JOURNAL OF BACTERIOLOGY, June 1991, p. 3644–3655
0021-9193/91/123644-12$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 12

Cloning, Sequencing, and Expression of *Bacillus subtilis* Genes Involved in ATP-Dependent Nuclease Synthesis

JAN KOOISTRA* AND GERARD VENEMA

Department of Genetics, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

Received 12 December 1990/Accepted 5 April 1991

The genes encoding the subunits of the *Bacillus subtilis* ATP-dependent nuclease (*add* genes) have been cloned. The genes were located on an 8.8-kb *SalI-SmaI* chromosomal DNA fragment. Transformants of a *recBCD* deletion mutant of *Escherichia coli* with plasmid pGV1 carrying this DNA fragment showed ATP-dependent nuclease activity. Three open reading frames were identified on the 8.8-kb *SalI-SmaI* fragment, which could encode three proteins with molecular masses of 135 (AddB protein), 141 (AddA protein), and 28 kDa. Only the AddB and AddA proteins are required for ATP-dependent exonuclease activity. Both the AddB and AddA proteins contained a conserved amino acid sequence for ATP binding. In the AddA protein, a number of small regions were present showing a high degree of sequence similarity with regions in the *E. coli* RecB protein. The AddA protein contained six conserved motifs which were also present in the *E. coli* helicase II (UvrD protein) and the Rep helicase, suggesting that these motifs are involved in the DNA unwinding activity of the enzyme. When linked to the T7 promoter, a high level of expression was obtained in *E. coli.*

In *Escherichia coli*, *recB*, *recC*, and *recD* genes encode the subunits for a complex enzyme, designated exonuclease V (56), which is an ATP-dependent single- and double-stranded DNA exonuclease (23, 39, 40), an ATP-stimulated single-stranded DNA endonuclease (22, 23, 40), an ATP-dependent DNA helicase (40, 46, 48, 55), and an ATP-dependent ATPase (23, 40). Chi DNA sequences, which stimulate bacteriophage lambda recombination, are also recognized and cleaved by this enzyme (9, 43, 54). The three genes encode nonidentical subunits with molecular weights (MWs) of 134,000, 129,000, and 67,000 (16, 17, 19).

Mutations in the *recB* and *recC* genes lead to reduction in homologous recombination in conjugation and generalized transduction, reduction in cell viability, and loss of repair of DNA damage (11). *recD* mutants, however, are not impaired in recombination and repair of DNA damage (3). Absence of the RecD subunit in the enzyme complex leads to a complete loss of the exonucleolytic functions of exonuclease V, but significant levels of helicase, endonuclease, and ATPase activities are still present, indicating that the exonucleolytic activities of the RecBCD enzyme are not required for genetic recombination and the repair of DNA damage caused by agents such as UV irradiation and mitomycin (MC) (40). The RecBCD enzyme can stimulate heteroduplex formation in the presence of the *E. coli* RecA protein (47, 60), but it is not yet clear which one of the various properties of the enzyme is critical for recombination (2).

ATP-dependent DNase-deficient mutants (*add* mutants) of *Bacillus subtilis* also exhibit reduced recombination in transformation with chromosomal DNA (28), reduced cell viability, and impaired ability to repair DNA damage caused by UV irradiation or MC treatment (28). The *B. subtilis* ATP-dependent DNase is also a multifunctional enzyme, possessing ATP-dependent double-stranded DNA exonuclease activity, DNA-dependent ATPase activity (13), and ATP-dependent helicase activity (51).

With respect to the composition of the *B. subtilis* ATP-dependent nuclease, contradictory results have been reported. According to Doly and Anagnostopoulos (13), the enzyme consists of five subunits with MWs of 81,000, 70,000, 62,000, 52,000, and 42,000. In contrast, Chestukhin et al. (10) reported that the enzyme was composed of only two subunits with MWs of 155,000 and 140,000.

To resolve this contradiction, and to study the function of the ATP-dependent nuclease in recombination in *B. subtilis*, we attempted to clone the *add* genes encoding the subunits of the enzyme. In a previous paper we reported the cloning of one of the *B. subtilis add* genes (*addA*) (28). The present paper describes the isolation of a second *add* gene (*addB*), the sequencing of both genes, the expression of both genes in an *E. coli* strain in which the *recBCD* genes were deleted, and the synthesis of the *B. subtilis add* gene products in *E. coli* under control of the T7 promoter.

MATERIALS AND METHODS

Strains and plasmids. The strains and plasmids used are listed in Table 1.

Chemicals and enzymes. The chemicals used were of analytical grade and were obtained from E. Merck AG (Darmstadt, Germany) or BDH (Poole, England). Restriction enzymes, T4 DNA ligase, and T4 DNA polymerase were used as recommended by the manufacturer (Boeringher GmbH, Mannheim, Germany).

Media. *B. subtilis* minimal medium consisted of Spizizen minimal salts (52) supplemented with glucose (0.5%) and casein hydrolysate (0.02%; Difco Laboratories, Detroit, Mich.). Amino acids, nucleotides (20 μg of each per ml), or vitamins (0.4 μg of each per ml) were added if required. Minimal agar consisted of minimal salts supplemented with 0.5% glucose, the required growth factors, and 1.5% agar. TY medium and TY agar were prepared as described by Biswal et al. (4). Trypticase agar was prepared by the method of Frischauf et al. (20).

Isolation of DNA. Plasmid DNA was isolated by the method of Ish-Horowicz and Burke (27). Radioactive DNA was isolated from *B. subtilis* 2G8 grown in minimal medium

* Corresponding author.

TABLE 1. Strains and plasmids

Strain or plasmid	Genotype, phenotype, or plasmid marker[a]	Source or reference
B. subtilis		
0G1	Prototrophic	6
8G5	*trpC2 tyr-1 his ade met rib ura nic*	6
2G8	*tyr-1 thy*	Laboratory collection
GSY2258(*add5*), GSY2270(*add71*), GSY2266(*add72*)	*hisH2 metB5*; *add* deficient	3a, 14
E. coli		
JM101	*supE thi* Δ(*lac proAB*) [F′ *traD36 proAB*$^+$*lacI*^{q}ZΔM15]	33
JM107	*endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) [F′ *traD36 proAB lacI*^{q}ZΔM15]	63
NM539	*sup*F *hsdR* (P2 *cox-3*)	20
AA 102F′	*recA pro supE* F′ *lac*$^+$ *thi endA hsdR*^{K} Δ*1*(*gal-chlD-pgl-att*)	41
AC 113(V186)	Δ(*argA-thyA*)*232*	8
BL21DE3	F-*hsd Sgal* T7 RNA polymerase gene (*lacUV5*)	53
Plasmids		
pUC7, pUC9, pUC18, pUC19	Ap^r, *E. coli* replicon	34, 58
pGV1	Km^r, *E. coli*/*B. subtilis* replicon	59
pGV88	Km^r, carrying the 8.8-kb *Sal*I-*Sma*I fragment	This paper
pE194	Em^r, *B. subtilis* replicon	26
pHV60	Cm^r, *E. coli* replicon	35
pAA-pZ618/619	Ap^r, *E. coli* replicon	41
pSPT18/19	Ap^r, *E. coli* replicon	Pharmacia LKB, Uppsala, Sweden

[a] Ap, ampicillin; Cm, chloramphenicol.

supplemented with 20 μg of tyrosine per ml, 2 μg of thymidine per ml, and [methyl-^{3}H]thymidine (10 μCi; specific activity, 20.7 mCi/mg; Amersham International plc, Amersham, United Kingdom). The specific activity was approximately 2.0×10^5 cpm/μg of DNA. DNA from bacteriophage lambda EMBL12 was isolated by the method described by Frischauf et al. (20).

Competence and transformation. *B. subtilis* cells were grown to competence as described by Bron and Venema (6). In transformations with plasmid DNA, the cells were exposed to DNA for 30 min and the cultures were then diluted twice with TY medium, incubated for 1 h at 37°C, and plated on TY agar plates containing kanamycin (20 μg/ml) or erythromycin (1 μg/ml). When erythromycin-resistant (Em^r) transformants were selected, the cultures were diluted twice with TY medium containing 0.05 μg of erythromycin per ml, incubated for 1 h at 37°C, and then plated.

E. coli was made competent and transformed by the method of Mandel and Higa (31).

Sensitivity to MC. Kanamycin-resistant (Km^r) transformants and Em^r transformants were tested for sensitivity to MC by transferring the transformants to TY agar plates containing kanamycin (20 μg/ml) or erythromycin (1 μg/ml) and MC (50 ng/ml). Transformants which failed to grow at an MC concentration of 50 ng/ml were considered to be MC sensitive. Wild-type cells were still able to grow on plates with 70 ng of MC per ml.

Assay of ATP-dependent exonuclease activity. Cultures (5 ml) of *B. subtilis* or *E. coli* grown in TY broth to an A_{600} of 1.0 were washed twice with 5 ml of 0.05 M Tris hydrochloride buffer (pH 8.0) containing 0.1 mM EDTA and 0.1 mM dithiothreitol, and the cells were resuspended in 0.5 ml of the same buffer. Cell lysates were prepared by using a French press (Amicon) at 1,000 lb/in^2. Samples (0.1 ml) of the lysate were added to 0.4 ml of a reaction mixture consisting of 0.1 M glycine-NaOH buffer (pH 9.2), 2.5 mM 2-mercaptoethanol, 0.05 M $MgCl_2$, 0.4 mg of bovine serum albumin per ml, and 75 μM ATP. Subsequently, 10 μl of ^{3}H-labeled DNA (40 μg/ml; specific activity, 2.0×10^5 cpm/μg), isolated from *B. subtilis* 2G8, was added, and, after incubation for 10 min at 40°C, the reaction was stopped by chilling on ice and the addition of 0.2 ml of calf thymus DNA solution (2 mg/ml) and 0.15 ml of trichloroacetic acid (50% [wt/vol]). After 30 min at 0°C, the mixtures were centrifuged in an Eppendorf centrifuge for 10 min and 0.6 ml of the supernatant was added to 5 ml of Hydroluma scintillation fluid. Radioactivity was counted for 5 min.

Isolation of a fragment with the Em^r marker from pE194. A *Taq*I fragment of plasmid pE194 was inserted into the *Acc*I site of plasmid pUC7. From the resulting plasmid, an *Eco*RI or *Bam*HI fragment containing the Em^r marker was isolated.

