGCTGCAAGGTATCGAAGTGTGCTGGATGACCGCAAGAGCGTGGCGGTGATGTTT
2201 GCTGATGGAAGTATCGGCTATTCGCGACACATATTGCTGGCGGACCGTAACCTGACACGACGATATGAAATATAATATGCTGCAACGCGGAGA
2301 AACAGTTAATTAAGACTGGTGACATCTCGAATATCTGGTTGAAACAGATTAAGGCTGATGCCAGAAAGGCTCTGAATTTCAAGGCCCTTTTTTACA
2401 TGGATTGTATAAATGGATAAACCAGGCGTACTGCCGCTGTTTCTCCCGACTTCTTTTACTTTCCGCTGTACCGTTGATATCAGCCAGCCAGAC
2501 CCATCAGCCACAGCAGTAGATGCCGAGGCAAAACATGGGCTGTCAAAATCCAGCATCAAGCTCTTTCCACGAACATCAATAAAGAAATAACGGCAC
2601 CTGATCTCAAAACCGCGGATTGCTGTTCTCCTCAAGCCTTGGGGTAACCTCATTTGGAATCCGCGTCTTCAGCACTTCTCTGTGAGTCACGTTGAAT
2701 CTTTCTGGGTGATAAATCGTTGCAAGCGACAGCGCGTGGCGTCCAGATCGTTTCCCTTAAAGAGCATGAAGCATAGTGATAAGTTT

FIG. 2. Sequence of the 2.8-kb fragment containing the *drpA* gene. The salient features of this sequence mentioned in the text are indicated by underlined boldface letters. These include the following: (i) the -35 and -10 regions of a sigma-70 promoter beginning at position 551 and extending to position 571 that was identified on the basis of sequence and physical mapping; (ii) the 5' end of the transcript measured by the S1 mapping extending from position 583; (iii) the deduced amino acid sequence of the DrpA protein (given below the ORF that begins at position 645); (iv) between positions 1617 and 1652, a canonical nucleotide-binding pocket; (v) the deduced factor-independent transcription terminator, which begins at position 2367. The sequence was determined from the set of fragments shown in Fig. 1. Sequencing was usually initiated from the universal primers that anneal to Mp18 and Mp19, though in one case a synthetic oligonucleotide was used to prime synthesis from the middle of fragment 18-2. This ensured that each part of the sequence of the *drpA* gene was based on at least two different sequencing runs. The sequences 5' and 3' to the *drpA* gene are usually based on two runs, but some parts are based on a single run.

TTATCCACA (12). Besides this sequence match, we have direct evidence that DrpA transcription is regulated by the DnaA protein. In *dnaA* mutants of *E. coli*, the expression of β -galactosidase in a *drpA'-lacZ* protein fusion is eight times higher than its expression in a cell that overproduces DnaA protein (T. Luby, unpublished data).

DrpA transcription terminator. Examination of the DNA sequence reveals that the *drpA* gene is followed by a sequence that appears to be a factor-independent transcription terminator (16). Beginning at bp 2367, the sequence AAAGGGTCTGAATTTCAAGGCCCTTTTTTTT has a stem-loop structure in which the only mismatch in the 12-bp stem is a G-T and there is a loop size of 3. In addition to the stem and loop, there is the characteristic sequence of T's 3' to the stem.

Other features of the sequence. The sequenced DNA fragment from pKFL2 presumably contains part of the coding sequences for another gene that is located on the 3'

side of the *drpA* gene and is translated in the same direction as the *drpA* gene. This ORF begins at position 2360, just to the left of the transcription terminator, and continues across the *Hind*III site at position 2791. The deduced amino acid sequence from this partial ORF is not homologous with any of the protein sequences maintained by the National Biomedical Research Foundation and available through GenBank.

