Purification, Characterization, and Reconstitution of DNA-dependent RNA Polymerases from *Caulobacter crescentus*

(Received for publication, March 17, 1997, and in revised form, June 4, 1997)

Jianguo Wu, Noriko Ohta, Andrew K. Benson, Alexander J. Ninfa, and Austin Newton

From the Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544

Cell differentiation in the *Caulobacter crescentus* cell cycle requires differential gene expression that is regulated primarily at the transcriptional level. Until now, however, a defined *in vitro* transcription system for the biochemical study of developmentally regulated transcription factors had not been available in this bacterium. We report here the purification of *C. crescentus* RNA polymerase holoenzymes and resolution of the core RNA polymerase from holoenzymes by chromatography on single-stranded DNA cellulose. The three RNA polymerase holoenzymes Eσ54, Eσ73, and Eσ70 were reconstituted exclusively from purified *C. crescentus* core and sigma factors. Reconstituted Eσ73 initiated transcription from the σ54-dependent fljK promoter of *C. crescentus* in the presence of the transcription activator FlbD, and active Eσ73 specifically initiated transcription from the σ70-dependent promoter of the *C. crescentus* heat-shock gene dnaK. For reconstitution of the Eσ73 holoenzyme, we overexpressed the *C. crescentus* rpoD gene in *Escherichia coli* and purified the full-length σ54 protein. The reconstituted Eσ73 recognized the σ70-dependent promoters of the *E. coli* lacUV5 and neo genes, as well as the σ70-dependent housekeeping promoters of the *C. crescentus* pleC and rsaA genes. The ability of the *C. crescentus* Eσ73 RNA polymerase to recognize *E. coli* σ70-dependent promoters is consistent with relaxed promoter specificity of this holoenzyme previously observed *in vivo*.

*Caulobacter crescentus* is a dimorphic, Gram-negative bacterium with a well defined cell cycle that generates two different daughter cells, a motile swarmer cell with a single polar flagellum and a nonmotile stalked cell. Formation of the new swarmer cell and its subsequent differentiation into a stalked cell result from a series of discrete morphogenetic events, including flagellar biosynthesis, flagellum rotation, loss of motility, and stalk formation at one pole during the cell cycle (reviewed in Refs. 1 and 2). Early experiments demonstrated that this sequence of developmental events depends on de novo RNA synthesis (3) and suggested that differential gene transcription, presumably involving RNA polymerase and its accessory proteins, plays a central role in regulating the developmental program.

Bacterial RNA polymerases (RNAP) are multi-subunit enzyme complexes that can be purified as the core polymerase (E) and the holoenzyme (Eσ, reviewed in Refs. 4 and 5). The core RNAP, composed of the α, β, and β′ subunits, carries out RNA chain elongation, whereas the holoenzyme, which also contains the sigma subunit (σ), recognizes specific promoter sequences. Multiple sigma factors with unique promoter specificities have been identified in many eubacteria, and the use of alternative sigma subunits is a fundamental mechanism for reprogramming RNAP specificity and controlling complex patterns of gene transcription (reviewed in Refs. 6–8).

The most extensive study of transcription regulation in *C. crescentus* has been carried out on the genes in the flagellar gene hierarchy (reviewed in Ref. 9). Flagellum formation requires the temporally controlled transcription of approximately 50 genes (10) that are organized in a regulatory hierarchy containing four classes of genes (I to IV). The Class II genes, which are expressed early in the cell cycle and encode basal body and switch components, contain a unique promoter consensus. Recent results have shown that transcription from the Class II promoters is regulated *in vivo* by the response regulator CtrA (11). The Class II gene products are required, in turn, for transcription of class III and class IV genes that are transcribed from σ54-dependent promoters late in the cell cycle. The σ54 factor and transcription activator FlbD, which are encoded by Class II genes rpoN and flbD, are required for transcription of the Class III and IV genes (reviewed in Ref. 9).

Biochemical studies of transcription *in vitro* have employed either the Eσ54 holoenzyme reconstituted from purified *Escherichia coli* components (12, 13), the heterologous Eσ54 holoenzyme reconstituted from the *C. crescentus* Eσ54 and the *E. coli* core RNAP (14), or a partially purified *C. crescentus* RNAP (15). The genes encoding three *C. crescentus* sigma factor subunits, σ54 (rpoH; Refs. 16 and 17), σ70 (rpoN; Refs. 14 and 18), and σ73 (rpoD; Ref. 19) have been cloned and sequenced. Although sigma factors σ54 (14) and σ32 (16) have now been overexpressed in *E. coli* and purified, the lack of a purified core RNAP has prevented the reconstitution of a transcription system exclusively from purified *C. crescentus* components. In addition, the principal *C. crescentus* sigma factor, which is required for transcription from the housekeeping promoters (19) and predicted to contain 653 amino acids with a molecular mass of 72,623 Da (σ73, 20), had not been isolated.