Construction of a *B. subtilis* genome bank in phage lambda EMBL12. A *B. subtilis* genome bank was constructed by cloning 12- to 17-kb chromosomal DNA fragments from strain 0G1, prepared by partial *Sau*3A digestion and isolated from an agarose gel, in *Bam*HI-digested phage lambda EMBL12 DNA (37) essentially as described by Vosman et al. (59). Test plating on *E. coli* NM539 showed that the bank contained 33,000 recombinants, jointly representing approximately 110 times the *B. subtilis* genome. The bank was amplified in *E. coli* NM539 as described by Maniatis et al. (32) and stored with a few drops of chloroform at 4°C.

Screening of the *B. subtilis* genome bank. To screen the *B. subtilis* genome bank, plasmid pHV60, carrying the 1.8-kb *Eco*RI-*Hin*dIII fragment (Fig. 1), which contains a part of the *addB* gene (28), was used as a probe. This plasmid was nick translated with [α-^{32}P]dCTP by the method of Rigby et al. (45). To isolate phages hybridizing with this probe, recombinant lambda EMBL12 was plated on Trypticase agar plates with *E. coli* NM539 as indicator. The plaques, approximately 1,000 per plate, were transferred to a GeneScreen Plus membrane (Dupont, NEN Research Products, Dreieich, Germany) as described by Maniatis et al. (32) and then hybridized against the probe by the method recommended by the manufacturer. Hybridizing plaques were picked from

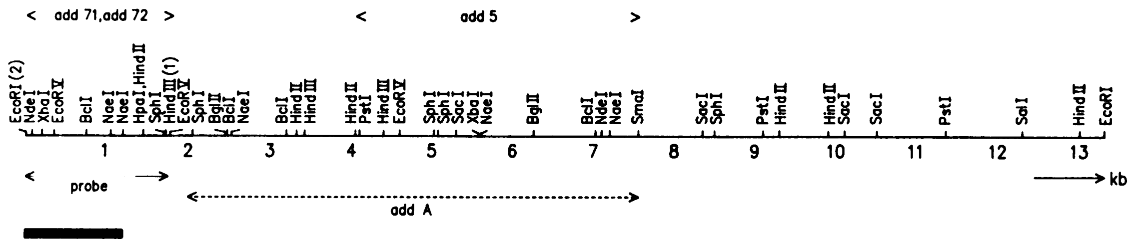

FIG. 1. Restriction map of the 13.3-kb *B. subtilis* chromosomal *Eco*RI fragment. The 1.8-kb *Eco*RI(2)-*Hin*dIII(1) fragment (<---->) inserted in plasmid pHV60 was used as a probe for screening the *B. subtilis* chromosomal DNA bank in phage lambda EMBL12. The region indicated by the bar is identical to that shown in Fig. 2.

the plates and tested for purity by a second cycle of hybridization with the probe.

Nucleotide sequence analysis. DNA restriction fragments were either subcloned in M13mp18 and M13mp19 (38) and transformed into *E. coli* JM101 or subcloned in the transposon-promoted deletion vectors pAA-pZ618 and pAA-pZ619 (41; Gold Biotechnology, St. Louis, Mo.) by using *E. coli* JM107 as the recipient. In the latter plasmids, the modified Tn9 transposon deletes the chromosomal DNA insert starting from a fixed site in the insertion sequence IS1-L region of the vector. Selection for deleted fragments was performed after introduction of the subclones in *E. coli* AA102F′ (41; Gold Biotechnology) by the procedure recommended by the manufacturer. The plasmids containing deleted fragments were finally transformed in *E. coli* JM107, and, for sequencing, single-stranded DNA was generated by superinfection with the helper phage M13K07 (Gold Biotechnology). Sequencing of DNA was performed by the dideoxynucleotide chain termination method (49) with [α-^{35}S]dATP. Sequence reactions were performed with T7 DNA polymerase (T7 sequencing kit; Pharmacia LKB, Uppsala, Sweden) and by use of the M13 universal primer when subcloning was done in M13mp18 and M13mp19 or of a primer complementary to the −23 region of IS1-L when plasmids pAA-pZ618 and pAA-pZ619 were used for subcloning. In those cases, in which there was no overlap between sequenced regions, the gaps were bridged by using 17-mer synthetic oligonucleotides based on known sequences. The sequencing reaction products were resolved on 6% polyacrylamide-8 M urea sequencing gels. Gels were run at 1,500 V, vacuum dried, exposed to Fuji RX X-ray films, and visualized by autoradiography. Both strands of the DNA were sequenced.

DNA and protein sequences were analyzed by using the Microgenie Sequence Analysis Program (44). The FASTP algorithm of Lipman and Pearson (30) was used for protein comparisons in the Swiss Protein Database (release 11).

Expression of *add* genes under control of the T7 promoter. The 8.8-kb *Sal*I-*Sma*I fragment, carrying the *add* genes, was linked to the T7 promoter in plasmids pSPT18 and pSPT19 (Pharmacia LKB). The resulting plasmids were introduced into an *E. coli* BL21DE3 strain containing a single copy of the gene for T7 RNA polymerase in the chromosome under control of the inducible *lacUV5* promoter (53). An overnight culture of this strain, carrying the plasmids with the *add* genes, was diluted 100-fold in 10 ml of TY medium containing ampicillin (100 μg/ml). The culture was grown at 37°C to an A_{600} of 0.5, and 0.1 ml of an isopropyl-β-D-thiogalactopyranoside solution (40 mM) was added to induce synthesis of the T7 RNA polymerase. At 1 h after addition of isopropyl-β-D-thiogalactopyranoside, 0.4 ml of rifampin (5 mg/ml) was added to inhibit the host *E. coli* cell RNA polymerases, and, after incubation for another hour, the cells were harvested and resuspended in 0.35 ml of a 0.5 M Tris hydrochloride buffer (pH 6.8) containing 0.4% sodium dodecyl sulfate (SDS) and 0.1% bromophenol blue. After heating for 5 min at 100°C, 8-μl samples were electrophoresed on an SDS-polyacrylamide (7.5%) gel by the method of Laemmli (29) by using a Mini-Protean II Vertical Electrophoresis System (Bio-Rad Laboratories, Inc., Richmond, Calif.). MW marker proteins consisted of myosin (MW, 200,000), β-galactosidase (MW, 116,250), phosphorylase *b* (MW, 92,500), bovine serum albumin (MW, 66,200), and ovalbumin (MW, 42,700) (SDS-PAGE Standards, high molecular weight; Bio-Rad). Electrophoresis was conducted at room temperature for about 1.5 h at a constant current of 20 mA. After electrophoresis, the gels were stained with Coomassie blue (Bio-Rad).

Nucleotide sequence accession number. The sequence

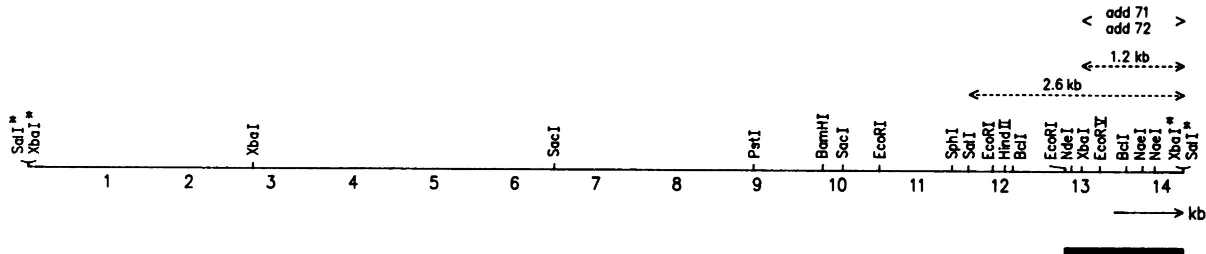

FIG. 2. Restriction map of the 14.2-kb *B. subtilis* chromosomal DNA fragment in phage lambda EMBL12. *Sal*I* and *Xba*I* are sites present in the multiple cloning site of lambda EMBL12; these sites are not present on the *B. subtilis* chromosomal DNA. The region indicated by the bar is identical to that shown in Fig. 1.

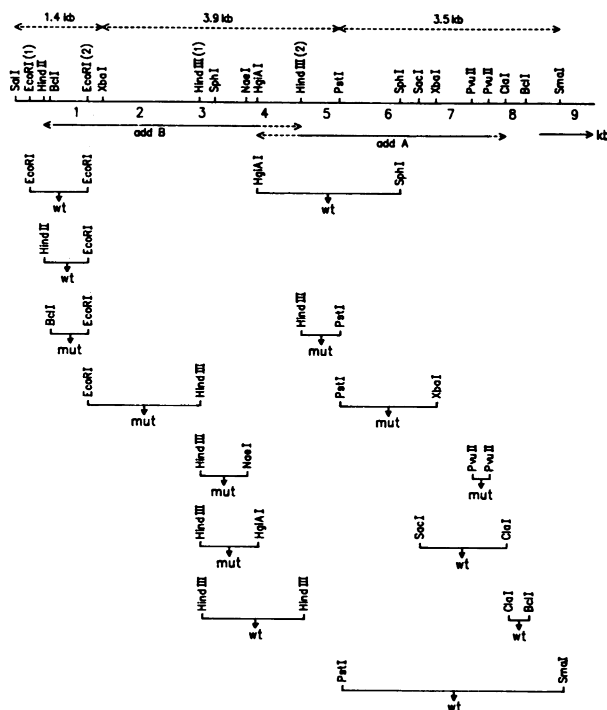

FIG. 3. Localization of *add* genes. The restriction fragments indicated were inserted into pUC plasmids. These plasmids were used to transform an 8G5 (add^+) strain of *B. subtilis*. The fragments which transformed the recipient to the mutant *add* phenotype (mut) and those which did not (wt) are indicated. The dotted parts of *addB* and *addA* contain the ends of the genes.

shown in Fig. 4 has been assigned GenBank accession number M63489.

RESULTS

Cloning of a second gene involved in ATP-dependent DNase synthesis. We reported previously that a complete transcription unit involved in ATP-dependent DNase synthesis (*addA*) was located between an *Sph*I site and an *Sma*I site on a 13.3-kb *Eco*RI fragment of the *B. subtilis* chromosome (Fig. 1) and that this fragment also contained part of a second gene (*addB*) involved in the synthesis of this enzyme. It was shown that the three previously isolated *add*-deficient mutations of *B. subtilis*, *add5*, *add71*, and *add72* (3a, 14), were all located on this 13.3-kb *Eco*RI fragment (28). A more precise localization revealed that the *add5* mutation was located on the *addA* gene and that the *add71* and *add72* mutations were located on the *addB* gene (Fig. 1) (28).

To clone the entire *addB* gene, the *B. subtilis* chromosomal DNA library in phage lambda EMBL12 was screened by using as a probe plasmid pHV60 carrying a 1.8-kb *Eco*RI(2)-*Hin*dIII(1) chromosomal DNA fragment (Fig. 1). One of the phages hybridizing with this probe contained a *B. subtilis* chromosomal DNA fragment of 14.2 kb. The restriction map of this fragment is shown in Fig. 2.

