Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 2035–2039, April 1984
Biochemistry

Cloning and expression of the gene for bacteriophage T7 RNA polymerase

(pBR322/*lac* UV5 promoter/enzyme purification/protease/directed transcription)

PARICHEHRE DAVANLOO*, ALAN H. ROSENBERG, JOHN J. DUNN, AND F. WILLIAM STUDIER†

Biology Department, Brookhaven National Laboratory, Upton, NY 11973

Communicated by Richard B. Setlow, December 23, 1983

ABSTRACT **The complete coding sequence of the gene for bacteriophage T7 RNA polymerase (T7 gene *1*) has been cloned in the plasmid pBR322. Large amounts of active enzyme can be accumulated in *Escherichia coli* when the cloned gene is transcribed from the *lac* UV5 promoter. A protease activity that apparently can nick the protein without causing it to fall apart can be a problem during purification, but a procedure is described that gives good yields of essentially homogeneous, highly active enzyme suitable for biochemical and physical studies. T7 RNA polymerase has a stringent specificity for its own promoters and will selectively transcribe DNA that has been linked to such a promoter. This specificity makes the enzyme useful both for producing specific RNAs *in vitro* and for directing the expression of selected genes inside the cell. Having the cloned gene also makes possible a detailed mutational analysis of the functioning of T7 RNA polymerase.**

Bacteriophage T7 RNA polymerase (EC 2.7.7.6) is produced early in T7 infection and plays a central role in regulating gene expression (for a review, see ref. 1). A single-chain enzyme with a molecular weight close to 100,000, T7 RNA polymerase has a stringent specificity for its own promoters, which contain a highly conserved sequence of 23 continuous base pairs including the start site for the RNA. Seventeen such promoters are found in T7 DNA, but few if any such promoters are found in host DNA or in any other DNAs unrelated to T7. Once T7 RNA polymerase has been produced during infection, other T7 gene products inactivate the host RNA polymerase, leaving all transcription in the cell directed to T7 DNA.

The great specificity of T7 RNA polymerase for its own promoters is not only central to the strategy of T7 infection but is also interesting and potentially useful to biochemists. Understanding the basis for this specificity is a challenge in its own right. Furthermore, the purified enzyme can be used to produce large amounts of specific RNAs simply by transcribing DNA that has been joined to a promoter for T7 RNA polymerase. Such RNAs could be useful as hybridization probes, mRNAs for *in vitro* translation, substrates for analyzing processing reactions or RNA splicing, or for any purpose requiring a specific RNA. T7 RNA polymerase, together with a suitably positioned promoter, can also be used to direct the transcription of selected genes inside the cell. Target genes potentially could be expressed at very high levels, using the same strategy employed by T7 itself: once T7 RNA polymerase had been made, the host RNA polymerase could be inactivated, thereby eliminating the synthesis of competing mRNAs. The T7 RNA polymerase made during T7 infection has already been shown to be capable of directing the transcription of genes cloned in plasmids (2–4).

The yield of purified T7 RNA polymerase from infected cells is not particularly good, because the enzyme is synthesized for only a few minutes during infection and does not accumulate to high levels. Nor does T7 infection provide an especially good source of T7 RNA polymerase for directing the transcription of specific genes inside the cell, because there is competition from promoters in T7 DNA itself, and the cells lyse within a short time. However, a clone of the active T7 RNA polymerase gene in a plasmid or other vector could provide a source of enzyme without these limitations.

One previous attempt to clone the T7 RNA polymerase gene has been reported, by Stahl and Zinn (5). They obtained a clone of the entire gene except for the last nucleotide of the termination codon, but they did not report any T7 RNA polymerase activity from the clone. Loss of the termination codon would generate a protein having additional amino acids at its carboxyl terminus, presumably an inactive enzyme.

The nucleotide sequence of T7 DNA is completely known, as are the locations of the coding sequence for T7 RNA polymerase and the genetic signals surrounding it (6, 7). This information, together with our previous experience in cloning T7 genes in pBR322 (3) and our collection of well-characterized deletion mutants of T7 (ref. 8 and unpublished data), allowed us to develop a successful strategy for cloning and expressing the T7 RNA polymerase gene.

MATERIALS AND METHODS

Bacteria, Phage Strains, and Plasmids. *Escherichia coli* HMS174 (r_{K12}^- m_{K12}^+ *recA1* *rif*^{R} nonsuppressor) (2) was used as the host for plasmid strains. Bacteriophage T7, suitable hosts, and techniques for growing and manipulating them have been described (9–11). DNA fragments were cloned in pBR322 (12). The source of *lacI-lac* UV5 promoter DNA was the plasmid pMC1 (13). Transcription from the *lac* UV5 promoter was induced by adding 0.4 mM isopropyl β-D-thiogalactoside (IPTG) to growing cultures.

