Analysis of Bacteriophage SP82 Major "Early" In Vitro Transcripts

ANTONITO T. PANGANIBAN AND H. R. WHITELEY*

Department of Microbiology and Immunology, University of Washington, Seattle, Washington 98195

A restriction map was constructed for the 17.5-kilobase *SalI* C fragment of SP82 DNA. Unfractionated SP82 DNA, the *SalI* C fragment, and restriction fragments derived from *SalI*-C were used as templates for in vitro synthesis by the *Bacillus subtilis* RNA polymerase, and the resulting transcripts were analyzed by gel electrophoresis. Comparison of the RNA species obtained from *SalI*-C with those produced from unfractionated DNA indicated that most of the RNAs and all of the major transcripts originate from the *SalI* C fragment; this fragment contains one copy of the terminally redundant portion of the genome. Seven major transcripts, a bidirectional terminator site, and 5 of the 13 minor transcripts were located within the 13-kilobase redundant region. The binding of polymerase to fragments of DNA produced by digestion of *SalI*-C with *HpaII* and *HhaI* was used to identify promoter-bearing regions within *SalI*-C.

Infection of *Bacillus subtilis* by either SP82 or SP01 results in the ordered appearance of phage transcripts and proteins (4, 5, 8, 14, 15, 19). This ordered expression of the phage genome is controlled largely by modification of the subunit structure of the host RNA polymerase after phage infection (15, 24). Whereas the temporal changes in the classes of RNA produced during infection have been well characterized (4, 5, 11, 19), relatively little is known about the order of "early," "middle," and "late" transcriptional units on the SP82 genome. An earlier study (12) showed that the ends of the SP82 genome are transcribed immediately after infection and are also transcribed in vitro by the unmodified *B. subtilis* RNA polymerase. Middle and late transcriptions were found to occur at many regions throughout the genome, and transcription of these regions in vitro was catalyzed by modified forms of RNA polymerase.

This paper reports on the number and sizes of the primary transcripts synthesized in vitro by the *B. subtilis* RNA polymerase from unfractionated SP82 DNA and from *SalI*-C, a 17.5-kilobase (kb) DNA fragment derived from one end of the genome. The locations of seven major RNA species on a detailed map of the *SalI*-C region were determined by transcribing mixtures of DNA fragments and individual fragments obtained by digesting *SalI*-C with *HaeII* and *HhaI*. In addition, binding studies were performed to identify the *Hpa*-*Hha* fragments possessing promoters recognized by the *B. subtilis* RNA polymerase.

MATERIALS AND METHODS

Restriction endonuclease digestion of DNA.

Endonucleases *HpaII*, *HhaI*, *HaeII*, *HaeIII*, *KpnI*, *XbaI*, *EcoRI*, *BglI*, and *BglII* were purchased from New England Biolabs, Beverly, Mass. *SalI* was purified by the method of Greene et al. (7) from a strain of *Streptomyces albus* G obtained from R. J. Roberts. One unit of enzyme activity was defined as that amount of protein with digests 1 μ g of phage lambda DNA in 1 h at 37°C. Digestions with *HpaII*, *HhaI*, *HaeII*, *BglII*, and *HaeII* were performed with 1 U/ μ g of SP82 DNA. A total of 5 U/ μ g of DNA were needed for digestion by *KpnI* and *XbaI*, and 15 U/ μ g of DNA were required for *EcoRI* digestion. Reaction mixtures containing all enzymes except *SalI* were incubated for 4 h at 37°C; *SalI* digestions were performed by incubating 0.5 U of enzyme per μ g of DNA for 1.5 h at 37°C. When two endonucleases were used, the enzymes were added simultaneously if their salt requirements were the same. If the salt optima were different, the first digestion was done with the enzyme requiring the lower salt concentration; the concentration of salt was then increased, and the second enzyme was added. After digestion, the fragments were extracted twice with phenol-chloroform (1:1) which had been pre-equilibrated with 10 mM Tris-1 mM EDTA, pH 7.5 (TE 7.5 buffer); fragments were desalted by passage through Bio-Gel P20, with TE 7.5 buffer being used for elution, and stored at 0°C.

In vitro RNA synthesis and marker RNAs. Reaction mixtures for the synthesis of RNA contained 40 mM Tris (pH 7.9), 20 mM MgCl₂, 0.14 M NaCl, 200 μ M GTP, 200 μ M CTP, 200 μ M UTP, 100 μ M ATP, 2 μ g of DNA, and 2.5 μ g of RNA polymerase in a volume of 50 μ l. All components except the nucleotides were combined and incubated at 37°C for 5 min. The four nucleotides and 4 μ Ci of [α -³²P]ATP (10 Ci/mmol; Amersham Corp., Arlington Heights, Ill.) were added; after 15 min of incubation at 37°C, the reaction mixtures were supplemented with a second addition of unlabeled ATP, GTP, CTP, and UTP to give a final concentration of 50 mM for each nucleotide and incubated for 5 min. Synthesis was stopped by adding

an equal volume of a solution containing 1% sodium dodecyl sulfate, 10 mM Tris (pH 7.4), 10 mM EDTA, 40% glycerol, and 0.1% bromophenol blue, and the preparation was heated in a boiling water bath for 1.5 min and stored at -20°C ; there was no detectable degradation of RNA after 1 week of storage.

The sizes of the transcripts were determined for a curve obtained by plotting the electrophoretic mobilities of known markers: ^{32}P -labeled T7 RNA extracted from *Escherichia coli* infected with wild-type T7 and with the ΔH1 mutant (22) and ^{32}P -labeled rRNA from *E. coli*. These preparations were contributed by Susan Strome. The accuracy of these determinations was ca. ± 50 bases.

