JOURNAL OF VIROLOGY, Oct. 1984, p. 9–15
0022-538X/84/100009-07$02.00/0

Vol. 52, No. 1

Nucleotide Sequence of the Temperate *Bacillus subtilis* Bacteriophage SPO2 DNA Polymerase Gene *L*

BENNY RÅDÉN AND LARS RUTBERG*

Department of Bacteriology, Karolinska Institutet, S-104 01 Stockholm, Sweden

Received 15 February 1984/Accepted 6 May 1984

Temperate *Bacillus subtilis* phage SPO2 codes for a phage-specific DNA polymerase. The polymerase gene has been cloned, and its nucleotide sequence has been determined. Within the sequence there is an open reading frame starting with a TTG and ending with three consecutive translational stop codons. Ten base pairs upstream from the proposed TTG initiation codon there is a probable ribosome-binding site with a calculated free energy of interaction with the 3′ end of *B. subtilis* 16S rRNA of −15 kcal (−63 kJ)/mol. Based on the sequence and the expression of the polymerase gene in three different hybrid plasmids, we conclude that this open reading frame is the structural gene for SPO2 DNA polymerase. The predicted molecular weight of the polymerase is 72,486. In hybrid plasmid pJB74, the terminal triplet of an open reading frame with coding capacity for a protein of ca. 10 kilodaltons overlaps with the translational initiation triplet TTG of the polymerase gene. We speculate that transcription and translation of this open reading frame can influence the amount of phage DNA polymerase made in SPO2-infected bacteria.

SPO2 is a temperate *Bacillus subtilis* bacteriophage (3). Some 20 genes have been identified in SPO2 by complementation and ordered on a linear genetic map (38). SPO2 and most virulent *B. subtilis* phages grow in the presence of 6-(*p*-hydroxyphenylazo)-uracil, an inhibitor of *B. subtilis* DNA polymerase III (4). It is proposed that 6-(*p*-hydroxyphenylazo)-uracil-resistant phages produce their own DNA polymerase and replicate independently of DNA polymerase III (12). A new DNA polymerase has been found in SPO2-infected bacteria; it is the product of a complementation group called gene *L* (28).

Recently we cloned the wild-type allele of gene *L* mutation *susL244* in plasmid pC194 (30). Three hybrid plasmids were isolated (originally called pC194-74, pC194-91, and pC194-96 [30]); one of these, pJB96, causes a 10- to 20-fold increase in the level of in vitro DNA polymerase activity in a *B. subtilis polA*(Ts) *polC*(Ts) mutant, indicating that it contains a complete gene *L*. For the other two plasmids a smaller increase (pJB74) or no increase (pJB91) in DNA polymerase activity was found. Gene *L* from pJB96 has been subcloned in *Escherichia coli* (31), where it is also expressed, which demonstrates that the gene codes for a new DNA polymerase rather than an activity which modifies a *B. subtilis* enzyme.

Restriction site maps of pJB74, pJB91, and pJB96 show that all three plasmids contain a complete gene *L* (this paper). To understand the different levels of expression of gene *L* in the hybrid plasmids, we determined its nucleotide sequence. The predicted molecular size of the enzyme is 72 kilodaltons (kd). The results indicate that the polymerase is translated from a polycistronic messenger and that its synthesis may be controlled both at the transcriptional and translational level.

MATERIALS AND METHODS

Bacteria and phage. The following *B. subtilis* strains were used: 3G18 (*ade met trpC2*), BD294 (*trpC2 thr-5 polA59 polC25*), SR135 (*trp-7 spoOA9 sup-3*), and W168 (prototrophic). Strain BD294 was obtained from T. A. Trautner.

Phage SPO2c_1 was originally obtained from J. Marmur. SPO2 *susL244* (38) was obtained from T. A. Trautner. The bacteria were kept on TBAB plates (tryptic blood agar base; Difco Laboratories). For strains carrying pC194 or pUB110, chloramphenicol (20 μg/ml) or kanamycin (5 μg/ml), respectively, was added to the plates. SPO2 was assayed on NY plates as described (28). Liquid cultures were grown in NY broth with added antibiotic when appropriate. Phage stocks were prepared in liquid culture or on plates as described (28). *Escherichia coli* JM103 [Δ(*lac pro*) *thiE strA supE endA sbcB hsdR*$^-$ F′ *traD36 proAB lacI*^{Q} *z*Δ*M15*] and M13mp8 and M13mp9 were obtained from L.-O. Hedén. The methods for growing M13 and preparing RF forms were essentially as described by Messing et al. (19–21).

Plasmids and preparation of plasmid DNA. pC194 is a 2.9-kilobase-pair *Staphylococcus aureus* plasmid (14, 16) which can be maintained in *B. subtilis* (8); it confers chloramphenicol *Cm* resistance. The isolation of pC194-SPO2 hybrids has been described (30). pUB110 is a 4.5-kilobase-pair plasmid originally described by Gryczan et al. (10, 11); it confers kanamycin *Km* resistance. Plasmid DNA was prepared as described by Canosi et al. (5), except that preparation of a clear lysate was omitted and instead the whole lysate was used in the first cesium chloride-ethidium bromide centrifugation. In our hands this considerably improves the recovery of covalently closed circular plasmid DNA. Minipreparations of plasmid were prepared as described by Birnboim and Doly (2). DNA concentrations were determined on agarose plates containing 1 μg of ethidium bromide per ml with calf thymus DNA as the standard (17).

Transformation. For transformation of *B. subtilis* with plasmid DNA, competent cells (1) or protoplasts (6) were used. *E. coli* transfection was done as described by Maniatis et al. (17).

Electrophoresis. DNA restriction fragments were analyzed by using 0.8 to 2% agarose gels or 5% polyacrylamide-bisacrylamide (29:1) gels. For sequencing, the samples were electrophoresed on 8 and 6% polyacrylamide-bisacrylamide (19:1) gels with 8 M urea. The buffer system used was 90 mM Tris base–90 mM boric acid–1 mM EDTA (pH 8.3).

