Nucleic Acids Research, 1996, Vol. 24, No. 22 **4565–4571**

Molecular analysis of RNA polymerase alpha subunit gene from *Streptomyces coelicolor* A3(2)

Eun-Jung Cho[+], Jae-Bum Bae, Ju-Gyeong Kang and Jung-Hye Roe*

Department of Microbiology, College of Natural Sciences, and Research Center for Molecular Microbiology, Seoul National University, Seoul 151-742, Korea

Received July 2, 1996; Revised and Accepted October 1, 1996

EMBL/GenBank/DDBJ accession no. X92107

ABSTRACT

The *rpoA* gene, encoding the α subunit of RNA polymerase, was cloned from *Streptomyces coelicolor* A3(2). It is preceded by *rpsK* and followed by *rplQ*, encoding ribosomal proteins S11 and L17, respectively, similar to the gene order in *Bacillus subtilis*. The *rpoA* gene specifies a protein of 339 amino acids with deduced molecular mass of 36 510 Da, exhibiting 64.3 and 70.7% similarity over its entire length to *Escherichia coli* and *B.subtilis* α subunits, respectively. Using T7 expression system, we overexpressed the *S.coelicolor* α protein in *E.coli*. A small fraction of this protein was found to be assembled into *E.coli* RNA polymerase. Antibody against *S.coelicolor* α protein crossreacted with that of *B.subtilis* more than with the *E.coli* α subunit. The ability of recombinant α protein to assemble β and β′ subunits into core enzyme *in vitro* was examined by measuring the core enzyme activity. Maximal reconstitution was obtained at α_2:β+β′ ratio of 1:2.3, indicating that the recombinant α protein is fully functional for subunit assembly. Similar results were also obtained for natural α protein. Limited proteolysis with endoproteinase Glu-C revealed that *S.coelicolor* α contains a tightly folded N-terminal domain and the C-terminal region is more protease-sensitive than that of *E.coli* α.

INTRODUCTION

DNA dependent RNA polymerases of prokaryotes are multisubunit enzymes composed of core enzyme with a subunit composition of $\alpha_2\beta\beta'$ and σ subunit specifying promoter recognition (1,2). The α subunit is the most characterized of the *Escherichia coli* RNA polymerase subunits, serving several functions. Initially, α subunit has been known to play a role in the assembly of multisubunit RNA polymerase complex, providing a scaffold for the assembly of β and β′ in the following order; $2\alpha \rightarrow \alpha_2 \rightarrow \alpha_2\beta \rightarrow \alpha_2\beta\beta'$ (3). It contains determinants for protein–protein interaction for its dimerization, for α–β interaction, and possibly, for α–β′ interaction. These determinants reside in the N-terminal two-thirds of α protein (4–7). The C-terminal third is involved in the interaction with various transcriptional activators (8–10), binding to specific DNA sequences (11,12), as well as its own dimerization (12). It has been demonstrated that the two regions constitute two separate domains connected by a flexible linker (12,13). The structure of the C-terminal domain (CTD) has been determined by NMR and was shown to contain a recognition loop and helices for both transcriptional activators and the DNA UP element, distinctly different from other DNA binding proteins (14). It has also been shown that the contact sites for both DNA and transcriptional activators occupy the same surface on αCTD (14–16).

Gram-positive bacteria of the genus *Streptomyces* undergo a complex morphological and physiological differentiation (17). When the growth of the substrate mycelial colony on a solid surface nearly ceases, usually triggered by nutrient limitation, the colonies develop aerial mycelia, utilizing nutrients primarily provided from the hydrolysis of substrate mycelia. The aerial mycelia further develop into chains of spores. The biosynthesis of various antibiotics by Streptomycetes usually occurs concomitantly with the development of aerial mycelia and spores. These differentiation processes involve a wide range of regulatory mechanisms including alternative σ factors (18). Even though several lines of growing evidence indicate that the interaction with various regulatory factors is central to this regulation, only a few transcription factors were identified until now (19–22). A better understanding of *Streptomyces* gene regulation will require not only the analysis of *cis*- and *trans*-acting regulatory factors, but also the detailed characterization of *Streptomyces* RNA polymerase as a whole as well as its subunits. In this study we present our work on cloning and characterization of the gene for the α subunit. The overproduced recombinant α protein was demonstrated to contain full activity for core enzyme assembly, and a somewhat different structure from *E.coli*.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Escherichia coli DH5α was used for all initial transformation and propagation of plasmids. Strain BL21 (DE3) pLysS, a derivative of lambda lysogen carrying an IPTG inducible gene for T7 RNA polymerase, was used to overexpress the α protein (23). The *Streptomyces coelicolor* strain used was M145, a prototrophic, SCP1$^-$ SCP2$^-$ derivative of the wild-type A3(2) strain. For RNA polymerase preparation, *S.coelicolor* A3(2) M145 was grown in 3.5 l YEME medium (24) containing 34% sucrose and 5 mM $MgCl_2$ under vigorous aeration in a fermenter for 23 h.

*To whom correspondence should be addressed. Tel: +82 2 880 6706; Fax: +82 2 888 4911; Email: jhroe@alliant.snu.ac.kr