Gel electrophoresis. RNA samples were electrophoresed through gels containing 3% polyacrylamide–0.13% *N,N'*-methylenebisacrylamide and 0.2% sodium dodecyl sulfate with 40 mM Tris–40 mM acetate–2 mM EDTA (pH 7.6) as the buffer (6). The gels were poured, allowed to polymerize overnight at room temperature, and pre-electrophoresed at 100 V for 1 h before samples were loaded. Electrophoresis of samples was at 70 V for 0.5 h and at 115 V for an additional 4 h or until the dye front just migrated off the gel. Gels were removed from the plates, soaked for 15 min in 10% glycerol, and dried on Whatman 3MM filter paper. RNA bands were detected by autoradiography of the gels for 2 days with Kodak X-Omat R film. DNA fragments used as templates for RNA synthesis were isolated by electrophoresis through 4% polyacrylamide–0.07% *N,N'*-methylenebisacrylamide slab gels with the Tris buffer described previously (13). Gels were stained with 0.005% ethidium bromide in electrophoresis buffer and visualized with short-wavelength UV light.

Isolation of individual restriction fragments. To isolate the *SalI* C fragment, samples containing 100 μg of SP82 DNA digested with *SalI* were layered on 12-ml linear gradients of 5 to 20% sucrose containing 50 mM Tris (pH 7.9), 2 mM EDTA, and 1 M NaCl and centrifuged at $200,000 \times g$ for 3.5 h with an SB 283 rotor of an International B60 ultracentrifuge. Fractions were collected by puncturing the bottom of the tube and analyzed by electrophoresis through agarose gels (13). Fractions containing the *SalI* C fragment were pooled and dialyzed against 4 liters of TE 7.5 buffer for 18 h. The DNA was concentrated by evaporation of liquid from the bags (13) or by extraction with 2-butanol (21), dialyzed for 6 h against TE 7.5 buffer, and stored at 0°C . To isolate smaller fragments, segments of acrylamide gels containing desired fragments were cut out, placed in a dialysis bag with a small amount of buffer, eluted electrophoretically, and extracted with phenol-chloroform as described above. NaCl was added to a final concentration of 0.5 M, and DNA was precipitated at -60°C with 3 volumes of ethanol and collected by centrifugation with a Beckman model B microfuge; the pellet was rinsed with 1 ml of 95% ethanol and dried, and the DNA was dissolved in TE 7.5 buffer.

Miscellaneous techniques. Methods used for the growth of *B. subtilis* (11), infection with SP82 (11), the preparation of CsCl-purified phage, and the extraction of phage DNA (13) have been described previously. Promoter-bearing DNA fragments were identified by the polymerase-binding assay of Jones et al.

(9) as modified by Spiegelman and Whiteley (20). Procedures have been described previously for the isolation of *B. subtilis* RNA polymerase (18).

RESULTS

Physical map of the *SalI* C fragment and the SP82 terminal redundancy. RNA synthesized either in vivo immediately after SP82 infection or in vitro by the *B. subtilis* RNA polymerase hybridizes to restriction fragments located near or at the ends of the phage genome (12). A physical map of the terminal region was constructed to examine transcription by the unmodified polymerase in greater detail. The restriction endonuclease *SalI* cleaves SP82 DNA three times, resulting in four fragments (13). The 17.5-kb *SalI* C fragment is derived from one end of the phage DNA and comprises about 12.5% of the 140-kb genome. Nine endonucleases were found to cleave this fragment, generating from 2 to 13 fragments. Conventional methods were used to determine the cleavage sites for these enzymes within the *SalI* C fragment (Fig. 1).

Restriction enzyme analysis of the other end of the genome revealed that SP82 is terminally redundant (data not shown). The SP82 redundancy, like that of the related phage SP01 (16), is nonpermuted and constitutes a direct, rather than an inverted, repeat. Summation of the 9 *HpaII*, 13 *HhaI*, and 21 *HpaII-HhaI* fragments that make up the SP82 redundant region indicates lengths of 13.5, 12.7, and 13.4 kb, respectively; the length of the SP01 terminal redundancy was determined to be 12.4 kb, based on the sizes of 4 *EcoRI** fragments (16). Since the *SalI* C fragment is derived from one end of the SP82 genome, it contains one copy of the redundant region plus an adjacent 4.3 kb of DNA (Fig. 1).

Conditions of transcription. RNAs were synthesized in vitro under conditions which were shown earlier to provide good specificity of transcription; i.e., RNAs were synthesized in the presence of 0.14 M NaCl with RNA polymerase preparations containing a high content of the sigma subunit and the 21,000-molecular-weight delta peptide. Under these conditions, RNAs produced in vitro contained less than 2% anti-mRNA. The effects of various conditions of in vitro synthesis on the fidelity of transcription have been assessed previously by analyzing RNAs for: strand selection (10), anti-mRNA content (18), translatability in an *E. coli* cell-free system (8, 18), hybridization to unfractionated DNA (19) and to isolated strands of DNA (24), and the ability to form binary complexes with promoter-bearing restriction fragments (20; E. C. Achberger and H. R. Whiteley, *J. Biol. Chem.*, in press). The effects of the delta peptide

FIG. 1. Physical map of the *SalI* C fragment of SP82 DNA. Length (kb) and the extent of the terminal redundancy are shown at the bottom.

on the formation of binary complexes will be reported elsewhere (E. C. Achberger and H. R. Whiteley, manuscript in preparation). The conditions of in vitro synthesis selected for these studies inhibited the formation of nonspecific polymerase-DNA complexes; they also decreased complex formation and transcription from weak promoters (Achberger and Whiteley, in press). In vitro translation of RNAs synthesized under these conditions yielded many peptides having the same electrophoretic mobilities as peptides synthesized in vivo early in infection (8).

