Overproduction, Purification, and Characterization of *Bacillus subtilis* RNA Polymerase σ^A Factor

BAN-YANG CHANG AND ROY H. DOI*

Department of Biochemistry and Biophysics, University of California, Davis, California 95616

Received 15 December 1989/Accepted 23 March 1990

By use of a T7 expression system, large amounts of active *Bacillus subtilis* RNA polymerase σ^A factor were produced in *Escherichia coli* cells. This overproduced protein was found in the form of inclusion bodies and constituted 40% of the total cellular protein. Because of the ease of isolation of the inclusion bodies and the acidic properties of σ^A , the protein was purified to more than 99% purity and the yield was about 90 mg/liter of culture. Gel mobility, antigenicity, specificity of promoter recognition, and N-terminal amino acid sequence of the overproduced σ were found to be the same as those of native σ^A . Partial proteolysis analysis of σ^A protein suggested the presence of a protease-sensitive surface region in the C-terminal part of the σ^A protein. The promoter -10 binding region of σ^A was less sensitive to proteases and was probably involved in a hydrophobic, tightly folded domain of σ^A protein.

The procaryotic RNA polymerase σ factor plays a role in the specificity of promoter recognition and the initiation of transcription (2, 3). Two specific functions of σ are to bind to the RNA polymerase core enzyme (18) and to recognize promoter sequences (3, 5, 9, 28, 31). Although considerable studies have been carried out with *Escherichia coli* σ^{70} , the structural difference between it and *Bacillus subtilis* σ^A (GenBank accession no. M10089) (10), the incomplete understanding of the structure and function of σ factors, and the absence of crystallographic data on σ factors have made the further analysis of σ^A a compelling subject.

One of the limitations in studying the functions and structure of *B. subtilis* σ^A factor has been difficulty in obtaining sufficient quantities of σ^A . *E. coli* σ^{70} has been overproduced to about 5% of the total cell protein (11, 16). Our previous attempts to overexpress the *B. subtilis* *sigA* gene by use of the lambda p_L promoter, *lac* promoter, and a two-plasmid T7 promoter system were not successful, since they were too "leaky" and the σ^A produced was lethal to the host *E. coli* cells.

By use of a newly developed T7 expression system (25, 30), we have been able to produce large amounts of σ^A and to develop a method for obtaining highly purified σ^A protein. The availability of large quantities of σ^A has allowed us to test the accessibility of different regions of σ^A to proteases, which may enable us to obtain some information about the conformation of σ^A . The experimental data indicate the existence of a relatively protease-sensitive region in the C-terminal part of σ^A protein, and this region may be located on the surface of σ^A protein. Interestingly, the promoter -10 binding region of σ^A , which is supposed to recognize promoters, was less accessible to proteases. Since this region was also found at a hydrophobic region of σ^A protein, we suggest that the exposure of this region probably involved a change of σ^A protein conformation or the removal of the overlying C-terminal region.

MATERIALS AND METHODS

Bacterial strain and media. *E. coli* BL21 (DE 3) ($F^- ompT r_B^- m_B^-$), which has a copy of the T7 RNA polymerase gene on the chromosome (F. W. Studier, personal communica-

tion), was used as the host for overexpression of the *sigA* gene.

Construction of the plasmid for overproduction of σ^A . To construct the overproducing plasmid pCD2, the pET3 vector (Studier, personal communication), which contains the T7 promoter, was digested with *Bam*HI at 37°C for 1 h, and the 1.3-kilobase *Eco*RI-*Acc*I fragment, which contains the entire *sigA* gene, was excised from a pBM plasmid containing the *sigA* gene (unpublished result). Both DNA fragments were mixed and end filled by Klenow enzyme reaction at 25°C for 30 min, heated to 70°C for 5 min, and precipitated with ethanol. After being vacuum dried, the DNA was dissolved in ligation buffer and ligated with T4 DNA ligase at 20°C for 24 h (Fig. 1). Competent *E. coli* BL21 (DE 3) cells were made by the standard method (20). The cells were incubated at 4°C overnight before transformation experiments were carried out. After transformation, the cells were plated on Luria-Bertani medium containing 100 μ g of ampicillin per ml.