Structure of pJB74, pJB91, and pJB96. SPO2 gene *L* was originally cloned from partially *Hin*dIII-digested phage DNA (30). Four *Hin*dIII fragments are common to all plasmids;

* Corresponding author.

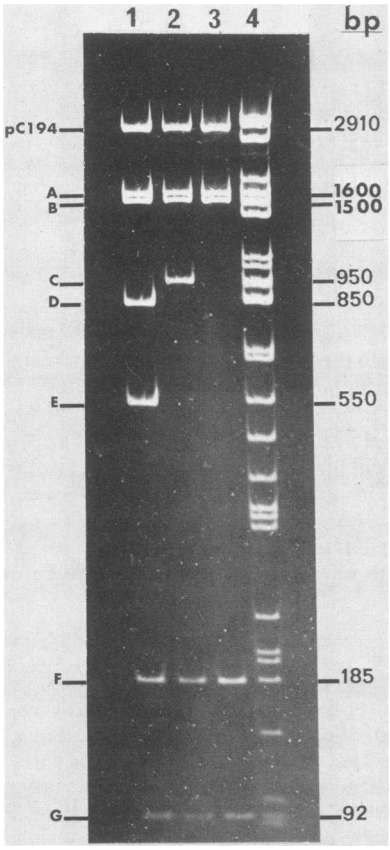

FIG. 1. Polyacrylamide gel electrophoresis pattern of *Hin*dIII-cleaved plasmid DNA. Lane 1, pJB74; lane 2, pJB91; lane 3, pJB96; lane 4, SPO2 DNA. *Eco*RI-cleaved ϕ105 DNA (33) and *Hae*III-cleaved pBR322 (35) were used as size markers.

pJB91 contains one and pJB74 contains two additional fragments (Fig. 1). The order of the *Hin*dIII fragments was determined from single and double digests with *Bgl*I, *Hin*cII, *Hin*dIII, and *Pvu*II (Fig. 2). The order of the four common *Hin*dIII fragments is the same in the three plasmids, but their orientation relative the vector pC194 is inverted in pJB96 compared with pJB74 and pJB91. Unique *Eco*RI and *Pst*I sites are located in pJB74 *Hin*dIII fragment E.

Sequencing. DNA sequencing was done by the dideoxy method (32) by using the M13 cloning system (19). Template single-stranded DNA was prepared by infecting exponentially growing JM103 with phage and incubating with shaking at 37°C for 12 to 18 h. The bacteria were then pelleted by centrifugation in an Eppendorf centrifuge, and M13 single-stranded DNA prepared from the supernatant as described by Messing (21).

The sequencing strategy was based on the structure of pJB74 (see Fig. 2). *Hin*dIII fragments E, B, F, and G were isolated from agarose or polyacrylamide gels, electroeluted, and then purified on a Whatman DE-52 column. The 2,950-base-pair (bp) *Pvu*II fragment, the 1,950-bp *Hin*cII fragment, and the 445-bp *Pst*I-*Pvu*II fragment (which overlaps *Hin*dIII fragments E and B) were similarly isolated and purified. Fragments E, B, F, G, the 2,950-bp *Pvu*II fragment, and the 1,950-bp *Hin*dII fragment were then cloned in both directions into M13mp8. The two *Eco*RI-HindIII subfragments (E′ and E″) of *Hin*dIII fragment E and the 445-bp *Pst*I-*Pvu*II fragment were cloned into M13mp9. With the above constructions, the areas indicated by dotted lines in Fig. 3 could be sequenced, but the order of fragments F and G could not be determined.

To sequence the rest of gene *L*, a modification of the method of Poncz et al. (26) was used. Thirty micrograms of the RF form of M13mp8 carrying the 2,950-bp *Pvu*II fragment was linearized by cleaving with *Eco*RI. The DNA was