We report here the first purification and characterization of the *C. crescentus* core RNAP, as well as the two holoenzymes Eσ73 and Eσ32. We also describe the purification of the *C. crescentus* principal sigma factor, σ73, after overexpression of rpoD in *E. coli*. The Eσ73 Eσ54, and Eσ32 holoenzymes have been reconstituted exclusively from purified *C. crescentus* pro-

---

* This work was supported in part by Public Health Service Grant GM22299 from the National Institutes of Health (to A.N.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

 ‡ Present address: Dept. of Food Science and Technology, University of Nebraska, Lincoln, NE 68583-0919.

 ¶ To whom correspondence should be addressed. Tel.: 609-258-3854; Fax: 609-258-6175; E-mail: anewton@molecular.princeton.edu.

1 The abbreviations used are: RNAP, RNA polymerase; E, core polymerase; Eσ, holoenzyme; nt, nucleotide; bp, base pair; PAGE, polyacylamide gel electrophoresis.
teins, and the transcriptional specificity of these RNAP preparations has been examined. The availability of a defined, reconstituted transcription system will allow detailed analysis of the roles of RNAP and accessory factors in the transcriptional regulation of developmental genes during cell differentiation and division in this bacterium.

**MATERIALS AND METHODS**

**Bacterial Strains, Media, and Materials—**E. coli strain DH5α was used for propagating plasmids and cultured in ML medium supplemented with ampicillin (100 μg/ml) or tetracycline (10 μg/ml) as necessary. *C. crescentus* wild-type strain CB15 (ATCC19809) was used for the purification of RNAP and grown in PYE (peptone yeast extract; Ref. 21) medium at the fermentation facility of the Waksman Institute. Restriction enzymes were purchased from either New England Biolabs or Boehringer Mannheim. T4 DNA ligase and T4 DNA polymerase were obtained from Boehringer Mannheim. In−32P]UTP was obtained from Amersham Corp. Oligonucleotides were synthesized by the Princeton University SynSeq facility. Heparin-agarose, single-stranded DNA-cellulose, and DEAE-cellulose were purchased from Bio-Rad. Poly(dA-T) was purchased from Sigma. E. coli core RNAP was purchased from Epicentre Technologies.

**RNA Polymerase Purification—**Most procedures of the purification are described in the methods described by Berg et al. (27). All steps were carried out at 4 °C unless noted otherwise. A batch of 100 g of frozen *C. crescentus* cells were broken into small pieces and placed in a 1-liter Waring Blender with 300 ml of grinding buffer (0.05 M Tris-HCl (pH 7.9), 5% (v/v) glycerol, 2 mM EDTA, 0.1 mM dithiothreitol, 1 mM β-mercaptoethanol, 0.253 mM NaCl, 23 μg/ml phenylmethylsulfonyl fluoride, and 130 μg/ml lysozyme). The cells were blended to allow lysis and shearing of the DNA. The sample was diluted with 500 ml of TGED (0.02% Triton-HCl (pH 7.9), 5% (v/v) glycerol, 0.1 mM EDTA, and 0.1 mM dithiothreitol) + 0.2% NaCl, blended, and then centrifuged for 30–40 min at 7000 rpm. The supernatant was collected as crude extract.

Crude extract was precipitated with Polymun P at a final concentration of 0.3% and centrifuged at 7000 rpm in Sorvall to collect the pellet. Proteins were eluted from the pellet with TGED + 1% NaCl after washing the pellet once with TGED + 0.2% NaCl. The Polymun P extract was then precipitated with ammonium sulfate at 50% saturation. The pellet obtained after centrifugation was resuspended in TGED and dialyzed twice against TGED + 50 mM NaCl.

Dialyzed sample was subject to column chromatography as follows. It was first applied to a heparin-agarose column pre-equilibrated with TGED + 50 mM NaCl and eluted with a NaCl gradient of 50–600 mM in TGED. Fractions were examined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and peak fractions containing RNA polymerase (β and α subunits), as visualized by Coomassie Blue staining (Fig. 1A), pooled and dialyzed as above. Dialyzed peak fractions were further purified on a DEAE-cellulose column by fractionating with a linear NaCl gradient of 50–600 mM. After dialysis, the pooled peak fractions (Fig. 1B) were finally fractionated on a single-stranded DNA-cellulose column also with a linear NaCl gradient of 50–600 mM.