Each of the *Xba*I and *Sal*I restriction fragments of this 14.2-kb fragment were inserted into plasmid pGV1 (59), which can replicate in both *B. subtilis* and *E. coli* and contains a Km^r marker. To test whether one of these fragments overlapped with the *add71* and *add72* mutations, which are located between the *Eco*RI and the *Hin*dIII site in the probe DNA (Fig. 1), the *add71* and *add72* mutants were transformed with pGV1, containing the various *Xba*I or *Sal*I fragments, and the resulting Km^r transformants were tested for sensitivity to MC. In transformations with pGV1 carrying the 1.2-kb *Xba*I fragment and with pGV1 carrying the 2.6-kb *Sal*I fragment (Fig. 2), transformants were formed which had acquired wild-type resistance to MC, indicating that these fragments were overlapping the *add71* and *add72* mutations. Comparison of the restriction map between the *Eco*RI(2) and *Hin*dIII(1) sites of the 13.3-kb *Eco*RI fragment (Fig. 1) with that of the 1.2-kb *Xba*I and 2.6-kb *Sal*I fragments (Fig. 2) showed complete correspondence. One of the *Xba*I sites in the 1.2-kb *Xba*I fragment and one of the *Sal*I sites of the 2.6-kb *Sal*I fragment were not present on the chromosome, so that, in all probability, they represent sites in the multiple cloning site of the recombinant phage lambda EMBL12.

To determine the ends of the *addB* gene and to delineate more precisely the *addA* gene previously localized between the *Sph*I site and the *Sma*I site on the 13.3-kb *Eco*RI fragment (Fig. 1), various restriction fragments of the 13.3-kb *Eco*RI fragment (Fig. 1) and of the 14.2-kb fragment (Fig. 2) were inserted into plasmid pUC18. Subsequently, an *Eco*RI fragment carrying the Em^r marker from plasmid pE194 (see Materials and Methods), which is expressed in *B. subtilis*, was inserted into the pUC18 recombinants. With these composite plasmids, a wild-type strain of *B. subtilis* was transformed and the Em^r transformants were tested for sensitivity to MC. Em^r transformants resulting from a Campbell-like integration will have the *add* mutant phenotype, which is sensitive to MC, if the plasmid contains an internal part of an *add* gene. On the other hand, transformants with the wild-type *add* phenotype (wild-type resistance to MC) will be obtained if the inserted DNA fragment does not contain a part of an *add* gene or if it contains at least one of the ends of the gene. Figure 3 shows the internal fragments and those containing one of the ends of the gene. From these results it can be concluded that the second *add* gene (*addB*) is located between the *Hin*dII site at the left and the *Hin*dIII(2) site at the right and that the *addA* gene is located between the *Hgi*AI site at the left and the *Cla*I site on the right of the fragment shown in Fig. 3.

Construction of a plasmid (pGV88) jointly containing the *addA* and *addB* genes. The construction of a plasmid in which both *add* genes are present was performed as follows. First, the 1.4-kb *Sal*I-*Xba*I fragment (Fig. 3) and the 3.9-kb *Xba*I-*Pst*I fragment (Fig. 3) were ligated together into plasmid pGV1, resulting in plasmid pGV53 containing the 5.3-kb *Sal*I-*Pst*I fragment. This 5.3-kb *Sal*I-*Pst*I fragment and the 3.5-kb *Pst*I-*Sma*I fragment (Fig. 3) were then ligated together in plasmid pGV1, resulting in plasmid pGV88. In this

TABLE 2. ATP-dependent nuclease activity in *E. coli* V186(pGV88) transformants

E. coli strain	ATP-dependent nuclease activity (degradation of ^{3}H-DNA to acid-soluble products [cpm] at various times of incubation)					
	3 min		5 min		10 min	
	+ATP	−ATP	+ATP	−ATP	+ATP	−ATP
V186	187	206	258	290	369	419
V186(pGV88) 1[a]	749	213	743	361	872	532
V186(pGV88) 2[a]	1,052	211	1,060	369	1,700	576

[a] Two assays were done with two individual transformants of V186 carrying pGV88.

```
   1 GTCGACGGTTTCTTGTTCATCTTTAAAAGCGGCGTCCAGTCCTCTTGTGAAATTGTCCGGGCATTCTTCCGTTTGTTTCGGCATGATGTGTTTGCCGGTA

 101 ATTGAACCATACAGTGTGGAAAATTCACGGAGGTGGCGAATCTCGTCACGCCTGATTTCTTCAATCTGTTTTCTGACCTCATCGCTGTGGCGAGTTCGG

 201 CAAGCTTGCGGTAGCATTGAATAGCACTGAATTCTCCGTTGATGGCCTTCTGCAAATTCCGAATCAAAACAGGATCTTCCCGGTATGGCACCGGATAGGG

 301 ATAGTAGCTGTAATGCAAGGGGATTCAACTCCATTCTTAGGAATTGGGCTTCTGTAAAAGTATATGTTGGGCAGGGAGAGATGTGCGGAGATAATCAGCT

 401 TTTTATATGTGAAAAGGCCGTTTTTACCAATAGATCAGATTGGTCATTTTCGTCAACATTCGATAAAATATAGAGAGATAAAAAGAGAGGGGTCTTCTAA
                                                                                   S.D.
     ORF1(addB)
      M  G  A  E  F  L  V  G  R  S  G  S  G  K  T  K  L  I  I  N  S  I  Q  D  E  L  R  R  A  P  F  G  K
 501 TTTGGGAGCAGAGTTTTTAGTAGGCAGGTCCGGGAGTGGAAAAACGAAGCTGATCATCAACAGCATTCAGGATGAATTGCGCCGGGCTCCATTCGGGAAG

      P  I  I  F  L  V  P  D  Q  M  T  F  L  M  E  Y  E  L  A  K  T  P  D  M  G  G  M  I  R  A  Q  V  F  S
 601 CCGATCATTTTTCTAGTCCCGGATCAAATGACGTTTTTAATGGAATACGAGCTTGCTAAAACGCCAGATATGGGCGGGATGATACGCGCTCAAGTGTTCA

       F  S  R  L  A  W  R  V  L  Q  H  T  G  G  M  S  R  P  F  L  T  S  T  G  V  Q  M  L  L  R  K  L  I
 701 GTTTTTCACGATTGGCCTGGCGCGTCCTCCAGCATACGGGAGGAATGAGCAGGCCGTTTCTGACGAGCACCGGCGTACAAATGCTCCTGCGGAAGCTCAT

       E  E  H  K  Q  E  F  K  V  Y  Q  K  A  S  D  K  S  G  F  T  A  Q  V  E  R  M  L  T  E  F  K  R  Y
 801 TGAGGAGCATAAACAGGAGTTCAAAGTCTATCAAAAAGCGAGTGACAAAAGCGGGTTTACCGCACAGGTTGAGCGGATGCTGACAGAGTTTAAGCGCTAT

      C  L  E  P  E  D  I  R  R  M  A  E  S  G  T  A  S  E  Y  R  G  E  R  V  L  S  E  K  L  H  D  L  S  I
 901 TGTCTGGAACCGGAGGATATCCGCCGGATGGCGGAAAGCGGAACGGCTTCCGAGTATCGCGGAGAACGTGTTTTATCTGAAAAGCTCCATGACTTATCAA

       L  Y  Q  Q  M  E  K  S  L  A  D  Q  Y  L  H  S  E  D  Y  L  T  L  L  A  E  H  I  P  L  A  E  D  I
1001 TTCTGTATCAGCAAATGGAAAAAAGCCTCGCAGATCAATATCTTCACTCTGAGGATTATTTGACATTGCTGGCAGAGCATATCCCGCTTGCGGAAGATAT

       K  G  A  H  I  Y  V  D  G  F  Y  Q  F  T  P  Q  E  F  R  V  L  E  Q  L  M  V  H  A  E  H  I  T  F
1101 AAAAGGCGCCCATATCTATGTGGATGGCTTTTATCAGTTTACCCCGCAGGAATTCAGGGTGTTGGAGCAGCTTATGGTTCATGCGGAGCATATCACGTTT

      S  L  T  A  D  K  P  S  Y  E  R  E  P  H  E  L  E  L  F  R  M  T  G  K  T  Y  Y  R  L  H  Q  K  A  K
1201 TCGCTCACAGCGGACAAGCCGTCATATGAGCGGGAGCCGCATGAACTGGAATTGTTCAGAATGACGGGGAAAACCTATTACCGCCTGCATCAGAAGGCGA

       E  L  N  L  D  I  T  Y  K  E  L  S  G  T  E  R  H  T  K  T  P  E  L  A  H  L  E  A  Q  Y  E  A  R
1301 AGGAACTGAATTTGGACATTACGTATAAAGAGCTGAGCGGGACTGAGCGGCATACAAAGACGCCGGAATTGGCGCATCTAGAGGCACAGTATGAAGCGCG

       P  A  I  P  Y  A  E  K  Q  E  A  L  T  V  M  Q  A  A  N  R  R  A  E  L  E  G  I  A  R  E  I  H  A
1401 TCCGGCCATTCCATACGCAGAAAAACAAGAAGCCCTTACTGTGATGCAGGCCGCAAACAGACGAGCTGAGCTGGAAGGCATTGCTCGGGAAATTCACGCC

      L  V  R  E  K  G  Y  R  Y  K  D  V  A  I  L  A  R  Q  P  E  D  Y  K  D  M  V  K  E  V  F  A  D  Y  E
1501 CTAGTCAGAGAGAAGGGATATCGCTATAAGGATGTAGCAATTCTTGCGCGCCAGCCGGAAGACTACAAGGATATGGTGAAGGAAGTTTTCGCAGATTACG

       I  P  Y  F  I  D  G  K  A  S  M  L  N  H  P  L  I  E  F  I  R  S  S  L  D  V  L  K  G  N  W  R  Y
1601 AGATTCCTTATTTCATTGACGGAAAAGCATCTATGCTGAACCATCCGTTAATTGAATTTATCCGGTCGAGCCTTGATGTTCTGAAAGGGAA[illegible]TGGCGTTA

       E  A  V  F  R  C  V  K  T  E  L  L  F  P  L  N  E  P  K  A  K  V  R  E  Q  V  D  Q  L  E  [illegible]  Y  C
1701 TGAAGCGGTGTTTCGCTGCGTGAAAACCGAACTGCTATTCCCGCTCAATGAACCGAAGGCCAAAGTGAGAGAACAGGTCGATCAGCTCGA [illegible]ATTACTGT

      I  A  Y  G  I  K  G  D  R  W  T  K  G  D  R  F  Q  Y  R  R  F  V  S  L  D  D  D  F  A  Q  T  D  Q  E
1801 ATCGCCTATGGTATTAAAGGCGACCGCTGGACAAAGGGCGATCGGTTCCAATACAGGCGTTTTGTGTCATTGGATGATGATTTTGCGCAGACTGATCAGG

       I  E  M  E  N  M  L  N  D  T  R  D  W  I  V  P  P  L  F  Q  L  Q  K  R  M  K  K  A  K  T  V  Q  E
1901 AAATCGAAATGGAAAACATGTTGAATGACACCCGCGATTGGATTGTTCCGCCGCTTTTTCAGCTTCAAAAACGCATGAAAAAAGCGAAGACGGTTCAAGA

       K  A  E  A  L  Y  R  Y  L  E  E  T  D  V  P  L  K  L  D  Q  E  R  Q  R  A  E  D  D  G  R  I  I  E
2001 GAAGGCGGAGGCGCTCTATCGTTATTTAGAAGAGACGGATGTGCCGCTGAAGCTGGATCAGGAAAGGCAGCGTGCTGAGGATGACGGCAGAATCATTGAA

      A  Q  Q  H  Q  Q  A  W  D  A  V  I  Q  L  L  E  E  F  V  E  M  M  G  D  D  E  I  S  L  D  L  F  Q  Q
2101 GCGCAGCAGCATCAGCAGGCGTGGGACGCAGTCATTCAGCTGCTTGAGGAGTTTGTTGAAATGATGGGAGATGATGAGATTTCTCTTGATTTGTTTCAGC

       M  I  E  A  G  A  E  S  L  T  F  S  L  I  P  P  A  L  D  Q  V  F  V  G  N  M  D  L  S  R  M  Y  G
2201 AAATGATAGAAGCCGGCGCGGAGTCGCTTACATTTTCTCTAATTCCGCCTGCGCTTGACCAGGTGTTTGTCGGCAATATGGATTTGTCCAGAATGTATGG

       T  S  C  T  F  V  L  G  A  N  D  G  V  L  P  A  R  P  D  E  N  G  V  L  S  D  D  D  R  E  W  L  K
2301 CACCTCCTGCACCTTTGTGCTCGGGGCAAACGACGGCGTTCTGCCGGCACGCCCTGATGAAAACGGGGTCCTGTCGGATGATGACCGGGAATGGCTGAAA

      T  I  G  V  E  L  S  S  G  G  R  E  R  L  L  D  E  H  F  L  I  Y  M  A  F  S  S  P  S  D  R  L  Y  V
2401 ACGATTGGGGTTGAGCTATCCTCAGGCGGACGAGAGCGTTTGCTTGATGAGCACTTCCTCATCTACATGGCGTTTTCAAGTCCGTCTGACCGGCTTTACG

       S  Y  P  I  A  D  A  E  G  K  T  L  L  P  S  M  I  V  K  R  L  E  E  L  F  P  H  H  K  E  R  L  L
2501 TATCGTATCCGATTGCTGATGCGGAAGGAAAAACGCTTTTGCCGTCGATGATCGTTAAGCGGCTGGAAGAACTGTTTCCGCATCATAAGGAGCGCCTGTT

       T  N  E  P  E  Q  V  S  D  E  E  Q  L  M  Y  V  V  N  K  S  V  A  Q  S  F  T  A  S  Q  L  R  L  W
2601 AACAAATGAACCTGAACAGGTCAGCGATGAGGAACAGCTGATGTATGTTGTGAATAAAAGCGTGGCACAGTCCTTTACCGCGAGCCAGCTCAGGTTATGG

      T  R  E  Y  D  I  S  D  V  W  W  S  T  Y  N  V  L  M  S  E  Q  D  R  L  Q  S  K  K  L  F  S  S  L  F
2701 ACTCGGGAATACGACATCAGCGACGTCTGGTGGAGCACGTACAATGTGCTGATGAGTGAGCAGGACAGGCTGCAATCGAAAAAGCTGTTCTCAAGCCTGT

       F  R  N  E  V  K  Q  L  E  R  S  V  S  R  Q  L  Y  G  E  R  I  Q  G  S  V  S  R  M  E  T  F  N  A
2801 TTTTCCGGAATGAAGTGAAGCAGCTTGAACGCAGCGTGTCGAGACAGCTCTATGGTGAACGTATTCAGGGCAGTGTATCGAGAATGGAAACCTTTAACGC

       C  P  F  S  H  F  A  S  H  G  L  H  L  K  E  R  Q  F  F  K  L  E  A  P  D  I  G  Q  L  F  H  S  S
2901 ATGCCCGTTTTCCCATTTTGCGTCACACGGGCTGCATTTGAAGGAACGGCAATTCTTCAAGCTTGAAGCACCGGATATCGGCCAGCTGTTTCATTCCAGC

      L  K  L  I  S  D  R  L  R  D  E  K  L  D  W  R  D  L  T  K  E  Q  C  E  L  F  S  Y  D  A  V  E  R  L
3001 TTAAAGCTGATTTCGGACAGGCTGCGTGACGAAAAGCTGGATTGGCGCGATTTAACGAAGGAGCAGTGCGAGCTGTTTTCCTATGATGCGGTAGAGCGGC

       A  P  K  L  Q  K  E  I  L  L  S  S  N  R  H  Y  Y  V  K  E  K  L  Q  K  I  V  T  R  V  S  G  I  L
3101 TGGCACCGAAACTGCAAAAGGAAATTCTGCTCAGCTCAAACCGGCATTATTATGTGAAGGAAAAACTGCAAAAAATTGTGACACGTGTGTCCGGCATTTT

       S  E  H  A  K  A  S  G  F  V  P  I  G  L  E  L  G  F  G  G  K  G  P  L  P  P  L  T  F  Q  L  K  N
3201 AAGCGAGCATGCGAAAGCGAGCGGATTCGTACCGATCGGGCTTGAACTGGGCTTTGGAGGAAAAGGGCCGCTTCCGCCGCTGACCTTTCAACTGAAAAAC
```