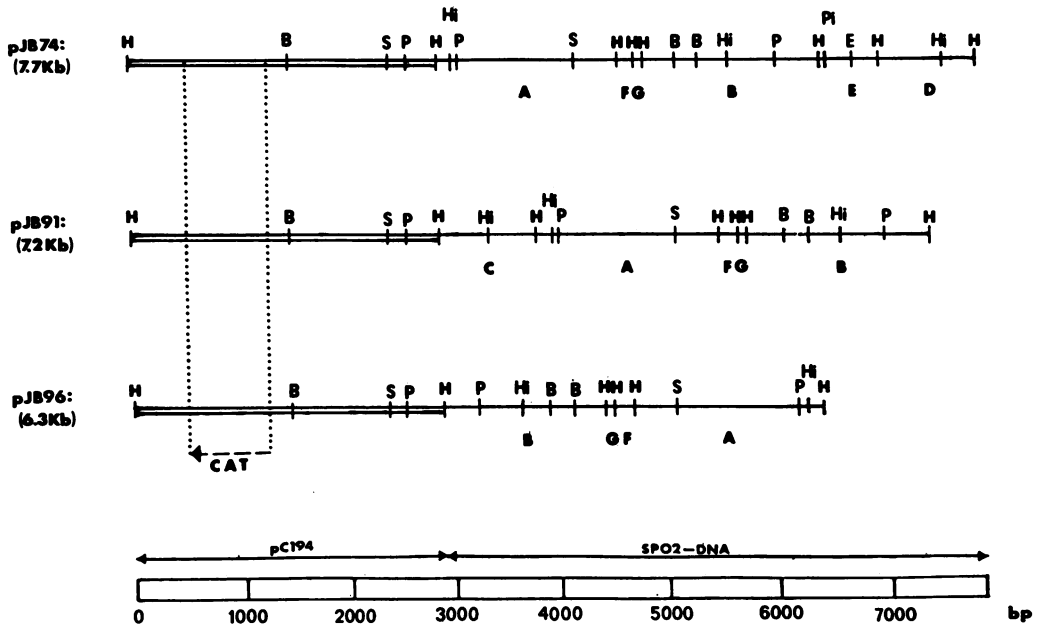

FIG. 2. Restriction maps of pC194-SPO2 hybrid plasmids. The double line represents pC194, and the single line represents SPO2 DNA. Symbols used for restriction enzyme sites are: B, *Bgl*I; E, *Eco*RI; Hi, *Hin*cII; H, *Hin*dIII; Pi, *Pst*I; P, *Pvu*II; and S, *Sau*3A. The *Pvu*II site in pC194 is actually two sites separated by 40 bp as determined by restriction mapping and sequencing (L. Melin, unpublished data). The *Sau*3A mapping is not complete, but the two sites relevant for the construction of pTB9 are shown (see Fig. 4). The relative order of *Hin*dIII fragments F and G was determined by nucleotide sequencing. The capital letters below the lines refer to the SPO2 *Hin*dIII fragments. CAT, Chloramphenicol transacetylase; the arrow indicates the transcriptional direction (13).

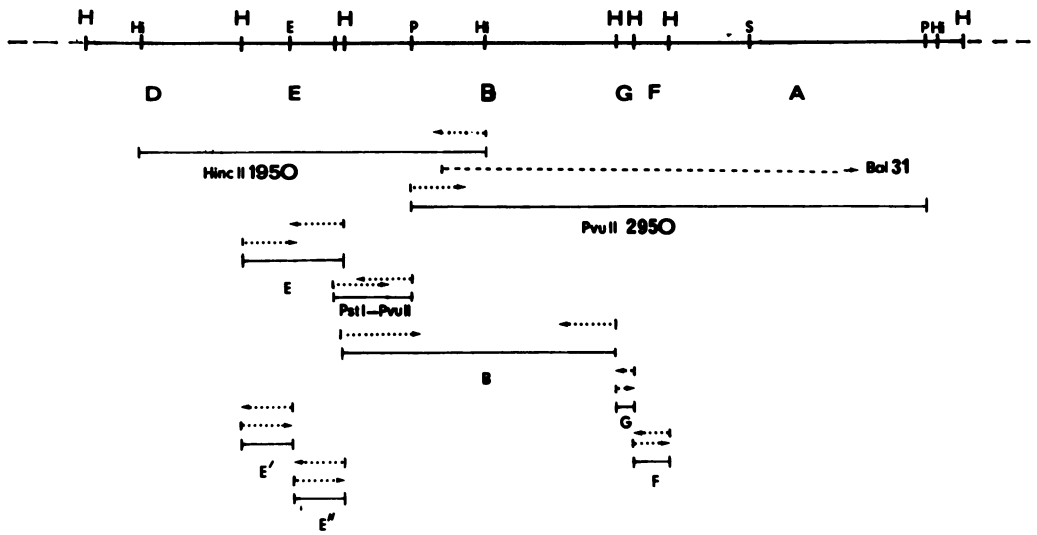

FIG. 3. Strategy for sequencing gene *L* from pJB74. The upper line represents the SPO2 part of pJB74. Above the line are shown the restriction enzymes sites (symbols are as described in the legend to Fig. 2). The capital letters below the line denote the SPO2 *Hin*dIII fragments. The thin continuous lines represent restriction fragments cloned in M13mp8 and M13mp9. The dotted lines and the arrows denote the parts of the fragments that were sequenced and the direction of sequencing. The dotted line marked Bal 31 denotes the part of the 2,950-bp *Pvu*II fragment sequenced from the M13mp9 clones derived from the partially Bal 31-digested fragment.

extracted with phenol, ethanol precipitated, and then suspended in 20 mM Tris-hydrochloride (pH 8.0), 12 mM $CaCl_2$, 12 mM $Mg Cl_2$, 600 mM NaCl, and 1 mM EDTA. Six units of nuclease Bal 31 was added to the sample at 37°C. At 1-min intervals, 2-μl samples were withdrawn into 1 μl of 0.5 M EDTA on ice to stop the reaction. The extent of the reaction was monitored by electrophoresing aliquots on 1% agarose gels. The samples were pooled, extracted with phenol, ethanol precipitated, and suspended in 50 μl of 10 mM Tris-hydrochloride (pH 7.5)–10 mM $MgSO_4$–1 mM dithiotreitol–50 mM NaCl–1 mM each of the four deoxyribonucleotides (dATP, dCTP, dGTP, dTTP). One unit of the Klenow fragment of *E. coli* DNA polymerase I was added, and the sample was incubated at 25°C for 15 min. The sample was then heated to 70°C for 5 min to destroy the DNA polymerase, and after being cooled to 37°C, it was cleaved with *Pst*I. The DNA was electrophoresed on a vertical 0.8% agarose gel. Fragments smaller than 2,950 bp were cut out of the gel and sliced into eight consecutive fractions. DNA was electroeluted from each fraction, purified, and cloned into M13mp9 cleaved with *Pst*I and *Hin*cII. The DNA was used to transfect JM103, and phage carrying inserts was identified on plates containing isopropyl-β-D-thiogalactopyranoside and 5-mono-4-chloro-3-indolyl-β-D-galactoside (22). In this manner we could generate a series of clones carrying progressively shortened fragments of the 2,950-bp *Pvu*II fragment in a fixed orientation. A preliminary characterization of the clones could be done by hybridizing DNA from the clones with DNA from some of the previously cloned fragments, e.g., *Hin*dIII-G or the 1,950-bp *Hin*cII fragment. Sequence analysis of clones from the eight groups gave overlapping segments (10 to 99% overlap) throughout the whole of the 2,950-bp *Pvu*II fragment. A total of 10 Bal 31-derived fragments were sequenced, covering a region extending 781 bp into the A fragment. From the overlaps the relative order of *Hin*dIII fragments G and F could be determined.