Two peak fractions, peak 1 and peak 2 (Fig. 1C), were pooled separately and dialyzed against storage buffer (TGED + 0.05% NaCl with 50% (v/v) glycerol). The dialyzed enzyme preparations were aliquoted and stored at −80 °C for future use.

**Isolation and Purification of Sigma Factors from RNA Polymerase holoenzymes—**Samples of RNAP holoenzyme were applied to SDS-PAGE gels (10% acrylamide), and after staining with KCl the gel was sectioned into 8 slices corresponding to the positions of visible proteins. Proteins were eluted from the gel slices, and SDS was removed by acetone precipitation, dissolved in 6 M guanidine HCl, and renatured, as described by Hager and Burgess (22).

**Construction of DNA Templates Used for in Vitro Transcription—**Plasmid pNEO containing the promoter of neomycin phosphotransferase neo gene was constructed by cloning the *HindIII*/BglII fragment of transposon Tn5 (24, 25) in pUC18 restricted with *HindIII* and *BamHI*. The linear fragment obtained by restricting with *HindIII* and EcoRI was used as a template in the in vitro transcription experiment, and the *BamHI* fragment of the expected transcript is 84 nucleotides (Fig. 2). The fragment containing the lacUV5 promoter-operator region of *E. coli lac* gene was cloned from the expression vector pVINIIA2 (26) as a 305-bp *HinII*/BamHI fragment. The DNA fragment from plasmid pVINIIA2 was subcloned into pUC18 to yield plasmid pLACUV5. The DNA template was obtained from pLACUV5 with *HindIII* and *BamHI* that should produce a 79-nt transcript (Fig. 2, 26). Plasmid pAACK8 was constructed as described previously (13). The fliK transcription start site lies 451 bp upstream of the T7 terminator site (Fig. 2). Supercopied plasmid DNA was used in the one-cycle transcription assay to determine the α43 activity (13). Plasmid pJW012 containing the *dnaK* α43-dependent promoter has been described previously (18). The SacI/Sall DNA fragment used as template in the run-off transcription assay to test the α43 activity. Plasmid pSSA41 was a gift of John Smit. The *HindIII*/ClaI DNA fragment containing the rsaA gene promoter was inserted into *HindIII*/*AccI* sites of pUC19. The *HindIII*/*NolI* fragment from this plasmid was used as template in the run-off transcription assay to determine the α73 activity. A 600-bp *HindIII*/BamHI fragment containing the *C. crescentus pleC* gene promoter was subcloned into *SalI*/*BamHI* and *BamHI* sites of *pBluescript* to yield plasmid pJW014. The 636-bp *HindIII*/SalI fragment from this plasmid was used as a DNA template in the run-off transcription assay to test the α73 activity.

In *Vitro Transcription Assays—*The formation of open complexes at different promoters from *C. crescentus* and *E. coli* was measured in single cycle or run-off transcription assays, as described previously (13, 16). The RNA polymerase holoenzymes were reconstituted by the addition of purified *C. crescentus* sigma factors to the purified *C. crescentus* core RNA polymerase (final concentration of 1 nmol) and incubated for 10 min at 4 °C. Transcripts obtained were fractionated on 7% urea-PAGE with end-labeled Sau3A fragments of pUC18 as DNA size markers and visualized by autoradiography.

**Assay for Core RNAP—**The core RNAP activity was determined using poly(dA-T) as template based on the method described by Berg et al. (27). Each reaction of a 100-μl volume contained 20 μl of assay solution (200 mM Tris-HCl (pH 8.0), 50 mM MgCl2, and 50 mM β-mercaptoethanol), 100 μM ATP, 10 μM UTP, and 5 μl of [α−32P]UTP. RNA synthesis was initiated by adding RNA polymerase and terminated after 10 min at 37 °C with 3 ml of ice-cold 3.5% perchloric acid containing 0.1 M sodium phosphate. The precipitates were collected on Whatman cellulose filter paper (3MM) and washed three times with cold 1 M HCl containing 0.1 M sodium phosphate and finally with cold ethanol. The radioactivity on the dry filter was determined by scintillation counting.

