In Vitro Transcription of the Early Region of *Caulobacter* Phage φCd1 Deoxyribonucleic Acid by Host RNA Polymerase†

Kei Amemiya and Lucille Shapiro*

ABSTRACT: Transcription of the *Caulobacter crescentus* phage φCd1 genome requires both the host RNA polymerase and a phage-encoded, rifampicin-resistant RNA polymerase. Transcription of the early region of the φCd1 genome was examined in vitro with *C. crescentus* RNA polymerase. Four transcripts, A, B, C, and D, which ranged in size from 2.9×10^6 to 0.53×10^6 daltons, were synthesized in vitro by the holoenzyme. Transcript A appeared to be the major transcript since (a) it was the size of the entire 20% of the genome shown in vivo to code for the early phage mRNA, (b) it was one of the first transcripts synthesized at low enzyme-to-DNA molar ratios, and (c) it was synthesized in approximately 3 times the molar equivalent observed for the other transcripts. The A transcript initiated primarily with GTP although a portion was also labeled with ATP. The B, C, and D transcripts were present in equivalent molar ratios, were all smaller than transcript A, and were found to yield RNase III digestion products that were subsets of each other as well as of transcript A. Each of these transcripts proved to be a de novo transcript since (a) each could be pulse labeled during the initial 20 s of the reaction and (b) each transcript contained a triphosphate at its 5′ terminus. Evidence is presented that suggests that the B and C transcripts initiate at or near the major A promoter but terminate at different termination or pause sites within the early region of the phage genome. Transcript D appears to initiate at a minor promoter within the terminally redundant region of the genome preceding the A promoter.

The bacterium *Caulobacter crescentus* carries out a series of morphogenetic events that include the formation of a flagellum, a pili, and a stalk at specific sites on the cell surface at specific times during each cell division cycle (Poindexter, 1964; Shapiro, 1976; Shapiro & Agabian, 1970). It has been demonstrated that several protein components of these surface structures are synthesized coincident with their assembly at the cell surface (Cheung & Newton, 1977; Lagenaur & Agabian, 1978; Poindexter, 1964; Shapiro & Maizel, 1973) and that their synthesis is dependent on de novo RNA synthesis (Newton, 1972). These observations suggest that regulation of gene transcription may be a factor in the temporal control of *C. crescentus* cell differentiation. In order to determine the manner by which transcription can be regulated in *C. crescentus*, we are examining the interaction of the *C. crescentus* DNA-dependent RNA polymerase (Amemiya et al., 1977) with defined genomic clones and phage DNA templates. Transcription of well-characterized phage DNA templates was undertaken in order to assess the regulatory signals, such as promoter sites, pause sites, and terminator sequences in *Caulobacter* DNA templates of high G+C content. The *Caulobacter* phage φCd1 has, like its host, a G+C content of 67% (West et al., 1976). It has previously been reported that *Caulobacter* phage, φCd1, bears a striking resemblance to phage T7 (Raboy et al., 1980; West et al., 1976). T7 and φCd1 are morphologically similar and have linear genomes of 26.5×10^6 and 29×10^6 daltons (Raboy et al., 1980; West et al., 1976), respectively. Both contain terminally repetitious sequences (Raboy et al., 1980; Ritchie et al., 1967). The *C. crescentus* host RNA polymerase has been shown to be required for the transcription of the early region of the phage DNA, and a phage-encoded RNA polymerase appears responsible for expression of the late region (Amemiya et al., 1980). Infection of *C. crescentus* by φCd1 was also shown to induce a phage-specific protein kinase (West et al., 1976) that phosphorylates the host RNA polymerase (D. A. Hodgson and K. Amemiya, unpublished results), as well as other bacterial and phage-encoded proteins. Although *Caulobacter* phage φCd1 DNA and coliphage T7 DNA were found not to have nucleotide sequence homology by Southern hybridization analysis, we show here that φCd1 and T7 are similar with respect to transcription and the processing of transcripts from the early region of the phage genome.

Initial studies on the ability of the *C. crescentus* RNA polymerase to recognize specific initiation and termination signals on phage T7 and *Escherichia coli* DNA showed that the *C. crescentus* RNA polymerase recognizes the same promoters and terminators in the early region of T7 DNA as the *E. coli* enzyme (Wiggs et al., 1979; K. Amemiya and L. Shapiro, unpublished results). Furthermore, it was demonstrated by Farnham & Platt (1980) that both the *C. crescentus* and the *E. coli* RNA polymerases recognize the same termination site within an altered attenuator in the *E. coli* tryptophan operon. We have found, however, that when the *E. coli* RNA polymerase was used to transcribe φCd1 DNA, the specific transcripts synthesized differed from those obtained when this template was transcribed by the homologous *C. crescentus* enzyme (K. Amemiya and L. Shapiro, unpublished results). In order to understand the types of transcription signals differentially read by the *C. crescentus* and the *E. coli* RNA polymerases, we have first analyzed the in vitro transcription of φCd1 DNA by the *C. crescentus* enzyme. We described here the RNA transcripts synthesized in vitro by the host RNA polymerase.

Materials and Methods

Materials. Unlabeled nucleoside triphosphates were obtained from Schwarz/Mann. [α-^{32}P]CTP, [α-^{32}P]ATP, [γ-^{32}P]GTP, and [γ-^{32}P]ATP were obtained from ICN or Amersham. φCd1 DNA was prepared as described previously (Raboy et al., 1980). Spermidine and heparin were purchased from Sigma. Agarose and polyacrylamide were obtained from Miles Laboratories and Bio-Rad Laboratories, respectively. Restriction enzymes *Hin*dIII and *Hin*cII were obtained from

† From the Department of Molecular Biology, Division of Biological Sciences, Albert Einstein College of Medicine, Bronx, New York, New York 10461. *Received February 26, 1982.* This investigation was supported by Grant PCM 800-7508 from the National Science Foundation and Grants GM 11301 and NIH/NCI P30 CAI-13330 from the National Institutes of Health.