Enzymes. Restriction enzymes were obtained from New England Biolabs or Boehringer Mannheim Biochemicals. The reaction conditions were as described by Maniatis et al. (17), T4 ligase was obtained from New England Biolabs, and the Klenow fragment of *E. coli* DNA polymerase I was obtained from Boehringer. SPO2 DNA polymerase was assayed as previously described (30).

Chemicals. Calf thymus DNA, ethidium bromide, bovine pancreas DNase, lysozyme, isopropyl-β-D-thiogalactopyranoside, 5-mono-4-chloro-3-indolyl-β-D-galactopyranoside, and 2-mercaptoethanol were obtained from Sigma Chemical Co. Unlabeled deoxyribonucleotides and dideoxyribonucleotides were obtained from Sigma or P-L Biochemicals. [^{3}H]dTTP was from New England Nuclear Corp.; [α-^{32}P]dCTP was from New England Nuclear or Amersham Corp. Agarose, polyacrylamide, and bisacrylamide were from Bio-Rad Laboratories. Polyethylene glycol 6000 was from BDH. CsCl was from BDH or Merck & Co., Inc. All

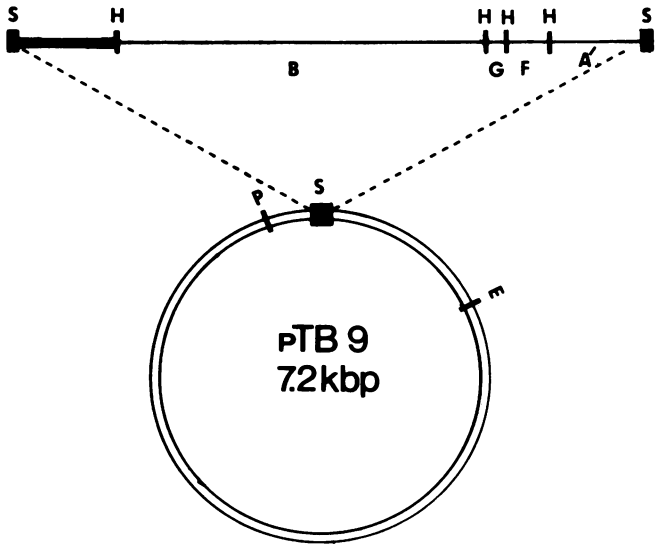

FIG. 4. Structure of pTB9. The unfilled and filled double lines are pUB110 and pC194, respectively. The *Pvu*II site and the *Eco*RI site in pUB110 are according to Jalanko et al. (15). Symbols for restriction sites are as described in the legend to Fig. 2.

other chemicals used were of purest commercially available quality.

RESULTS

Properties of strain BD294 carrying pJB74, pJB91, or pJB96. Some properties of BD294(pJB74), BD294(pJB91), and BD294(pJB96) are summarized in Table 1. Extracts from bacteria carrying pJB74 or pJB96 have increased levels of DNA polymerase activity, whereas only background levels are found in BD294(pJB91). The increased activity is not a gene dosage effect. The copy numbers of all three plasmids are similar and about half that of the vector pC194. The efficiency of plating (EOP) of SPO2 *susL244* is 1 on BD294(pJB74) and BD294(pJB96) compared with 0.1 for BD294(pJB91). After one cycle of growth of SPO2 *susL244* in the former strains, 10 to 20% of the progeny are wild-type recombinants, compared with 50 to 80% for BD294(pJB91). This probably reflects the fact that in strains carrying pJB74 or pJB96, SPO2 *susL244* is complemented by DNA polymerase already present in the bacteria, whereas for pJB91 phage DNA replication can only occur after a functional gene *L* has been reconstituted by recombination between phage and plasmid. However, since BD294(pJB96) extracts contain high levels of DNA polymerase activity, gene *L* must be contained within the four *Hin*dIII fragments common to all three plasmids.

Location of the *susL244* mutation. *Hin*dIII fragments A and B were subcloned in pC194. The EOP of SPO2 *susL244* is ca. 20 times above the background level in BD294 carrying the B fragment (data not shown). Marker rescue of gene *L* from *Eco*RI fragments is known to be severely reduced compared with marker rescue of other genes (9). Part of gene *L* is thus located in *Hin*dIII fragment B.