**Overexpression and Purification of C. crescentus α73 Protein from E. coli—**The cloned *ropD* gene was identified and cloned by colony hybridization to its *Myxococcus xanthus* homologue (28). DNA sequence analysis confirmed that the open reading frame in this recombinant plasmid was identical to that published for the *C. crescentus ropD* gene (20). A Smal-SalI DNA fragment from plasmid pGIR210, which contains the *ropD* gene, was subcloned into the mutagenesis vector *pAl-1* (Promega), and an *NdeI* restriction site was introduced at the first codon of the *ropD* gene open reading frame by site-directed mutagenesis. The 2.7-kilobase pairs DNA fragment was then subcloned into the *NdeI* and *HindIII* sites of the expression vector *pSET/A* to yield plasmid pJW41 in which the entire open reading frame of *ropD* gene is translationally fused to the first codon of the *T7* gene. *E. coli* strain BL21 (DE3) carrying the plasmid pJW41 was used to overproduce α73 protein in the presence of isopropyl-β-D-thiogalactoside.

The overexpressed α73 protein was purified by a previously described method (29, 30). The RpoD protein was not soluble and formed inclusion bodies that were solubilized with 6 M guanidine HCl in TGED buffer. The solubilized protein was renatured by dialysis against the TGED buffer and then further purified by chromatography on a DEAE-cellulose column. This method yielded α73 protein that was greater than 95% pure, as judged by Coomassie Blue staining of SDS-PAGE gels.

**RESULTS**

**Purification of RNA Polymersases—**Cellular RNAP from *C. crescentus* was purified by fractionation of cell extracts with Polymun P, ammonium sulfate precipitation, and chromatography on heparin-agarose (Fig. 1A) and DEAE-cellulose (Fig. 1B), which removed many of the contaminating proteins. Peak fractions from the DEAE-cellulose column containing the RNAP β, β′, and α subunits were pooled and applied to a single-stranded DNA-cellulose column (Fig. 1C). As shown in Fig. 1C, the majority of the RNAP β, β′, and α subunits eluted in two distinct peaks along with several minor proteins. The fractions in peaks 1 and 2 were pooled separately (“Materials and Methods”).

2 G. Ramakrishnan and A. Newton, unpublished observations.
Identification of E₇₃ and E₃₂ RNAP Holoenzymes—We assayed RNAP activity in peak 1 and peak 2 using E. coli DNA templates containing either the s₇₀-dependent neo promoter or the lacUV5 promoter (Fig. 2; “Materials and Methods”). Both of these promoters are recognized in vivo by C. crescentus (see below). A third DNA template (Fig. 2) contained the s₃₂-dependent, dnaK P₁ heat-shock promoter of C. crescentus (31). The RNAP preparation from peak 2 recognized all three promoters, and specific transcripts of the predicted sizes were obtained from each template (Fig. 3, lanes 2, 4, and 6). Thus the peak 2 preparation contained s₃₂ holoenzyme (E₃₂) and s₇₃ holoenzyme (E₇₃) activities. No detectable transcripts were observed when the RNAP preparation from peak 1 was assayed using the same DNA templates (Fig. 3, lanes 1, 3, and 5). These data suggest the possibility that peak 1 contained either inactive RNAP holoenzyme or only core RNAP.

Identification of C. crescentus Core RNAP—We assayed for RNAP activity directly in an in vitro transcription assay using poly[d(A-T)] as template as described by Berg et al. (27). The results summarized in Table I indicate that both peak 1 and peak 2 contained active core polymerase, although the specific activity was ~3-fold higher in the peak 2 RNAP. Interestingly, peak 2 RNAP was also more active on the poly[d(A-T)] template than purified E. coli core polymerase.

We next examined if the peak 1 RNAP core preparation could be reconstituted to give active holoenzyme. C. crescentus σ₅₄ protein was used in the initial reconstitution experiments because earlier work had demonstrated that Er₅₄ RNAP holoenzyme reconstituted from E. coli RNAP core and purified C. crescentus σ₅₄ specifically recognizes σ₅₄-dependent promoters from both E. coli and C. crescentus (14). The reconstituted Er₅₄ holoenzyme also required the activator protein FlbD for initiation of transcription from C. crescentus σ₅₄-dependent promoters, as had been observed both in vitro and in vivo (13).

The work described here demonstrates that in the presence of C. crescentus σ₅₄ and FlbD, the purified RNAP preparation from peak 1 specifically recognized the σ₅₄-dependent fljK promoter and produces a transcript of the expected size from this template (Fig. 2 and Fig. 4, lanes 2 and 4). This result confirms that peak 1 fractions contain an active RNAP core enzyme that can be used for assays of sigma factor activity. In the absence of added σ₅₄, however, RNAP in neither peak 1 nor peak 2 recognized the fljK promoter (Fig. 4, lanes 1 and 3), indicating that none of these fractions contained an active Er₅₄ holoenzyme. Consequently, we refer to the peak 1 pool as core RNAP. The peak 2 pool, which appears to contain the E₃₂ and E₇₃ holoenzymes (Fig. 3), as well as excess core RNAP (see below), we refer to as peak 2 RNAP.