0006-2960/82/0421-4707$01.25/0

both Bethesda Research Laboratories and New England Biolabs. *C. crescentus* RNase III was prepared in our laboratory by V. Bellofatto. BA85 nitrocellulose filter paper was purchased from Schleicher and Schuell.

RNA Polymerase Preparations. RNA polymerase holoenzyme was purified from *C. crescentus* CB13 by the method of Amemiya et al. (1977) with the addition of Bio-Gel A1.5m (Bio-Rad Laboratories) column chromatography after fractionation of the enzyme on a phosphocellulose column. The Bio-Gel column (1.5 × 67 cm) was equilibrated with buffer C [0.05 M Tris-HCl,[1] pH 8.0, 10^{-4} M dithiothreitol, and 10^{-4} M EDTA containing 0.5 M KCl and 20% glycerol (v/v)]. Column fractions containing enzyme activity were combined, concentrated by Amicon filtration, and dialyzed against storage buffer (Amemiya et al., 1977). The *C. crescentus* core polymerase was purified as previously described (Amemiya et al., 1977) and was then passed twice through a phosphocellulose column before being dialyzed against storage buffer. Holoenzyme was 60–80% saturated with the σ subunit, while the core polymerase preparation contained less than 1% of the σ subunit. The specific activities (Amemiya et al., 1977) of the enzymes with ϕCd1 DNA as template were approximately 1100 and 360 units/mg of protein for the holoenzyme and core polymerase, respectively. The percentage of active molecules in the holoenzyme preparation was determined to be approximately 25% according to the method of Chamberlin et al. (1979).

RNA Polymerase Assays. The standard RNA polymerase assay (0.1 mL) contained, unless stated otherwise, 40 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 1 mM dithiothreitol, 2 mM spermidine, 0.5 mM ATP, GTP, and UTP and 0.1 mM [α-^{32}P]CTP (200–2000 cpm/pmol), 5–20 μg of ϕCd1 DNA, and enzyme. In most cases the enzyme and DNA were incubated for 10 min at 37 °C before the addition of the ribonucleoside triphosphates and heparin (50 μg/mL), and incubation was continued for 15 min. The reaction was stopped by the addition of 25 μL of a solution containing 27 mM Tris, 27 mM boric acid, 0.1 M EDTA, 0.1% sodium dodecyl sulfate, 0.1% bromophenol blue dye, and 30% glycerol (v/v). A 5–10-μL aliquot was withdrawn to determine the amount of acid-precipitable counts incorporated. In order to study the effect of heparin (50 μg/mL) on the time course of RNA synthesis, we followed the procedure described by Chamberlin et al. (1979), except the reaction volume was 0.5 mL and 50-μL aliquots were removed.

RNase III Assays. ϕCd1 RNA transcripts were digested with *C. crescentus* RNase III in a reaction mixture (0.1 mL) that contained 50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 60 mM KCl, 0.4 mM dithiothreitol, 0.1 mM EDTA, 50 μg/mL bovine serum albumin, RNA transcript, and enzyme. Incubation was at 37 °C for the indicated periods. The amount of enzyme used varied from approximately 1 to 2 units per assay. One unit of enzyme is defined as that amount of enzyme that produced a decrease of 1 nmol of acid-insoluble material in 1 h at 37 °C. ϕCd1 transcripts used for RNase III digestion were synthesized as described above. The RNA transcipts were separated and purified by passing the reaction mixture through a Sephadex G-50 column (0.7 × 27 cm) equilibrated with 20 mM sodium acetate (pH 5.4), 1 mM EDTA, and 0.1% $NaDodSO_4$, preceding electrophoresis through a polyacrylamide gel. The transcripts were cut out of the gels, and the gel slices were crushed in 20 mM sodium acetate and 1 mM EDTA and then eluted from the gel overnight. The extracted transcripts were extracted with phenol twice, precipitated with ethanol, and resuspended in 10 mM Tris-HCl (pH 8.0), 20 mM NaCl, and 0.1 mM EDTA.

[1] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; $NaDodSO_4$, sodium dodecyl sulfate.

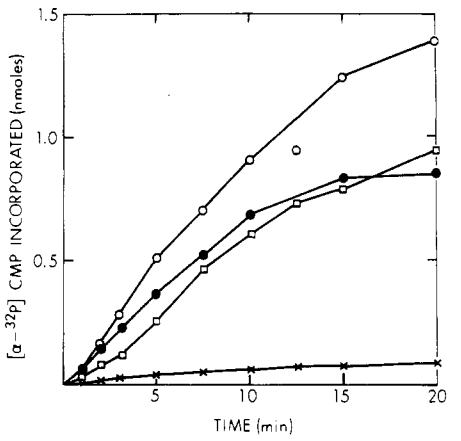

FIGURE 1: Effect of heparin on time course of RNA synthesis by *C. crescentus* RNA polymerase holoenzyme and core polymers on a phage ϕCd1 DNA template. RNA synthesis was determined as described under Materials and Methods in the presence or in the absence of heparin (50 μg/mL). Aliquots of 50 μL were removed from the reaction mixtures at the indicated times. Reaction mixtures (0.5 mL) were as described under Materials and Methods and contained 30 μg of ϕCd1 DNA with either holoenzyme (8.4 μg) or core polymerase (21.9 μg). The enzyme-to-DNA molar ratio was 16 and 50, respectively. RNA polymerase holoenzyme activity is shown with (●) or without (○) heparin. Core enzyme activity is shown with (×) or without (□) heparin.