Subcloning of SPO2 gene *L* in pUB110. The purpose of this

A

```
HindIII E  10         20         30         40         50         60         70         80
AAGCTTTTTAGATTAAAAGAACTATGAGGGGACGAGAGTTTCCCCTTTTAATAGATTGGGAGTGAAAACATGTTTAA

       90        100        110        120        130        140        150        160
TCCATCAGACTGGTATATCACACCTCAAGAATACGAACGGGCAGCTAAAAACGGTATTGACAAAATACGCCTTGAAAGAA

      170        180        190        200        210        220        230        240
GAATACGTTTCTCAGGATGGGATAAAGAAAGGGCCATAACAACACCAGTAAAAAAACGCAAGCCAAAGTCAAAATGGGGC

      250        260 EcoRI   270        280        290        300        310        320
TGATGTTGCAGTATCAAACGGAATTCACCGAAACTTATTTTTTGCCCGTGTGAATAGGCTGGGCTGGGATGAAGAAAAGG

      330        340        350        360        370        380        390        400
CGGCGACAACCCCGGCATCTGACGTGAGGGAGAACATGAAAAAAGTACAAAAGCATTCGTCTCGTAATAAACACTATCGT
                      rbs E      Start E
      410        420        430        440        450        460        470        480
CATTCCCCTGAGTTGATAGCTCTTGCTGAGTCGAACGGCATTAAATACCTGACCTATATTAACCGGGTCAATAACTGCGG

      490        PstI       510        520        530        540        HindIII B  560
AATGGACCCTTACGAAGCTGCAGTTACACCCCTTAGATCACGTCAAGAGGTTGCCCAACTCGGCCGTAAAGCTTTCCAAG
                                                          Stop E
      570        580        590        600        610    ...620        630        640
AAAAATACGGGGATGTAAATGGGCTATTTTTTCAAAAAAGGCGGGCTTCTAATTGAAAACTTTATCAATTGATATAGAAA
                          rbs Pol                Start Pol
      650        660        670        680        690        700        710        720
CATTCTCAAGCGTTGATCTTTTAAAAGCAGGCGTTTACGCGTACACAGAAGCCCCTGACTTTGAAATTTTGTTGTTTGCT

      730        740        750        760        770        780        790        800
TATGCTTTCGATGATGATCCTGTAAAAATAATTGATCTGGCTCAGGGGGATACGCTACCGCATGAGGTGCTTGTGGCGCT

      810        820        830        840        850        860        870        880
GACAAGTTCGAAGGTAATCAAGACGGCATATAACGCTAACTTTGAAAGAACGTGTATTGCAAAGCATTTTAATTTAATGC

      890        900   PvuII    920        930        940        950        960
TGCTTCCGGCGCAGTGGAGGTGTACAGCTGTCCACGCTACTACTTTAGGGCTTCCTGGGAATCTTGACGGAGTGGCTAAA

      970        980        990       1000       1010       1020       1030       1040
GCGTTAAAGCTATCAGCGCAGAAGGATAAAGCAGGTAAAGCCCTTATTCGTTACTTTTCAGTACCGTGTAAACCAACCAA

     1050       1060       1070       1080       1090       1100       1110       1120
GGCGAATGGGCAAAGGGTTAGGAATCTACCCGAGCACGACCCGGAGAAATGGGAGAAATTCAAAGTCTATTGTATCCAGG

     1130       1140       1150       1160       1170       1180       1190       1200
ATGTTGAAGTAGAACGTGCAATCAAAAACCGCATTTCCAAGTTTGAACCTTTGGAGTCTGAACATAAATTATGGGCGTTA

     1210       1220       1230       1240       1250       1260       1270       1280
GATCAGGAAATAAATGATAGGGGTGTTCGGATAGATGTGGACCTAGTCAAGCATGCAATAGCCTGTGATGAACAATATCA

     1290       1300   HincII   1320       1330       1340       1350       1360
GGCCGGCTTAATAGCAGAGGCTAAGAAGTTGACCGGTCTGCCCAATCCCAACAGTACGGCGCAATTAAAAAAATGGCTTG

     1370       1380       1390       1400       1410       1420       1430       1440
AAGAGAAGGGGCTTACTATTTCCAGTCTAGCTAAAGACAAAATAGAGGAACTTATTGAAAACACAAACGATGAAACTGTT

     1450       1460       1470       1480       1490       1500       1510       1520
CACAGGGTTTTGAGATTAAGGCAAGAAATGGCGAAAACCTCAGTTAAAAAGTATCTGGCAATGGAGAAAGCTCTTTGCCC
```

FIG. 5. Nucleotide sequence of gene *L* in pJB74. The nucleotide sequence of the anti-sense strand of pJB74 from the start of *Hin*dIII fragment E over the terminal *Sau*3A of pTB9 is shown. The positions of some pertinent restriction sites are indicated; the capital letter after *Hin*dIII denotes the start of the various SPO2 *Hin*dIII fragments (see Fig. 2). The inverted repeat in *Hin*dIII-E is denoted by arrows. rbs E and rbs Pol indicate the proposed ribosome-binding sequences for the coding sequences for the proposed E peptide and the polymerase, respectively. Start E and Start Pol are the respective translational initiation triplets. Triplets marked with dots are the translational stop codons for the two open reading frames. One possible −35, −10 structure is indicated in *Hin*dIII fragment E by dashed lines.

TABLE 1. Some properties of strain BD294 carrying pJB74, pJB91, or pJB96

Strain and plasmid	Relative DNA polymerase activity	Copy number[a]	EOP[b] of SPO2 *susL244*
BD294	1		<0.001
BD294(pC194)	1	28, 39	<0.001
BD294(pJB74)	2–3	10	1
BD294(pJB91)	1	10, 21	0.1
BD294(pJB96)	10–30	10, 11	1

[a] Plasmid copy numbers were determined by cesium chloride-ethidium bromide centrifugation of [^{3}H]thymidine-labeled bacteria as described by Tanaka et al. (36).

[b] The EOP measured on different strains was related to the EOP measured on SR135, which carries the *sup-3* suppressor gene.

experiment was to determine the maximum size of gene *L*. pJB96 DNA was partially digested with *Sau*3A and ligated to *Bam*HI-cleaved pUB110. The mixture was used to transform strain 3G18 selecting for kanamycin resistance. Among 1,000 transformants tested, 11 were found which plated SPO2 *susL244* with an EOP of 1. Plasmids from these transformants were analyzed by agarose gel electrophoresis. The smallest plasmid observed, called pTB9, was transformed into BD294. Extracts of BD294(pTB9) have 10- to 20-fold increased levels of DNA polymerase activity. pTB9 contains SPO2 *Hin*dIII fragments B, F, and G and the proximal 429 bp of the 1,700-bp A fragment (Fig. 4). Adjacent to fragment B is a 481-bp pC194 fragment. We examined five other plasmids obtained by cloning partially digested pJB96 DNA into pUB110 and which express gene *L*. In addition to SPO2 DNA, they all contain the 481-bp pC194 fragment. pTB9 contains about 2,197 bp of SPO2 DNA, which represents coding capacity for a protein with a molecular size of 80 to 90 kd.