FIG. 4. Sequence of the 8.8-kb *Sal*I-*Sma*I region. Shine-Dalgarno (S.D.) sequences are underlined. Term, probable transcription terminator.

G C T M E L V G R I D R V D K A E S S K G L L L R I V D Y K S S D K

3301 GGCTGTACGATGGAACTCGTCGGGCGAATTGACCGTGTTGATAAGGCTGAAAGCTCAAAAGGCCTGCTCCTCAGGATTGTCGATTATAAATCAAGCGACA

G L D L A E V Y Y G L A L Q M L T Y L D L S I T H S A D W L G M R

3401 AAGGCCTTGACTTAGCGGAAGTATATTACGGATTGGCACTGCAAATGCTGACGTACCTTGATTTATCGATTACACATTCAGCTGACTGGCTCGGGATGAG

A T P A G V L Y F H I H D P M I Q S N L P L G L D E I E Q E I F K

3501 GCGACGCCTGCCGGAGTGCTGTATTTCCATATTCATGACCCGATGATTCAATCTAATCTCCCGCTTGGGCTTGACGAGATTGAACAGGAGATCTTTAAG

K F K M K G L L L G D Q E V V R L M D T T L Q E G R S N I I N A G L

3601 AAATTTAAAATGAAGGGCTTGCTCCTCGGTGATCAGGAAGTGGTTCGCCTCATGGATACAACCCTTCAAGAGGGACGTTCAAATATCATAAACGCCGGCT

K K D G S L R S D S A A V G E K E F D L L T K H V R R T F Q E A G

3701 TGAAAAAAGACGGCTCTCTCAGATCAGACTCAGCAGCAGTCGGTGAAAAGGAATTTGATCTTTTGACAAAGCATGTGCGCCGCACCTTCCAAGAAGCGGG

E Q I T D G R V S I E P Y K M K N K T P C T Y C A F K S V C Q F D

3801 CGAACAAATCACCGACGGGCGCGTATCCATTGAGCCGTACAAAATGAAGAACAAGACGCCGTGCACATACTGTGCGTTCAAATCAGTATGCCAATTTGAT

ORF2(addA)

M N I P

E S L E E N E Y R P L K A E K D K T I L E W I K K E A D G N E H S *

3901 GAATCACTGGAAGAAAACGAGTATCGCCCATTAAAGGCTGAAAAGGACAAGACAATACTTGAGTGGATAAAAAAGGAGGCGGATGGCAATGAACATTCCT

S.D.