Polyacrylamide Gel Electrophoresis. ϕCd1 RNA transcripts were analyzed by electrophoresis on 1.5% polyacrylamide–0.5% agarose (De Wachter & Fiers, 1972) slab gels (13 × 35 cm) as described by Golomb & Chamberlin (1974). In order to compare the transcripts synthesized under various conditions, we applied equal amounts of acid-precipitable counts to each lane. The electrophoretically separated transcripts were visualized by autoradiography with Kodak NS-5T No-Screen X-ray film.

Hybridization Analysis. ϕCd1 DNA was digested with *Hin*dIII or *Hin*cII as previously described (Amemiya et al., 1980), and restriction fragments were resolved by agarose gel electrophoresis (Amemiya et al., 1980) on 1.0% or 1.4% agarose, respectively. ϕCd1 *Hin*dIII restriction fragments were transferred to nitrocellulose filter paper as previously described (Amemiya et al., 1980), and ϕCd1 *Hin*cII restriction fragments were transferred to nitrocellulose filter paper by the method of Southern (1975). ^{32}P-Labeled ϕCd1 RNA transcripts were hybridized to transferred restriction fragments as previously described (Amemiya et al., 1980).

Results

ϕCd1 Transcripts Synthesized in Vitro by C. crescentus Holoenzyme and Core Polymerases. The rate and extent of RNA synthesis by the holoenzyme and core polymerases were affected by the presence of heparin (Figure 1). The incorporation of [α-^{32}P]CMP by the core enzyme in the presence of heparin was very low even with relatively high enzyme-to-DNA molar ratios and after 20 min was only about one-tenth of that incorporated by the core enzyme in the absence of heparin. The rate of [α-^{32}P]CMP incorporation by the holoenzyme decreased substantially after 10 min in the presence of heparin and then leveled off, whereas in the absence of heparin RNA synthesis continued.

Polyacrylamide gel electrophoresis of the transcripts produced by holoenzyme in the presence or absence of heparin

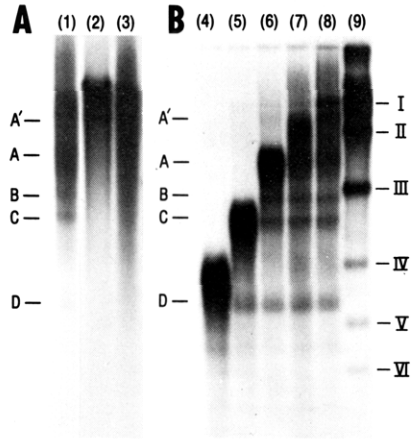

FIGURE 2: Electrophoretic analysis of RNA transcripts made by *C. crescentus* RNA polymerase. The conditions for synthesis of transcripts and analysis by polyacrylamide gel electrophoresis were as described under Materials and Methods. Panel A shows the effects of heparin (50 μg/mL) on RNA transcripts synthesized by holoenzyme or core polymerase. The reaction mixtures contained 12.3 μg of φCd1 DNA and 3.4 μg of RNA polymerase holoenzyme (E/DNA = 16) or 8.8 μg of core polymerase (E/DNA = 52). Lane 1 contains transcripts made by holoenzyme in the presence of heparin. Lanes 2 and 3 contain transcripts made by core polymerase with and without heparin, respectively. Equal amounts of acid-precipitable counts were applied to each lane. The total amount of [α-^{32}P]CMP incorporated into RNA in each reaction was as follows: (1) 1.8, (2) 0.15, and (3) 1.5 nmol. Panel B shows pulse-labeled RNA transcripts synthesized by RNA polymerase holoenzyme. The reaction mixtures contained 4.7 μg of RNA polymerase and 20 μg of φCd1 DNA (E/DNA = 14). Enzyme and DNA were incubated for 10 min at 37 °C, and then [α-^{32}P]ATP (8370 cpm/pmol) and the other ribonucleoside triphosphates were added to initiate the reaction. After a 20-s incubation cold ATP (5.0 mM) was added and the incubation continued for (4) 1, (5) 2, (6) 5, (7) 10, and (8) 15 min. Equal amounts of acid-precipitable counts were applied to each lane. The total amount of [α-^{32}P]AMP incorporated into RNA in each reaction was as follows: (4) 0.09, (5) 0.15, (6) 0.17, (7) 0.13, and (8) 0.14 nmol. Lane 9 contains RNA transcripts made by T7-specific RNA polymerase with T7 DNA for molecular weight standards.

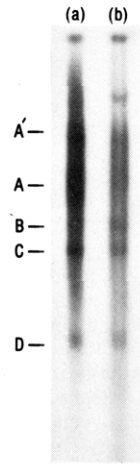

FIGURE 3: Electrophoretic analysis of [γ-^{32}P]GTP-labeled RNA transcripts made by RNA polymerase holoenzyme. The reaction mixtures (0.1 mL) were as described under Materials and Methods and contained 4.3 μg of RNA polymerase and 16 μg of φCd1 DNA (E/DNA = 16). The enzyme and DNA were incubated at 37 °C for 10 min and were then incubated for 2 min with either [α-^{32}P]CTP (1950 cpm/pmol) at a CTP concentration of 0.1 mM or [γ-^{32}P]GTP (9730 cpm/pmol) at a GTP concentration of 0.05 mM in the presence of heparin. After the labeling period the reaction mixtures were incubated for an additional 15 min with 5.0 mM cold CTP or GTP, respectively. The reaction mixture was then passed through a Sephadex G-50 column (0.7 × 27 cm), and the transcipts were resolved by polyacrylamide gel electrophoresis as described under Materials and Methods. The total amount of acid-precipitable counts recovered was 0.33 nmol for [α-^{32}P]CMP-labeled RNA and 1.32 pmol for [γ-^{32}P]GTP-labeled RNA. Lane a contained [α-^{32}P]CMP-labeled transcripts (20 000 cpm) and lane b contained [γ-^{32}P]GTP-labeled transcripts (13 000 cpm).