Growth of the σ^A -overproducing cells. Cells containing the pCD2 plasmid were grown in 2YT (16 g of Bacto-Tryptone, 10 g of yeast extract, and 5 g of NaCl per liter) plus 0.4% glucose at 37°C. When the cell density reached 160 Klett units, isopropyl- β -D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM. After 50 min, rifampin was added to a final concentration of 100 μ g/ml. The cells were then harvested after 2 h of incubation at 37°C.

Detection of overproduced σ^A protein by SDS-PAGE and Western blot analysis. A sample (1 ml) of harvested cell was centrifuged and the cell pellet was suspended in 100 μ l of 0.5 $\times$ SET buffer (20% sucrose, 50 mM Tris hydrochloride [pH 7.6], 50 mM EDTA) which contained 0.2 mg of lysozyme per ml. The cell suspension was vortexed and kept on ice for 1 h; 200 μ l of 2 $\times$ SAB buffer (0.125 M Tris hydrochloride [pH 6.8], 4% sodium dodecyl sulfate [SDS], 20% glycerol, 0.002% bromophenol blue) was added. The suspension was then stirred with a minimixer for 10 min and heated to 90°C for 3 min. The cell debris was spun down. Various amounts (5 to 30 μ l) of the lysate were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (10% polyacrylamide). The running system was the same as the standard Laemmli system (17). The gels were stained with 0.05% (wt/vol) Coomassie brilliant blue R and destained with 10% acetic acid. For Western blot (immunoblot) analysis, the

* Corresponding author.

FIG. 1. Construction of overexpression plasmid pCD2. The 1.3-kb *sigA* gene fragment and *Bam*HI-linearized pET3 plasmid were treated with Klenow fragment and then ligated. Symbols: →, T7 phage gene 10 promoter; ■, T7 phage terminator. The *Bam*HI site is a unique cloning site in the plasmid.

protein bands were transferred to nitrocellulose paper (0.45- μ m pore size BA 85; Schleicher & Schuell, Inc.) and blocked with blocking buffer (3% bovine serum albumin, 0.15 M NaCl, 0.001 M EDTA, 0.05 M Tris hydrochloride [pH 7.4], 0.05% Tween 20) for 5 h in a plastic bag. Afterward, the buffer was poured away and the primary antibody (polyclonal anti- σ^A) reaction was carried out in fresh blocking buffer. The reaction mixture was incubated at room temperature overnight and then washed four times with NET buffer (150 mM NaCl, 5 mM EDTA, 50 mM Tris hydrochloride [pH 7.4], 0.25% gelatin, and 0.05% Nonidet P-40) with gentle shaking. Secondary antibody reaction was carried out by adding peroxidase-conjugated goat anti-rabbit antibody to the fresh blocking buffer and the nitrocellulose paper contained in a bag. The reaction was incubated at room temperature for 3 h and then washed four times with NET buffer and once with TBS buffer (20 mM Tris hydrochloride [pH 7.4], 500 mM NaCl). The nitrocellulose paper was dipped in horseradish peroxidase color development solution (0.02% 4-chloro-1-naphthol and 0.12% hydrogen peroxide in TBS buffer) until the blue bands appeared.