Location of early transcripts. Comparison of the RNAs synthesized under the above conditions from the *SalI* C fragment with those produced from the entire phage genome revealed that both templates code for the same array of RNAs (Fig. 2, lanes 1 and 2). It should be noted that transcription of a small portion of the genome outside the *SalI*-C region can be detected under other less restrictive conditions of in vitro synthesis, i.e., in the presence of low concentrations of NaCl or $MgCl_2$ or at high polymerase/DNA ratios (Achberger and Whiteley, in press; Achberger, unpublished data); a low level of in vivo transcription of the same portion of the genome outside the *SalI* C region can also be detected in the early stages of infection (J. M. Lawrie, Ph.D. dissertation, University of Washington, Seattle, 1978). The existence of a weakly transcribed early region outside the terminal redundancy has also been reported pre-

viously for phage SP01 (16). However, as shown in Fig. 2, all of the major and minor transcripts produced from the entire genome under the conditions of synthesis used in these studies are also produced from *SalI*-C.

Based on densitometric tracings of the bands shown in Fig. 2, the transcripts produced from *SalI*-C can be roughly divided into two groups: the major RNAs, which account for 85% of the synthesis, and the minor RNAs, which contain the remaining 15% of the radioactivity. There are 7 major transcripts synthesized by *B. subtilis* polymerase (A, D, E, F, I, N, and Q) and 13 minor transcripts (B, C, G, H, J, K, L, M, O, P, R, S, and T). To determine the location and direction of synthesis for the major RNA species, the *SalI* C fragment was digested by several restriction enzymes, and the resulting fragments were used to direct in vitro RNA synthesis (Fig. 2 and 3). By correlating the loss of major RNAs after digestion with endonucleases, it was possible to determine the initiation and termination sites for RNA synthesis. Table 1 summarizes the effect of nine endonucleases on the synthesis of each major RNA.

RNA A. The following data indicate that RNA A is produced from the left end of *SalI*-C: RNA A disappears after digestion with *HaeII* and *HhaI*, but it is synthesized intact after digestion with the other seven enzymes tested (Fig. 2 and 3). Since the length of RNA A (ca. 3,000 bases) is longer than the distance between any of the cleavage sites for these seven enzymes

FIG. 2. RNA species synthesized by the *B. subtilis* RNA polymerase. RNAs were synthesized from the following templates: (1) SP82 DNA; (2, 4, 6, 8, and 11) isolated *SalI*-C; (3) *SalI*-C digested with *HaeII*; (5) *SalI*-C digested with *HhaI*; (7) *SalI*-C digested with *HpaII*; (9) *SalI*-C digested with *HaeIII*; and (10) *SalI*-C digested with *BglII*. The letter designations of the major (bold print) and minor (light print) RNA species are shown in the margin, and the base lengths of truncated derivatives of major RNAs are indicated on the gels. The base lengths of the major and minor RNAs are given in Tables 1 and 2, and cleavage sites for the restriction endonucleases are shown in Fig. 1.

(maximum distance of 1,300 base pairs between the *XbaI* site and the nearest *KpnI* site [Fig. 1]), RNA A must be synthesized entirely to the left of the *XbaI* site.

The two enzymes affecting synthesis of RNA A, *HaeII* and *HhaI*, both cleave DNA at the same tetranucleotide sequence (5'-GCGC-3'), but cleavage by *HaeII* is limited to only those sites with an adjacent purine and pyrimidine at the 5' and 3' ends of the sequence (PuGCGCPy). Consequently, *HaeII* cleavage occurs only at a subset of the sites recognized by *HhaI*. Of the five *HhaI* sites to the left of the *XbaI* site, three are recognized by *HaeII* (Fig. 1). Initiation of RNA A and elongation toward sites recognized by both endonucleases would lead to a shortened derivative of RNA A of identical length. Alternatively, synthesis could occur through a proximal *HhaI* site and a more distal *HaeII*-*HhaI*

FIG. 3. RNA species synthesized by the *B. subtilis* RNA polymerase. RNAs were synthesized from the following templates: (1, 3, 5, and 7) isolated *SalI*-C; (2) *SalI*-C digested with *EcoRI*; (4) *SalI*-C digested with *XbaI*; (6) *SalI*-C digested with *BglII*; and (8) *SalI*-C digested with *KpnI*. The base lengths of truncated derivatives of major RNAs are indicated on the gels.

site, resulting in an *HhaI*-truncated derivative shorter than that produced by *HaeII* digestion. Examination of RNAs produced from the *HaeII*- and *HhaI*-digested *SalI* C fragment (Fig. 2, lanes 3 and 5) revealed that the shortened RNAs were of different lengths, indicating that synthesis of RNA A proceeds first through an *HhaI* site not recognized by *HaeII*. Furthermore, transcription of *HhaI* fragments i, h, and d (Fig. 1) resulted in no detectable RNA species (data not shown). Initiation of RNA A from near the end of the genome should result in an *HhaI*-truncated derivative of RNA A that is shorter than that produced by *HaeII* digestion. When the *HaeII* d fragment (Fig. 1) was used as a template for RNA synthesis, a 1,350-base RNA was observed (Fig. 4, lane 1). This indicates that RNA A starts at or near the genome end and is elongated

TABLE 1. Observed and predicted lengths of truncated RNAs transcribed from *SalI*-C digested with restriction enzymes^a