showed five RNA species (A′, A, B, C, and D) (Figure 2). The same pattern of transcripts was observed when rifampicin (25 μg/mL), instead of heparin, was used to inhibit reinitiation by holoenzyme. Core enzyme produced a very heterogeneous population of transcripts in the absence of heparin (Figure 2A), which ranged in size from greater than 5.5 × 10^6 to as small as 2 × 10^5. These core-catalyzed transcripts hybridized to the entire φCd1 genome, unlike the transcripts produced by the holoenzyme which, as shown below, hybridized specifically to the early region (20%) of the φCd1 genome. In the presence of heparin, however, core enzyme synthesized a large transcript similar in size to the A′ transcript produced by holoenzyme (Figure 2A). Because of the very low level of synthesis by the core enzyme the amount of enzyme was significantly increased in order to see the product on gels. The large transcript was found to hybridize to the same regions of the φCd1 genome as the holoenzyme A′ transcript, suggesting that the core enzyme catalyzed a small amount of a large read-through product when reinitiation was prevented by heparin. This low level of specific RNA synthesis by the core enzyme in the presence of heparin may be due to the presence of a small amount of σ in the core enzyme preparation which was amplified by the high enzyme-to-DNA molar ratio required for φCd1 transcription.

Analysis of Pulsed-Labeled and End-Labeled φCd1 Transcripts. Reaction mixtures containing φCd1 DNA and holoenzyme in the presence of heparin were incubated with [α-^{32}P]CTP for 20 s in the presence of the other ribonucleoside triphosphates. This was followed by a chase with cold CTP for varying time periods to allow the transcripts to elongate and terminate (Figure 2B). A transcript that comigrated with transcript D was first detected 2 min after the pulse. Transcripts B and C and a large heterogeneous population of RNA approximately the size of transcript A could be seen after 5 min. At 10 min transcript A was clearly discernible, and a larger transcript (A′) was detected. On the basis of transcript size and their time of synthesis, we estimate the average chain elongation rate to be approximately 19 nucleotides/s. In some cases a transcript larger than A′ could be seen as in Figure 4. From the size of this transcript, it suggests that there may be another read-through product that terminates in the late region of the φCd1 genome.

The observation that all the transcripts contained labeled [α-^{32}P]CMP after 15 min of incubation, although label was present in the reaction mixture for only the first 20 s, argues that all the transcripts initiated within the 20-s time period and that none of the transcripts represent cleavage products from the 3′ end of the RNA molecules. This result predicts that each of the φCd1 transcripts should have a triphosphate at the 5′ terminus of the molecule. RNA synthesized in the presence of [γ-^{32}P]GTP or [γ-^{32}P]ATP was separated by polyacrylamide gel electrophoresis, and it was found that all five transcripts were terminally labeled with [γ-^{32}P]GTP (Figure 3). Transcript A was also partially labeled with [γ-^{32}P]ATP (data not shown). In comparison to [γ-^{32}P]GTP incorporation, only one-tenth to one-fifth of the amount of [γ-^{32}P]ATP was incorporated into acid-precipitable counts.

The apparent molecular weights of the φCd1 transcripts ranged from 4.7 × 10^6 to 0.5 × 10^6 (Table I). The molar yield of transcript A was 3 times that of transcript C, while transcripts A′, B, and D were present in an amount equivalent to transcript C. Because RNAs migrating as the A transcript

Table I: Molecular Weights and Molar Yields of φCd1 in Vitro Transcripts[a]

φCd1 transcript	$M_r \times 10^{-6}$	relative yield (%)	molar ratio
A′	4.7	31.5	1.2
A	2.9	45.7	3.1
B	1.8	12.4	1.4
C	1.4	7.5	1.0
D	0.53	2.9	1.0

[a] Molecular weights were estimated from electrophoretic mobilities relative to RNA transcripts made by T7-specific RNA polymerase on T7 DNA (Golomb & Chamberlin, 1974). φCd1 in vitro transcripts were made in the presence of heparin at an enzyme-to-DNA ratio of 16. After synthesis of the transcripts the reaction mixture was passed through a Sephadex G-50 column (0.7 × 27 cm), the peak fractions were collected and concentrated, and transcripts were resolved by polyacrylamide gel electrophoresis as described under Materials and Methods. The transcripts were cut out of the gel, the radioactivity was counted, and the relative yield was calculated by dividing the amount of radioactivity in each transcript by the total yield of radioactivity obtained for all the transcripts and multiplying by 100. The molar ratios were calculated relative to the molar yield of transcript C. This was determined by dividing the radioactivity of the transcript by the molecular weight of the transcript and dividing the result by the molar yield obtained for transcript C.

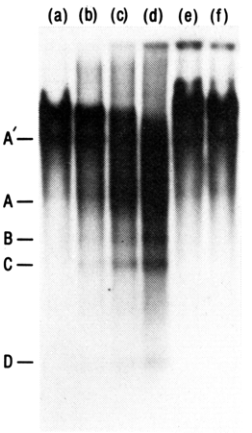

FIGURE 4: Electrophoretic analysis of RNA transcripts made at increasing enzyme-to-DNA molar ratios. The conditions for synthesis of transcripts and analysis by polyacrylamide gel electrophoresis were as described under Materials and Methods. The reaction mixtures contained 12 μg of φCd1 DNA and the following amounts of enzyme: (a) 0.85 μg of holoenzyme (E/DNA = 4), (b) 1.7 μg of holoenzyme (E/DNA = 8), (c) 3.4 μg of holoenzyme (E/DNA = 16), (d) 5.1 μg of holoenzyme (E/DNA = 25), (e) 8.8 μg of core enzyme (E/DNA = 51), and (f) 18 μg of core enzyme (E/DNA = 104). Equal amounts of acid-precipitable counts were applied to each lane. The total amount of [α-^{32}P]CMP incorporated into RNA in each reaction was as follows: (a) 0.15, (b) 0.67, (c) 1.5, (d) 1.8, (e) 0.10, and (f) 0.13 nmol.

were found to initiate with GTP as well as some ATP and because 3 molar equiv of transcript A were synthesized under conditions in which reinitiation was blocked, transcript A may consist of separate species that are relatively close in size.