Isolation and Identification of Sigma Factor Subunits from the RNAP Holoenzyme Preparation—Individual proteins in the peak 2 RNAP preparation (Fig. 1C; peak 2 pool) were isolated after electrophoresis on preparative SDS-PAGE gels as described under “Materials and Methods.” SDS-PAGE gels are shown of fractions eluted from a heparin-agarose column (A), a DEAE-cellulose column (B), and a single-stranded DNA-cellulose column (C). Gels were stained with Coomassie Blue. L, protein preparation loaded onto the column; W, column wash. The positions of core RNAP subunits are indicated as β, β’, and α.

Identification of E₅₄ and Er₅₄ RNAP Holoenzymes—We assayed RNAP activity in peak 1 and peak 2 using E. coli DNA templates containing either the σ₇₀-dependent neo promoter or the lacUV5 promoter (Fig. 2; “Materials and Methods”). Both of these promoters are recognized in vivo by C. crescentus (see below). A third DNA template (Fig. 2) contained the σ₃₂-dependent, dnaK P₁ heat-shock promoter of C. crescentus (31). The RNAP preparation from peak 2 recognized all three promoters, and specific transcripts of the predicted sizes were obtained from each template (Fig. 3, lanes 2, 4, and 6). Thus the peak 2 preparation contained σ₃₂ holoenzyme (Er₃₂) and σ₇₃ holoenzyme (Er₇₃) activities. No detectable transcripts were observed when the RNAP preparation from peak 1 was assayed using the same DNA templates (Fig. 3, lanes 1, 3, and 5). These data suggest the possibility that peak 1 contained either inactive RNAP holoenzyme or only core RNAP.

Identification of C. crescentus Core RNAP—We assayed for RNAP activity directly in an in vitro transcription assay using poly[d(A-T)] as template as described by Berg et al. (27). The results summarized in Table I indicate that both peak 1 and peak 2 contained active core polymerase, although the specific activity was ~3-fold higher in the peak 2 RNAP. Interestingly, peak 2 RNAP was also more active on the poly[d(A-T)] template than purified E. coli core polymerase.

We next examined if the peak 1 RNAP core preparation could be reconstituted to give active holoenzyme. C. crescentus σ₅₄ protein was used in the initial reconstitution experiments because earlier work had demonstrated that Er₅₄ RNAP holoenzyme reconstituted from E. coli RNAP core and purified C. crescentus σ₅₄ specifically recognizes σ₅₄-dependent promoters from both E. coli and C. crescentus (14). The reconstituted Er₅₄ holoenzyme also required the activator protein FlbD for initiation of transcription from C. crescentus σ₅₄-dependent promoters, as had been observed both in vitro and in vivo (13).

The work described here demonstrates that in the presence of C. crescentus σ₅₄ and FlbD, the purified RNAP preparation from peak 1 specifically recognized the σ₅₄-dependent fljK promoter and produces a transcript of the expected size from this template (Fig. 2 and Fig. 4, lanes 2 and 4). This result confirms that peak 1 fractions contain an active RNAP core enzyme that can be used for assays of sigma factor activity. In the absence of added σ₅₄, however, RNAP in neither peak 1 nor peak 2 recognized the fljK promoter (Fig. 4, lanes 1 and 3), indicating that none of these fractions contained an active Er₅₄ holoenzyme. Consequently, we refer to the peak 1 pool as core RNAP. The peak 2 pool, which appears to contain the Er₇₃ and Er₃₂ holoenzymes (Fig. 3), as well as excess core RNAP (see below), we refer to as peak 2 RNAP.
scribed by Hager and Burgess (23). The gel was sectioned into 8 slices corresponding to the positions of bands visualized with KCl, with slice 1 containing bands at the top of the gel. Proteins were eluted from the gel slices, and a portion of each eluted protein sample was then analyzed on a second SDS-PAGE gel (Fig. 5). The peak 2 RNAP preparation contained several proteins in addition to core subunits $\beta$ and $\beta'9$ isolated from slice 1 and $\alpha$ isolated from slice 4. Potential sigma factors were the 75-kDa protein in slice 2 and the 34-kDa protein in slice 5. Unidentified proteins A and B of molecular masses 55 and 50 kDa, respectively, were found in slice 3, and a third unknown protein C of 28 kDa was detected in slice 6.