Major RNA	Base length	Observed/predicted length (bases) of truncated RNA when <i>SalI</i> -C was digested with:								
		<i>HaeII</i>	<i>HhaI</i>	<i>HpaII</i>	<i>HaeIII</i>	<i>BglII</i>	<i>XbaI</i>	<i>BglI</i>	<i>EcoRI</i>	<i>KpnI</i>
A	3,000	1,350/1,300	770/760	N	N	N	N	N	N	N
D	2,200	470/510	220/250	850/860	1,150 ^b	N	N	1,150	—/<200	1,350/1,360
E	1,770	N	—/<250	N	N	N	1,420 ^b	N	N	N
F	1,640	1,350/1,290	770/760	N	N	N	N	N	N	N
I	1,280	N	800/810	N	N	N	730 ^b	N	N	N
N	730	N	N	N	N	N	N	N	<200 ^b	—/<200
Q	430	N	N	N	N	N	N	N	N	—/<200

^a The *SalI* C fragment was digested with restriction enzymes, transcribed by *B. subtilis* RNA polymerase, and analyzed by electrophoresis as described in the text. The observed length is the size determined by gel analysis; the predicted length is based on the distances between restriction sites. N, No truncation of RNAs was found or predicted; —, RNA not detected.

^b Observed length used as a basis for map position and predicted lengths.

rightward (Fig. 5). When the *HhaI* g fragment (Fig. 1) from the genome end and the *HhaI*-digested *HaeII* d fragment were used as templates, a 770-base RNA was synthesized, confirming that RNA A initiates near the genome end (Fig. 4, lanes 2 and 3).

RNA D. The *HaeIII* site and one of the *BglI* sites are located at about the same place on the *SalI* C fragment (Fig. 1), and digestion of *SalI*-C with either endonuclease before RNA synthesis resulted in the loss of RNA D and the appearance of a new 1,150-base truncated derivative (Fig. 2, lane 9, and Fig. 3, lane 6). These results indicate that initiation of RNA D occurs 1,150 base pairs from the *HaeIII* and *BglI* sites. Digestion of *SalI*-C with endonuclease *HpaII* also resulted in the disappearance of RNA D and the synthesis of a new 850-base RNA (Fig. 2, lane 7), indicating that the RNA D initiation site is closer to the *HpaII* site than to the *HaeIII* or *BglI* site. Since the *HaeIII* and *BglI* sites are to the left of all *HpaII* sites in the *SalI* C fragment (Fig. 1), synthesis of RNA D must begin on the *HpaII* d fragment and proceed leftward through the *HpaII* and then the *HaeIII* and *BglI* sites. Leftward synthesis of RNA D should also result in transcription through *HaeII*, *HhaI*, *EcoRI*, and *KpnI* sites on *SalI*-C (Fig. 1). The predicted and observed base lengths of the truncated RNA D after digestion with these endonucleases were consistent with leftward synthesis of RNA D (Table 1; Fig. 2 and 3).

RNA E. Synthesis of RNA E was affected by prior digestion of *SalI*-C with *XbaI* and *HhaI*, but synthesis of the transcript was uninterrupted by digestion with all other endonucleases (Fig. 2 and 3). *XbaI* digestion resulted in the loss of two major RNA species (E and I) and the appearance of two new truncated RNAs (1,420 and 730 bases; Fig. 3, lane 4). Since RNA I (1,280

FIG. 4. RNAs synthesized from isolated *HhaI* and *HaeII* fragments. RNAs were synthesized from the following templates: (1) *HaeII*-d; (2) *HhaI*-g; (3) *HaeII*-d digested with *HhaI*; (4) *HhaI*-e; (5) *HhaI*-f; (6) *HhaI*-g; (7) *HhaI*-c; and (8) *HaeII*-b. Identification of intact major and minor RNAs and base lengths of the truncated derivatives are given in the margin.

bases) is shorter than the 1,420-base RNA generated by *XbaI* digestion, the 1,420-base RNA must be derived from RNA E, and the 730-base transcript would then be from RNA I. *HaeII* digestion of *SalI*-C did not affect the synthesis of RNA E (Fig. 2, lane 3), so initiation of this transcript could take place about 1,400 bases to either the right or the left of the *XbaI* site (Fig. 1). Leftward transcription beginning 1,420 bases to the right of the *XbaI* site should cross a *KpnI* site and possibly an *EcoRI* site (Fig. 1), but digestion with these endonucleases had no effect on the synthesis of RNA E (Fig. 3, lanes 2 and

FIG. 5. Location of the seven major RNAs produced by transcription of *SalI-C* by *B. subtilis* RNA polymerase.

8). Rightward synthesis beginning left of the *XbaI* site would place the initiation site for RNA E in the *HhaI* l fragment at the very left end of the *HaeII* b fragment (Fig. 1). As predicted, transcription of the *HaeII* b fragment resulted in the synthesis of RNA E (Fig. 4, lane 8). Digestion with *HhaI* would result in the loss of RNA E and the synthesis of a truncated derivative shorter than the *HhaI* l fragment (250 base pairs [Table 1]). The gel system used in the analysis of the RNAs cannot detect transcripts shorter than about 200 bases. It is likely, therefore, that initiation of RNA E begins near the left end of the *HaeII* b fragment and proceeds rightward through an *HhaI* site less than 200 base pairs "downstream" and then through the more distally located *XbaI* site (Fig. 5).

RNA F. Like RNA A, RNA F was truncated after digestion of *HhaI* and *HaeII*, but digestion by the other seven enzymes had no effect on the transcript (Fig. 2 and 3). For the same reasons as those discussed for the placement of RNA A, synthesis of RNA F must begin near the genome end and proceed rightward first through an *HhaI* site and then a site recognized by both *HhaI* and *HaeII* (Fig. 1). Transcription of the *HaeII* d and *HhaI* g fragments and the *HhaI*-digested *HaeII* d fragment (Fig. 4, lanes 1 through 3) indicates that transcription begins at only one site in this region. This result can be explained if RNA F and RNA A are initiated at

a single promoter (or closely adjacent promoters) near the genome end. Subsequent transcriptional read-through of the RNA F terminator 1,640 base pairs downstream would generate the longer A transcript, which terminates 3,000 bases from the end of the genome (Fig. 5).