Nucleotide sequence of gene *L*. The above data suggest a simple model for expression of gene *L* in the three plasmids (see below). From this model we decided to sequence from pJB74 *Hin*dIII fragments E, B, G, and F, and the proximal 429 bp of fragment A. The strategy for sequencing is outlined above and is shown in Fig. 3. The sequence obtained is shown in Fig. 5. We identified two open reading frames. The first one starts at 356 bp in fragment E and proceeds 63 bp into fragment B. It has coding capacity for a hypothetical protein of ca. 10 kd, which we will call the E peptide. Six base-pairs upstream from the initiation codon ATG, there is a potential ribosome-binding site. It has a calculated free energy of interaction (ΔG) with the 3′ end of *B. subtilis* 16S RNA of −16 kcal (−67 kJ)/mol. This value is similar to those reported for a number of other gram-positive ribosome-binding sites (23). The second open reading frame

B

```
      1530       1540       1550       1560 BglI       1580       1590       1600
TGATAATCGTGTTCGCGGCTTACTTCAATTTTACGGTGCCAGCCGTACTGGCAGATGGGCCGGACGTTTGGTTCAAGTTC

      1610       1620       1630       1640       1650       1660       1670       1680
AGAATTTGCCACAAAACAAGATAGAGGACCTGGACACAGCCAGGAACTTATTGAAGGGCGGACACTATGAAGCGATCGAA

      1690       1700       1710       1720       1730       1740       1750       1760
TTGCTTTATGGCCAAGTTCCTTTTGTACTGTCTCAGCTTGTTAGAACAGCTTTTATCCCGTCAGAAGGTAATGAGTTTTA

      1770 BglI  1790       1800       1810       1820       1830       1840
TGTCTCTGACTTCTCGGCCATCGAGGCCCGTGTAATTGCTTGGCTTGCAGGAGAAGAATGGCGTTTGGAAGTGTTCAACA

      1850       1860       1870       1880       1890       1900       1910       1920
CTCACGGGAAGATTTACGAGGCTTCGGCAGCGCAAATGTTCAAGGTTCCGGTTGAGTCAATCACTAAAGGCAGCCCATTA

      1930       1940       1950       1960       1970       1980       1990       2000
AGACAAAAAGGTAAAGTTGCCGAGTTAGCTTTGGGCTACCAGGGCGGAAAAGGTGCGCTTATTCAGATGGGTGCCCTAAA

      2010       2020       2030 HindIII G 2050       2060       2070       2080
CATGGGCTTAGCAGAGGGTGAACTTCCTGAGCTTGTAAAAGCTTGGCGGACAGCAAACAAAAAGATAGTTAAGTTTTGGT

      2090       2100       2110       2120 2130HindIII F      2150       2160
ATGACGTAGAAGCCGCAGCAATTAAAGCAGTTAAAGAGAGAAAGCCAGTTAAGCTTCAGCACGGCCTAACCTTCTTATAC

      2170       2180       2190       2200       2210       2220       2230       2240
GAATCTGGCATTTTGTTTGTACAGCTACCGTCAGGCAGGCGCCTGGCATATGCGAAGCCAAAACTTGAATTGGATGAACG

      2250       2260       2270       2280       2290       2300       2310 HindIII A
CTTTGGAAAAGAAGCACTTACTTATGAGGGGAAACTTGAGTCTGGTAAGTGGGGGCGGTTGAACACATACGGCGGCAAGC

      2330       2340       2350       2360       2370       2380       2390       2400
TTGTTGAGAATATCGTTCAGGCAACTGCACGGGATTGTCTGGCTATTACACTTATGCGTTTAGACAATGCTGGCTATAAA

      2410       2420       2430       2440       2450       2460       2470       2480
ACTGTCATGCACGTTCATGATGAAGCTGTCCTGGATGTTCCCCGAGGGAAAAATGAATTGGATAAAGTAGAGGCCATTAT

      2490       2500       2510       2520       2530       2540       2550 Stop Pol
GGGAGAGCCTATTTCTTGGGCTAAGGGTCTACCCTTGACCGCGGACGGCTTTGTTACAGATTACTATAAAAAAGATTAAT

..... 2570       2580       2590       2600       2610       2620       2630       2640
AGTAGGGGCTGAGATTATGAGTTTTATAAGAAGTCTAATAGCCGTGTTCGGGCTTTTATTCTCAGTGAGGAAAAAGGAAG

      2650       2660       2670       2780       2690       2700       2710       2720
TCGAATATTACGAATGGATAGAAAGGGATGGTAAGTGATGGCTATTATTAAAAATTATTCACTTGGGGATTTTGGCTGCA

      2730       2740 Sau3A 2760       2770       2780       2790       2800
TTGGGACCGAAACCGAGTTTATTTTGATCATCAAAAATAAAATTTATGGGCCTTGGGGTAATGAGCGCTTTGAGTCCGTT

      2810       2820       2830       2840       2850       2860       2870       2880
TTAAATACTTTAGAAGGCGTACTTGAGAAGCCTCGAAAATCAAAAAGATCGGGGGCCCAAATGTTTATGAGTGATGATAA

      2890       2900       2910       2920       2930       2940       2950       2960
GGAACTGGCTCGATTAAAAAGAAAGGTCGATAGACTCATTGAAAAATGCGATAGAAAAGGGCATGAGTTTAATGATTTTG

      2970       2980       2990       3000       3010       3020       3030       3040
AAATCGGGACCATTAGGAGAATAGGGCACGCCAAAAGCATGGATGATCTCAACTATTTGGTAGGAATTTCAATACAAATG
```

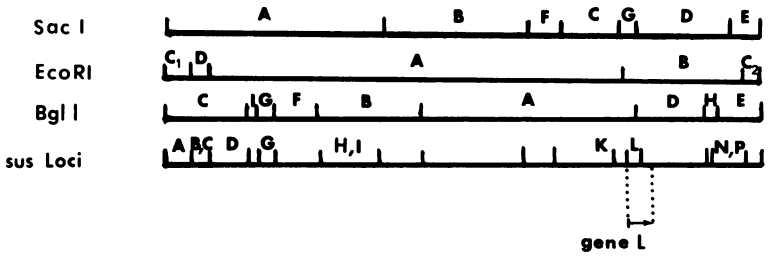

FIG. 6. Position of gene *L* on the SPO2 chromosome. The map and the restriction sites are redrawn from those Graham et al. (9). The *Bgl*I site described by them in the *L* gene is two sites separated by 215 bp. The arrow indicates the direction of transcription of gene *L*.