Enzymes and Linkers. Restriction endonucleases and enzymes used in cloning DNA were obtained from New England BioLabs, Bethesda Research Laboratories, or Boehringer Mannheim. Synthetic deoxynucleotide linkers containing a *Bam*HI site (C-G-G-G-A-T-C-C-C-G) or a *Bgl* II site (C-A-G-A-T-C-T-G) were obtained from New England BioLabs.

Cloning. Preparation and cloning of DNA fragments was by standard techniques (14, 15), essentially as outlined previously (3).

T7 Deletion Mutants. A number of deletions lying to the left or right of gene *1* have been described and characterized (8). These deletions apparently arose by genetic crossovers between short repeated sequences. The crossover for D159 is at the sequence A-A-T-G-C-T-G-A, located at nucleotides 975 and 3023 in the nucleotide sequence of T7 DNA, and the crossover for C74 is at the sequence G-T-G-G-C-C-T, located at nucleotides 1458 and 3128 (6, 8). The likely crossover

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: UOP Inc., Corporate Research Center, Ten UOP Plaza, Algonquin & Mt. Prospect Roads, Des Plaines, IL 60016.
†To whom reprint requests should be addressed.

sequences for LG4 (A-A-T-A-C-G-A-C-T-C-A-C-T-A at 5832 and 7879) and LG26 (G-G-T-A-A-G-A-A at 7165 and 8658) were deduced from restriction mapping and the locations of repeated sequences in this region of T7 DNA (unpublished data). Heteroduplexes between the DNAs of double-deletion mutants D159,LG4 and C74,LG26, when digested with the single-strand-specific nuclease S1, would be expected to produce a fragment extending from nucleotides 3128 to 5845, which would contain the entire coding sequence for T7 RNA polymerase (see Fig. 1). The double-deletion strains were constructed by conventional genetic crosses, and the presence of both deletions in each double-deletion strain was confirmed by restriction analysis of the DNA.

Isolation of Gene *1* Fragments. Heteroduplexes between the DNAs of the double-deletion strains D159,LG4 and C74,LG26 were prepared directly from concentrated stocks of phage particles that had been purified by isopycnic banding in CsCl solution containing small amounts of Tris and EDTA (9). The stocks in CsCl solution were diluted 1:50 in 30 mM NaOH and left for 10 min at room temperature to release and denature the DNA. The final DNA concentration was 100 μg/ml. The DNA mixture was neutralized by adding 0.1 vol of 0.5 M Tris·HCl, pH 6.8, and the DNA was renatured by incubating for 10 min at 65°C. The renatured DNA was precipitated with ethanol and redissolved in 5 mM Tris·HCl, pH 6.8. The solution was adjusted to one-third the original volume and to a composition of 0.3 M NaCl, 4.5 mM $ZnSO_4$, and 30 mM sodium acetate, pH 4.6, for treatment with S1 nuclease. After digestion sufficient to release the double-stranded fragments from the heteroduplexes, the DNA mixture was subjected to electrophoresis on a 1% agarose gel, where the gene *1* fragment migrated as a sharp band well resolved from the other digestion products.

Assay for T7 RNA Polymerase Activity. Reaction mixtures contained 40 mM Tris·HCl at pH 7.9, 8 mM $MgCl_2$, 5 mM dithiothreitol, 4 mM spermidine·HCl, and 0.4 mM each of ATP, GTP, CTP, and UTP (16). To monitor activity in crude extracts and during purification, incorporation of [α-^{32}P]UTP (30 μCi/μmol) into RNA in 15 min at 37°C in a 50-μl reaction mixture containing 2.5 μg of T7 DNA was measured by insolubility in 5% trichloroacetic acid. When assaying cruder fractions, streptolydigin (100 μg/ml) was added to inhibit *E. coli* RNA polymerase and 1 mM K_2HPO_4 to inhibit polynucleotide phosphorylase.