RNA I. Digestion of *SalI-C* with *XbaI* resulted in the truncation of RNA I to a new 730-base RNA having the same mobility as RNA N (determined by densitometric scanning of lane 4 of Fig. 3), so initiation of RNA I must be 730 bases to either the right or the left of the *XbaI* site. Rightward synthesis of RNA I predicts that *HhaI* digestion of *SalI-C* should result in the loss of RNA I and the production of a new 810-base RNA (Table 1). Digestion with *HhaI* resulted in the truncation of RNA I as expected; as shown in lane 5 of Fig. 2, a derivative was produced that migrated closely with the 770-base derivative produced from RNA A and RNA F. Furthermore, transcription of the *HhaI* f fragment (Fig. 1) yielded a single 800-base RNA (Fig. 4, lane 5). RNA I was also synthesized intact from the isolated *HaeII* b fragment (Fig. 4, lane 8). If rightward elongation of RNA I takes place as suggested by the analysis, RNA I and RNA E must overlap: RNA I is initiated at a point on the DNA within the region serving as template for RNA E, and the two RNAs share a termination site in the *HhaI* e fragment (Fig. 5).

RNA N. Digestion of *SalI*-C with *KpnI* resulted in the loss of RNA N, and no shortened RNA could be detected (Fig. 3, lane 8); all other endonucleases had no effect on the synthesis of RNA N. In addition, transcription of the *HhaI* e and *HaeII* b DNA fragments resulted in the synthesis of two major RNAs, RNA N and RNA Q, and several minor transcripts discussed below (Fig. 4, lanes 4 and 8). Since initiation of RNA N began at a point very near one of the *KpnI* sites in the *HhaI* e fragment (Fig. 1), transcription of RNA N cannot be rightward; rightward transcription would cross *HhaI*, *HaeII*, *HaeIII*, *BglI*, and *EcoRI* sites. Leftward transcription would result in the complete synthesis of RNA N without crossing any of these sites (Table 1). Initiation of RNA N could begin near either of the *KpnI* sites in the *HhaI* e DNA fragment. However, if synthesis begins near the site located closest to the *HaeII* site (Fig. 1 and 5), transcription would then overlap with RNA D and termination of both transcripts would occur at the same point in the *HhaI* e fragment.

RNA Q. As described above, RNA N and RNA Q were synthesized from the *HhaI* e and *HaeII* b fragments. In addition, digestion with *KpnI* resulted in the loss of RNA Q, and no shortened derivative was detected (Fig. 3, lane 8). Rightward elongation of RNA Q would result in transcription through an *EcoRI* site, but this endonuclease had no effect on the synthesis of RNA Q (Fig. 3, lane 2). These results are consistent with initiation of the transcript to the right of the leftmost *KpnI* site (Fig. 1). Subsequent leftward elongation would end at the terminator shared by RNA D and RNA N.

Organization of the major RNAs. The positions and direction of synthesis of the seven major RNAs are summarized in Fig. 5. The major RNAs produced by the *B. subtilis* polymerase all map within the terminally redundant region of the phage DNA. Synthesis of the RNA A and RNA F transcripts begins very near the genome end; initiation of these two RNAs begins near or at the same point on the DNA, and elongation of RNA A results in read-through of the RNA F terminator. The remaining major RNAs are organized in such a way that initiation begins at several places on the DNA and ends at a bidirectional termination site in the *HhaI* e fragment. All major transcripts overlap with at least one other transcript.

Minor RNA species. As with the major RNAs, endonuclease digestion of *SalI*-C should result in the loss of minor RNA species and the appearance of new truncated derivatives. However, the relatively low level of synthesis of these RNAs (15% of the total) and the variable volumes applied to the gels (Fig. 2 and 3) made it

difficult to detect minor RNAs shorter than 1,800 bases. Consequently, mapping of some minor RNAs was attempted by transcription of individual *HhaI* and *HaeII* DNA fragments. Evidence indicated that 5 of the 13 minor RNAs (B, C, J, O, and P) are located near the bidirectional terminator and that at least 1 minor RNA is derived from a region of the terminal redundancy that does not code for any of the major RNA species. Information concerning the minor RNAs is summarized in Table 2.

The 2,530-base RNA B was lost after digestion by the same endonucleases that resulted in the loss of major RNA D (Fig. 2, lanes 3, 5, 7, and 9; and Fig. 3, lanes 2, 6, and 8). This indicates that the two RNAs overlap, and the length of RNA B suggests that initiation of the transcript occurs about 300 base pairs "upstream" from the RNA D initiation site. Elongation would be leftward to the bidirectional termination site.