Purification of σ^A protein from overproduction system. The method for protein purification was modified from that used to purify *E. coli* major sigma factor (11). Cells from 2 liters of culture (about 12 g [wet weight] of cells) were suspended in 35 ml of disruption buffer (0.05 M Tris hydrochloride [pH 8.0], 2 mM EDTA, 0.1 mM dithiothreitol (DTT), 1 mM 2-mercaptoethanol, 0.233 M NaCl, 5% glycerol, 130 μ g of lysozyme per ml) and incubated on ice for 30 min. Sodium deoxycholate was added to a 0.05% final concentration, and the cells were then sonicated until they were lysed. The lysate was diluted with 35 ml of TEDG buffer (10 mM Tris hydrochloride [pH 7.9], 0.1 mM EDTA, 0.1 mM DTT, 5% glycerol) containing 0.2 M NaCl. The cell debris was spun down with a Sorvall GSA rotor at 4°C at 8,000 rpm. The cell debris was suspended in 20 ml of TEDG buffer containing 6 M guanidine hydrochloride and spun at 15,000 rpm for 30

min at 4°C, and the supernatant was diluted gradually to 1,200 ml with TEDG buffer. The protein suspension was centrifuged, and the supernatant was diluted again to 2,200 ml with TEDG buffer. The protein in the suspension was then absorbed with 20 g of DEAE-cellulose which was equilibrated with TEDG buffer containing 0.1 M NaCl. The DEAE-cellulose column was packed, and an NaCl salt gradient from 0.15 to 0.45 M was applied; fractions were checked by SDS-PAGE, and those with σ protein were pooled and precipitated with 40% (wt/vol) ammonium sulfate. The protein pellet was dissolved in 3 ml of TEDG-0.2 M NaCl buffer and run on a Sephadex G-75 column equilibrated with the same buffer. Fractions containing σ were pooled and dialyzed against storage buffer (20 mM Tris hydrochloride [pH 7.9], 10 mM MgCl_2 , 1 mM EDTA, 0.1 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride, 30% glycerol). The amount of protein was determined with a Bio-Rad protein determination kit with plasma gamma globulin as the standard.

In vitro transcription with reconstituted holoenzyme. (i) **Procedures for autoradiography.** Core was obtained as described previously (4, 8, 13). In order to form holoenzyme, core enzyme (3 μ g) was mixed with 1.5 μ g of purified σ in a microcentrifuge tube and incubated in an ice bath for 15 min. To this mixture, 0.3 μ g of *Hind*III-digested ϕ 29 DNA was added. The enzyme and DNA solution were incubated on ice for another 10 min, and 50 μ l of reaction cocktail mixture was added to start the transcription reaction (the final concentrations were 40 mM Tris hydrochloride [pH 7.9]; 10 mM MgCl_2 ; 125 mM KCl; 20 mM DTT; 0.2 mM each UTP, CTP, GTP, and ATP; 5% glycerol; and 3 μ Ci of [α - 32 P]ATP). The transcription reaction was incubated at 37°C for 10 min and then placed immediately into an ice bath. *E. coli* tRNA (4 μ g) was then added to the reaction mixture, which was extracted twice with phenol. Solution (40 μ l) in upper aqueous phase was then pipetted into a new tube, and the RNA transcripts were precipitated by the addition of 5 μ l of 3 M sodium acetate and 2 volumes of 100% ethanol at -70°C for 1 h. The pellet was then dissolved in a solution containing 8 M urea, 0.01% bromophenol blue, and 0.015% xylene cyanol solution. After treatment at 65°C for 3 min, the samples were applied to an 8 M urea denaturing gel and run at 170 V. Autoradiography was done after the electrophoresis.

(ii) **Procedures for scintillation counting of 32 P-labeled RNA transcript.** Holoenzyme reconstitution and transcription reaction were as described above. The reaction mixture contained 40 mM Tris hydrochloride [pH 7.9], 10 mM MgCl_2 , 100 mM KCl, 4 mM DTT, 0.4 mM nucleoside triphosphates, 1 μ Ci of [3 H]UTP, 0.1 mg of gelatin per ml, 10% glycerol, 0.016 pmol of *Hind*III-digested ϕ 29 DNA, and 6 pmol of core polymerase with increasing amounts of σ^A factor (see Fig. 5). After the reaction, 50 μ g of calf thymus DNA was added to the mixture and the transcripts were precipitated with 10% trichloroacetic acid by keeping the mixture on ice for 20 min. The transcript was filtered through GF/C filter paper and washed with 5% trichloroacetic acid. The filter paper was rinsed with 95% ethanol, and the radioactivity was counted with a scintillation counter.