Purification of T7 RNA Polymerase. The following procedure produced about 15 mg of essentially pure T7 RNA polymerase. A 100-ml culture of HMS174/pAR1219 was grown in a shaking flask at 37°C in M9TB medium (9) containing 20 μg of ampicillin per ml. When the culture reached an optical density of 0.6 at 600 nm, the *lac* UV5 promoter was induced. After an additional 4 hr of growth, the cells were harvested by centrifugation, yielding about 1 g. All buffers used in the purification contained 20 mM Tris·HCl at pH 8.0, 1 mM dithiothreitol, 1 mM EDTA, and 5% (vol/vol) glycerol; only the additional components are named. All operations were carried out at 0–4°C; centrifugations were done in a Sorvall SS-34 rotor.

The cells were lysed by suspending in 27 ml of 22 mM NH_4Cl, adding 9 mg of egg white lysozyme in 3 ml of the same buffer, and freezing and thawing three times. The viscosity was reduced by passing the extract five times through a 20-gauge needle. This lysate was centrifuged 20 min at 18,000 rpm and the pellet was extracted with another 10 ml of buffer. To the combined supernatants was added 1.6 ml of 5% streptomycin sulfate (a final concentration of 0.2%), and the mixture was stirred on ice for 10 min to allow precipitation. The precipitate was collected by centrifuging 10 min at 15,000 rpm and extracted twice with 15 ml of 0.2 M NH_4Cl. To this extract was added 2 vol of saturated ammonium sulfate, which precipitated the T7 RNA polymerase.

The ammonium sulfate precipitate was dissolved in 20 ml of buffer and loaded onto a 15-ml column of Affi-Gel blue (Bio-Rad) that had been equilibrated with 0.2 M NH_4Cl. The column was washed with 0.4 M NH_4Cl/0.1 M sodium phosphate, pH 8.0, and the protein was eluted with the same buffer containing 1.5 M NH_4Cl. The peak fractions were dialyzed against 50 mM NH_4Cl, loaded on a 10-ml column of DEAE-cellulose (Whatman DE52), washed with the same buffer, and eluted with a 100-ml gradient from 50 to 500 mM NH_4Cl. These peak fractions were diluted with buffer to a conductivity corresponding to 0.1 M NH_4Cl, loaded on a 1-ml column of GTP-agarose (agarose-hexane-GTP, type 4, P-L Biochemicals), washed with 0.1 M NH_4Cl, and eluted with 0.2 M NH_4Cl. The resulting enzyme was both very pure and highly active (see Fig. 3).

RESULTS

Cloning the Entire Coding Sequence for T7 RNA Polymerase. Our strategy was to try cloning gene *1* under conditions in which its expression would be minimal, in case T7 RNA polymerase should be lethal. We had been successful in cloning many T7 genes in the *Bam*HI site of pBR322, even genes for enzymes that might be expected to be lethal if expressed to any significant extent (ref. 3 and unpublished data). Clones of potentially lethal genes were obtained in one orientation only (referred to as the silent orientation), where transcription in the counterclockwise direction would be needed to produce mRNA. Little if any transcription has been found to cross the *Bam*HI site of pBR322 in this direction (17–19). Therefore, we tried cloning gene *1* in this site.

We also expected that any promoters in T7 DNA itself that could direct transcription of the cloned gene would have to be removed. A minor promoter for *E. coli* RNA polymerase (the C promoter) lies just ahead of gene *1* in T7 DNA (20), and a promoter for T7 RNA polymerase itself lies immediately past the end of the coding sequence (Fig. 1). The presence of the latter promoter is potentially catastrophic: T7 RNA polymerase, by transcribing completely around pBR322 DNA (3, 4), would be able to direct the synthesis of its own mRNA from a cloned fragment that contained both this promoter and intact gene *1*. This would lead to an autocatalytic increase both in the level of T7 RNA polymerase and in the rate of transcription of the plasmid, which almost certainly would be lethal to the cell. Potentially, a single molecule of active T7 RNA polymerase would be sufficient to trigger this response, so such a construction would be stable only if there were absolutely no expression of the cloned gene, a situation that may be difficult or impossible to achieve.

Search of the nucleotide sequence revealed no restriction sites that would permit convenient isolation of a fragment containing all of the gene *1* coding sequence but not the promoters. We therefore isolated an appropriate fragment from heteroduplexes formed between the DNAs of different deletion mutants of T7 (Fig. 1). The strategy was similar to that employed by Stahl and Zinn (5), except that we prepared heteroduplexes from overlapping double-deletion mutants in order to provide an optimal substrate for the single-strand-specific endonuclease S1.