Transcription of the *HaeII* b fragment (Fig. 1) yielded minor RNA C and major RNAs, RNA E, RNA I, and RNA N (Fig. 4, lane 8). The length of RNA C suggests that initiation begins either at the promoter for RNA N or a short distance upstream, near the end of the *HhaI* e fragment (this fragment is contained within *HaeII*-b [Fig. 1]). Transcription would proceed leftward through the bidirectional termination site, resulting in a transcript containing a segment of antimessage (Fig. 5). The amount of antimessage would represent a very small proportion of the total RNA since, as stated above, the 13 minor RNAs together account for only 15% of the total in vitro RNA synthesis. Read-through of the terminator is proposed below to account for some of the RNAs produced by

TABLE 2. Minor RNAs synthesized by the *B. subtilis* RNA polymerase

Minor RNA	Length (bases)	Position
B	2,530	Starts 300 base pairs upstream from RNA D initiation site and stops at bidirectional terminator
C	2,500	Within <i>HaeII</i> -b; starts at promoter for RNA L or near end of <i>HhaI</i> -e
G	1,350	— ^a
H	1,320	—
J	1,240	Within <i>HhaI</i> -e
K	1,130	—
L	930	—
M	840	—
O	650	Within <i>HhaI</i> -e
P	560	Within <i>HhaI</i> -e
R	410	—
S	380	—
T	270	—

^a —, Insufficient data to position the RNA.

transcription of the isolated *HhaI* e fragment. As documented elsewhere (J. S. Downard and H. R. Whiteley, submitted for publication), most of the larger RNAs, such as RNA C, can be processed to smaller RNAs by extracts of *B. subtilis*. Conceivably, the coding portions of such read-through transcripts could be translated after processing; alternatively, sequences preceding the terminator may code for a translational stop signal.

Transcription of the 1.5-kb *HhaI* e DNA fragment revealed that at least three other minor RNAs are located near the bidirectional terminator: RNAs J, O, and P. The location of the initiation sites for these RNAs could not be determined because *HhaI* fragments extracted from gels were refractory to digestion by *KpnI*, presumably because of inhibitory substances carried over in the isolation of the fragments. As shown in lane 4 of Fig. 4, transcription of the *HhaI* e DNA fragment yielded a total of seven RNAs: two major RNAs, RNA N and RNA Q, three minor RNAs as mentioned above, plus two additional RNAs of 1,450 and 1,290 bases. These bands, plus two truncated RNAs of 770 and 220 bases described earlier (Table 1), account for all of the bands seen in lane 5 of Fig. 2. The prominent 1,450-base RNA is similar in size to the entire *HhaI* e DNA fragment. This transcript was not synthesized after double digestion of *SalI*-C with *HhaI* and *KpnI*, and no truncated derivative was detected (data not shown), suggesting that transcription is leftward from a site at or near the end of the *HhaI* e fragment. However, the origin of this RNA is not known. It may be speculated that it arises by transcription from a promoter near the right end of the *HhaI* e fragment (possibly a promoter for one of the unmapped minor RNAs) through the terminator to the end of the fragment. In fact, there appears to be an initiation site near the right end of the *HhaI* e fragment: digestion of *HhaI*-e with *EcoRI* (Fig. 1) yields a 105-base-pair fragment which forms a stable binary complex with RNA polymerase (data not shown). The prominence of the RNA in lane 4 of Fig. 4 may have been a result of one or more of the following factors: (i) transcription of an individual fragment eliminated competition for polymerase by promoters on other fragments; (ii) reaction mixtures containing an isolated fragment had a higher polymerase/DNA ratio with respect to an individual fragment than did reaction mixtures containing a digest of *SalI*-C; and (iii) initiation near the end of the fragment may have been facilitated by local unwinding of the DNA (such unwinding is required for the formation of the open promoter complex [1]). Thus, the relative intensity of this band may have been arti-

factual, resulting from transcription of an isolated fragment and the proximity of a promoter to the end of this fragment.

Read-through of the bidirectional terminator could also account for the synthesis of the 1,290-base RNA. Thus, an RNA of 1,290 bases could be produced by leftward transcription from the promoter for major RNA N to the end of the fragment. In one experiment, an additional RNA of 650 bases was synthesized from the *HhaI* e DNA fragment and from the *HhaI* digest of *SalI*-C; possibly, this RNA also arose by read-through from a promoter for one of the minor RNAs. DNA sequence analysis of the *HhaI* e fragment may reveal the structure of the bidirectional terminator, locate exactly the promoters for the major and minor RNAs, and confirm the presence of a promoter near the right end of the fragment.

Transcription of the *HhaI* c DNA fragment (Fig. 1) resulted in the synthesis of a 770-base RNA (Fig. 4, lane 7). This shortened RNA would comigrate with the truncated derivatives of major RNAs, RNA A and RNA F in analysis of RNAs produced from *SalI*-C digested with *HhaI* (Fig. 2, lane 5). Since none of the major species initiates within the *HhaI* c DNA fragment, the 770-base RNA is probably a derivative of one of the unmapped minor RNAs. Because there is only one RNA derived from the *HhaI* c fragment, the amount of RNA on the gel cannot be taken as an indication of promoter strength.

Promoter-bearing fragments within *SalI*-C. Using the filter-binding method of Jones et al. (9), Spiegelman and Whiteley (20) were able to identify three *HpaII* fragments of SP82 that contain promoters recognized by the unmodified RNA polymerase, and all three fragments are derived from near the ends of the genome (Lawrie, Ph.D. dissertation). The ordering of the 10 *HpaII*, 13 *HhaI*, and 22 *HpaII*-*HhaI* fragments within the *SalI* C fragments (Fig. 1) allows greater resolution of initiation sites for the unmodified RNA polymerase. When these fragments were tested for the ability to complex the unmodified RNA polymerase, it was found that only a subset contained promoters (Fig. 6). The following nine *HhaI* fragments formed binary complexes with *B. subtilis* RNA polymerase: b, c, d, e, f, g, h, k, and l. Achberger and Whiteley (in press) have examined *HhaI* fragments produced from unfractionated SP82 DNA to identify early promoters and have classified promoters as "strong" or "weak" on the basis of complex formation at decreasing polymerase/DNA ratios and at increasing salt concentrations. Of the approximately 60 *HhaI* fragments produced from SP82 DNA, only fragments b, c, d, e, f, g, and h were found to contain

strong promoters (fragments 12, 15, 18, 26, 32, 35, and 37, respectively, produced from the digest of unfractionated SP82 DNA [Achberger and Whiteley, in press]). The relative strengths of complex formation with *HhaI* fragments k and l were not determined since these fragments migrated off the gel in the electrophoretic system used in the earlier investigation (Achberger and Whiteley, in press). The strong promoter-bearing fragments all mapped within the SP82 terminal redundancy (Fig. 7). Five other fragments identified as having weaker promoters all map outside the redundant region (Achberger and Whiteley, in press). Correlation of these data with the transcription map (Fig. 5) revealed that all major RNA species are initiated on fragments found to contain strong promoters.