Effect of Enzyme-to-DNA Molar Ratio and Ionic Strength on Transcription of φCd1 DNA. Analysis of the size distribution of RNA transcripts as a function of enzyme concentration showed that transcripts A′, A, and C were synthesized at relatively low enzyme-to-DNA molar ratios (E/DNA = 8) and transcripts B and D were synthesized only at higher enzyme concentrations (Figure 4). At the highest enzyme concentration examined (E/DNA = 25) the resolution between transcripts A and A′ became less apparent. In contrast to the holoenzyme, increasing the core enzyme concentration (E/DNA from 51 to 104) did not result in the appearance of transcripts other than transcript A′. These results indicate that there are a limited number of heparin-resistant sites on φCd1 DNA used by the holoenzyme for the initiation of transcripts and that some of these heparin-resistant sites (those responsible for transcripts A and C) are preferentially used.

The optimum salt concentration for the incorporation of [α-^{32}P]CMP into acid-precipitable material was found to be 50 mM KCl. At higher salt concentrations the enzyme activity decreased. The relative specific activity of each transcript was determined as a function of salt concentration, and it was found that transcripts C and D were significantly more sensitive to salt than transcripts A and B (Table II). The response of transcript A′ to increasing ionic strength was similar to that of transcript A. Increasing ionic strength did not alter the transcripts synthesized by the core system. It thus appears that the early φCd1 transcripts synthesized by the holoenzyme are differentially affected by both enzyme-to-DNA molar ratios and ionic strength.

RNase III Digestion of φCd1 in Vitro Transcripts. Analysis of RNA extracted from φCd1-infected cells in the presence of chloramphenicol revealed that the early in vivo transcripts were not the same size as the φCd1 transcripts synthesized in vitro (V. Bellofatto, K. Amemiya, and L. Shapiro, unpublished results). An RNase III like activity has been purified to homogeneity from *C. crescentus* (V. Bellofatto, K. Amemiya, and L. Shapiro, unpublished results) and appears to be involved in the processing of the early φCd1 in vivo transcripts. In order to determine if the φCd1 in vitro transcripts are processed and to assist us in mapping the transcripts, we digested the φCd1 transcripts with *C. crescentus* RNase III and analyzed them by polyacrylamide gel electrophoresis (Figure 5). *C. crescentus* RNase III cleaved the φCd1 transcripts into a limited number of discrete products that ranged from 333 to 4545 nucleotides in length. Although in the figure shown the digestion of the RNA transcripts was not complete, it is evident that the four transcripts, A, B, C, and D, yield RNase III cleavage products of similar electrophoretic mobility, suggesting that the cleavage products are derived from overlapping transcripts. The smaller cleavage products and some of the intermediate size products appear to be limit products because further digestion with RNase III did not change their mobility, and these transcripts correspond in size to several of the φCd1 transcripts synthesized in vivo (V. Bellofatto, K. Amemiya, and L. Shapiro, unpublished results). It appears, therefore, that the φCd1 in vitro and in vivo

Table II: Characteristics of φCd1 in Vitro Transcripts

φCd1 transcript[a]	effect of E/DNA ratio		effect of ionic strength			DNA fragment of transcription initiation	
	low	high	50 mM KCl	150 mM HCl	5′ nucleotide	*Hin*dIII	*Hin*cII
A	+	−	−	−	GTP (ATP)	F	D2
B	−	+	+	+	GTP	F	D2
C	+	+	+	−	GTP	F	D2
D	−	+	+	−	GTP	F	J3

[a] (+) indicates presence of transcript; (−) indicates absence of transcript.

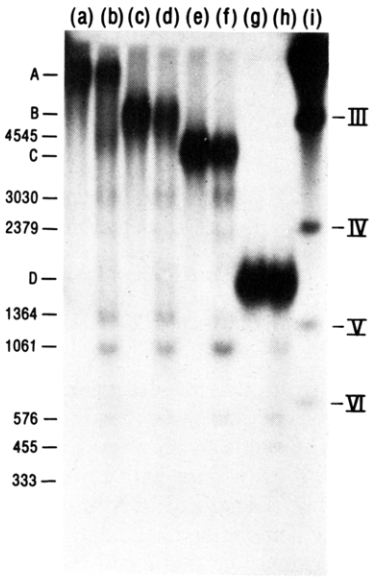

FIGURE 5: Electrophoretic analysis of ϕCd1 RNA transcripts cleaved by *C. crescentus* RNase III. ϕCd1 RNA transcripts were synthesized (E/DNA = 16), resolved by polyacrylamide gel electrophoresis, and eluted from the gel. Each transcript (10000 cpm) was digested with approximately 1 unit of *C. crescentus* RNase III for 30 min as described under Materials and Methods and analyzed on a 2% polyacrylamide–0.5% agarose slab gel. Lanes a, c, e, and g show ϕCd1 RNA transcripts A, B, C, and D, respectively, incubated without *C. crescentus* RNase III. Lanes b, d, f, and h show ϕCd1 RNA transcripts A, B, C, and D, respectively, incubated with *C. crescentus* RNase III. Lane i shows T7 RNA transcripts synthesized by the phage T7 RNA polymerase for molecular weight markers.