[+]Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

Oligonucleotide primers, PCR and cloning

For cloning the *rpoA* gene, oligonucleotide primers α1 (5′-GT(G/C) CT(G/C) CAC GA(G/A) TAC AGC AC-3′) and α2 (5′-CTT (G/C)AG GCA GTT GTA GCT (G/C)CG-3′) were designed as to contain mixed compositions of 50% each of G and C or G and A at each degenerate codon position, and synthesized by Oligos, Etc. (USA). The PCR was performed in the reaction mixture (100 µl) containing 20 pmol each of α1 and α2 primers, four dNTPs at 0.2 mM each, 10 mM Tris–HCl (pH 9.0), 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 100 ng *S.coelicolor* genomic DNA and 2.5 U *Taq* DNA polymerase (Poscochem, Korea). The reaction mixture was subjected to 30 cycles of polymerization consisting of 1 min denaturation at 94°C, 1 min annealing at 37°C, and 1.5 min extension at 72°C. The 4 kb *Pst*I/*Bam*HI genomic DNA fragment hybridizing with the PCR product was cloned into pUC18 and further subcloned into pGEM-7zf(+). To construct α-expressing pET plasmid (pET-α), an oligonucleotide primer α*Nde*I (5′-ACTGAAGGATCCC-CATATGCTGAT-3′) was designed as to create *Nde*I site (underlined) at the start codon (ATG). PCR was performed with α*Nde*I and universal forward sequencing primers (#1212, NEB) using 100 ng pGEM-7 plasmid containing the *rpoA* gene as a template. Amplification was carried out for 20 cycles (denaturation at 94°C for 1 min, annealing at 42°C for 1 min, and extension at 72°C for 1 min). The PCR product was cloned into pET-3a plasmid to create pET-α.

DNA sequencing and sequence analysis

DNA sequencing was carried out on double-stranded DNA fragments cloned into pGEM-7zf(+) by both manual and automated sequencing (ALFexpress, Pharmacia). The sequence was deposited to EMBL under accession number X92107. Sequence comparison with the database was done using BLAST program (25), and CLUSTAL V for multiple alignments (26).

Overexpression and purification of *S.coelicolor* α protein

Escherichia coli BL21 (DE3) pLysS cells containing pET-α plasmid were grown to OD_{600} of 0.6. Following induction by 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG), cells were harvested and stored at –70°C until use. The frozen cells were suspended in the lysis buffer [10 mM Tris–HCl (pH 7.9 at 4°C), 1 mM EDTA, 0.1 M NaCl, 0.3 mg/ml of lysozyme and 1 mM phenylmethylsulfonyl fluoride (PMSF)] for 20 min on ice and sodium deoxycholate was added to a final concentration of 0.2% and further incubated for 20 min before sonication. After removing cell debris by centrifugation, proteins were precipitated with 60% saturated ammonium sulfate and resolubilized in and dialyzed against buffer TGED [10 mM Tris–HCl (pH 7.9 at 4°C), 0.1 mM EDTA, 0.1 mM DTT, 5% glycerol]. The proteins were then subjected to DEAE–Sepharose column chromatography, eluted with a linear gradient of 0–0.5 M NaCl. Fractions containing α protein were eluted at 0.3 M NaCl and were further purified on the preparative polyacrylamide gel electrophoresis system (BioRad).

Western blot analysis

Purified α protein was electrophoresed on 10% SDS–PAGE. Gel slice containing α was crushed and resuspended in phosphate buffered saline (0.8% NaCl, 0.02% KCl, pH 7.2). A total of 250 µg protein was used to immunize two mice for antibody production. The Western blot analysis was performed as described by Blake *et al.* (27). Blotted filters were incubated overnight with polyclonal mouse antibody against α (1:1000 dilution) in TBS buffer containing 20 mM Tris–HCl (pH 7.5), 500 mM NaCl, 0.5% BSA and 0.05% Tween 20. Bound antibody was detected by anti-mouse IgG goat antibodies conjugated with alkaline phosphatase.

Purification of RNA polymerase from *S.coelicolor*

RNA polymerase was prepared by a combined modification of two procedures developed for the purification of holoenzymes from *E.coli* and *S.coelicolor* (28,29). In brief, *S.coelicolor* mycelial cells (40 g wet weight) were suspended in 400 ml lysis buffer [10 mM Tris–HCl (pH 7.9 at 4°C), 2 mM EDTA, 1 mM PMSF, 1 mM 2-mercaptoethanol, 0.2 M NaCl, 1 mM DTT, 5% glycerol] containing 1 g lysozyme and were incubated at 4°C for 30 min. The cells were disrupted by grinding with aluminum oxide in a commercial mixer. Crude extracts were rapidly mixed with Polymin P at a final concentration of 0.35% (v/v) and subjected to salt extraction, ammonium sulphate precipitation, DEAE–Sephacel, and heparin–Sepharose chromatographies. For further purification, the fractions containing intact β and β′ subunits were pooled and applied via a 1 ml superloop to a MonoQ HR 1/1 anion exchange column on FPLC (Pharmacia). The column was washed with 4 ml TGED plus 0.2 M NaCl and then a 20 ml linear gradient of 0.2–0.6 M NaCl was applied at a flow rate of 0.5 ml/min. Fractions containing pure RNA polymerase were diluted with two times concentrated storage buffer (TGED with 50% glycerol and 0.15 M NaCl) and stored at –70°C.

Small scale purification of RNA polymerase from *E.coli*

Escherichia coli BL21 cells containing pET-α plasmid were grown to OD_{600} of 0.6 and treated with 0.01 mM IPTG for 6 h. Cells from 200 ml culture were harvested and lysed in the same buffer used for large scale preparation of RNA polymerase (28). Following Polymin P fractionation, the polymerase fraction was loaded on 1 ml heparin–Sepharose column and proteins were eluted with step gradient of KCl from 0.3 to 0.8 M. The pooled fractions containing RNA polymerase were further fractionated on glycerol gradient centrifugation as described by Glass *et al.* (30).