K P A D S T W T D D Q W N A I V S T G Q D I L V A A A A G S G K T A

4001 AAACCGGCAGACAGCACATGGACAGATGACCAATGGAATGCCATTGTTTCAACCGGCCAGGATATTCTTGTGGCAGCGGCTGCGGGCTCTGGTAAAACCG

V L V E R M I R K I T A E E N P I D V D R L L V V T F T N A S A A

4101 CTGTGCTCGTTGAACGAATGATTCGGAAAATCACCGCGGAGGAAAACCCAATAGATGTAGACCGTCTTCTCGTTGTGACATTTACAAACGCCTCAGCGGC

E M K H R I A E A L E K E L V Q R P G S L H I R R Q L S L L N R A

4201 AGAGATGAAGCACCGAATCGCAGAAGCCTTGGAAAAAGAGCTTGTACAGCGCCCCGGCTCGCTGCATATTAGACGCCAGCTGTCTCTTTTAAACCGGGCC

S I S T L H S F C L Q V L K K Y Y Y L I D L D P G F R I A D Q T E G

4301 AGCATTTCGACGCTCCATTCCTTTTGCCTGCAAGTGCTGAAGAAATATTACTACTTGATCGATCTTGATCCGGGCTTTCGGATAGCTGATCAGACGGAAG

E L I G D E V L D E L F E D E Y A K G E K A F F E L V D R Y T T D

4401 GCGAGCTGATCGGGGATGAAGTGCTTGACGAGCTGTTTGAAGACGAATACGCAAAGGGCGAAAAGGCGTTTTTTGAGCTTGTTGACCGCTATACGACAGA

R H D L D L Q F L V K Q V Y E Y S R S H P N P E A W L E S F V H L

4501 CCGGCATGATCTGGATCTGCAATTTCTCGTTAAACAGGTGTACGAGTATTCCCGATCCCATCCCAACCCGGAGGCGTGGCTGGAAAGCTTTGTTCATTTG

Y D V S E K S A I E E L P F Y Q Y V K E D I A M V L N G A K E K L L

4601 TATGATGTATCAGAAAAGAGCGCCATCGAGGAGCTGCCGTTTTATCAATATGTCAAAGAAGATATTGCAATGGTGCTTAACGGGGCGAAGGAAAAGCTCT

R A L E L T K A P G G P A P R A D N F L D D L A Q I D E L I Q H Q

4701 TGCGCGCGCTTGAGCTGACGAAAGCGCCGGGCGGCCCGGCGCCGCGTGCTGACAATTTTCTTGATGATCTTGCTCAGATTGATGAACTGATTCAGCATCA

D D F S E L Y K R V P A V S F K R A K A V K G D E F D P A L L D E

4801 GGACGATTTCAGTGAACTATATAAGCGGGTGCCCGCCGTCTCTTTTAAGCGTGCCAAAGCAGTAAAAGGGGATGAGTTCGATCCAGCGCTCCTTGATGAG

A T D L R N G A K K L L E K L K T D Y F T R S P E Q H L K S L A E M

4901 GCGACAGATTTGAGGAACGGCGCAAAAAAACTGCTTGAAAAGCTCAAAACCGACTACTTCACGCGAAGTCCTGAACAGCACTTGAAAAGCCTAGCCGAGA

K P V I E T L V Q L V I S Y G K R F E A A K Q E K S I I D F S D L

5001 TGAAGCCTGTCATTGAAACGCTCGTACAGCTTGTCATCAGCTATGGAAAACGATTCGAAGCTGCGAAACAGGAAAAATCAATCATCGATTTTTCGGATTT

E H Y C L A I L T A E N D K G E R E P S E A A R F Y Q E Q F H E V

5101 GGAGCATTACTGTTTAGCGATTTTGACAGCTGAGAATGACAAAGGTGAACGTGAGCCGAGCGAGGCTGCAAGGTTTTATCAGGAACAGTTTCATGAGGTG

L V D E Y Q D T N L V Q E S I L Q L V T S G P E E T G N L F M V G D

5201 CTCGTTGACGAATATCAGGATACCAACCTCGTGCAGGAATCGATTCTGCAGCTCGTCACAAGCGGTCCGGAGGAGACTGGTAACCTGTTTATGGTAGGAG

V K Q S I Y R F R L A E P L L F L S K Y K R F T E S G E G T G R K

5301 ATGTCAAACAGTCGATTTATCGATTCAGGCTTGCGGAGCCGCTTCTTTTCCTCTCTAAATACAAACGGTTTACAGAGAGCGGAGAAGGCACGGGCGGAA

I D L N K N F R S R A D I L D S T N F L F K Q L M G G K I G E V D

5401 AATCGATTTAAATAAAAATTTCCGAAGCCGGGCTGATATTTTAGACAGCACAAACTTTTTATTTAAACAGCTGATGGCGGCAAAATCGGTGAGGTCGAT

Y D E Q A E L K L G A A Y P D N D E T E T E L L L I D N A E D T D A

5501 TATGACGAGCAGGCTGAGCTGAAGCTTGGTGCAGCGTATCCGGACAATGACGAGACGGAAACAGAGCTGCTGCTGATCGACAACGCAGAAGATACGGATG

S E E A E E L E T V Q F E A K A I A K E I R K L I S S P F K V Y D

5601 CAAGCGAGGAAGCAGAAGAGCTTGAAACGGTGCAGTTTGAGGCAAAAGCCATCGCTAAGGAAATTCGTAAGCTGATTTCATCGCCGTTTAAGGTGTATGA

G K K K T H R N I Q Y R D I V I L L R S M P W A P Q I M E E L R A

5701 CGGAAAAAAGAAAACACATCGCAATATTCAATACCGAGATATCGTGATTTTGCTCCGTTCGATGCCGTGGGCTCCGCAAATCATGGAGGAGCTGAGAGCA

Q G I P V Y A N L T S G Y F E A V E V A V A L S V L K V I D N P Y Q

5801 CAGGGCATACCGGTTTACGCCAATTTAACGTCAGGCTATTTTGAAGCGGTCGAAGTCGCCGTCGCGCTTTCTGTGCTGAAGGTGATTGATAATCCGTATC

D I P L A S V L R S P I V G A D E N E L S L I R L E N K K A P Y Y

5901 AGGATATACCGCTTGCCTCTGTGCTGCGCTCACCGATTGTCGGAGCAGATGAAAACGAGCTGTCTTTGATCCGGCTTGAAAATAAAAAAGCGCCGTACTA

E A M K D Y L A A G D R S D E L Y Q K L N T F Y G H L Q K W R A F

6001 TGAGGCGATGAAAGACTACCTGGCTGCTGGTGACCGGAGCGATGAGCTTTATCAAAAGTTAAATACGTTTTACGGACATCTGCAAAAATGGCGCGCGTTT

S K N H S V S E L I W E V Y R D T K Y M D Y V G G M P G G K Q R Q A

6101 TCGAAAAACCACTCAGTATCTGAGCTGATTTGGGAAGTGTACCGCGACACCAAATATATGGATTATGTCGGCGGCATGCCGGGCGGAAAACAGCGCCAGG

N L R V L Y D R A R Q Y E S T A F R G L F R F L R F I E R M Q E R

6201 CCAATTTGCGTGTTCTTTATGACAGGGCGCGTCAATATGAATCAACGGCATTTCGCGGCTTGTTCCGTTTCCTGCGGTTTATCGAACGCATGCAGGAGCG

G D D L G T A R G L S E Q E D V V R L M T I H S S K G L E F P V V

6301 GGGCGATGATCTTGGTACGGCGAGAGGCCTCAGCGAGCAGGAGGATGTTGTCCGCTTAATGACGATCCACAGCAGCAAAGGGCTCGAATTTCCAGTCGTG

FIG. 4—*Continued.*

```
      F V A G L G R N F N M M D L N K S Y L L D K E L G F G T K Y I H P Q
6401 TTTGTAGCAGGTCTCGGCCGGAATTTCAACATGATGGATTTGAACAAATCGTACCTGCTGGATAAGGAGCTCGGATTTGGCACGAAGTATATTCATCCGC

       L R I S Y P T L P L I A M K K K M R R E L L S E E L R V L Y V A L
6501 AATTACGCATCAGCTATCCGACACTTCCGCTCATTGCGATGAAGAAAAAAATGCGCAGGGAGCTGCTGTCAGAGGAATTGCGTGTGCTCTATGTTGCATT

      T R A K E K L F L I G S C K D H Q K Q L A K W Q A S A S Q T D W L
6601 AACGAGAGCGAAGGAAAAGCTGTTTCTGATTGGCTCATGTAAGGATCATCAGAAACAGCTTGCAAAATGGCAGGCATCCGCGTCCCAAACTGATTGGCTT

     L P E F D R Y Q A R T Y L D F I G P A L A R H R D L G D L A G V P A
6701 CTGCCGGAATTTGACCGCTATCAGGCGAGAACGTATCTAGATTTCATTGGGCCGGCTCTTGCCAGGCACAGAGACTTGGGGGATTTGGCTGGTGTGCCAG

      H A D I S G H P A R F A V Q M I H S Y D L L D D D L E E R M E E K
6801 CACACGCTGACATCTCAGGTCACCCGGCTCGTTTTGCCGTTCAAATGATCCATTCCTATGATTTGCTTGATGATGATCTGGAAGAAAGAATGGAAGAAAA

      S E R L E A I R R G E P V P G S F A F D E K A R E Q L S W T Y P H
6901 AAGCGAGCGCCTAGAAGCGATCCGCCGAGGTGAACCAGTTCCCGGCTCGTTTGCGTTTGATGAAAAAGCCCGCGAGCAGCTGAGCTGGACCTACCCGCAT

     Q E V T Q I R T K Q S V S E I K R K R E Y E D E Y S G R A P V K P A
7001 CAAGAAGTGACGCAAATTCGGACAAAGCAATCAGTTTCTGAGATCAAGAGAAAAAGAGAGTACGAGGATGAATACAGCGGCAGGGCCCCTGTAAAACCGG

      D G S I L Y R R P A F M M K K G L T A A E K G T A M H T V M Q H I
7101 CTGATGGAAGCATTCTGTACAGACGTCCCGCTTTTATGATGAAAAAAGGCCTGACAGCGGCAGAGAAAGGACTGCCATGCATACGGTTATGCAGCATAT

      P L S H V P S I E E A E Q T V H R L Y E K E L L T E E Q K D A I D
7201 CCCGCTGTCACATGTGCCGTCGATAGAAGAAGCTGAGCAGACGGTTCACAGGCTTTATGAAAAAGAGCTTCTCACTGAAGAACAAAAAGACGCTATTGAT

     I E E I V Q F F H T E I G G Q L I G A K W K D R E I P F S L A L P A
7301 ATAGAAGAAATCGTGCAATTTTTCCATACAGAAATCGGCGGACAGCTGATCGGTGCTAAGTGGAAGGACCGGGAAATACCATTCAGCTTAGCGCTTCCGG

      K E I Y P D A H E A D E P L L V Q G I I D C L Y E T E D G L Y L L
7401 CCAAGGAGATCTATCCTGATGCACACGAGGCAGATGAGCCGCTTTTAGTGCAGGGTATTATTGACTGTCTCTATGAAACTGAGGACGGATTATATCTATT

      D Y K S D R I E G K F Q H G F E G A A P I L K K R Y E T Q I Q L Y
7501 GGATTATAAGTCGGACCGGATTGAGGGCAAATTCCAGCATGGATTTGAAGGAGCGGCCCCGATCTTGAAGAAACGATATGAAACGCAAATTCAGCTGTAC

      T K A V E Q I A K T K V K G C A L Y F F D G G H I L T L *
7601 ACGAAGGCAGTGGAGCAAATTGCAAAAACAAAGGTAAAGGGATGTGCGCTTTATTTCTTTGACGGAGGGCACATTCTGACATTATAGCGAGATCCATAAG
                                                                                          -------------
                                                                 ORF3
                                                                 M R I L H T A D W H L G K T
7701 CTCCGGAATTTCAGGCGGAGCGGGTCTCGTTTTCTATATAAAAGAGAGGTGAAAGCATTTGCGGATTTTACATACGGCTGACTGGCATCTTGGAAAAACG
     ------>  Term <------------                 S.D.

      L E G R S R L S E Q A D V L D E L N T I V K D E Q I D A I V M A G D
7801 CTTGAAGGAAGAAGCAGGCTGAGTGAACAGGCCGATGTGCTGGATGAACTGAATACGATTGTAAAGGATGAGCAGATCGATGCCATTGTAATGGCGGGGG

       A F D T V N P P A L A E Q L F Y E S L S A L S D R G K R P I V V I
7901 ATGCATTTGACACCGTAAACCCGCCAGCTTTAGCCGAGCAGCTGTTTTATGAAAGCCTGTCTGCGCTTAGCGACAGAGGAAAGCGCCCGATCGTCGTCAT

      A G N H D N P D R L S A A S P L T H E N G I H L I G Y P T T E P I
8001 TGCCGGAAATCACGATAATCCTGACCGTTTGTCCGCCGCTTCACCGCTGACACATGAAAACGGCATTCATTTAATCGGTTATCCGACAACAGAGCCGATT

     H I E V P S A G E L L A V G A L A Y P S E A R L N E V L S D T F D E
8101 CATATTGAAGTGCCTTCAGCAGGAGAGCTTTTGGCGGTTGGAGCGCTGGCGTACCCATCTGAAGCGCGGTTAAATGAAGTGCTATCCGATACGTTTGATG

      K L L R D H Y D V K I R Q A F E H M T S R F R T D A V K I A A S H
8201 AAAAGCTGCTCCGTGATCACTATGATGTGAAAATCAGGCAGGCGTTTGAGCATATGACAAGCCGGTTCCGGACCGATGCAGTGAAGATTGCCGCCAGCCA

      I Y V A G E T K P I Q T G R L K S A V H T R W L L K A C R Q M P L
8301 TATTTATGTGGCAGGCGAAACCAAACCGATTCAGACGGGCCGATTGAAGTCGGCGGTGCATACACGGTGGCTGCTGAAAGCCTGCCGGCAGATGCCGCTT

     T L R S A I C I A R K R S S G R G R L R V I Q D L R S P T A F L K R
8401 ACGTTGCGCTCGGCCATTTGCATCGCCCGCAAACGATCAAGCGGGCGCGGACGCTTGCGCGTTATTCAGGATCTCCGCTCGCCTACAGCTTTTCTGAAGC

      A T L N Q *
8501 GGGCTACGCTAAATCAGTGACGATTGTGGACGCAAAGCCTGGGGAAGAGGCCACTTGGCAAGAGGTGTTATTATCAAGCGGCAAGCCACTTGTGAAGTGG

8601 AAAGCGGCAAATGGATTAAGTGAGGTGTACAGCTGGCTGGATGAAGGCAGAGATCAGAATGCTTGGATTGACTTGGAAATCCGTGTCGCGGACCAGCTGT

8701 CACTTGAAGAGATTCACAGGCTGCGGAAGGCG
```