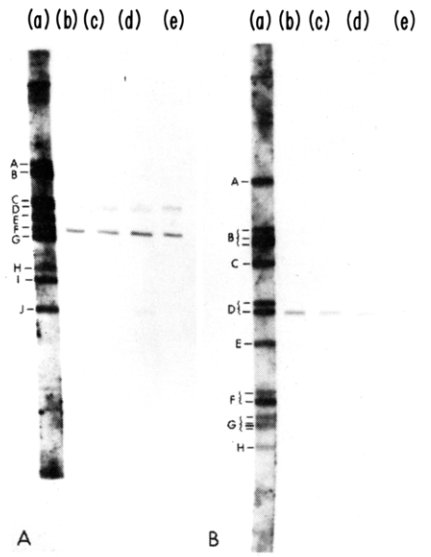

FIGURE 6: Hybridization of ϕCd1 RNA transcripts to ϕCd1 DNA restriction fragments. RNA transcripts were prepared with [α-^{32}P]CTP (5500 cpm/pmol) as described in Figure 2, except they were labeled for 60 s. After the transcripts were resolved by slab gel electrophoresis, they were eluted from the gel and digested with *C. crescentus* RNase III. Each transcript (15000 cpm) was digested with 1.6 units of RNase III for 60 min and phenol extracted before being used as hybridization probes. Hybridization was performed as described under Materials and Methods. RNA transcripts hybridized to ϕCd1 DNA *Hin*dIII restriction fragments are shown in panel A and to ϕCd1 DNA *Hin*cII restriction fragments in panel B. The RNA probes used in panels A and B were the following: (a) ϕCd1 RNA synthesized in vivo, (b) transcript A, (c) transcript B, (d) transcript C, and (e) transcript D.

transcripts differ because the in vivo transcripts are processed by RNase III. As described below, we have taken advantage of the processing of ϕCd1 transcripts by RNase III to map the position of the in vitro transcripts.

Location of in Vitro Transcripts on ϕCd1 Restriction Map. The regions on the ϕCd1 genome where the holoenzyme initiates transcription were determined by hybridizing pulse-labeled transcripts to Southern blots (Southern, 1975) of ϕCd1 DNA restriction fragments. The transcripts were prepared by pulse labeling with [α-^{32}P]CTP for 60 s and then allowing elongation to proceed in the presence of excess cold CTP. The pulse-labeled transcripts were then separated by polyacrylamide gel electrophoresis, eluted from the gel, and digested with *C. crescentus* RNase III. This enzyme was shown above to yield a limited number of RNA fragments (Figure 5). Since only the nucleotides near the 5′ ends of the transcripts were labeled in the 60-s pulse, treatment with this enzyme yields short, discrete RNAs carrying label from the 5′ region of the original transcript. On the basis of the rate of transcription calculated above (19 nucleotides/s), the labeled portion of the transcripts would be approximately 1200 nucleotides in length. Transcripts A, B, C, and D labeled and prepared in this manner hybridized primarily to ϕCd1 DNA restriction fragment *Hin*dIII F and *Hin*cII D2 (Figure 6). The ϕCd1 *Hin*dIII F restriction fragment is located at the left end of the ϕCd1 *Hin*dIII restriction map (see Figure 8). The ϕCd1 *Hin*cII D2 restriction fragment is located very near the left end of the ϕCd1 *Hin*cII restriction map. There was some light hybridization of ϕCd1 transcripts A, B, and C to ϕCd1 fragment *Hin*dIII D and *Hin*cII C and also light hybridization of transcript C to fragment *Hin*dIII J. It is not clear why this light hybridization was observed, since the *Hin*dIII D and J restriction fragments are located near the middle of the ϕCd1 genome (see Figure 8). Hybridization of the ϕCd1 transcripts to the *Hin*dIII restriction fragments suggests that *Hin*cII fragment C may reside at or near the *Hin*dIII fragments D and J. Transcript D also exhibited some hybridization to ϕCd1 restriction fragments *Hin*dIII C and *Hin*cII B2. Both of these restriction fragments are located at or very close to the right end of their respective restriction maps. Since we have shown that ϕCd1 DNA is terminally redundant (Raboy et al., 1980), it is possible that RNA transcribed from one end would hybridize to a homologous region at the other end. Results of the hybridization of transcript D to fragment *Hin*cII B2, which is located near the right end of the ϕCd1 genome, would suggest that the redundant region of the genome might extend into fragment *Hin*cII B2. We do not know if the redundant regions extend through fragments *Hin*cII J3 and I3. Our results, therefore, suggest that transcripts A, B, C, and possibly D are initiated near the left end of the ϕCd1 genome, within the *Hin*dIII fragment F, and are transcribed toward the right. It can be argued that transcript D initiates at an independent promoter within the terminally redundant region at the left end of the genome, which contains sequences homologous to the right end. This conclusion is based on the observations that transcript D hybridized to both ends of the ϕCd1 genome whereas transcripts A, B, and C hybridized only to the right end, yet the D transcript yielded RNase III cleavage products that appeared to be a subset of the RNase III fragments produced from transcripts A, B, and C. Furthermore, we show below that transcript D is transcribed from the same DNA strand as the other three transcripts.