Isolation of subunits from purified RNA polymerase

The α, β and β′ subunits were separated from the core RNA polymerase as described by Lill *et al.* (31) with slight modifications. Core enzyme (3 mg) in 3 ml of storage buffer was diluted with 6 ml of deionized 9 M urea (final concentration 6 M) and adjusted to 10 mM DTT. Following 15 min incubation at 30°C, the mixture was dialyzed for 3 h against buffer $TG_{20}ED$ [TGED with 20% (v/v) glycerol] containing 6 M urea. The dissociated enzyme was applied to a phosphocellulose column (0.9 × 4 cm) equilibrated with the same buffer. The column was then washed with $TG_{20}ED$ plus 6 M urea. Subunit α came out in the flow-through fraction whereas subunits β and β′ were eluted at 0.2 and 0.7 M KCl, respectively. The fractions containing each subunit were dialyzed overnight against the renaturation buffer [50 mM Tris–HCl (pH 7.9 at 4°C), 0.1 mM EDTA, 1 mM DTT, 10 mM $MgCl_2$, 0.3 M KCl, 20% glycerol] to remove urea. They were concentrated by ammonium sulphate precipitation. The protein pellet was redissolved in TGED buffer and dialyzed overnight against the storage buffer and kept at –20°C.

Reconstitution of RNA polymerase core enzyme

To reconstitute core enzyme *in vitro*, either natural or recombinant α protein was mixed with β and β′ subunits at various stoichiometric molar ratios in reconstitution buffer containing 7–10 mM Tris–HCl (pH 7.9 at 25°C), 7 mM $MgCl_2$, 25 mM $MnCl_2$, 0.07–0.1 mM EDTA, 0.3 M KCl, 10 mM DTT and 40% glycerol. The final concentration of total protein was adjusted between 0.2 and 0.5 mg/ml. The mixture was incubated at 30°C for 30 min to 1 h. The RNA synthesizing activity of the resulting enzyme was determined as described by Burgess and Jendrisak (32).

RESULTS

Cloning of the *rpoA* gene from *S.coelicolor*

Comparison of amino acid sequences of α proteins from *E.coli*, *Bacillus subtilis* and plant chloroplasts revealed several conserved regions (33–37). Based on this observation, oligonucleotide primers were designed from two such conserved sequences to amplify the intervening region of *S.coelicolor* α gene. Degenerate oligonucleotide primers corresponding to *E.coli* residues 64–70 (α1) and 265–271 (α2) were synthesized (Fig. 2). PCR reaction on chromosomal DNA from *S.coelicolor* A3(2) M145 produced a single species of product of 620 bp, the size of which is as predicted from the sequence of *rpoA* gene from *E.coli*. The PCR product was cloned and partially sequenced. The translated amino acid sequence was 35 and 41% identical to the sequences of the corresponding regions of *E.coli* (34 of 98 amino acid residues) and *B.subtilis* (40 of 98 amino acid residues) α subunit, respectively. This PCR product was used as a probe for the Southern hybridization of *S.coelicolor* chromosomal DNA digested with *Kpn*I, *Pst*I, *Bam*HI or *Pst*I/*Bam*HI. A single distinct band was detected for each digest (data not shown). The hybridizing *Pst*I–*Bam*HI fragment of 4 kb was cloned into pUC18 and further subcloned into pGEM-7zf(+).

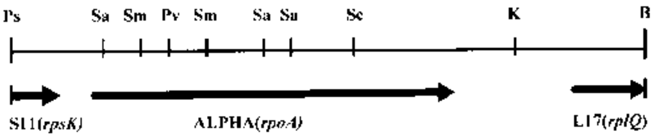

Figure 1. Restriction map and gene order in the DNA fragment containing *rpoA* gene of *S.coelicolor*. Restriction map of the 1.6 kb DNA fragment containing genes for ribosomal protein S11 (*rpsK*), α (*rpoA*) and ribosomal protein L17 (*rplQ*) was presented. The thick arrows indicate the position of genes identified. Only the C-terminal 99 nucleotides of *rpsK* gene and the N-terminal 213 nucleotides of *rplQ* gene were included. The restriction enzyme sites identified are as follows: Pv, *Pvu*II; Sa, *Sal*I; Ps, *Pst*I; K, *Kpn*I; B, *Bcl*I; Sc, *Sac*I; Sm, *Sma*I. The accession number for the nucleotide sequence of this fragment is EMBL X92107.

Sequence identification of the *rpoA*, *rpsK* and *rplQ* genes

Nucleotide sequencing identified an open reading frame for the α subunit (*rpoA*), preceded by a truncated sequence from ribosomal protein S11 (*rpsK*) with 132 bp of intercistronic region, and followed by another truncated sequence from ribosomal protein L17 (*rplQ*) with 191 bp of intercistronic region. The restriction map and the gene organization of a 1.6 kb *Pst*I–*Bcl*I fragment were shown in Figure 1. The C-terminal 30 residues encoded by *S.coelicolor rpsK* exhibited high similarity to the C-terminal residues of *E.coli* ribosomal protein S11 (20 identical residues) (33). The N-terminal 71 residues encoded by *S.coelicolor rplQ* also exhibited 54% similarity to the N-terminal residues of *E.coli* ribosomal protein L17 (33).

The *rpoA* open reading frame encodes a protein of 339 amino acids with a deduced molecular mass of 36 510 Da. A potential Shine–Dalgarno (SD) sequence, GAAGGA, was found between nucleotides (nt) 212 and 217, consistent with the average distance (6–10 nt) between the ATG and the SD sequences in eubacteria. Amino acid sequences of *S.coelicolor* α subunit was compared with other known bacterial α sequences (Fig. 2). Between *E.coli*