FIG. 4—*Continued.*

plasmid, both *add* genes are located on an 8.8-kb chromosomal DNA fragment between a unique *Sal*I site and a unique *Sma*I site.

The *addA* and *addB* genes restore ATP-dependent exonucleolytic activity in *E. coli recBCD* mutants. To test whether the *addA* and *addB* genes could restore the ATP-dependent nuclease activity, an *E. coli recBCD* deletion mutant was transformed with plasmid pGV88 and the resulting transformants were assayed for ATP-dependent nuclease activity. Table 2 shows that ATP-dependent nuclease activity was present in the transformants, indicating (i) that the *B. subtilis add* genes can be expressed in *E. coli*, (ii) that the *add* genes must be present in an intact form on plasmid pGV88, and (iii) that all *B. subtilis* genes required for ATP-dependent nuclease activity are present on the 8.8-kb *Sal*I-*Sma*I fragment in pGV88.

In the V186 *E. coli* strains, transformed with pGV88, the viability, resistance to UV irradiation, and recombination in conjugation were also restored to the wild-type level (data not shown).

Nucleotide sequence of the *add* genes. For sequencing, we first attempted to isolate a series of overlapping subclones by insertion of fragments in the transposon-promoted deletion vectors pAA-pZ618 and pAA-pZ619 (41; Gold Biotechnology). However, since this method frequently yielded identical deletions, which is in accord with findings by others (1, 41), the larger part of the sequences was obtained by subcloning in M13. The 8.8-kb *Sal*I-*Sam*I fragment, carrying

TABLE 3. Alignment of putative ATP binding sequences in *E. coli* and *B. subtilis* enzymes[a]

Protein	Residues	Sequence[a]
E. coli		
UvrA	24–45	D K L I V V T G L S G S G K S S L A F D T L
	633–654	G L F T C I T G V S G S G K S T L I N D T L
UvrB	32–53	L A H Q T L L G V T G S G K T F T I A N V I
UvrD	22–43	R S N L L V L A G A G S G K T R V L V H R I
RecA	59–80	G R I V E I Y G P E S S G K T T L T L Q V I
RecB	16–37	Q G E R L I E A S A G T G K T F T I A A L Y
RecD	164–185	R R I S V I S G G P G T G K T T T V A K L L
B. subtilis		
AddB	1–22	L G A E F L V G R S G S G K T K L I I N S I
AddA	23–44	G Q D I L V A A A A G S G K T A V L V E R M

[a] Identical or similar residues are boxed.

the two *add* genes, was completely sequenced on both strands. The sequence is shown in Fig. 4. Three open reading frames (ORF1, ORF2, and ORF3) were present, all preceded by a putative Shine-Dalgarno sequence (24) at nucleotides (nt) 482 to 494 (sequence AAAGaGAGGGgTC), 3972 to 3984 (sequence AAAGGAGGagGAT), and 7741 to 7752 (sequence AAAGaGAGGTGA), respectively. The ΔG's of these ribosome binding sites were -18.2, -21, and -22.6 kcal (1 cal = 4.184 J), respectively, as calculated by Tinoco et al. (57). These putative ribosome binding sites were used to assign presumptive initiation codons to the three ORFs, namely, TTG at nt 502, ATG at nt 3989, and TTG at nt 7759. The spacings between the ribosome binding sites and the start codon were 13, 11, and 11 nt, respectively, which is in good agreement with other functional spacings reported in *B. subtilis* (24). The three ORFs with termination codons at nt 4000, 7685, and 8515, could encode polypeptides with 1,166, 1,232, and 253 amino acids, having MWs of 134,629, 141,090, and 28,143, respectively. Since transformation of a *B. subtilis* wild-type strain with plasmid pUC18, carrying the Em^r marker in addition to the *Hin*dII-*Eco*RI fragment from nt 452 to 1150, and with plasmid pUC18, carrying the *Hgi*AI-*Sph*I fragment from nt 3861 to 6174, resulted in Em^r transformants with the Add^+ phenotype, the *addB* gene and the *addA* gene must be controlled by their own promoter. However, upstream of the ribosome binding sites of ORF1 and ORF2, corresponding with the *addB* and *addA* genes, no consensus *B. subtilis* promoters (12, 36, 61) were detected. This suggests that the promoters of the *add* genes are weak. The findings (Fig. 3) that the end of the *addB* gene at the left is located between the *Hin*dII site (nt 452) and the *Bcl*I site (nt 550) and that the other end of this gene is located between the *Hgi*AI site (nt 3861) and the *Hin*dIII(2) site (nt 4585), and, further, that the end of the *addA* gene at the left is located between the *Hgi*AI site (nt 3861) and the *Hin*dIII(2) site (nt 4585) and that the other end of that gene is located between the *Pvu*II site (nt 7592) and the *Cla*I site (nt 7876) are in good agreement with the proposed positions of ORF1 and ORF2.

The *addA* gene is closely followed by a stem-loop structure from nt 7688 to 7729, with a ΔG of -24.6 kcal (57), which may function as a terminator (5). This region is followed by a thymine-rich region immediately downstream of the stem-loop structure, suggesting that it represents a rho-independent transcription terminator site (42).

The *addA* gene is closely followed by a third ORF. No consensus promoter sequence for ORF3 could be detected. When a *B. subtilis* wild-type strain was transformed with plasmid pUC18, jointly carrying the *Cla*I-*Bcl*I fragment from nt 7876 to 8214 and the Em^r marker from pE194, all of the resulting transformants tested (a total of 25) showed wild-type levels of ATP-dependent exonuclease activity (data not shown). Since this *Cla*I-*Bcl*I fragment represents an internal fragment of the third ORF, this result suggests that the third ORF is not involved in ATP-dependent exonuclease activity. Both protein AddB and protein AddA contain a putative conserved region for ATP binding (Table 3).

The amino acid sequences of the three ORFs were also compared with those of the *E. coli* RecB, RecC, and RecD proteins. With the exception of the conserved ATP-binding sites, no further significant sequence similarity, as defined by Lipman and Pearson (30), was observed between sequences of the *B. subtilis* AddB protein and the third ORF and those of the *E. coli* RecB, RecC, or RecD protein. However, in the amino acid sequence of the AddA protein, eight small regions of considerable sequence similarity with the *E. coli* RecB protein were present (Fig. 5). The positions of these regions of similarity in the AddA and the RecB protein

```
Region  Protein Residue                                     Residue
        AddA     30   A A A G S G K T                          37
  1                   : . : : . : : :
        RecB     23   A S A G T G K T                          30

        AddA     57   V D R L L V V T F T N A S A A E          72
  2                   : .   : : : : : : :   : . . : :
        RecB     56   V E E L L V V T F T E A A T A E          71

        AddA    406   L V D E Y Q D T                         413
  3                   . . : : . : : :
        RecB    382   M I D E F Q D T                         389

        AddA    437   G D V K Q S I Y R F R L A               449
  4                   : :   : : . : :   : :   :
        RecB    413   G D P K Q A I Y A F R G A               425

        AddA    584   D I V I L L R S                         591
  5                   : :   . : . : :
        RecB    552   D I S V L V R S                         559

        AddA    792   T I H S S K G L E F P V V               804
  6                   : : :   : : : : : . : . :
        RecB    742   T I H K S K G L E Y P L V               754

        AddA    858   E L L S E E L R V L Y V A L T R A       874
  7                   :   : . : . : : . : : : : : : : .
        RecB    795   E R L A E D L R L L Y V A L T R S       811

        AddA   1168   Y L L D Y K S                          1175
  8                   : : : : : : :
        RecB   1076   Y L L D Y K S                          1081
```

FIG. 5. Sequence similarity between regions of the *B. subtilis* AddA protein and the *E. coli* RecB protein. A double dot indicates that the amino acids are identical; single dots indicate conservative replacements.