A portion of the proteins eluted from the SDS-PAGE gel slices were renatured (see "Materials and Methods") and combined with $C.\text{crescentus}$ core RNAP (peak 1) to determine their ability to direct transcription. Assay of the 34-kDa protein renatured from gel slice 5 (Fig. 5) on the dnaK P1 template produced a run-off transcript of the expected size (62 nt; Fig. 2), which is consistent with the identification of this protein as the $\sigma^{32}$ factor (16). When renatured proteins from gel slices 1–8

A portion of the proteins eluted from the SDS-PAGE gel

slices were renatured (see "Materials and Methods") and combined with $C.\text{crescentus}$ core RNAP (peak 1) to determine their ability to direct transcription. Assay of the 34-kDa protein renatured from gel slice 5 (Fig. 5) on the dnaK P1 template produced a run-off transcript of the expected size (62 nt; Fig. 2), which is consistent with the identification of this protein as the $\sigma^{32}$ factor (16). When renatured proteins from gel slices 1–8

slices were renatured (see "Materials and Methods") and combined with $C.\text{crescentus}$ core RNAP (peak 1) to determine their ability to direct transcription. Assay of the 34-kDa protein renatured from gel slice 5 (Fig. 5) on the dnaK P1 template produced a run-off transcript of the expected size (62 nt; Fig. 2), which is consistent with the identification of this protein as the $\sigma^{32}$ factor (16). When renatured proteins from gel slices 1–8

slices were renatured (see "Materials and Methods") and combined with $C.\text{crescentus}$ core RNAP (peak 1) to determine their ability to direct transcription. Assay of the 34-kDa protein renatured from gel slice 5 (Fig. 5) on the dnaK P1 template produced a run-off transcript of the expected size (62 nt; Fig. 2), which is consistent with the identification of this protein as the $\sigma^{32}$ factor (16). When renatured proteins from gel slices 1–8
Isolation and Purification of the C. crescentus rpoD Gene Product—To further characterize the principal C. crescentus sigma factor identified in the reconstitution experiments (Fig. 6), we overexpressed the rpoD gene in E. coli and purified the full-length $\sigma^{37}$ protein to near-homogeneity (see “Materials and Methods”; Fig. 7A; lane 3). The C. crescentus rpoD gene has been shown to encode a predicted polypeptide of 653 amino acids with a molecular mass of 72,623 Da and designed as $\sigma^{37}$ (20).

The size of the overexpressed protein (Fig. 7A, lane 3) is similar to the very faint protein band at $\sim$75 kDa in the peak 2 RNAP preparation (Fig. 7A, lane 2) and close to the predicted 72,623-Da size of the rpoD gene product (20). The purified protein was also examined by Western blot analysis. An anti-E. coli $\sigma^{70}$ antibody cross-reacted with the major protein band at $\sim$75 kDa, as well as with several smaller bands that presumably result from proteolysis of RpoD (Fig. 7B, lane 3). The anti-$\sigma^{70}$ antibody can also recognize the protein at $\sim$75 kDa present in the peak 2 RNAP (Fig. 7B, lane 2) but failed to recognize any proteins in the peak 1 RNAP preparation (Fig. 7B, lane 1). These results further support our assignment of peak 1 as core enzyme and the 75-kDa protein present in the peak 2 RNAP as $\sigma^{73}$. The $\sigma^{73}$ present in the peak 2 RNAP (Fig. 7B, lane 2) displays a slightly different mobility from that overproduced from the C. crescentus rpoD gene in E. coli (Fig. 7B, lane 3), perhaps as a result of $\sigma^{70}$ modification in one of the bacteria.

Functional Analysis of C. crescentus $\sigma^{73}$ in Vitro—We examined the sigma factor activity of the purified rpoD gene product in reconstitution experiments with core RNAP using in vitro transcription assays (Fig. 8). Purified $\sigma^{73}$ protein directed transcription from the $\sigma^{70}$-dependent promoter of E. coli neo gene in the presence of peak 1 core RNAP (Fig. 8, lane 2), whereas the core enzyme alone did not (Fig. 8, lanes 1). These results and those in Fig. 6 demonstrate that the C. crescentus RpoD protein isolated either from the peak 2 RNAP preparation or E. coli cells overexpressing the cloned C. crescentus rpoD gene recognizes E. coli $\sigma^{70}$-dependent promoters, suggesting that the C. crescentus $\sigma^{73}$ is a functional homologue of the principal E. coli sigma factor $\sigma^{70}$.