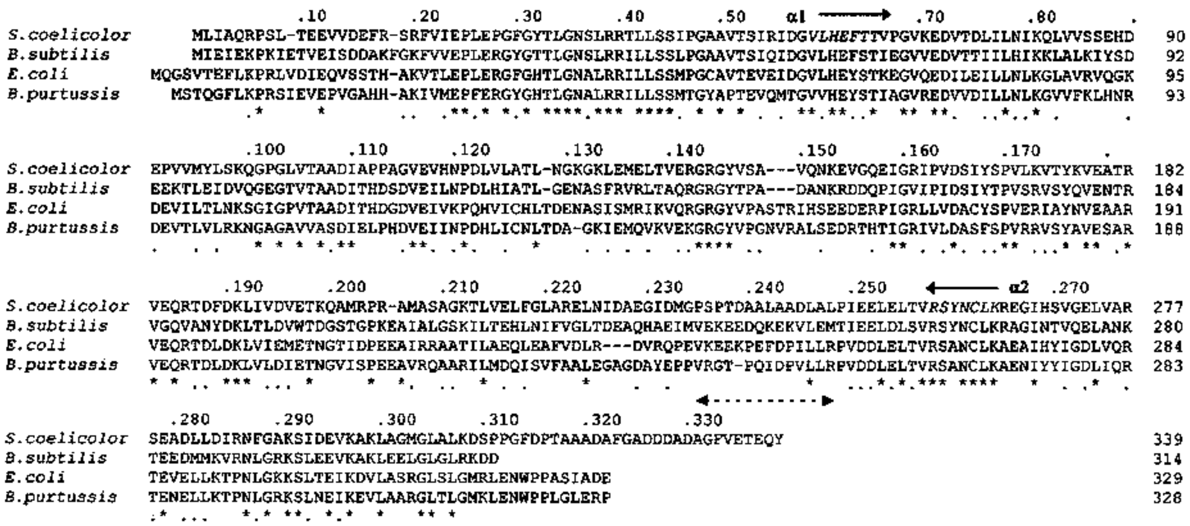

Figure 2. Comparison of amino acid sequences of *S.coelicolor* α protein with those from other bacteria. The comparison was made with sequences from *B.subtilis*, *E.coli* and *B.pertussis* (33,34,38). The amino acid residue identical at a given position in all of these organisms is marked by an asterisk, whereas the conserved substitution is marked by a dot. The gap which is introduced to allow optimal alignment is denoted with a dash. The residues corresponding to the PCR primers used to amplify the intervening region are overlined with arrows. Protease-sensitive linker region in *E.coli* α is demarcated by an underlying arrow (12,13).

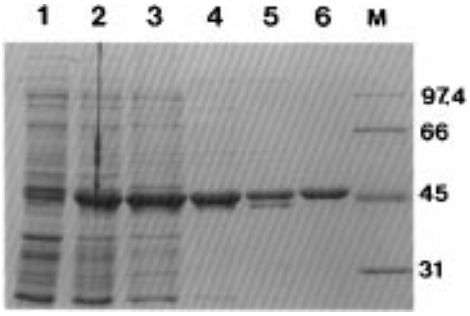

Figure 3. Overexpression and purification of recombinant α protein from *E.coli*. Samples from progressive stages of purification process were subjected to SDS–PAGE and stained with Coomassie blue. Lane 1, crude lysate of *E.coli* cells harboring pET-3a plasmid; lane 2, crude lysate of *E.coli* cells harboring pET-α plasmid; lane 3, resuspended ammonium sulfate precipitate; lanes 4 and 5, DEAE–Sepharose fractions containing α; lane 6, purified α protein following preparative electrophoresis; lane M, molecular weight size markers.

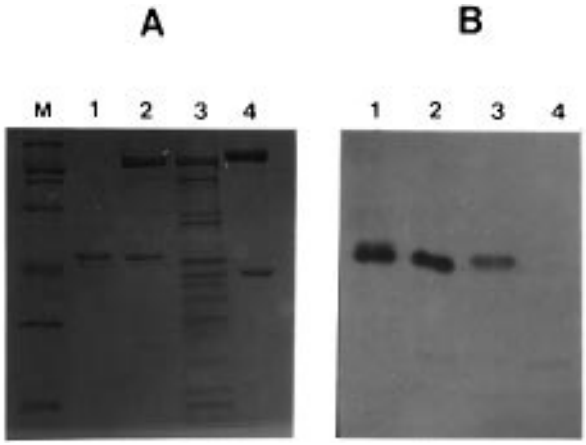

Figure 4. Immunoblot analysis of α from various sources. (**A**) Purified recombinant α protein (lane 1) and RNA polymerase from *S.coelicolor* (lane 2) were subjected to gel electrophoresis, along with partially purified RNA polymerase preparation from *B.subtilis* (lane 3) and *E.coli* (lane 4). Protein bands were visualized by staining with Coomassie blue. (**B**) Immunoblot of the gel in A with polyclonal antibodies raised against purified *S.coelicolor* recombinant α protein.

and *S.coelicolor*, 39.8% of residues were identical and 64.3% were conserved considering similar residues. When compared with *B.subtilis*, the degree of conservation was 48.2 and 70.7% for identical and similar matches, whereas with *Bordetella pertussis* it was 40.3 and 65.2% (38).

Several conserved domains can be predicted from the distribution of homologous residues. One long stretch of conserved residues from 23 to 46 near the N-terminus and a segment from 250 to 264 are most prominent. These correspond well with the positions known to be involved in subunit assembly, and in contacting with transcriptional factors and DNA, respectively (16).

Expression and purification of α protein

Using the T7 expression system, as described in Materials and Methods, we were able to obtain a large amount of *S.coelicolor* α protein in *E.coli*. Following induction with 0.4 mM IPTG, *S.coelicolor* α protein made up to 35% of the total protein in *E.coli* cell extracts. The recombinant α protein was purified to >95% homogeneity as shown in Figure 3. The apparent molecular mass was determined to be ~46 kDa, substantially larger than the predicted one. We compared the antigenicity among α proteins from *S.coelicolor*, *B.subtilis*, and *E.coli* by immunoblotting (Fig. 4). The polyclonal antibodies raised against the recombinant *S.coelicolor* α protein cross-reacted with the natural α protein in purified RNA polymerase, demonstrating that the recombinant α protein was antigenically indistinguishable from the natural one (Fig. 4B, lanes 1 and 2). We found that the antibody against *S.coelicolor* α protein cross-reacted better with *B.subtilis* α than that of *E.coli* (Fig. 4B, lanes 3 and 4), suggesting that the structural difference in α subunit between *S.coelicolor* and *E.coli* is rather substantial. The deduced pI value of *S.coelicolor* α from the amino acid composition is 4.36, which is most acidic when compared with those of *E.coli* subunits; β′ (6.85), β (5.30), α (4.65) and σ^{70} (4.40) (39). This acidic property of the *S.coelicolor* α protein was confirmed by two-dimensional gel electrophoresis performed in parallel with *E.coli* RNA polymerase under the denaturing condition (8 M urea) (data not shown).