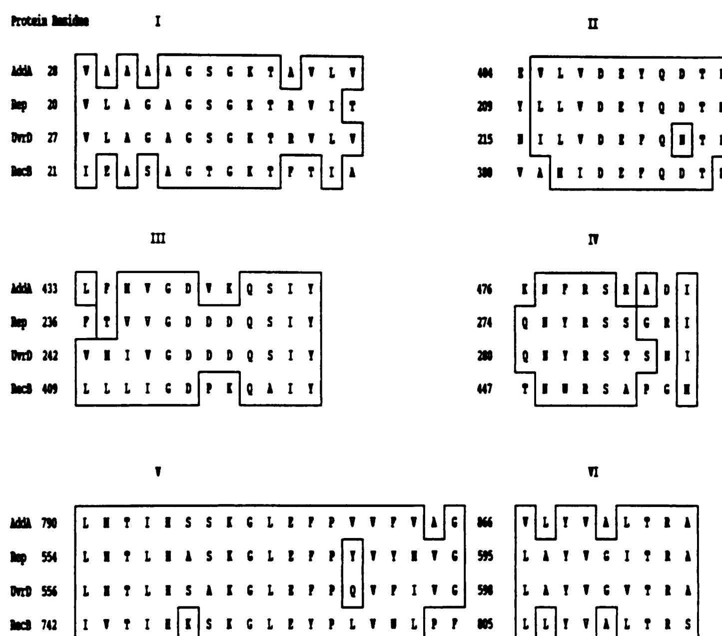

FIG. 6. Conserved regions in the AddA protein, the UvrD helicase, the Rep helicase, and the RecB protein. Identical or similar residues are boxed.

corresponded with each other. Region 1 of similarity (residues 30 to 37 in the AddA protein) contains the conserved region for ATP binding (Table 3).

Recently, six conserved motifs were reported to be present in various proteins with DNA helicase activity (25). These regions in two *E. coli* DNA helicases, i.e., the UvrD and Rep proteins (7, 15, 21, 62), and in the *E. coli* RecB protein are presented in Fig. 6 (17). This figure shows that the conserved motifs are also present in the *B. subtilis* AddA protein, suggesting that these regions are involved in the DNA unwinding activity of the *B. subtilis* ATP-dependent nuclease. The first region of similarity in Fig. 6 contains the conserved region for ATP binding (Table 3). The percent similarity between the conserved regions of the AddA protein and those of the DNA helicases, the UvrD and Rep proteins, is of the same order of magnitude as that between the regions of the UvrD and Rep proteins (Table 4). The large stretch of similarity (19 amino acids; Fig. 6, region V) also appeared to be highly conserved in the AddA, UvrD, and Rep helicases (Table 4), suggesting that this region may be essential for DNA helicase activity.

Expression of *add* genes under control of the T7 promoter. To achieve a high degree of expression of the *add* genes, the 8.8-kb *Sal*I-*Sma*I fragment carrying the *add* genes was linked to the T7 promoter in plasmids pSPT18 and pSPT19. These plasmids were introduced into the *E. coli* T7 RNA polymerase expression system (strain BL21DE3) (53). The results (Fig. 7) show that if the *add* genes were linked to the T7 promoter in the proper orientation, as in pSPT19, large quantities of three proteins were produced, which were absent in extracts of cells, carrying plasmid pSPT18 or pSPT19 lacking the 8.8-kb *Sal*I-*Sma*I fragment. These proteins also were not seen in cells carrying plasmid pSPT18 in which the *add* genes on the 8.8-kb *Sal*I-*Sma*I fragment were linked to the T7 promoter in the reverse orientation. The molecular masses estimated from the positions of the three proteins on the polyacrylamide gel, 112, 128, and 36 kDa, were of the same order of magnitude as those calculated from the amino acid sequences of the three ORFs on the 8.8-kb *Sal*I-*Sma*I fragment.

DISCUSSION

An 8.8-kb *B. subtilis* chromosomal DNA fragment was isolated which contained all of the genes required for ATP-dependent nuclease activity. By Campbell-type integration of plasmid pUC18 carrying various restriction fragments of this 8.8-kb fragment, it was established that only two *add*

TABLE 4. Percent identity between conserved motifs in the *B. subtilis* AddA protein and the *E. coli* UvrD protein, Rep protein, and RecB protein

Proteins	Identity (%)[a] with domain:					
	I (14 aa)	II (11 aa)	III (12 aa)	IV (9 aa)	V (19 aa)	VI (9 aa)
AddA-Rep	64 (71)	82 (91)	58 (67)	44 (56)	68 (89)	56 (78)
AddA-UvrD	79 (79)	64 (82)	58 (75)	44 (67)	74 (89)	56 (78)
AddA-RecB	43 (71)	45 (73)	58 (83)	33 (56)	53 (89)	78 (100)
Rep-UvrD	93 (93)	64 (82)	75 (83)	67 (78)	74 (95)	89 (100)

[a] The values are the percentages of amino acids (aa) which are identical in the various motifs. The percentages of similarity (identical plus conservative replacements) are given in parentheses.

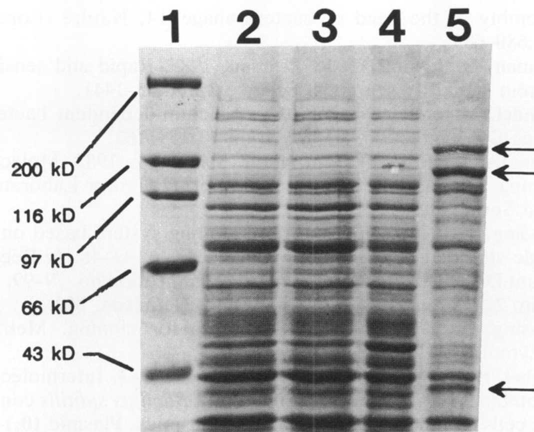

FIG. 7. Synthesis of proteins in *E. coli* BL21DE3 carrying plasmid pSPT18 or pSPT19 with and without the 8.8-kb *B. subtilis* *SalI-SmaI* fragment. Proteins were separated by polyacrylamide gel electrophoresis and stained with Coomassie blue. Lane 1, molecular size markers; lane 2, *E. coli*(pSPT18); lane 3, *E. coli*(pSPT19); lane 4, *E. coli*(pSPT18) (carrying the 8.8-kb *SalI-SmaI* fragment in which the *add* genes are linked to the T7 promoter in the reverse orientation); lane 5, *E. coli*(pSPT19) (carrying the 8.8-kb *SalI-SmaI* fragment in which the *add* genes are linked to the T7 promoter in the proper orientation). The arrows indicate the positions of the proteins encoded by the three genes on the 8.8-kb *SalI-SmaI* fragment.

genes were located on the 8.8-kb DNA fragment. Sequencing of the 8.8-kb DNA fragment revealed that those genes could encode two proteins, AddB and AddA. Although both *add* genes must be controlled by their own promoters, they slightly overlap, suggesting that they may constitute an operon. This has been suggested with respect to the *E. coli* *prt*, *recB*, and *recD* genes, which also show a limited degree of sequence overlap (16–18), notwithstanding available evidence that the *recB* and *recD* genes are controlled by their own promoters (3, 40, 50). No regions with high similarity to consensus sequences of known *B. subtilis* promoters (12, 36, 61) were present, suggesting that the *add* genes are controlled by weak promoters. This is in agreement with the low level of expression which was observed in transcriptional fusions of the *add* genes with a promoterless *lacZ* gene and measuring of β-galactosidase activity, which was not inducible by MC treatment of the cells (data not shown). In *E. coli*, expression of *recBCD* genes is also weak. It has been suggested that this weak expression may be due to either a weak promoter activity of the *recBCD* genes (19) or to a low efficiency of translation (16, 17).

In *E. coli*, mutations in the *recB* and *recC* genes, but not in the *recD* gene (3), result in a reduced recombination and repair of DNA damage caused by UV irradiation (11). Mutations in both the *B. subtilis addB* and *addA* genes result in reduced transformation and an increased sensitivity to UV irradiation and MC treatment (28). Thus, it would seem that the *B. subtilis add* genes are functionally similar to the *E. coli recB* and *recC* genes. Also, the molecular masses of the *B. subtilis* AddB and AddA proteins (135 and 141 kDa, respectively), calculated from the amino acid sequences, are of the same order of magnitude as those of the *E. coli* RecC and RecB proteins (129 and 134 kDa, respectively) (17, 19), whereas their mobilities on a polyacrylamide gel are approximately the same as those of the RecB and RecC proteins (3).

In many respects, the region involved in synthesis of the ATP-dependent nuclease in *B. subtilis* is different from that in *E. coli*. In *E. coli*, the *recBC* region is interrupted by the presence of a protease (*prt*) gene (18), whereas in *B. subtilis*, the *add* genes are contiguous. Furthermore, in *E. coli*, three genes encode subunits of the ATP-dependent exonuclease, whereas only two genes in *B. subtilis* are required for this enzyme activity. The *E. coli recB* gene is followed by the *recD* gene. In *B. subtilis*, the *addA* gene, showing a slight overall degree of similarity with the *E. coli recB* gene, is also followed by a third ORF. The molecular mass of the protein encoded by this open reading frame is 28 kDa, which is much smaller than the 67-kDa molecular mass of the *E. coli* RecD protein (16). The amino acid sequence of the *B. subtilis* protein does not show significant sequence similarity with that of the RecD protein. All of these differences suggest that this *B. subtilis* protein is not involved in ATP-dependent exonuclease activity. This is supported by the observation that interruption of this ORF by plasmid insertion did not abolish ATP-dependent exonuclease activity in *B. subtilis*, measured as breakdown of DNA to acid-soluble products, whereas it has been shown that, for DNA hydrolysis, the RecD subunit in the RecBCD complex is indispensable (40).

Both *B. subtilis* Add proteins contain a conserved region for ATP binding which is also present in the *E. coli* RecB protein (17). The AddB protein shows no significant amino acid sequence similarity with the *E. coli* RecC protein. Although the overall amino acid sequence similarity between the AddA protein and the *E. coli* RecB protein is low, both proteins share a number of small regions of significant sequence similarity. One of these regions contains the conserved region for ATP binding. The AddA protein shares six conserved motifs with *E. coli* helicases (the UvrD protein and Rep protein). This suggests that these regions play an important role in the DNA unwinding activity of the enzyme. Five of the eight regions of amino acid sequence similarity in the AddA protein and the *E. coli* RecB protein (Fig. 5, regions 1, 3, 4, 6, and 7) contain a conserved motif of DNA helicases (cf. Fig. 6). The three other conserved regions of similarity (Fig. 5, regions 2, 5, and 8) may be involved in other properties of the ATP-dependent nuclease. The molecular masses of the AddB and AddA proteins calculated from the amino acid sequences were 135 and 141 kDa, respectively. These molecular masses deviate greatly from the data presented by Doly and Anagnostopoulos (13), who reported that the ATP-dependent nuclease of *B. subtilis* consists of five subunits, with molecular masses of 81,000, 70,000, 62,000, 52,000, and 42,000 Da. Our data agree reasonably well with those of Chestukhin et al. (10), who found that the enzyme was composed of only two subunits, with molecular masses of 140 and 155 kDa as estimated from polyacrylamide gel electrophoresis.