DISCUSSION

Hybridization of RNAs synthesized in vivo

FIG. 6. Promoter-bearing fragments of SP82 DNA and *SalI*-C. The left lane of each panel shows the fragments generated by digestion with *HpaII*, *HhaI*, or the *HpaII*-*HhaI* double digest. The right lane of each panel shows the fragments complexed by *B. subtilis* RNA polymerase. Promoter-bearing fragments of *SalI*-C are identified by their letter designations (Fig. 1).

and in vitro to *HpaII* and *HhaI* fragments of SP82 DNA has been used previously to identify regions of DNA transcribed at intervals throughout infection (12, 13; Lawrie, Ph.D. dissertation). The ordering of these fragments (Fig. 1) reveals that most of the transcripts produced in vitro by *B. subtilis* polymerase come from *HhaI* fragments g through c, located in the 13.1-kb SP82 terminal redundancy; the same fragments are transcribed early in infection (Lawrie, Ph.D. dissertation).

Terminal repetition in phage DNA is generally considered to play a role during DNA replication (23), and the presence of concatemeric intermediates in SP01-infected cells (2) suggests that the SP82 and SP01 redundancies also serve this function. However, the length of the terminal redundancy in these phages (12 to 13 kb) appears excessive solely for a role in DNA concatemerization. For example, 100 base pairs of terminal repetition is sufficient for T7 DNA concatemerization (17). A second possible function of the SP82 redundancy may be to provide two copies of the early phage genes.

The present data indicate that the major transcripts originate from several different promoters and that all of the major RNAs overlap with at least one other transcript. Transcription occurs in both directions; this is compatible with earlier data showing that RNA synthesized in vivo and in vitro by the *B. subtilis* polymerase from SP82 DNA hybridizes to approximately the same extent to each of the strands (11). The mapping data suggest that synthesis of five of the major RNAs is terminated at a bidirectional terminator and that two are terminated at other sites. Synthesis of one minor RNA (RNA C) is postulated to involve read-through of the bidirectional terminator; production of one major RNA (RNA A) apparently occurs by read-through of the terminator for RNA F. Read-through of the bidirectional terminator has also been proposed to explain the production of two RNAs from an isolated restriction fragment. Judging from the relative intensities of RNA bands detected on gels, transcripts were made in different relative amounts. It appears unlikely that elongation would selectively favor certain transcripts. It seems more likely that differences

FIG. 7. Location of promoter-bearing fragments within *SalI*-C. Bold lines indicate promoter-bearing fragments.

in the initial interaction of polymerase with promoters or, possibly, differences in the rates of initiation may be responsible for differences in synthesis. Alternatively, undetected weak termination sites could result in diminution of specific RNAs. Some minor RNAs synthesized from individual fragments had a higher intensity, relative to other RNAs on the gels, suggesting that the efficiency of promoter utilization was greater when competition by other promoters on other fragments was eliminated.

Binding studies identified nine fragments within *SalI*-C capable of forming stable complexes with polymerase (Fig. 7). Synthesis of each of the major transcripts mapped in the present study is apparently initiated within a fragment identified as possessing a strong promoter capable of binding *B. subtilis* polymerase (Achberger and Whiteley, in press). Three promoter-containing fragments (*HhaI*-h, -d, and -k) were found for which there are no corresponding major transcripts. Perhaps these fragments contain promoters for unmapped minor RNA species.

Several considerations suggest that the longest major transcripts code for more than one polypeptide. First, the lengths of some of these RNAs are sufficient to code for several average-sized polypeptides and, in addition, many of the early phage peptides are relatively small in size (8). Second, the organization of the major RNAs is such that all transcripts overlap with one or more of the other transcripts (Fig. 5). This implies that some promoters are located at intergenic sites within longer transcriptional units in the DNA. Finally, some of the major RNAs synthesized by RNA polymerase in vitro can be processed to RNAs similar in size to those synthesized in vivo (Downard and Whiteley, submitted for publication). These observations taken together support the idea that the primary transcripts produced in vitro are polycistronic and suggest that similar polycistronic RNAs are produced in vivo and are cleaved at noncoding sites on the RNA in a manner analogous to the processing of early mRNA's of phage T7 (3).

Middle and late transcriptions by the modified forms of RNA polymerase take place both within and outside the terminal redundancy (12, 16). The major transcripts synthesized by the unmodified polymerase completely span some *HhaI* fragments of the redundancy (*HhaI*-g, -i, -f, -e, and -j [Fig. 5]), and some of these same fragments hybridize with RNA made during the middle and late stages of infection. This indicates either that some of the polymerase remains unmodified during phage infection or that some regions of the phage genome are transcribed by both the unmodified and modified forms of po-

lymerase.

The mapping of the in vitro transcripts on the *SalI* C fragments of SP82 provides a means of locating promoter and terminator sequences and should be of use in isolating fragments containing these sequences, thus permitting detailed studies on the regulation of transcription. Mapping of transcripts which are known to be produced from *SalI*-C during the middle of the infection will permit similar studies of transcription by a modified form of the polymerase.