starts with TTG, which is a common initiation triplet in gram-positive bacteria (18). This TTG overlaps with the terminal codon and the termination triplet for the E peptide. Ten base pairs upstream from the TTG is a potential ribosome-binding site ($\Delta G = -15$ kcal [-63 kJ]). This second open reading frame has coding capacity for a 72-kd protein (72486D). It extends through *Hin*dIII fragment B, across the fragments G and F, and 238 bp into fragment A. It terminates with three consecutive termination triplets: TAA TAG TAG. The *Sau*3A site used in the construction of pTB9 (Fig. 4) is located 181 bp downstream from the translational stop. Transcription of the E peptide open reading frame should start in *Hin*dIII fragment E as possibly also transcription of the 72-kd open reading frame. Between positions 27 and 48 in fragment E there is a guanine-cytosine rich inverted repeat followed by five consecutive pyrimidines. This structure is typical of some transcriptional stop signals (25, 27, 39). Several potential -35 and -10 promoter sequences can be found between the inverted repeat and the proposed ribosome-binding site for the E peptide transcript, e.g., a canonical -35 sequence (TTGACA) is found at positions 137 through 142, followed 16 bp downstream by a possible -10 sequence (AAGAAT) (27, 39). However, studies of transcripts from the cloned SPO2 DNA are required to identify promoter sequences with any certainty (7, 23, 34).

DISCUSSION

B. subtilis BD294 carrying either pJB96 or its derivative pTB9, from which ca. 1,270 bp of *Hin*dIII fragment A have been deleted, are high SPO2 DNA polymerase producers. From this fact, and the fact that a suppressible DNA polymerase mutation is located in *Hin*dIII fragment B, we conclude that the structural gene for SPO2 DNA polymerase is contained with *Hin*dIII fragments B, G, and F and the proximal 429 bp of fragment A (A'; Fig. 4). Starting with a TTG in *Hin*dIII fragment B there is an open reading frame proceeding through fragments G and F and ending 238 bp into fragment A' with three consecutive translational termination triplets. It has the coding capacity for a 72-kd protein. All other open reading frames, reading in the same direction, have coding capacities for proteins of less than 10 kd. Also, none of those open reading frames is preceded by a reasonable ribosome-binding site. There is no open reading frame in the reversed order, starting in fragment A', which has coding capacity for a polypeptide larger than 18 kd. All three plasmids, pJB74, pJB91, and pJB96, carry a complete SPO2 DNA polymerase gene, but the gene is not expressed in pJB91. The fact that all plasmids contain *Hin*dIII fragment A and that ca. 1,270 bp of fragment A, distal to *Hin*dIII fragment B, can be deleted in pJB96 without impairing expression of the gene excludes the possibility that the polymerase gene is read from a promoter located in fragment A. We conclude that the open reading frame coding for a 72-kd protein is the structural gene for SPO2 DNA polymerase.

The present data suggest a simple model for transcription of the SPO2 DNA polymerase gene, which accounts for the properties of the three plasmids studied here. Since pJB91 makes no polymerase although it contains a complete structural gene, transcription of the polymerase gene must be initated from a promoter located outside *Hin*dIII fragment B, i.e., more than ca. 60 bp upstream from the initiation triplet TTG. In pJB74 the gene is transcribed from its natural promoter. We suggest that this promoter is located in *Hin*dIII fragment E. The presence of a possible transcriptional stop signal at positions 27 through 48 in fragment E (see Fig. 5) lends some speculative support to the idea that an SPO2 transcript is initiated in fragment E.

In pJB96 the polymerase gene is removed from its natural promoter and instead the gene is transcribed from a promoter in the 482-bp *Hin*dIII-*Sau*3A fragment from the vector pC194. Within this fragment there is an open reading frame coding for a hypothetical D protein (13); in pJB96 this is fused in phase with the terminal part of the proposed E peptide gene.

In pJB91, the polymerase gene is not transcribed. The amount of SPO2 DNA polymerase activity found in BD294(pJB96) is 5 to 10 times greater than that found in BD294(pJB74); this might simply reflect differences in promoter strengths. However, the structure of pJB74 suggests the possibility of a more complex control of DNA polymerase in SPO2. Starting at 336 bp in *Hin*dIII fragment E, there is an open reading frame, preceded by a typical gram-positive ribosome-binding site, coding for a proposed E peptide of ca. 10 kd. The reading frame for the E peptide overlaps with the proposed ribosome-binding site for the polymerase gene and with part of the initiation triplet TTG. If the E peptide and the polymerase genes are transcribed from different promoters, with the polymerase promoter located in the E peptide open reading frame, the amount of polymerase gene transcript could be modulated by E transcription. On the other hand, if both the E peptide and polymerase are translated from a polycistronic mRNA, the amount of polymerase produced could be controlled also at the translational level. We are presently trying to examine these possibilities.

Finally, Young and coworkers have established a correlated genetic and restriction map of phage SPO2 (9). By combining their data with the results presented in this paper, the location, size, and direction of transcription of the phage polymerase gene can be determined on the SPO2 genome (Fig. 6).

ACKNOWLEDGMENTS

We thank Sven-Åke Franzén and Tove Harde for expert technical assistance.

This work was supported by grants from the Swedish Medical Research Council, KabiGen AB, Karolinska Institutets forskningsfonder, and the fund of Emil and Vera Cornell.