REFERENCES

1. **Ahmed, A.** 1985. A rapid procedure for DNA sequencing using transposon-promoted deletions in *Escherichia coli*. Gene **39:** 305–310.
2. **Amundsen, S. K., A. M. Neiman, S. M. Thibodeaux, and G. R. Smith.** 1990. Genetic dissection of the biochemical activities of the RecBCD enzyme. Genetics **126:**25–40.
3. **Amundsen, S. K., A. F. Taylor, A. M. Chaudhury, and G. R. Smith.** 1986. *Rec*D: the gene for an essential third subunit of exonuclease V. Proc. Natl. Acad. Sci. USA **83:**5558–5562.

3a. **Anagnostopoulos, C.** Personal communication.

4. **Biswal, N., A. K. Kleinschmidt, H. C. Spatz, and T. A. Trautner.** 1967. Physical properties of the DNA of bacteriophage SP50. Mol. Gen. Genet. **100:**39–55.
5. **Brendel, V., and E. N. Trifonov.** 1984. A computer algorithm for testing potential prokaryotic terminators. Nucleic Acids Res. **12:**4411–4427.

6. **Bron, S., and G. Venema.** 1972. Ultraviolet inactivation and excision repair in *Bacillus subtilis*. I. Construction and characterization of a transformable eightfold-auxotrophic strain and two ultraviolet-sensitive derivatives. Mutat. Res. **15**:1–10.
7. **Chao, K., and T. M. Lohman.** 1990. DNA and nucleotide-induced conformational changes in the *Escherichia coli* Rep and Helicase II (UvrD) proteins. J. Biol. Chem. **265**:1067–1076.
8. **Chaudhury, A. M., and G. R. Smith.** 1984. *Escherichia coli recBC* deletion mutants. J. Bacteriol. **160**:788–791.
9. **Cheng, K. C., and G. R. Smith.** 1987. Cutting of Chi-like sequences by the RecBCD enzyme of *Escherichia coli*. J. Mol. Biol. **194**:747–750.
10. **Chestukhin, A. V., M. F. Shemyakin, N. A. Kalinina, and A. A. Prozorov.** 1972. Some properties of ATP dependent deoxyribonucleases from normal and *rec*-mutant strains of *Bacillus subtilis*. FEBS Lett. **24**:121–125.
11. **Clark, A. J.** 1973. Recombination deficient mutants of *Escherichia coli* and other bacteria. Annu. Rev. Genet. **7**:67–86.
12. **Doi, R. H., and L.-F. Wang.** 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. Microbiol. Rev. **50**:227–243.
13. **Doly, J., and C. Anagnostopoulos.** 1976. Isolation, subunit structure and properties of the ATP-dependent deoxyribonuclease of *Bacillus subtilis*. Eur. J. Biochem. **71**:309–316.
14. **Doly, J., E. Sasarman, and C. Anagnostopoulos.** 1974. ATP-dependent deoxyribonuclease in *Bacillus subtilis* and a mutant deficient in this activity. Mutat. Res. **22**:15–23.
15. **Finch, P. W., and P. T. Emmerson.** 1984. The nucleotide sequence of the *uvrD* gene of *E. coli*. Nucleic Acids Res. **12**:5789–5799.
16. **Finch, P. W., A. Storey, K. Brown, I. D. Hickson, and P. T. Emmerson.** 1986. Complete nucleotide sequence of *recD*, the structural gene for the α subunit of exonuclease V of *Escherichia coli*. Nucleic Acids Res. **14**:8583–8594.
17. **Finch, P. W., A. Storey, K. E. Chapman, K. Brown, I. D. Hickson, and P. T. Emmerson.** 1986. Complete nucleotide sequence of the *Escherichia coli recB* gene. Nucleic Acids Res. **14**:8573–8582.
18. **Finch, P. W., R. E. Wilson, K. Brown, I. D. Hickson, and P. T. Emmerson.** 1986. Complete nucleotide sequence of the *Escherichia coli ptr* gene encoding protease III. Nucleic Acids Res. **14**:7695–7703.
19. **Finch, P. W., R. E. Wilson, K. Brown, I. D. Hickson, A. E. Tomkinson, and P. T. Emmerson.** 1986. Complete nucleotide sequence of the *Escherichia coli recC* gene and of the *thyA-recC* intergenic region. Nucleic Acids Res. **14**:4437–4451.
20. **Frischauf, A. M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
21. **Gilchrist, C. A., and D. T. Denhardt.** 1987. *Escherichia coli rep* gene: sequence of the gene, the encoded helicase, and its homology with *UvrD*. Nucleic Acids Res. **15**:465–475.
22. **Goldmark, P. J., and S. Linn.** 1970. An endonuclease activity from *Escherichia coli* absent from certain *rec*$^-$ strains. Proc. Natl. Acad. Sci. USA **67**:434–441.
23. **Goldmark, P. J., and S. Linn.** 1972. Purification and properties of the *recBC* DNase of *Escherichia coli* K-12. J. Biol. Chem. **247**:1849–1860.
24. **Hager, P. W., and J. C. Rabinowitz.** 1985. Translational specificity in *Bacillus subtilis*, p. 1–31. *In* D. Dubnau (ed.), The molecular biology of the bacilli, vol. 2. Academic Press, Inc., New York.
25. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333**:22–23.
26. **Iordanescu, S.** 1977. Relationships between cotransducible plasmids in *Staphylococcus aureus*. J. Bacteriol. **129**:71–75.
27. **Ish-Horowicz, D., and F. J. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2999.
28. **Kooistra, J., B. Vosman, and G. Venema.** 1988. Cloning and characterization of a *Bacillus subtilis* transcription unit involved in ATP-dependent DNase synthesis. J. Bacteriol. **170**:4791–4797.
29. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
30. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
31. **Mandel, M., and A. Higa.** 1970. Calcium-dependent bacteriophage DNA infection. J. Mol. Biol. **53**:159–162.
32. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Messing, J.** 1979. A multipurpose cloning system based on the single-stranded DNA bacteriophage M13, p. 43–48. *In* Recombinant DNA Technical Bulletin, NIH publication no. 79–99, vol. 2, no. 2. National Institutes of Health, Bethesda, Md.
34. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
35. **Michel, B., B. Niaudet, and S. D. Ehrlich.** 1983. Intermolecular recombination during transformation of *Bacillus subtilis* competent cells by monomeric and dimeric plasmids. Plasmid **10**:1–10.
36. **Moran, C. P., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186**:339–346.
37. **Natt, E., and G. Scherer.** 1986. EMBL12, a new lambda replacement vector with sites for *Sal*I, *Xba*I, *Bam*HI, *Sst*I and *Eco*RI. Nucleic Acids Res. **14**:7128.
38. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligo deoxynucleotide-directed mutagenesis. Gene **26**:101–106.
39. **Oishi, M.** 1969. An ATP-dependent deoxyribonuclease from *Escherichia coli* with a possible role in genetic recombination. Proc. Natl. Acad. Sci. USA **64**:1292–1299.
40. **Palas, K. M., and S. R. Kushner.** 1990. Biochemical and physical characterization of exonuclease V from *Escherichia coli*. J. Biol. Chem. **265**:3447–3454.
41. **Peng, Z.-G., and R. Wu.** 1986. A simple and rapid nucleotide sequencing strategy and its application in analyzing a rice histone 3 gene. Gene **45**:247–252.
42. **Platt, T.** 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55**:339–372.
43. **Ponticelli, A. S., D. W. Schultz, A. F. Taylor, and G. R. Smith.** 1985. Chi-dependent DNA strand cleavage by RecBC enzyme. Cell **41**:145–151.
44. **Queen, C., and L. J. Korn.** 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
45. **Rigby, P. W. J., M. Dieckman, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
46. **Roman, L. J., and S. C. Kowalczykowski.** 1989. Characterization of the helicase activity of the *Escherichia coli* RecBCD enzyme using a novel helicase assay. Biochemistry **28**:2863–2873.
47. **Roman, L. J., and S. C. Kowalczykowski.** 1989. Formation of heteroduplex DNA promoted by the combined activities of *Escherichia coli* recA and recBCD proteins. J. Biol. Chem. **264**:18340–18348.
48. **Rosamond, J., K. M. Telander, and S. Linn.** 1979. Modulation of the action of the *recBC* enzyme of *Escherichia coli* K-12 by Ca^{2+}. J. Biol. Chem. **254**:8646–8652.
49. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
50. **Sasaki, M., T. Fujeyoshi, K. Shimada, and Y. Tokagi.** 1982. Fine structure of the *rec*B and *rec*C gene region of *Escherichia coli*. Biochem. Biophys. Res. Commun. **109**:414–422.
51. **Shemyakin, F. M., A. A. Grepachevsky, and A. V. Chestukhin.** 1979. Properties of *Bacillus subtilis* ATP-dependent deoxyribonuclease. Eur. J. Biochem. **98**:417–423.
52. **Spizizen, J.** 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. USA **44**:1072–1078.
53. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of

cloned genes. J. Mol. Biol. **189**:113–130.

54. **Taylor, A. F., D. W. Schultz, A. S. Ponticelli, and G. R. Smith.** 1985. RecBC enzyme nicking at Chi sites during DNA unwinding: location and orientation dependence of the cutting. Cell **41**:153–163.
55. **Taylor, A. F., and G. R. Smith.** 1980. Unwinding and rewinding of DNA by the recBC enzyme. Cell **22**:447–457.
56. **Telander-Muskavitch, K. M., and S. Linn.** 1981. *Rec*BC-like enzymes: exonuclease V deoxyribonucleases, p. 233–250. *In* P. D. Boyer (ed.), The enzymes, vol. 14. Academic Press, Inc., New York.
57. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.
58. **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.
59. **Vosman, B., J. Kooistra, J. Olijve, and G. Venema.** 1987. Cloning in *Escherichia coli* of the gene specifying the DNA-entry nuclease of *Bacillus subtilis*. Gene **52**:175–183.
60. **Wang, T. V., and K. C. Smith.** 1989. The roles of RecBCD, Ssb and RecA proteins in the formation of heteroduplexes from linear-duplex DNA in vitro. Mol. Gen. Genet. **216**:315–320.
61. **Wu, J. J., M. J. Howard, and P. J. Piggot.** 1989. Regulation of transcription of the *Bacillus subtilis spoIIA* locus. J. Bacteriol. **171**:692–698.
62. **Yamamoto, Y., T. Ogawa, H. Shinagawa, T. Nakayama, H. Matsuo, and H. Ogawa.** 1986. Determination of the initiation sites of transcription and translation of the *uvrD* gene of *Escherichia coli*. J. Biochem. **99**:1579–1590.
63. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC vectors. Gene **33**:103–119.