If S1 nuclease were to cut at the last unpaired nucleotides in the heteroduplex between the double-deletion DNAs D159,LG4 and C74,LG26, the gene *1* fragment would extend from nucleotide 3128 to 5845. The AUG initiation codon for T7 RNA polymerase begins at nucleotide 3171 and the UAA termination codon ends at nucleotide 5822, so this fragment would contain the entire coding sequence. The fragment would begin 11 base pairs ahead of the 5′ end of the natural gene *1* mRNA (6) and would therefore contain the signals needed to initiate protein synthesis, but it would not include

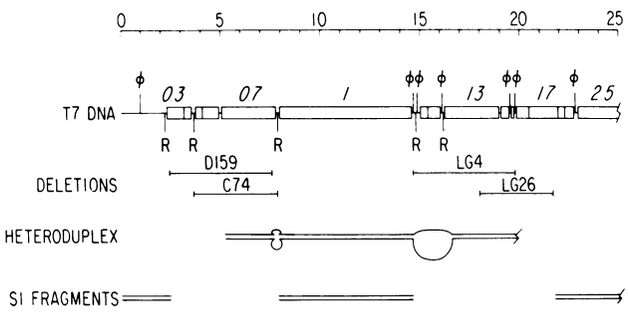

FIG. 1. Isolation of a gene *1*-containing fragment from T7 DNA. The left 25% of T7 DNA is represented (6): coding sequences for the T7 proteins in this region are shown as open boxes at their positions in T7 DNA, with gene numbers indicated for some; the sites of RNase III cleavages of T7 mRNAs are identified by R; promoters for T7 RNA polymerase are identified by ϕ. The locations of the deletions used to make the double-deletion strains are shown. The heteroduplex between D159,LG4 and C74,LG26 is represented by double lines to indicate double-stranded regions and single lines to indicate unpaired single-stranded regions, approximately to the same scale as the map. The double-stranded fragments expected after S1 nuclease digestion of the heteroduplex are placed according to their positions in the DNA.

the minor promoter for *E. coli* RNA polymerase, which initiates chains at nucleotide 3113. The conserved sequence of the promoter for T7 RNA polymerase that lies immediately to the right of the coding sequence covers nucleotides 5831 to 5853, so the fragment would contain only the first 15 of the 23 base pairs in the conserved sequence. It is not known whether this would be sufficent to allow promoter function if the fragment were ligated into the plasmid.

Because of the possibility that a functional promoter for T7 RNA polymerase might remain in the fragment produced by S1 nuclease digestion of the heteroduplex DNA, we also digested the fragments very lightly with BAL-31 exonuclease, to remove a few nucleotides from the ends of the fragment. The fragments generated by S1 alone, or by S1 followed by BAL-31 treatment, were purified by gel electrophoresis and adapted with *Bam*HI linkers for insertion into the *Bam*HI site of pBR322. The ligation mixture was used to transform *E. coli* HMS174 to ampicillin resistance, and the transformants were enriched for tetracycline-sensitive clones by treatment with cycloserine.

When individual transformants from several different ligation mixtures were tested by gel electrophoresis of the plasmid DNAs, very few inserts were found, and none contained a complete gene *1*. To enrich for plasmids containing an intact gene, plasmid DNA prepared from a mixed population of transformants obtained after cycloserine enrichment was subjected to gel electrophoresis, and DNA was recovered from the region of the gel where plasmids containing complete gene *1* would be expected to migrate. This DNA was then used to transform *E. coli* HMS174, and 48 new tetracycline-sensitive clones were isolated and analyzed for inserts. Four of these 48 plasmid DNAs appeared to carry an insert of the proper size to contain all of gene *1*. Gel electrophoresis of the fragments produced by cutting these plasmid DNAs with *Hin*cII and with *Kpn* I produced identical patterns from all four clones, which were essentially the patterns expected if the entire gene had been cloned in the silent orientation. The fragments cloned in this experiment had been treated with BAL-31.

One of these four plasmids, pAR1151 (Fig. 2), was chosen for further examination. To determine exactly where the cloned fragment begins and ends in the nucleotide sequence of T7 DNA, the fragment was released from the plasmid by cutting with *Bam*HI and the nucleotide sequence at each end was determined by the techniques of Maxam and Gilbert (21). The results showed that, after accounting for the sequence of the *Bam*HI linkers, the fragment that was cloned in pAR1151 begins at nucleotide 3146 of T7 DNA and ends at nucleotide 5840. Because the last nucleotide of the linker at the left end of the fragment happens to be the same as nucleotide 3145 of T7 DNA and the first nucleotide of the linker at the right end the same as nuceotide 5841, the actual limits of T7 DNA sequence are 3145–5841, that is, position 7.88–14.63 in T7 DNA. Therefore, the cloned fragment contains slightly more than the entire coding sequence for T7 RNA polymerase.