DISCUSSION

We have determined the complete nucleotide sequence of the *drpA* gene. This was done by showing that the cloned insert in pKFL2 carries the *drpA* gene, as shown by the fine-structure map in Fig. 1. The complete nucleotide sequence of that insert revealed only a single ORF that could code for the DrpA protein. The gene itself is remarkable for its simplicity. Preceding the gene is a promoter whose

FIG. 3. S1 protection mapping of the DrpA transcript. Shown is the autoradiogram resulting from the S1 nuclease protection experiment. After the DNA probe was annealed overnight to the mRNA preparation to yield the RNA-DNA hybrid, the mixture was reacted with the S1 endonuclease for the length of time (in minutes) shown above the lanes. The sequence shown (5' to 3' from top to bottom) corresponds to the sequence beginning at position 583 in Fig. 2. A transcript whose 5' end began with the A at position 583 would produce a band that migrated to the indicated position, and the correspondingly shorter transcripts would migrate to their indicated positions. Lane a contains a 180-bp size standard, and lane b contains a sequencing ladder of a known sequence that was used to calibrate the gel. Shown is a 12-cm vertical section from an 80-cm polyacrylamide gel. The band which corresponds to the probe reannealing to its complement migrates to a position about 25 cm above the major transcript bands.

transcript begins 62 nucleotides before the first AUG. The ORF is 516 codons long and codes for a protein with an MW of 57,300. Following the stop codon by 177 nucleotides, an excellent-looking factor-independent transcription terminator sequence is apparent from the sequence. The *drpA* gene appears to be a member of the *dnaA* regulon.

DrpA is an essential agent whose presence is needed for global RNA synthesis as well as DNA synthesis. The results in this paper clearly establish that *drpA1*(Ts), *drpA2*(Ts), and *drpA3*(Ts) are recessive to the wild-type gene, eliminating the possibility that the *drpA*(Ts) mutations directly inhibit RNA and DNA synthesis, and support the notion that temperature sensitivity results from the loss of a normally essential function.

Identifying the essential role of DrpA is the problem. The *drpA* gene does not code for any of the known RNA polymerase subunits, which have all been identified and sequenced (18). Furthermore, when cell extracts are prepared from thermally arrested *drpA*(Ts) strains, normal levels of RNA polymerase activity are seen. In addition, rabbit antibody prepared against purified *E. coli* RNA polymerase immunoprecipitates the same complex of polypeptides from extracts prepared from both wild-type and thermally arrested *drpA*(Ts) strains (Syvanen, unpublished data). On the basis of a number of lines of evidence, we

FIG. 4. Primer extension mapping of the DrpA transcript. Shown is the autoradiogram from the primer extension measurement of the 5' end of the DrpA transcript. Lane 1 shows the DNA fragment produced from the mRNA isolated from SY203(pKFL2); the fragment was annealed to the oligonucleotide (positions 780 to 760) and reacted with reverse transcriptase (2). Lanes 2 through 5 show the sequencing ladder (A, T, G, and C, respectively) generated by the same probe annealed to the appropriate single-stranded DNA and submitted to the dideoxy sequencing reaction. In lane 2, the sequence of A's corresponding to the eight T's at the 5' end of the mRNA can be seen. Shown is a 40-cm vertical section from an 80-cm polyacrylamide gel.

believe that *drpA* does not exert its effect on ribonucleotide precursor synthesis. (i) RNA synthesis stops 15 min sooner than DNA synthesis does when growth of the *drpA1*(Ts) strain is thermally arrested. (ii) The mutational defect is apparent in the toluene-treated cells when nucleotide precursors are added (7). (iii) The levels of ATP as visualized by polyethyleneimine chromatography appear to be normal in *drpA*(Ts) strains at 42°C. (iv) The genetic-map positions for all of the relevant biosynthetic genes map elsewhere.

There appear to be two general possibilities for a role for DrpA in RNA synthesis. The first possibility is that DrpA is directly required for RNA polymerase activity in the cell. The standard substrates for RNA polymerase *in vitro* are DNA templates free from their associated proteins (10, 17), though some of these proteins, when added to the purified system, are known to inhibit the reaction (5). There are a number of lines of evidence indicating that transcription is much more complex than is reflected by the standard RNA

polymerase assay (8, 19). The second possibility is that DrpA is involved in regulating transcription either by maintaining proper chromosome structure or by antagonizing natural transcriptional inhibitors. Besides *drpA*, there may be other only partially characterized genes whose products also contribute to transcription (3, 6, 9, 12).