LITERATURE CITED

1. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–746.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Boice, L. B., F. Eiserling, and W. R. Romig.** 1969. Structure of *Bacillus subtilis* phage SPO2 and its DNA: similarity of *Bacillus subtilis* phages SPO2, ϕ105, and SPP1. Biochem. Biophys. Res. Commun. **34**:398–403.
4. **Brown, N. C.** 1970. 6-(*p*-hydroxyphenylazo)-uracil: a selective inhibitor of host DNA replication in phage-infected *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. **67**:1454–1461.
5. **Canosi, U., G. Morelli, and T. A. Trautner.** 1978. The relationship between molecular structure and transformation efficiency of some *S. aureus* plasmids isolated from *B. subtilis*. Mol. Gen. Genet. **166**:259–267.
6. **Chang, S., and S. N. Cohen.** 1979. High frequency transformation of *B. subtilis* protoplasts by plasmid DNA. Mol. Gen. Genet. **168**:111–115.
7. **Davison, B. L., T. Leighton, and J. C. Rabinowitz.** 1979. Purification of *Bacillus subtilis* RNA polymerase with Heparin-Agarose. *In vitro* transcription of ϕ29 DNA. J. Biol. Chem. **254**:9220–9226.
8. **Ehrlich, S. D.** 1977. Replication and expression of plasmids from *Staphylococcus aureus* in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. **74**:1680–1682.
9. **Graham, S., S. Sutton, Y. Yoneda, and F. E. Young.** 1982. Correlation of the genetic map and the endonuclease site map of *Bacillus subtilis* bacteriophage SPO2. J. Virol. **42**:131–134.
10. **Gryczan, T. J., and D. Dubnau.** 1978. Construction and properties of chimeric plasmids in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U.S.A. **75**:1428–1432.
11. **Gryczan, T. J., A. G. Shivakamur, and D. Dubnau.** 1980. Characterization of chimeric plasmid cloning vehicles in *Bacillus subtilis*. J. Bacteriol. **141**:246–253.
12. **Hemphill, H. E., and H. R. Whiteley.** 1975. Bacteriophages of *Bacillus subtilis*. Bacteriol. Rev. **39**:257–315.
13. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. **150**:815–825.
14. **Iordanescu, S.** 1975. Recombinant plasmid obtained from two different compatible staphylococcal plasmids. J. Bacteriol. **124**:597–601.
15. **Jalanko, A., I. Palva, and H. Söderlund.** 1981. Restriction maps of plasmids pUB110 and pBD9. Gene **14**:325–328.
16. **Löfdahl, S., J. E. Sjöström, and L. Philipson.** 1978. Characterization of small plasmids from *Staphylococcus aureus*. Gene **3**:149–154.
17. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **McLaughlin, J. R., C. L. Murray, and J. C. Rabinowitz.** 1981. Unique features in the ribosome binding sequence of the gram-positive *Staphylococcus aureus* β-lactamase gene. J. Biol. Chem. **256**:11283–11291.
19. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
20. **Messing, J., B. Gronenborn, B. Müller-Hill, and P. H. Hofschneider.** 1977. Filamentous coliphage M13 as a cloning vehicle. Insertion of a *Hin*dII fragment of the *lac* regulatory region in M13 replicative form *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **74**:3642–3646.
21. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either strand of double-digest restriction fragments. Gene **19**:269–276.
22. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Moran, C. P., Jr., N. Lang, S. F. J. Le Grice, M. Stephens, A. L. Sonenschein, 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.
24. **Murray, C. L., and J. C. Rabinowitz.** 1982. Nucleotide sequences of transcription and translation initiation regions in *Bacillus subtilis* phage ϕ29 early genes. J. Biol. Chem. **257**:1053–1062.
25. **Platt, T.** 1981. Termination of transcription and its regulation in the tryptophan operon of *E. coli*. Cell **24**:10–23.
26. **Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey.** 1982. Non-random DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. U.S.A. **79**:4298–4302.
27. **Rosenberg, M., and D. Court.** 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319–353.
28. **Rutberg, L., and R. W. Armentrout.** 1972. Deoxyribonucleic acid polymerase activity in a deoxyribonucleic acid polymerase I-deficient mutant of *Bacillus subtilis* infected with temperate bacteriophage SPO2. J. Virol. **10**:658–660.
29. **Rutberg, L., R. W. Armentrout, and J. Jonasson.** 1972. Unrelatedness of temperate *Bacillus subtilis* bacteriophages SPO2 and ϕ105. J. Virol. **9**:732–737.
30. **Rutberg, L., B. Rådén, and J.-I. Flock.** 1981. Cloning and expression of bacteriophage SPO2 DNA polymerase gene L in *Bacillus subtilis*, using the *Staphylococcus aureus* plasmid pC194. J. Virol. **39**:407–412.
31. **Rådén, B.** 1983. Expression of temperate *Bacillus subtilis* phage SPO2 DNA polymerase gene in *Escherichia coli*. FEMS Microbiol. Lett. **17**:69–71.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
33. **Scher, B. M., D. H. Dean, and A. J. Garro.** 1977. Fragmentation of *Bacillus subtilis* bacteriophage ϕ105 DNA by restriction endonuclease *Eco*RI: evidence for complementary single-stranded DNA in the cohesive ends of the molecule. J. Virol. **23**:377–383.
34. **Shorenstein, R. G., and R. Losick.** 1973. Comparative size and properties of the sigma subunits of ribonucleic acid polymerase from *Bacillus subtilis* and *Escherichia coli*. J. Biol. Chem. **248**:6170–6173.
35. **Sutcliffe, J. G.** 1978. pBR322 restriction map derived from the DNA sequence: accurate DNA size markers up to 4361 nucleotide pairs long. Nucleic Acids Res. **5**:2721–2728.
36. **Tanaka, T., M. Kuroda, and K. Sakaguchi.** 1977. Isolation and characterization of four plasmids from *Bacillus subtilis*. J. Bacteriol. **129**:1487–1494.
37. **Tinoco, I., Jr., 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.
38. **Yasunaka, K., H. Tsukamoto, S. Okubo, and T. Horiuchi.** 1970. Isolation and properties of supressor-sensitive mutants of *Bacillus subtilis* bacteriophage SPO2. J. Virol. **5**:819–821.
39. **Yoshikawa, H., and J. Ito.** 1982. Nucleotide sequence of the major early region of bacteriophage ϕ29. Gene **17**:323–335.