Compared with the minimum fragment expected from S1 nuclease treatment of the heteroduplex DNA, the cloned fragment of pAR1151 has lost 19 base pairs from the left end and 5 from the right end, presumably because of the BAL-31 treatment. The first six nucleotides of the natural mRNA are missing, but the fragment does contain 26 base pairs ahead of the AUG initiation codon and therefore has the entire predicted ribosome-binding and initiation region. The last 46 nucleotides of the natural mRNA are also missing, but the fragment contains 19 base pairs beyond the end of the coding sequence.

The cloned fragment in pAR1151 contains only the first 11 of the 23 highly conserved base pairs of the promoter for T7 RNA polymerase that is located just past the end of the coding sequence; the fragment ends 7 base pairs ahead of the natural start site for this promoter. It therefore seemed quite unlikely that pAR1151 would have any promoter activity for T7 RNA polymerase. Indeed, no such promoter activity was found in pAR1151 DNA, testing either with purified T7 RNA polymerase or during T7 infection [the assay described by McAllister *et al.* (4)].

The Cloned Gene Produces Active T7 RNA Polymerase. When a cloned fragment of T7 DNA can express a functional gene, T7 amber mutants defective in that gene are usually able to plate with high efficiency on a restrictive host that carries the cloned fragment (2, 3). In the case of gene *1* amber mutants, active T7 RNA polymerase is absolutely required to produce any increase in plating efficiency above the reversion frequency. We expected that little if any active T7 RNA polymerase would be produced from pAR1151, because the cloned gene is in the silent orientation in the *Bam*HI site of pBR322, where little gene *1* mRNA should be produced. However, we found that gene *1* amber mutants form plaques on HMS174/pAR1151 with the same efficiency as on a host that carries an amber suppressor, although the plaques are somewhat variable in size. Clearly, active T7 RNA polymerase can be made in these cells. The variability of plaque size suggests that the level of active enzyme in the cell may be rather low, but high enough that plaques are eventually produced from virtually every infection.

The gene *1* mRNA that is produced from pAR1151 may be

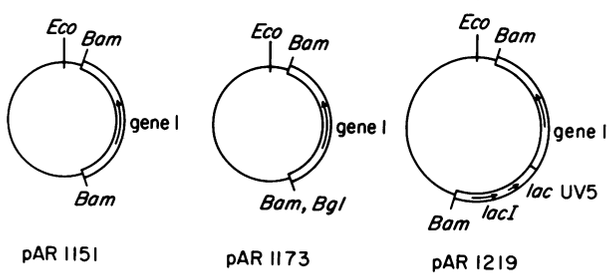

FIG. 2. Plasmids that contain T7 gene *1*. DNA lengths are shown approximately to scale: the single line represents pBR322 DNA; the double line represents the DNA inserted in the *Bam*HI site of pBR322. The *Eco*RI site of pBR322 is at the top of each representation, and the locations of all *Bam*HI and *Bgl* II sites are indicated. The arrows indicate the direction of transcription that produces mRNA for gene *1* and *lacI* and the direction of transcription from the *lac* UV5 promoter.

synthesized by readthrough of termination signals or by initiation at relatively weak promoters in the plasmid (17–19). Alternatively, a new promoter could have been generated in the joining of pBR322 and T7 DNA sequences, although this seems unlikely from comparing the junction sequence with sequences of known promoters for *E. coli* RNA polymerase (22). The gene *1* mRNA produced from this clone is almost certainly embedded in additional RNA sequences at one or both ends, but the information needed for synthesis of T7 RNA polymerase is clearly present and can be recognized. The level of transcription needed to allow gene *1* mutants to form plaques may in fact be very low, particularly if the gene *1* mRNA is relatively stable and able to compete efficiently for translation, and if only small amounts of active enzyme are needed.