Association of *S.coelicolor* α subunit with *E.coli* RNA polymerase *in vivo*

The incorporation of the *S.coelicolor* α subunit into *E.coli* RNA polymerase holoenzyme *in vivo* was examined to verify its ability to assemble into RNA polymerase. *E.coli* BL21 cells containing pET-α were slightly induced for *S.coelicolor* α protein with 0.01 mM IPTG instead of 0.4 mM. Under this condition, most of the recombinant α protein exists in soluble fractions. RNA polymerase holoenzyme was purified from these cells through Polymin-P fractionation, ammonium sulfate precipitation, heparin–Sepharose chromatography and glycerol gradient ultracentrifugation. Figure 5A demonstrated a SDS–PAGE profile of purified RNA polymerase from *E.coli* BL21 (lane 2) and *S.coelicolor* as a control (lane 1). In lane 2, a very faint band of 46 kDa was visible along with thicker band of *E.coli* α which migrates faster. The presence of *S.coelicolor* α protein in *E.coli* RNA polymerase preparation was more clearly demonstrated by immunoblotting with antibody against *S.coelicolor* α protein (Fig. 5B, lane 2). The result indicates that the *S.coelicolor* α subunit is able to assemble with *E.coli* β and β′ subunits to form hybrid core enzymes. We checked each fraction of glycerol gradient for the presence of *S.coelicolor* α by immunoblotting and found out that the *S.coelicolor* α protein was present only in the bottom fractions where *E.coli* β and β′ subunits cosedimented as well as in the top fractions existing as free subunits (data not shown). This rules out the possibility that *S.coelicolor* α might have been contaminated in the *E.coli* RNA polymerase preparation.

Reconstitution of *S.coelicolor* RNA polymerase core enzyme from the recombinant α protein *in vitro*

In order to test further the assembly function of the recombinant α protein, we performed *in vitro* reconstitution of *S.coelicolor* RNA polymerase core enzyme from its subunits. β and β′ subunits were dissociated from the purified RNA polymerase by

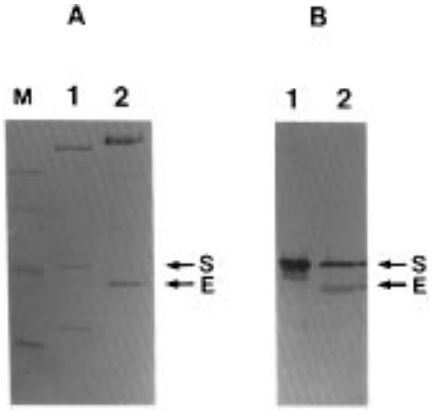

Figure 5. Incorporation of *S.coelicolor* α subunit into *E.coli* RNA polymerase. (**A**) SDS–PAGE of RNA polymerases purified from *S.coelicolor* (lane 1) and *E.coli* cells harboring pET-α (lane 2). (**B**) Immunoblot analysis of the gel in (A) with antibody against *S.coelicolor* α protein. Arrows indicate the position of the α protein of *E.coli* (E) and *S.coelicolor* (S).

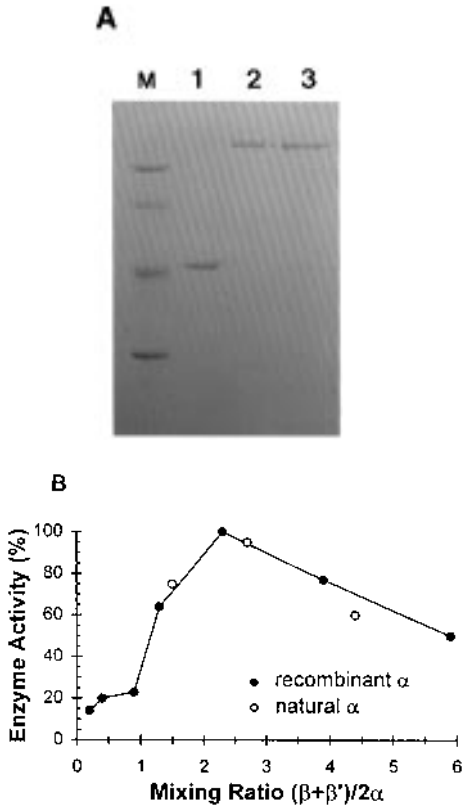

Figure 6. Reconstitution of *S.coelicolor* RNA polymerase core enzyme from purified subunits. (**A**) SDS–PAGE of purified subunits of RNA polymerase used in reconstitution. The subunits were separated by phosphocellulose chromatography of urea-dissociated core enzyme as described in the text. 1 µg each of α (lane 1), β (lane 2) and β′ (lane 3) proteins were loaded on 10% polyacrylamide gel. (**B**) Reconstitution efficiency at different mixing ratios. To a fixed amount of mixture containing 100–150 ng/µl each of β and β′, varying amounts of either recombinant or natural α protein were added at the indicated molar ratio. RNA polymerase activity was assayed as described in the text. The maximal activity at the ratio of 2.3 was taken as 100%.