The work presented here lays the foundation for more systematic biochemical investigations. We can already infer that DrpA may have an ATPase activity (as judged by the presence of a canonical phosphonucleotide-binding pocket), and this may open a useful avenue into its enzymatic properties.

ACKNOWLEDGMENTS

We thank John D. Hopkins and Karen Lech for helpful comments.

This work was supported by Public Health Service grant GM 28142 from the National Institutes of Health.

LITERATURE CITED

1. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNA's by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* 12:721-732.
2. Calzone, F. J., R. J. Britten, and E. R. Davidson. 1987. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. *Methods Enzymol.* 152:611-632.
3. Friedman, D., E. Olson, C. Georgopoulos, K. Tilly, I. Herskowitz, and F. Banuett. 1984. Interactions of bacteriophage and host macromolecules in the growth of bacteriophage λ . *Microbiol. Rev.* 48:299-325.
4. Hanks, S. K., A. M. Quinn, and T. Hunter. 1988. The protein kinase family: concerned features and deduced phylogeny of the catalytic domains. *Science* 241:42-52.
5. Hubscher, U., H. Lutz, and A. Kornberg. 1980. Novel histone H2A-like protein of *E. coli*. *Proc. Natl. Acad. Sci. USA* 77:5097-5101.
6. Lathe, R., and J. P. Lecocq. 1977. The *firA* gene, a locus involved in the expression of Rif^R in *E. coli*. *Mol. Gen. Genet.* 154:43-52.
7. Lech, K. F., C. H. Lee, R. R. Isberg, and M. Syvanen. 1985. New gene in *Escherichia coli* K-12 (*drpA*): does its product play a role in RNA synthesis? *J. Bacteriol.* 162:117-123.
8. Li, S. C., C. L. Squires, and C. Squires. 1984. Antitermination of *E. coli* rRNA transcription is caused by a control region segment containing lambda nut-like sequences. *Cell* 38:851-860.
9. Liebke, H. H., and J. Speyer. 1977. A new gene in *E. coli* RNA synthesis. *Mol. Gen. Genet.* 189:314-320.
10. McClure, W. R. 1985. Mechanism and control of transcription initiation in prokaryotes. *Annu. Rev. Biochem.* 54:171-204.
11. McCommas, S. A., and M. Syvanen. 1988. Temporal control of transposition in Tn5. *J. Bacteriol.* 170:889-894.
12. McMacken, R., L. Silver, and C. Georgopoulos. 1987. DNA replication, p. 564-612. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
13. Miller, J., E. Claeys, J. Kirschbaum, S. Nasi, S. Vanden Elsacker, R. Molholt, G. Gross, D. Fields, and E. Baulz. 1976. Altered RNA polymerases resulting from temperature-sensitive mutations in the *rif* region of the *E. coli* chromosome, p. 519-538. In M. Chamberlin (ed.), *RNA polymerase*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure. 1984. *E. coli* promoter sequences predict *in vitro* RNA polymerase selectivity. *Nucleic Acids Res.* 12:789-800.
15. Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* 26:101-106.
16. Platt, T. 1986. Transcription termination and the regulation of gene expression. *Annu. Rev. Biochem.* 55:339-372.
17. Roubiere-Yaniv, J., and M. Yaniv. 1979. *E. coli* DNA binding protein Hu forms nucleosome-like structure with circular double stranded DNA. *Cell* 17:265-274.
18. Scaife, J., M. Gomma, T. Linn, J. Herlig, L. Rowen, and R. Calendar. 1980. Organization of the genes for bacterial RNA polymerase subunits, p. 3-20. In S. Osawa, H. Ozeki, H. Uchida, and T. Ura (ed.), *Genetics and evolution of RNA polymerase, tRNA and ribosomes*. University of Tokyo Press, Tokyo.
19. Schmidt, M. C., and M. J. Chamberlin. 1987. *nusA* protein of *E. coli* is an efficient transcription termination factor for certain terminator sites. *J. Mol. Biol.* 195:809-818.