Expressing Cloned T7 RNA Polymerase at High Level. To make large amounts of T7 RNA polymerase from the cloned gene, it was necessary to increase the synthesis of the gene *1* mRNA. To facilitate construction of plasmids in which promoters could be placed ahead of gene *1*, we first made a derivative of pAR1151 in which a *Bgl* II site was inserted into the *Bam*HI site ahead of gene *1*. There are no *Bgl* II sites in gene *1* or pBR322, so the inserted *Bgl* II site provides a unique cloning site immediately ahead of gene *1*. The modification was made by partial digestion of pAR1151 with *Bam*HI, filling in the ends with DNA polymerase, attaching synthetic *Bgl* II linkers, digesting with *Bgl* II, and religating the exposed *Bgl* II sites. Because of the sequence of the linkers, a *Bam*HI site should be regenerated to each side of the *Bgl* II site. As it happens, the modified plasmid we selected, pAR1173 (Fig. 2), does not contain a *Bam*HI site between the *Bgl* II site and gene *1*, presumably due to some imperfection in the cutting and ligation reactions. This is a convenient result because sequences that are inserted in the *Bgl* II site will remain with gene *1* as a single fragment upon digestion with *Bam*HI, allowing promoter–gene *1* constructs made in pAR1173 to be moved easily to other vectors.

The promoter we placed ahead of gene *1* in pAR1173 is the inducible *lac* UV5 promoter, which remains sensitive to the *lac* repressor but is no longer subject to catabolite repression (23). We cloned a fragment that contains the *lacI* gene as well as the *lac* UV5 promoter, so that we would not have to rely on the chromosomal gene of the host cell to provide *lac* repressor; we also hoped that multiple copies of *lacI* would supply enough *lac* repressor to keep the promoter reasonably well shut off in the multicopy plasmid (24).

The plasmid pMC1 (13) was used as the source of *lac* UV5 promoter DNA. The 1724-base-pair fragment that contains the *lacI* gene, the *lac* UV5 promoter region, and the beginning of the *lacZ* gene (13) was isolated from pMC1 by partial digestion with *Hinc*II and inserted into the *Bam*HI site of pBR322 by using synthetic linkers. The *Bam*HI fragment isolated from this clone was then ligated into the *Bgl* II site of pAR1173 ahead of gene *1*, taking advantage of the identity of the 4-base extensions produced by *Bgl* II and *Bam*HI. A resulting plasmid in which the *lac* UV5 promoter is directed toward gene *1* is pAR1219 (Fig. 2).

When grown in the absence of inducer, HMS174/pAR1219 complements gene *1* amber mutants. In fact, pAR1219 seems to give larger and more uniform plaques than pAR1151 or pAR1173, suggesting that the repressed *lac* UV 5 promoter may produce a somewhat higher level of gene *1* mRNA than the background level produced in the absence of this promoter.

When cultures of HMS174/pAR1219 are induced, T7 RNA polymerase is produced at a rapid rate and can accumulate to levels where it is the major protein of the cell (Fig. 3). Enzyme activity, as assayed in crude extracts, increases along with accumulation of the protein. The rate of increase of turbidity of the culture is almost as high for the induced culture as for a parallel uninduced culture, and the number of cells that can form colonies when plated in the absence of inducer does not decrease for several hours, suggesting that T7 RNA polymerase may not be very toxic to the cells.