ACKNOWLEDGMENTS

This research was supported by Public Health Service grants GM-20784 and GM-26100 from the National Institute of General Medical Science. A.T.P. was supported by the Molecular and Cellular Biology Training Program, grant GM-07270. H.R.W. is a recipient of Research Career Award K6-GM-442 from the National Institute of General Medical Sciences.

LITERATURE CITED

1. Chamberlin, M. 1974. The selectivity of transcription. *Annu. Rev. Biochem.* 43:721-775.
2. Cregg, J. M., and C. P. Stewart. 1978. Terminal redundancy of "high frequency of recombination" markers of *Bacillus subtilis* phage SP01. *Virology* 86:530-541.
3. Dunn, J. J., and F. W. Studier. 1973. T7 early RNAs are generated by site-specific cleavages. *Proc. Natl. Acad. Sci. U.S.A.* 70:1559-1563.
4. Gage, L. P., and E. P. Geiduschek. 1971. RNA synthesis during bacteriophage SP01 development: six classes of SP01 RNA. *J. Mol. Biol.* 57:279-300.
5. Gage, L. P., and E. P. Geiduschek. 1971. RNA synthesis during bacteriophage SP01 development. II. Some modulations and prerequisites of the transcriptional program. *Virology* 44:200-210.
6. Gegenheimer, P., N. Watson, and D. Apirion. 1977. Multiple pathways for primary processing of ribosomal RNA in *E. coli*. *J. Biol. Chem.* 252:3064-3073.
7. Greene, P. J., H. L. Heynecker, F. Bolivar, R. L. Rodriguez, M. C. Betlach, A. A. Covarrubias, K. Backman, D. J. Russel, R. Tait, and H. W. Boyer. 1978. A general method for the purification of restriction enzymes. *Nucleic Acids Res.* 5:2373-2380.
8. Hiatt, W. R., and H. R. Whiteley. 1978. Translation of RNAs synthesized in vivo and in vitro from bacteriophage SP82 DNA. *J. Virol.* 25:616-629.
9. Jones, B. B., H. Chan, S. Rothstein, R. D. Wells, and W. S. Reznikoff. 1977. RNA polymerase binding sites in λ plac5 DNA. *Proc. Natl. Acad. Sci. U.S.A.* 74:4914-4918.
10. Lawrie, J. M., G. B. Spiegelman, and H. R. Whiteley. 1975. DNA strand specificity of transcripts produced in vivo and in vitro by RNA polymerase from SP82-infected *Bacillus subtilis*. *J. Virol.* 15:1286-1288.
11. Lawrie, J. M., G. B. Spiegelman, and H. R. Whiteley. 1976. DNA strand specificity of temporal RNA classes produced during infection of *Bacillus subtilis* by SP82. *J. Virol.* 19:359-373.
12. Lawrie, J. M., G. B. Spiegelman, and H. R. Whiteley. 1977. Localization of transcripts produced in vivo and in vitro on phage SP82 genome. *Gene* 2:251-262.
13. Lawrie, J. M., and H. R. Whiteley. 1977. A physical map of bacteriophage SP82 DNA. *Gene* 2:233-250.
14. Levinthal, C., J. Hosoda, and D. Shub. 1967. The control of protein synthesis after phage infection, p. 71-87. In J. S. Colter and M. Paranchych (ed.), *Molecular biology of viruses*. Academic Press, Inc., New York.
15. Losick, R., and J. Pero. 1976. Regulatory subunits of

- RNA polymerase, p. 227-246. In R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Pero, J., N. M. Hannett, and C. Talkington.** 1979. Restriction cleavage map of SP01 DNA: general location of early, middle, and late genes. *J. Virol.* **31**:156-171.
 17. **Ritchie, D. A., C. A. Thomas, L. A. MacHattie, and P. C. Wensink.** 1977. Terminal repetition in non-permuted T3 and T7 bacteriophage DNA molecules. *J. Mol. Biol.* **23**:365-376.
 18. **Spiegelman, G. B., W. R. Hiatt, and H. R. Whiteley.** 1978. Role of the 21,000 molecular weight polypeptide of *Bacillus subtilis* RNA polymerase in RNA synthesis. *J. Biol. Chem.* **253**:1756-1765.
 19. **Spiegelman, G. B., and H. R. Whiteley.** 1974. In vivo and in vitro transcription by ribonucleic acid polymerase from SP82-infected *Bacillus subtilis*. *J. Biol. Chem.* **249**:1483-1489.
 20. **Spiegelman, G. B., and H. R. Whiteley.** 1978. Bacteriophage SP82 induced modifications of *Bacillus subtilis* RNA polymerase result in recognition of additional RNA synthesis initiation sites on phage DNA. *Biochem. Biophys. Res. Commun.* **81**:1058-1065.
 21. **Stafford, D. W., and D. Bieber.** 1975. Concentration of DNA solutions by extraction with 2-butanol. *Biochim. Biophys. Acta* **378**:18-21.
 22. **Studier, F. W.** 1973. Analysis of bacteriophage T7 early RNAs and proteins on slab gels. *J. Mol. Biol.* **79**:237-248.
 23. **Watson, J. D.** 1972. Origin of concatemeric T7 DNA. *Nature (London) New Biol.* **239**:197-201.
 24. **Whiteley, H. R., G. B. Spiegelman, J. M. Lawrie, and W. R. Hiatt.** 1976. The in vitro transcriptional specificity of RNA polymerase isolated from SP82-infected *Bacillus subtilis*, p. 587-600. In R. Losick and M. Chamberlin (ed.), RNA polymerase. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.