---

$4$ J. Wu and A. Newton, unpublished observations.
We next examined the ability of purified σ73 to confer transcriptional specificity in the recognition of two promoters, rsaA (32, 33) and pleC (34; Fig. 2), that have been used to define the σ73 promoter consensus for C. crescentus (19). Both of these C. crescentus promoters have been characterized in vivo (19) and shown to contain a −35 consensus sequence similar to that in E. coli and a −10 consensus divergent from that in E. coli. The rsaA and pleC promoters were recognized in the in vitro transcription assays by the reconstituted Eσ73 holoenzyme (Fig. 8, lanes 4 and 6). The more efficient transcription from the neo promoter in these experiments (Fig. 8, lane 2) is consistent with measurements of promoter strength in vivo using transcription fusions. β-Galactosidase assays of C. crescentus wild-type strains carrying either the neo-lacZ (6876 Miller units) or the rsaAp-lacZ (1545 Miller units) fusions indicates that the neo promoter is 4- to 5-fold stronger than the rsaA promoter under these conditions.4

The sizes of the rsaA and pleC transcripts observed in vitro (Fig. 8) were those expected from the transcription start sites mapped in vivo for the two genes, i.e. 110 and 62 nt, respectively (Fig. 2; Ref. 19). Transcription initiation was RpoD-dependent, since the peak 1 core RNAP alone did not recognize either promoter (Fig. 8, lanes 3 and 5). These results confirm that the purified rpoD gene product, σ25, is the principal C. crescentus sigma factor and that it is capable of recognizing the C. crescentus housekeeping gene promoters, as well as E. coli σ73-dependent promoters.

FIG. 8. Reconstitution of C. crescentus Eσ73 holoenzyme in vitro. The Eσ73 was reconstituted by mixing the purified C. crescentus peak 1 core RNAP and purified C. crescentus σ73 protein in vitro. The Eσ73 activity was determined by run-off transcription assays on DNA templates containing the σ73-dependent neo promoter from E. coli (lanes 1 and 2) or the pleC (lanes 3 and 4) and rsaA (lanes 5 and 6) housekeeping gene promoters from C. crescentus.

**DISCUSSION**

Many developmental events in C. crescentus are dependent on differential gene expression regulated at the level of transcription. Unlike Bacillus subtilis, where an extensive cast of alternative sigma factors and regulatory proteins governing sporulation have been identified through genetic and biochemical analysis (reviewed in Ref. 8), a defined in vitro transcription system has not been available in C. crescentus. The purification of C. crescentus RNAP was described in early studies (35, 36), but the transcriptional specificity of these enzyme preparations was not characterized. More recently, a partially purified RNAP preparation was used for the study of class II flagellar gene regulation (15). However, this is the first report of the purification and resolution of the holoenzymes and core RNAP and the reconstitution of RNAP holoenzymes exclusively from C. crescentus components.

Heparin-agarose chromatography, which has been used successfully for isolation of RNA polymerase from a number of bacterial species, including E. coli, B. subtilis, B. steatorrhophilus, Lactobacillus casei, L. plantarum, and Clostridium pasteurianum (reviewed in Ref. 4), provided a great enrichment of C. crescentus RNAP (Fig. 1A). Chromatography on single-stranded DNA agarose (37) or on phosphocellulose (38) has also been reported to resolve E. coli holoenzyme and core RNAP. In our hands single-stranded DNA-cellulose chromatography was crucial for resolving C. crescentus RNAP into its core and holoenzyme fractions (Fig. 3C). Phosphocellulose and Bio-Rex 70 were not effective in resolving core and holoenzyme, although these earlier attempts were hampered by the lack of a purified sigma factor to assay for core activity.5

The fact that the first peak from the DNA-cellulose column (peak 1; Fig. 1C) contained active core RNAP and the second peak (peak 2; Fig. 1C) contained Eσ32 and Eσ73 holoenzymes, as well as core RNAP (see below), was demonstrated by in vitro transcription assays of the pooled fractions (Fig. 3; Table I) and the isolation of active Eσ32 and Eσ73 from these RNAP fractions of peak 2 (Fig. 6; 16). In the presence of purified C. crescentus Eσ54 and its activator protein FlbD, core RNAP from peak 1 recognized the Eσ73-dependent promoter of the C. crescentus flagellin gene fliK (Fig. 4), as observed previously for a heterologous holoenzyme containing either C. crescentus Eσ54 or E. coli Eσ54 and the E. coli core RNAP (13–15). Therefore, C. crescentus core RNAP appears to function interchangeably with its E. coli counterpart in this transcription assay.