Peptide mapping and N-terminal sequence determination. Trypsin and chymotrypsin were used in these experiments. Different amounts of protease (10 to 100 μ g) were added to 10 μ g of σ protein solution in TEDG-50 mM NaCl buffer. The reaction mixture was incubated at room temperature for 30 min; phenylmethylsulfonyl fluoride was then added to make a final concentration of about 1 mM. Afterwards, 2 $\times$ sample application buffer was added. After SDS-PAGE,

FIG. 2. Overexpression of σ^A protein. Lanes: A, D, and G, cells grown at 37°C in 2YT plus 0.4% glucose medium containing 100 μ g of ampicillin per ml; B, E, and H, as for lanes A, D, and G, except that IPTG was added to 0.4 mM when cell density reached 160 Klett units; C, F, and I, as for lanes B, E, and H, except that rifampin was added to 100 μ g/ml 1 h after IPTG induction. The *E. coli* strain and the plasmid in the cell are indicated at the bottom. All cells were harvested 2 h after rifampin was added to lanes C, F, and I. A sample (0.1 ml) from each cell lysate was run on SDS-PAGE (10% polyacrylamide). The numbers in the left margin indicate masses (in kilodaltons) of molecular weight markers. α , β , and β' are subunits of RNA polymerase core enzyme.

protein bands were transferred to polyvinylidene difluoride (PVDF) membrane and stained with Coomassie blue (21). Bands were cut out of the membrane, and their N-terminal sequences were analyzed by an automated gas-phase protein sequenator at the University of California, Davis, Protein Structure Laboratory.

Materials. Restriction enzymes, Klenow enzyme, and T4 DNA ligase were purchased from Bethesda Research Laboratories, Inc. Reagents for SDS-PAGE and DEAE-Affi-Gel Blue column were from Bio-Rad Laboratories. DEAE-cel-

lulose (DE 52) was from Whatman, Inc. Sephadex G-75 was from Pharmacia. Trypsin and chymotrypsin were from Worthington Diagnostics. Guanidine hydrochloride, 4-chloro-1-naphthol, and hydrogen peroxide were from Sigma Chemical Co. Yeast extract and Bacto-Tryptone were from Difco. Nitrocellulose paper was from Schleicher & Schuell. PVDF membrane was from Millipore Corp.

RESULTS

Cloning and overexpression of σ^A protein. From 24 mini-screens of the *E. coli* BL21 transformants, we found two clones with the *sigA* gene insert in the correct orientation. The two clones were stable in the absence of IPTG. However, they became unstable when IPTG was added to the culture. In liquid cultures about 12 min after IPTG was added, the growth of cells slowed and, more significantly, the viable cell number dropped rapidly as the cells started to accumulate σ^A protein (data not shown). The results readily show the overproduced sigma protein at a position of about 55 kilodaltons (kDa) (Fig. 2, lanes H and I). When the protein patterns are compared, it is obvious that the addition of rifampin decreases the synthesis of other cellular proteins. On the basis of the results of densitometer scanning of the gel, it was found that the overproduced sigma made up about 40% of the total cellular protein.

Purification of σ^A protein from the overproduction system. The overproduced σ was found to be in the form of inclusion bodies and was found in the cell debris when the cell was broken. In the first purification step, the σ protein in the cell debris was dissolved by treatment with 6 M guanidine hydrochloride. In this step, we found that most of the other cellular proteins were absent and the recovery rate of σ^A was about 62.5% (Fig. 3A and Table 1). The second step was to remove other proteins by using a DEAE-cellulose column. The protein with a molecular weight of around 43,000 (probably elongation factor Tu [EF-Tu] of *E. coli*) was one of the major contaminants that we wanted to remove in this step. By using a shallow NaCl gradient from 0.15 to 0.45 M in TEDG buffer, σ^A protein was found to be eluted at around 0.24 to 0.28 M NaCl and most of the contaminants were removed. At this step, the purity of sigma protein was about 95% and the recovery rate was about 57.5% (Fig. 3A and