6 M urea and isolated through phosphocellulose column chromatography as described in Materials and Methods. The separation of isolated subunits was checked by SDS–PAGE as shown in Figure 6A. We assigned the higher salt-eluted subunit as β′ as in *E.coli* (40) and noted that β′ migrated slightly faster than β, unlike in *E.coli*. The isolated β and β′ and the recombinant α proteins were then mixed in the absence of urea at different molar ratios of $\beta\beta'/\alpha_2$ and measured for RNA polymerizing activity (Fig. 6B). The active enzyme was recovered only when all the subunits were combined. Urea was not required in the reconstitution process as observed by Lill *et al.* (31) for *E.coli* RNA polymerase. The optimal recovery was achieved when the subunits were mixed at the α_2:ββ′ ratio of 1:2.3, instead of the stoichiometric 1:1 ratio. This indicates that the amount of 'competent' β and/or β′ molecules in these preparations was ~43% of the total β and β′ proteins. More excess amounts of α over β and β′ subunits were found to be inhibitory to enzyme assembly or activity. The maximum activity of reconstituted enzyme corresponded to 40–60% of the activity of core enzyme treated in parallel with 6 M urea without column separation of subunits, comparable to the efficiency with *E.coli* RNA polymerase (41). When the natural α subunit dissociated from the native core enzyme replaced the recombinant α protein, there was no difference in the efficiency of reconstitution within experimental error. This indicates that the recombinant α protein is indeed capable of assembling functional RNA polymerase with β and β′ subunits, just like natural α protein.

Proteolytic analysis of α protein

To obtain preliminary information on the subdomain structure of α subunit, the purified α protein was partially digested with endoproteinase Glu-C which cuts at Glu residues. The reaction products were run on SDS-polyacrylamide gel and compared with *E.coli* α protein subjected to the same treatment in parallel (Fig. 7). *Streptomyces coelicolor* α was cleaved into one major (33 kDa_{app}) and two minor (~28 kDa_{app}) fragments, while *E.coli* α was cleaved into two major fragments, 28 and 8 kDa_{app} as observed by Blatter *et al.* (12). Previous studies on domain mapping of *E.coli* α have demonstrated that the protease resistant 28 kDa fragment is from the highly structured N-terminal two-thirds of α, and the smaller 8 kDa fragment is from the C-terminal region suggesting that α consists of two major domains, N-terminal domain (NTD) and C-terminal domain (CTD), linked by protease-sensitive flexible linker from residue 235 to 248 (12,13). Both the 33 and 28 kDa fragments were derived from the N-terminal portion of *S.coelicolor* α, since various lengths of deletions from the C-terminal end up to residue 250 barely changed the pattern of proteolysis (data not shown). We were not able to observe any smaller fragments around 10 kDa for *S.coelicolor* α (Fig. 7, lane 2).

The distribution of Glu residues in the C-terminal region from residue 235 to 339 in *S.coelicolor* α is such that there are nine Glu residues altogether, and excluding Glu-335 and 337 which are very close to C-terminus, the remaining seven residues all correspond to either Glu or Asp residues in *E.coli*. There are no extra Glu residues in *S.coelicolor* C-terminal region which do not correspond to these positions. The partial cleavage pattern of *E.coli* α with V8 protease (same as Glu-C), under conditions where it cuts at both Glu and Asp, has been reported to be very similar to that of the Glu-C digest, predicting Glu-241 as the major cut site for both enzyme treatments (12,13). Glu and Asp residues C-terminal to Glu-241 in *E.coli* are not accessible to

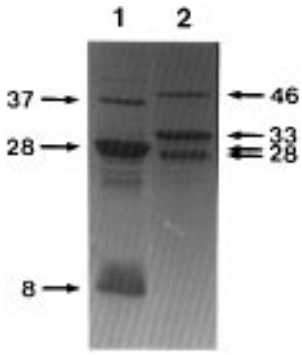

Figure 7. Limited proteolytic cleavage of *S.coelicolor* α protein. α proteins (10 µg) from *E.coli* (lane 1) or *S.coelicolor* (lane 2) were digested with 0.1 µg endoproteinase Glu-C (Sigma) in TGED buffer plus 150 mM NaCl at 25°C for 2 h. Digestion products were analyzed by SDS–PAGE, and visualized by Coomassie blue staining. The apparent molecular weights (in kDa) of the reaction products as well as the parent molecule are denoted on both sides.

proteolytic cleavage. If *S.coelicolor* αCTD assumes a similar structure to *E.coli* αCTD, then the corresponding position will not be accessible either in *S.coelicolor*. The absence of any discrete proteolytic product from *S.coelicolor* α C-terminal region indicates that the C-terminal remainder of Glu-C digest is very labile toward proteinase attack unlike the C-terminal portion of *E.coli* α. We have also treated *S.coelicolor* α with endoproteinases Lys-C and trypsin. They produced essentially the same cleavage pattern with *E.coli* α as reported previously (12,13), but no discrete CTD product with *S.coelicolor* α, confirming the result with Glu-C digest (data not shown). These results suggest that the C-terminal domain of *S.coelicolor* α may assume a looser structure compared with that of *E.coli*.

DISCUSSION

We isolated the gene for α subunit protein of *S.coelicolor* RNA polymerase and demonstrated that its gene product expressed in *E.coli* can indeed be assembled into a functional RNA polymerase core enzyme. The *rpoA* gene was found to be preceded by *rpsK* (S11) and followed by *rplQ* (L17). This gene order is the same as in *B.subtilis* and *Chlamydia trachomatis*. In *E.coli*, *rpoA* lies in an operon with four ribosomal protein genes in the order, *rpsM* (S13), *rpsK* (S11), *rpsD* (S4), *rpoA* and *rplQ* (L17) (33). The organization of *B.subtilis rpoA* operon differs from *E.coli* in that it lacks *rpsD* gene preceding *rpoA* (42). In the genomes of both *B.subtilis* and *C.trachomatis*, the *rpoA* gene is preceded by the *rpsK* gene (43). In this respect, *S.coelicolor* is more closely related with these groups than *E.coli*. The absence of *rpsD* is interesting because S4 has been shown to act as a translational repressor of the α operon in *E.coli*, inhibiting the translation of the three ribosomal protein cistrons preceding the α and L17 cistron distal to α (44,45). This autogenous regulation of the α operon by S4 results in differential expression of ribosomal proteins and α in *E.coli*. Further investigation is needed as to whether the ribosomal proteins and α are differentially regulated in *S.coelicolor*, and if so, what is the mechanism of action.