In order to determine if transcripts A, B, C, and D were transcribed from the same or different strands of the ϕCd1 genome, we isolated *Hin*dIII fragment F from the right end, heat denatured it, and resolved the separated strands by agarose gel electrophoresis (Figure 7). All four transcripts were found to hybridize to the strand designated s, clearly showing that transcripts A, B, C, and D are transcribed from

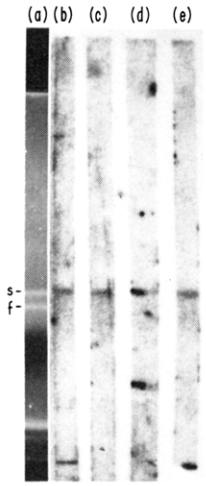

FIGURE 7: Hybridization of ϕCd1 RNA transcripts to separated strands of ϕCd1 DNA restriction fragment *Hin*dIII F. ϕCd1 DNA restriction fragment *Hin*dIII F was isolated and denatured in boiling water for 5 min and quick cooled. The heat-denatured restriction fragment was separated by electrophoresis on a 1% agarose slab gel and transferred to nitrocellullose filter paper as described under Materials and Methods. ϕCd1 RNA transcripts (12 500 cpm/transcript) were hybridized to the separated strands as described under Materials and Methods. The lanes show hybridization with the following transcripts: (b) transcript A, (c) transcript B, (d) transcript C, and (e) transcript D. Lane a shows the ethidium bromide stained agarose gel of heat-denatured ϕCd1 restriction fragment *Hin*dIII F.

the same strand of ϕCd1 DNA.

Discussion

The in vivo transcription of *Caulobacter* bacteriophage ϕCd1 DNA requires both the host RNA polymerase and a phage-specific rifampicin-resistant RNA polymerase (Amemiya et al., 1980). We have shown previously that the host RNA polymerase is responsible for the transcription of the early region (up to approximately 20 map units) while the phage-specific RNA polymerase is required for the expression of the late region. The size and organization of ϕCd1 DNA and its encoded gene products are remarkably similar to the coliphages T3 and T7. Because the G+C content of ϕCd1 DNA (67%) is considerably higher than either T3 or T7 (50%), there may be significant differences in the secondary structure of both pause sites and terminators in the *Caulobacter* phage DNA. Accordingly, we have analyzed the ϕCd1 DNA transcripts synthesized in vitro.

Transcription of ϕCd1 DNA in vitro by the *C. crescentus* RNA polymerase was found to yield four transcripts (A, B, C, and D) from the early region of the genome (Figure 8). Three observations suggest that transcript A is the major ϕCd1 transcript synthesized in vitro by the holoenzyme: (a) Transcript A was one of the first transcripts to be synthesized at low enzyme-to-DNA molar ratios. (b) The A transcript was synthesized in at least 3 molar equiv compared to that of the other ϕCd1 transcripts. (c) The size of transcript A, 2.9×10^6 daltons, is approximately the size of the early region of the ϕCd1 genome as determined by hybridization of RNA extracted from ϕCd1-infected cells to ϕCd1 DNA restriction fragments (Amemiya et al., 1980). These results suggest that the site where transcript A terminates is the major termination site between the early and late regions of the phage genome. Since the A transcript appears to initiate primarily with GTP, but to some extent with ATP as well, and since the A transcript is synthesized in 3 times the molar equivalent of the other transcripts, it may be that the A transcript is heterogeneous and initiates from three closely aligned promoters.

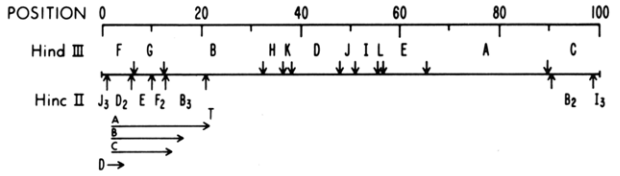

FIGURE 8: Restriction map of ϕCd1 DNA. The ϕCd1 *Hin*dIII restriction fragments were ordered as previously described (Raboy et al., 1980), and the same procedures were used to obtain the partial restriction map of the ϕCd1 *Hin*cII restriction fragments. The arrows beneath the restriction map indicate the position and direction of transcription of the early in vitro ϕCd1 RNA transcripts.

A larger in vitro transcript, designated A′, was found to hybridize to regions beyond those found to comprise the early region in vivo. The synthesis of the A′ transcript has similar properties to the A transcript with respect to enzyme concentration and ionic strength. The A′ transcript continued to elongate after the completion of the A transcript, suggesting that, in vitro, termination at the end of the early region is incomplete and that efficient termination at this site may require a *rho* termination factor. Furthermore, in the presence of low concentrations of the *E. coli rho* factor a decrease in the A′ transcript was observed (K. Amemiya and L. Shapiro, unpublished results). The A′ read-through transcript was present in only 1 molar equiv under conditions in which the A transcript was present in 3 molar equiv. This would imply that approximately one out of four RNA polymerase molecules that reach the major terminator are able to read through the terminator at the end of the early region. In fact, this was also found to be the case when *C. crescentus* RNA polymerase was used to transcribe T7 DNA in vitro (Wiggs et al., 1979; K. Amemiya and L. Shapiro, unpublished results).

The three other in vitro transcripts (B, C, and D) were found to initiate synthesis within the first 20 s of the reaction. These transcripts were present in equivalent molar ratios, contained a triphosphate, GTP, at their 5′ terminus, and were all smaller than transcript A. Digestion of these transcripts, as well as transcript A, with RNase III purified from *C. crescentus* yielded several fragments with the same mobility, suggesting that these transcripts have extensive overlap. These transcripts therefore result either from initiation at different early promoters or from termination at different termination or pause sites within the early region.

Although transcript D was found to initiate from the same restriction fragment within the early region as transcript A, it is likely that transcript D initiates at an independent promoter site because its synthesis required a relatively high enzyme-to-DNA molar ratio, it was more salt-sensitive than the A transcript, and it hybridized to ϕCd1 DNA both at the left end of the early region and at the right end of the phage genome. It can be argued that the D promoter resides in the terminally redundant region at the left end of the genome preceding the A promoter since (a) the D transcript hybridized to the same DNA strand of a restriction fragment from the left end of the genome as the other transcripts, (b) it contained a subset of the same RNase III cleavage products as the other transcripts, and (c) it contained sequences homologous to the right end of the genome, whereas the other transcripts did not.