FIG. 3. Purification of major σ^A protein (a) and Western blot analyses (b). Lanes: A, crude lysate of host cell, BL21 (DE 3); B, crude lysate of host cell harboring pET3 plasmid; C, crude lysate of host cell harboring pCD2 plasmid; D, diluted supernatant of guanidine hydrochloride extract of cell debris from cells carrying pCD2; E, DEAE-cellulose-purified fractions of σ^A protein; F, Sephadex G-75-purified fractions of σ^A protein. Purification steps and Western blot analyses were performed as indicated in Materials and Methods. α , β , and β' are subunits of RNA polymerase core enzyme.

TABLE 1. Purification of σ^A

Purification step	Yield ^a of σ^A (mg)	Recovery (%)	Purity (%)
Cell lysate	240	100	40
Guanidine hydrochloride extracts	150	63	75
DEAE-cellulose fractions	138	58	95
Sephadex G-75 fractions	93	39	99

^a From 1 liter of cell culture (6 g [wet weight] of cells).

Table 1). The last purification step (Sephadex G-75) removed contaminating *E. coli* RNA polymerase core enzyme and sigma factor. The protein purity was above 99%, as determined by densitometric analysis (Fig. 3A, lane F). By using this scheme, we have been able to purify about 90 mg of 99% pure σ protein from 1 liter of culture (about 6 g [wet weight] of cells).

Authenticity of purified σ^A protein. Several tests were run to determine the properties of the overproduced σ factor. The overproduced protein, when run on SDS-PAGE, showed the same mobility as the native σ^A (Fig. 2 and 3). Western blot analysis revealed that the purified σ had the same antigenicity as the native σ (Fig. 3B, lanes C through F) and the RNA polymerase (Fig. 3B, control lane). Also, a polyclonal antibody which was raised against the overproduced σ reacted with σ^A factor from *B. subtilis* cells after DEAE-Affi-Gel Blue column purification and host protein adsorption (data not shown). In order to test the promoter specificity of the overproduced sigma factor, we compared the transcription patterns of native holoenzyme and the reconstituted holoenzyme containing the overproduced sigma factor by using the early promoters of $\phi 29$ phage as the template. The transcription patterns are shown in Fig. 4. The four major bands (about 80, 130, 280, and 450 bases), which are transcripts from the early $\phi 29$ promoters, were present in lanes containing native RNA polymerase (lane A), native RNA polymerase plus σ^A (lane B), and core enzyme plus σ^A (lane E). There were no clear band patterns present in control lanes (lanes C and D). The effect of σ concentra-

FIG. 4. Runoff transcription of $\phi 29$ phage early promoters by using reconstituted holoenzyme. Lanes: A, RNA polymerase holoenzyme isolated from *B. subtilis*; B, RNA polymerase holoenzyme and σ^A ; C, σ^A only; D, core enzyme only; E, core enzyme and σ^A factor. The assay procedure is described in Materials and Methods. Band sizes (in bases) are shown on the right.

FIG. 5. Effect of σ^A concentration on transcription activity of reconstituted RNA polymerase holoenzyme. The activity of core polymerase on $\phi 29$ *Hind*III-digested DNA is 230 cpm. The assay procedure is described in Materials and Methods. The quantities of RNA polymerase core enzyme and σ^A used in this experiment were determined with a Bio-Rad protein determination kit.

tion on transcription was confirmed by observing in vitro transcription with the same template by using RNA polymerase holoenzymes with different molar ratios of σ to core enzyme. An increasing transcription activity was found (Fig. 5). These results indicate that the overproduced σ^A has the same specificity of promoter recognition as the native σ^A and that RNA polymerase transcription activity is enhanced by increasing concentrations of σ^A . Another line of evidence which demonstrated the authenticity of the purified σ^A was the analysis of the N-terminal sequences of the proteolytic fragments of σ^A . Two tryptic fragments (47 and 35 kDa) (Fig. 6) and one chymotryptic fragment (42 kDa) (Fig. 7) were found to have Ala, Asp, Lys, Gln, and Thr at their N termini. This sequence is the same as that revealed from DNA sequence analysis (10), except that the N-terminal methionine residue was absent from this fragment.