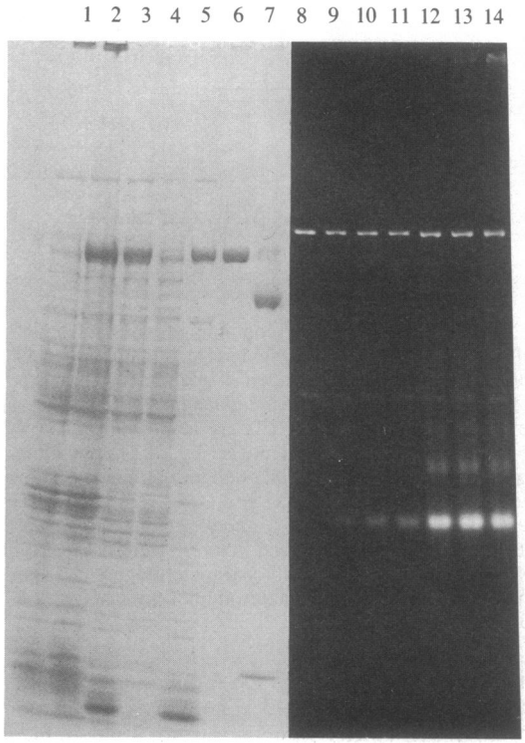

FIG. 3. Purification of T7 RNA polymerase and transcription by the purified enzyme. Proteins (lanes 1–7) were analyzed by electrophoresis in a discontinuous buffer system containing sodium dodecyl sulfate on a 10–20% polyacrylamide gradient gel (25), followed by staining with Coomassie blue. Nucleic acids (lanes 8–14) were analyzed by electrophoresis in 40 mM Tris/20 mM acetic acid/2 mM Na_3EDTA plus ethidium bromide (1 μg/ml) on a 1% agarose gel and were visualized by fluorescence. *Lanes 1–6* show the proteins observed upon induction and purification of T7 RNA polymerase. Lane 1 contained 10 μl of an uninduced culture of HMS174/pAR1219, lane 2 contained 10 μl of a parallel induced culture, and lanes 3–5 contained equivalent amounts of crude extract, supernatant from the streptomycin precipitation, and extracted streptomycin precipitate from the induced culture. Lane 6 contained about 5 μg of purified protein. Lane 7 shows an example of a preparation in which most of the protein had suffered nicks, as discussed in the text. *Lanes 8–14* show RNA synthesis by the purified T7 RNA polymerase. The template was DNA of the recombinant plasmid pAR436, which has a fragment of T7 DNA containing the *ϕ10* promoter for T7 RNA polymerase cloned in the *Bam*HI site of pBR322 (3). The plasmid DNA had been cut with *Bam*HI to separate the *ϕ10*-containing fragment from the pBR322 DNA (lane 8). Of this mixture of DNA fragments, 200 ng was transcribed in a volume of 20 μl at 37°C. Lanes 9–11 show 2.5 μl of reaction mixture after 15-, 30-, or 60-min incubation with 0.25 μg of purified polymerase, and lanes 12–14 the same times with 2.5 μg of polymerase.

If moderate to high levels of T7 RNA polymerase are tolerated quite well by *E. coli*, it may not have been necessary to remove the minor C promoter for *E. coli* RNA polymerase from ahead of the coding sequence to clone an active gene *1* fragment. In fact, it may be possible to place a constitutive promoter of appropriate strength ahead of gene *1* in pAR1173 to provide a convenient source of enzyme without induction.

Even though the cell may be able to tolerate relatively high levels of T7 RNA polymerase, it seems unlikely that a fragment containing both the translatable coding sequence and the promoter for T7 RNA polymerase could have been cloned: the autocatalytic response would require too many resources of the cell. Attempts to place a promoter for T7

RNA polymerase into a plasmid along with the gene *1* fragment have produced only arrangements in which the promoter directs transcription opposite to the direction needed to synthesize gene *1* mRNA.

Purification of T7 RNA Polymerase from Induced Cultures. Cultures of HMS174/pAR1219 that had been induced for 4 hr have been used as a source of T7 RNA polymerase. Perhaps 10–20% or more of the total protein in these cells is T7 RNA polymerase (Fig. 3, lane 2), and virtually all of it seems to be soluble and active. Typically, 10–15 mg of pure T7 RNA polymerase can be obtained from 100 ml of induced culture.

Initially, a problem with protease activity was encountered in purification of the enzyme. Electrophoresis of different preparations of purified protein on polyacrylamide gels in the presence of sodium dodecyl sulfate revealed that a variable fraction of the molecules had sustained one or more cuts, which released specific fragments upon electrophoresis. These fragments were not resolved from intact protein by gel filtration of the native enzyme, so it seems likely that the cleavage products remain associated under nondenaturing conditions. The most prevalent cut produced a large and a small fragment, although it is possible that other cuts may also be made. An example of a preparation in which most of the molecules had sustained this cut is shown in Fig. 3, lane 7. Nicked enzymes can be partly resolved from intact ones by chromatography on phosphocellulose or heparin agarose, and analysis of the fractionated protein indicates that the nicked enzymes retain substantial polymerase activity.

The gene *1* protein seemed intact in crude extracts, and the nicking became apparent only as the protein was fractionated from the nucleic acids. We reasoned that an association with DNA in the crude extract might protect the polymerase from a protease, and that removal of the protease before removal of the DNA might allow the purification of intact polymerase. The purification procedure described in *Materials and Methods* was designed to do this. As can be seen from the gel pattern (Fig. 3, lane 6), the protein prepared in this way had very little nicked material.

The purified enzyme is very active and specific in transcribing DNA that has a promoter for T7 RNA polymerase, and large amounts of RNA can be produced. This is demonstrated in Fig. 3, where two different concentrations of purified enzyme were used to transcribe a mixture of two DNA fragments, one of which contained a promoter for T7 RNA polymerase and the other of which did not. At the higher polymerase concentration, virtually all of the precursors were incorporated into RNA in 15 min (lanes 12–14); at the lower concentration, incorporation continued for at least an hour (lanes 9–11). Almost all of the RNA was the size expected if it had started at the promoter for T7 RNA polymerase and ended at the end of the DNA fragment, indicating that the transcription was specific and the enzyme was free of RNase activity. The DNA fragments remained intact, indicating that the enzyme was also free of DNase activity.