Our purified peak 2 RNAP preparation displayed activity only on σ32- and σ73-dependent promoters. Possible explanations for the failure to recover active Eσ54 include (i) an unstable σ54 protein that is inactivated during purification and (ii) low affinity of σ54 for binding to core RNAP, which results in its dissociation from the core and loss early in protein fractionation. Consistent with the latter possibility are two observations. First, some bacterial RNAP holoenzymes are quite unstable and dissociate early during procedures suitable for isolation of other RNAP holoenzymes (39), and second, full-length E. coli σ54 does not bind to E. coli core RNAP tightly (40).

Three proteins (Fig. 5; bands A, B, and C) of unknown function are also associated with the RNAP holoenzyme fractions. These proteins could represent additional sigma factors, breakdown products of σ32, other RNAP subunits, such as delta or omega, or proteins that fortuitously fractionate with RNAP. Protein C (Fig. 5), like σ32 (16) and σ73, has been isolated and subjected to amino acid sequencing, but unlike the latter two proteins, there is no similarity of its amino-terminal amino acid sequence to any sequences deposited in GeneBank.5 It will be interesting to determine whether any of these three proteins, A, B, and C, are RNAP subunits or accessory proteins that are involved in transcriptional regulation.

The relative amounts of σ32 and σ73 as visualized by the intensity of staining in SDS-PAGE gels displayed variability depending on the purified preparation examined, but we estimated that pooled peak 2 RNAP contained less than 0.4 mol eq of total σ factor relative to the core subunits (see Fig. 5). This result suggests that peak 2 RNAP also contains core enzyme.6

5 N. Ohta, A. Ninfa, and A. Newton, unpublished observations.
Consistent with the presence of excess core in the peak 2 RNAP fractions was the stimulation of transcription from the $\sigma^{54}$-dependent $fljK$ promoter by the addition of purified $\sigma^{54}$ and FlbD to the peak 2 pool (Fig. 4) and the high activity of this RNAP pool when assayed on the poly[d(A-T)] template (Table I).

The above results raise the question of why the core enzyme elutes from the single-stranded DNA column in two peaks. One possibility is that the more active core enzyme (Table I) binds more tightly to the column and elutes with the holoenzymes in peak 2. Alternatively, the core eluted in peak 2 could represent core that has bound tightly as the holoenzyme from which sigma factors elute at lower salt concentrations than the other polymersases in those complexes. Whatever the explanation of this fractionation, our experiments have depended critically on the isolation of the functional core RNAP. The inability to resolve different RNAP holoenzymes from one another$^5$ prompted us initially to overexpress and purify $C.\ crescentus$ $\sigma$ factors, including $\sigma^{54}$ (14), $\sigma^{32}$ (16), and $\sigma^{73}$ (Fig. 7) for the reconstitution of specific RNAP holoenzymes reported in this study.

$C.\ crescentus$ recognizes $E.\ coli$ $\sigma^{70}$-dependent promoters in vivo (20), and our results demonstrate that the purified (Figs. 3 and 6) and reconstituted (Fig. 8) $C.\ crescentus$ Ec$^{73}$ efficiently recognized the lacUV5 and neo promoters from $E.\ coli$, as well as promoters of the $C.\ crescentus$ housekeeping genes rsaA and pleC (Fig. 8). The $\sim$35 consensus sequence of the $C.\ crescentus$ biosynthetic and housekeeping gene promoters (19) is similar to the $\sim$35 consensus of $E.\ coli$ $\sigma^{70}$-dependent promoters, but the $\sim$10 sequences from these two bacteria align only poorly. Moreover, the $C.\ crescentus$ $\sim$10 and $\sim$35 sequences are more closely spaced than in most $E.\ coli$ promoters (19). These results suggest that the principal $C.\ crescentus$ $\sigma$ factor $\sigma^{73}$ has less promoter specificity than its $E.\ coli$ counterpart $\sigma^{70}$.

The availability of core RNAP from $C.\ crescentus$ and the capability of reconstituting active holoenzymes using purified $\sigma$ factors will permit biochemical analysis of gene regulation in vitro in great detail. These reagents will also be important in studying the role of RNAP and accessory factors in the temporal and spatial regulation of developmental gene transcription during the cell division and differentiation.

Acknowledgments—We thank G. Ramakrishnan for the original isolation of the $C.\ crescentus$ rpoD gene used in this study. We also thank R. R. Burgess for providing anti-$E.\ coli$ $\sigma^{70}$ antibody, S. Inouye for the $M.\ xanthus$ rpoD clone, and J. Smit for plasmid pSSA41 containing the rsaA promoter.