On the basis of the site of initiation of the B and C transcripts, the size of these transcripts, and the fact that their RNase III cleavage products overlap and correspond to those produced from transcript A, it is likely that these transcripts initiate at or near the major A promoter and terminate within the early region of the genome. Termination of the B and C transcripts could result from *rho*-dependent termination sites

within the early region that function weakly in the absence of *rho*. Although it has been demonstrated that both the T3 and T7 phage genomes contain *rho*-dependent transcription termination sites in the early region (Adhya et al., 1979; Darlix & Horaist, 1975; Kassavetis & Chamberlin, 1981), in a preliminary study we have found that low concentrations of the *E. coli rho* factor affected only the ϕCd1 in vitro transcripts A′ and A. Alternatively, termination could result from discrete RNA polymerase pause sites similar to those that have been found in the early region of T7 DNA under suboptimal substrate concentrations (Darlix & Horaist, 1975; Kassavetis & Chamberlin, 1981). It is unlikely, however, that the B and C transcripts observed in this study are due to pause-site termination since synthesis occurred in the presence of excess nucleoside triphosphate (0.5 mM and labeled nucleotide at 0.1 mM) and could not be chased into higher molecular weight transcripts once they were completed. The possibility remains that there are weak termination sites within the early region of the ϕCd1 genome that function to prevent complete read through of some enzyme molecules. These sites could be structurally similar to but independent of RNase III sites and because of the high G+C content could function differently with the *C. crescentus* RNA polymerase and the *E. coli* RNA polymerase or *rho* factor. In fact, we have recently demonstrated that although the *E. coli* RNA polymerase synthesizes the major A transcript from the ϕCd1 DNA in vitro, the C and D transcripts are replaced by transcripts of different sizes.

Acknowledgments

We thank Vivian Bellofatto for her assistance with the RNase III studies.

References

Adhya, S., Sarkar, P., Valenzuela, D., & Maitra, U. (1979) *Proc. Natl. Acad. Sci. U.S.A. 249*, 6901.
Amemiya, K., Wu, C. W., & Shapiro, L. (1977) *J. Biol. Chem. 252*, 4157.
Amemiya, K., Raboy, B., & Shapiro, L. (1980) *Virology 104*, 109.
Chamberlin, M. J., Nierman, W. C., Wiggs, J., & Neff, N. (1979) *J. Biol. Chem. 254*, 10061.
Cheung, K. K., & Newton, A. (1977) *Dev. Biol. 56*, 417.
Darlix, J. L., & Horaist, M. (1975) *Nature (London) 256*, 288.
De Wachter, R., & Fiers, W. (1972) *Anal. Biochem. 49*, 184.
Farnham, P. J., & Platt, T. (1980) *Cell (Cambridge, Mass.) 20*, 739.
Golomb, M., & Chamberlin, M. J. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 760.
Kassavetis, G. A., & Chamberlin, M. J. (1981) *J. Biol. Chem. 256*, 2777.
Lagenaur, C., & Agabian, N. (1978) *J. Bacteriol. 135*, 1062.
Newton, A. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 447.
Poindexter, J. S. (1964) *Bacteriol. Rev. 28*, 231.
Raboy, B., Shapiro, L., & Amemiya, K. (1980) *J. Virol. 34*, 542.
Ritchie, D. A., Thomas, C. A., Machattie, L. A., & Wensink, P. C. (1967) *J. Mol. Biol. 23*, 365–376.
Shapiro, L. (1976) *Annu. Rev. Microbiol. 30*, 377.
Shapiro, L., & Agabian, N. (1970) *Proc. Natl. Acad. Sci. U.S.A. 67*, 200.
Shapiro, L., & Maizel, Jr., J. V. (1973) *J. Bacteriol. 113*, 478.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503.
West, D., Lagenaur, C., & Agabian, N. (1976) *J. Virol. 17*, 568.
Wiggs, J. L., Bush, J. W., & Chamberlin, M. J. (1979) *Cell (Cambridge, Mass.) 16*, 97.

Enzymatic Synthesis of a 21-Nucleotide Coat Protein Binding Fragment of R17 Ribonucleic Acid†

Marc Krug, Pieter L. de Haseth,‡ and Olke C. Uhlenbeck*

ABSTRACT: An oligoribonucleotide with a sequence identical with the bacteriophage R17 replicase initiator region has been synthesized. The sequence also encompasses the binding domain of R17 coat protein, which is known to act as a translational repressor at this site. The 21-nucleotide fragment was synthesized entirely by enzymatic methods, T4 RNA ligase being used to join shorter oligomers. The resulting fragment has a secondary structure with the expected thermal stability. Since the synthetic fragment binds R17 coat protein with the same affinity as a 59-nucleotide fragment isolated from R17 RNA, we conclude that it has full biological activity.

The specific "recognition" of an RNA sequence by a protein is thought to be due to a number of specific contacts spread over the surface of both molecules. A promising approach toward the identification of these contacts has been to use an in vitro synthesized target sequence that can be modified at will. For example, this approach has been used to examine the details of the interactions between *lac* repressor and *lac* operator DNA (Caruthers, 1980). In this paper we describe the synthesis of a short RNA fragment that is involved in a specific RNA–protein interaction. The procedure permits the construction of variant RNA fragments, which should allow the identification of the nucleotides in contact with the protein.

About 10 min after infection, translation of the bacteriophage R17 replicase gene is repressed when the product of the coat protein gene binds to the replicase initiator region (Sugiyama & Nakada, 1968). This binding is specific and strong enough to permit the protection of a 59-nucleotide fragment from ribonuclease T_1 digestion of the genomic RNA (Bernardi

† From the Department of Biochemistry, University of Illinois, Urbana, Illinois 61801. *Received April 14, 1982.* This work was supported by a grant from the National Institutes of Health (GM 19059).

‡ Present address: Department of Chemistry, University of Colorado, Boulder, CO 80309.

0006-2960/82/0421-4713$01.25/0