The sequence comparison of *S.coelicolor* α protein with three other known bacterial α proteins revealed that there are several conserved regions. Mutation studies on *E.coli* α have demonstrated that the residues in the N-terminal portion are involved in the assembly of the core RNA polymerase (5–7). Point mutations within the C-terminal domain that affected transcriptional activation by regulatory factors or binding to UP element identified individual residues of α as potential contact sites (8,10,12). Most of these residues are conserved in *S.coelicolor* α as well. One example is the binding site for cyclic AMP (cAMP) receptor protein (CRP). The residues between 258 and 270, known to be involved in the interaction with CRP at class I CRP-dependent promoters (16,46,47), are very well conserved. These are residues at 258(D), 259(D), 260(L), 261(E), 262(L), 265(R), 268(N) and 269(C) in *E.coli*, corresponding to *S.coelicolor* residues at 251(E), 252(E), 253(L), 254(E), 255(L), 258(R), 261(N) and 262(C), respectively. The residues involved in the recognition of the UP element of *rrnB* P1 promoter in *E.coli*, 260(L), 262(L), 264(V), 265(R), 266(S), 268(N), 269(C), are all correspondingly conserved in *S.coelicolor* (14–16). One of the two residues at 297(K) and 298(K) in *E.coli*, known to be involved in both UP element and CRP recognition, are conserved in *S.coelicolor* at 291. Whether these conservations have any meaning in *S.coelicolor* is not known, since so far neither a homolog of CRP nor CRP-dependent promoters have been reported for *Streptomyces* spp. Furthermore, intracellular concentration of cAMP is relatively constant and does not appear to be involved in catabolite repression in *Streptomyces* spp. (48). The presence of UP element in *S.coelicolor* promoters has not been reported either. One of the two residues, Pro-322 and Pro-323 of *E.coli* α, which is involved in OmpR activation (49) is also conserved at 316. Lys-271 involved in interactions with the activators CysB, AraC and MelR in *E.coli* is conserved in *S.coelicolor* at residue 264 (50). Although no homologs of these activators have been described to date in *Streptomyces* spp., the work in *E.coli* to define the activator contact sites on α can help assign a role on *S.coeliclor* α that might interact similarly with different array of transcriptional activators. *S.coelicolor* α subunit protein contains an extra C-terminal tail of 17 amino acids. This extended C-terminal tail is also observed in the Chlamydial α protein, with complete lack of homology between the two (43). The prominent feature in the C-terminal tail of *S.coelicolor* α is the abundance of acidic residues (six out of 17).

Even though the peptide sequences are rather highly conserved between *S.coelicolor* and *E.coli* α proteins, the structural difference is manifested by low immuno-crossreactivity and the difference in susceptibility to endoproteinase. The less compact structure in C-terminal domain of *S.coelicolor* α can be postulated, and the different interaction between *S.coelicolor* α and DNA or other transcriptional regulators can be inferred. Further genetic and biochemical studies are necessary in this respect.

ACKNOWLEDGEMENTS

We thank Dr A. Ishihama for helpful discussions and providing pGEMAX185 for the overproduction of the *E.coli* α subunit; Dr D. Hopwood for providing *S.coelicolor* M145 A3(2) strain; Drs N. Fujita and J.-W. Suh for valuable discussions; Dr H.-D. Youn for technical assistance in protein purification; J.H. Lee for automatic sequencing; other members of our lab for the sequence data analysis and helpful discussions. This work was supported by a grant from the Korea Science and Engineering Foundation for SRC (Research Center for Molecular Microbiology) to J.-H. Roe.