DISCUSSION

The cloned gene for T7 RNA polymerase provides a convenient and ample source of this enzyme for biochemical and physical studies. The predicted amino acid sequence of the protein and the nucleotide sequences of its specific promoters are both known (6, 7) and should facilitate analysis of how this enzyme interacts so specifically with a relatively large promoter sequence. As one of the simplest RNA polymerases known, T7 RNA polymerase is a prime candidate for x-ray crystallographic studies, both alone and in association with its promoter sequence.

Accumulation of specific nicks in T7 RNA polymerase during purification has not been reported previously, nor observed by us, in enzyme purified from infected cells. This suggests the possibility that the protease responsible for these nicks may be inactivated during T7 infection. If so, it should be possible to identify the gene responsible for this inactivation. The apparent specificity of the nicks suggests that the T7 RNA polymerase may be folded into specific domains that leave protease-sensitive sites in limited regions. It will be interesting to determine the precise locations of these nicks and to probe the molecule with other proteases.

The ability to produce active enzyme from the cloned gene now makes it possible to map functional regions of the enzyme by random and specific mutational analysis. We have isolated T7 deletion mutants that completely lack the RNA polymerase gene, propagating them on hosts that express active polymerase from the cloned gene. These mutants, together with clones that express altered polymerases, should make it possible to analyze in great detail the functions of T7 RNA polymerase during infection. T7 RNA polymerase appears to be involved in replication as well as transcription of T7 DNA, and mutational analysis of the cloned gene may now make it possible to define these separate functions.

As outlined in the introduction, the stringent specificity of T7 RNA polymerase for a relatively large promoter sequence makes this enzyme very useful for producing specific RNAs. An example of the high level of specific transcription that can be generated by the purified enzyme is shown in Fig. 3. T7 RNA polymerase from the cloned gene should equally well be able to direct high levels of transcription from appropriate promoters inside the cell.

We thank Michele Calos for the kind gift of the plasmid pMC1. This research was carried out at Brookhaven National Laboratory under the auspices of the United States Department of Energy.

1. Studier, F. W. & Dunn, J. J. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 999–1007.
2. Campbell, J. L., Richardson, C. C. & Studier, F. W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2276–2280.
3. Studier, F. W. & Rosenberg, A. H. (1981) *J. Mol. Biol.* **153**, 503–525.
4. McAllister, W. T., Morris, C., Rosenberg, A. H. & Studier, F. W. (1981) *J. Mol. Biol.* **153**, 527–544.
5. Stahl, S. J. & Zinn, K. (1981) *J. Mol. Biol.* **148**, 481–485.
6. Dunn, J. J. & Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477–535.
7. Moffatt, B. A., Dunn, J. J. & Studier, F. W. (1984) *J. Mol. Biol.* **173**, 265–269.
8. Studier, F. W., Rosenberg, A. H., Simon, M. N. & Dunn, J. J. (1979) *J. Mol. Biol.* **135**, 917–937.
9. Studier, F. W. (1969) *Virology* **39**, 562–574.
10. Studier, F. W. (1973) *J. Mol. Biol.* **79**, 227–236.
11. Studier, F. W. (1975) *J. Mol. Biol.* **94**, 283–295.
12. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113.
13. Calos, M. P. (1978) *Nature (London)* **274**, 762–765.
14. Wu, R., ed. (1979) *Methods in Enzymology* (Academic, New York), Vol. 68.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
16. Carter, A. D., Morris, C. E. & McAllister, W. T. (1981) *J. Virol.* **37**, 636–642.
17. Stüber, D. & Bujard, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 167–171.
18. West, R. W., Jr., & Rodriguez, R. L. (1980) *Gene* **9**, 175–193.
19. West, R. W., Jr., & Rodriguez, R. L. (1982) *Gene* **20**, 291–304.
20. McConnell, D. J. (1979) *Nucleic Acids Res.* **6**, 525–544.
21. Maxam, A. M. & Gilbert, W. (1979) *Methods Enzymol.* **65**, 499–560.
22. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11**, 2237–2255.
23. Silverstone, A. E., Arditti, R. R. & Magasanik, B. (1970) *Proc. Natl. Acad. Sci. USA* **66**, 773–779.
24. Backman, K., Ptashne, M. & Gilbert, W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4174–4178.
25. Studier, F. W. (1973) *J. Mol. Biol.* **79**, 237–248.