REFERENCES

1 Burgess,R.R. (1976) In Losick.R. and Chamberlin,M.J. (eds), *RNA polymerase*. Cold Spring Harbor Laboratory Press, New York, pp 69–100.
2 Yura,T. and Ishihama,A. (1979) *Annu. Rev. Genet.*, **13**, 59–97.
3 Ishihama,A. (1981) *Adv. Biophys.*, **14,** 1–35.
4 Hayward,R., Igarashi,K. and Ishihama,A. (1991) *J. Mol. Biol.*, **221**, 23–29.
5 Kimura,M., Fujita,N and Ishihama,A. (1994) *J. Mol. Biol.*, **242**, 107–115.
6 Kimura,M. and Ishihama,A. (1995) *J. Mol. Biol.*, **248**, 756–767.
7 Kimura,M. and Ishihama,A. (1995) *J. Mol. Biol.*, **254**, 342–349.
8 Russo,F. and Silhavy,T. (1992) *J. Biol. Chem.*, **267,** 4515–4518.
9 Ishihama,A. (1992) *Mol. Microbiol.*, **6**, 3283–3288.
10 Ishihama,A. (1993) *J. Bacteriol.*, **175,** 2483–2489.
11 Ross,W., Gosink,K., Salomon,J., Igarashi,K., Zou,C., Ishihama,A., Severinov,K. and Gourse,R.L. (1993) *Science*, **262**, 1407–1413.
12 Blatter,E.E., Wilma,R., Tang,H., Gourse,R.L. and Ebright,R.H. (1994) *Cell*, **78,** 889-896.
13 Negishi,T., Fujita,N. and Ishihama,A. (1995) *J. Mol. Biol.*, **248**, 723–728.
14 Jeon,Y.H., Negishi,T., Shirakawa,M., Yamazaki, T., Fujita,N., Ishihama,A. and Kyogoku,Y. (1995) *Science*, **270**, 1495–1497.
15 Gaal,T., Ross,W., Blatter,E.E., Tang,H., Jia,X., Krishnan,V.V., Assa-Munt,N., Ebright,R.H. and Gourse,R.L. (1996) *Genes Dev.*, **10**, 16–26.
16 Murakami,K., Fujita,N. and Ishihama,A. (1996) *EMBO J.* **15**, 4358–4367.
17 Chater,K.F. (1989) *Trends Genet.*, **11**, 372–377.
18 Buttner,M.J. (1989) *Mol. Microbiol.*, **3,** 1653–1659.
19 Horinouchi,S., Kito,M., Furuya,K., Hong,S.K., Miyake,K. and Beppu,T. (1990) *Gene*, **95**, 49–56.
20 Fernandez-Moreno,M.A., Caballero,J.L., Hopwood,D.A. and Malpartida,F. (1991) *Cell*, **66**, 769–780.
21 Takano,E., Gramajo,H.C., Strauch,E., Andres,N., White,J. and Bibb,M.J. (1992) *Mol. Microbiol.*, **6**, 2797–2804.
22 Retzlaff,L. and Distler,J. (1995) *Mol. Microbiol.*, **18**, 151–162.
23 Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Duberdorff,J.W. (1990) *Methods Enzymol.*, **185,** 60–89.
24 Hopwood,D.A., Bibb,M.J., Chater,K.F., Kieser,T., Bruton,C.J., Kieser,H.M., Lydiate,D.J., Smith,C.P., Ward,J.M. and Schrempf,H. (1985) *Genetic Manipulation of Streptomyces: A Laboratory Manual.* The John Innes Foundation, Norwich, UK.
25 Altschul,S.F., Warren,G., Webb,M., Eugene,W.M. and Lipman,D.J. (1990) *J. Mol. Biol.*, **215,** 403–410.
26 Higgins,D.G., Bleasby,A.J. and Fuchs,R. (1992) *Comput. Appl. Biosci.*, **8**, 189–191.
27 Blake,M.S., Johnston,K.H., Russel-Jones,G.J. and Gotschlich,E.C. (1989) *Anal. Biochem.*, **136**, 175–179.
28 Hager,D.A., Jin,D.J. and Burgess,R.R. (1990) *Biochemistry*, **29**, 7890–7894.
29 Buttner,M.J. and Brown,N.L. (1985) *J. Mol. Biol.*, **185**, 177–188.
30 Glass,R.E., Ralphs,N.T., Fujita,N. and Ishihama,A. (1988) *Eur. J. Biochem.*, **176**, 403–407.
31 Lill,U., Behrendt,E.M. and Hartmann,G.R. (1975) *Eur. J. Biochem.*, **52**, 411–420.
32 Burgess,R.R. and Jendrisak,J.J. (1975) *Biochemistry*, **14**, 4634–4638.
33 Bedwell,D.M., Davis,G., Gosink,M., Post,L., Nomura,M., Kestler,H., Zengel,J.M. and Lindahl,L. (1985) *Nucleic Acids Res.*, **13,** 3891–3903.
34 Boylan,S.A., Suh,J.-W., Thomas,S.M. and Price,C.W. (1989) *J. Bacteriol.*, **171,** 2553–2562.
35 Ohyama,K., Fukuzawa,H., Kohchi,T., Shirai,H., Sano,T., Sano,S., Umesono,K., Shiki,Y., Takeuchi,M., Chang,Z., Aota,S., Inokuchi,H. and Ozeki,H. (1986) *Nature*, **322**, 572–574.
36 Shinozaki,K., Ohme,M., Tanaka,M., Wakasugi,T., Hayashida,N., Matsubayashi,T., Zaita,N., Chunwongse,J., Obokata,J., Yamaguchi-Shinozaki,K., Ohto,C., Torazawa,K., Meng,B.Y., Sugita,M., Deno,H., Kamogashira,T., Yamada,K., Kusuda,J., Takaiwa,F., Kato,A., Tohdoh,N., Shimada,H. and Sugiura,M. (1986) *EMBO J.*, **5**, 2043–2049.
37 Sijben-Müller,G., Hallick,R.B., Alt,J., Westhoff,P. and Herrmann,R.G. (1986) *Nucleic Acids Res.*, **14,** 1029–1044.
38 Carbonetti,N.H., Fuchs,T.M., Patamawenu,A.A., Irish,T.J., Deppisch,H. and Gross,R. (1994) *J. Bacteriol.*, **176**, 7267–7273.
39 Levin,B.J., Orphanos,P.D., Ficshmann,B.S. and Beychok,S. (1980) *Biochemistry*, **19,** 4808–4814.
40 Yarbrough,L.R. and Hurwitz,J. (1974) *J. Biol. Chem.*, **249**, 5400–5404.
41 Ishihama,A. and Ito,A. (1972) *J. Mol. Biol.*, **72**, 111–123.
42 Suh,J.-W., Boylan,S.A. and Price,C.W. (1986) *J. Bacteriol.* **168**, 65–71.
43 Gu,L., Wenman,W.M., Remacha,M., Meuser,R., Coffin,J. and Kaul,R. (1995) *J. Bacteriol.*, **177**, 2594–2601.
44 Dean,D. and Nomura,M. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3591–3594.
45 Nomura,M., Gourse, R. and Baughman,G. (1984) *Annu. Rev. Biochem.*, **53**, 75-117.
46 Tang,H., Severinov,K., Goldfarb,A., Fenyo,D., Chait,B. and Ebright,R.H. (1994) *Genes Dev.*, **8**, 3058–3067.
47 Zou,C., Fujita,N., Igarashi,K. and Ishihama,A. (1992) *Mol. Microbiol.*, **6**, 2599–2605.
48 Chatterjee,S. and Vining,L.C. (1982) *Can. J. Microbiol.*, **28**, 311–317.
49 Slauch,J.M., Russo,F.D. and Silhavy,T.J. (1991) *J. Bacteriol.*, **173**, 7501–7510.
50 Thomas,M.S. and Glass,R.E. (1991) *Mol. Microbiol.*, **5**, 2719–2725.