Peptide mapping of σ^A protein. To determine whether the σ^A protein contains domains which are resistant or sensitive to proteolytic attack, the purified σ^A protein was partially digested with trypsin and chymotrypsin. The peptide fragments were run on SDS-PAGE and transferred to PVDF membrane. Three tryptic fragments with apparent M_r s of about 47,000, 35,000, and 20,000 (Fig. 6) and three chymo-

FIG. 6. Proteolysis of the purified σ^A protein by trypsin. Lane A, Molecular weight markers; lanes B through E, tryptic fragments generated in the presence of different amounts of trypsin. The amounts of trypsin added for 10 μ g of σ protein are 10 ng (lane B), 25 ng (lane C), 50 ng (lane D), and 100 ng (lane E). The numbers in the left and right margins indicate apparent molecular masses in kilodaltons.

FIG. 7. Proteolysis of purified σ^A protein by chymotrypsin. Lane A, Molecular weight markers; lanes B through E, chymotryptic fragments generated in the presence of different amounts of chymotrypsin. The amounts of chymotrypsin added for 10 μ g of sigma protein were 10 ng (lane B), 25 ng (lane C), 50 ng (lane D), and 100 ng (lane E). The numbers in the left and right margins indicate apparent molecular masses in kilodaltons.

tryptic fragments with apparent M_r s of about 42,000, 22,000, and 18,000 (Fig. 7) were cut out of the membrane and sequenced at their N termini; their sequences were aligned with that of the native σ^A (Table 2 and Fig. 8). From these results (Fig. 6 and 7), we found that the 47-kDa tryptic fragment and the 42-kDa chymotryptic fragment appeared when limiting amounts of protease were added. These results indicated that some sequence in the C-terminal part of the σ^A protein was very accessible to proteases and was easily attacked. In addition, the 20-kDa tryptic fragment and the 18-kDa chymotryptic fragment were relatively resistant to protease digestion. The N termini of both fragments were within the proposed promoter -10 binding region (region 2.4) of σ proteins (15). They also seemed to comprise one of the most hydrophobic regions of the *B. subtilis* major σ

TABLE 2. N-terminal amino acid sequences of peptide fragments obtained after proteolysis of σ^A

Fragment size (kDa)	N-terminal sequence	Cleavage position ^a
55	Ala-Asp-Lys-Gln-Thr	2
47	Ala-Asp-Lys-Gln-Thr	2
35	Ala-Asp-Lys-Gln-Thr	2
20	Thr-Ile-Arg-Ile-Pro	208
42	Ala-Asp-Lys-Gln-Thr	2
22	Ala-Gln-Lys-Ile-Glu	124
18	Arg-Ala-Ile-Ala-Asp	200

^a Amino acid number from the amino acid sequence of σ^A protein.

(region 2.4 and the connecting sequence between regions 2 and 3) (Fig. 8 and 9).

DISCUSSION

The successful overexpression of the *sigA* gene in *E. coli* was dependent on eliminating a basic level of *sigA* gene expression which was lethal to the host cell. Previous attempts with the lambda p_L or *lac* promoter-controlled systems were not successful, since they were too leaky and the σ^A produced during growth was lethal for the cell. However, the use of a *lacUV5* promoter-controlled T7 RNA polymerase gene in the *E. coli* chromosome and a T7 RNA polymerase promoter-controlled *sigA* in the multicopy pCD2 plasmid allowed us to synthesize large amounts of σ^A . The fact that strain BL21 (DE 3) [pCD2] grew at the same rate as the parental strain BL21 (DE 3) indicates that very little or no leakiness of the *sigA* gene occurred prior to induction with IPTG. In most overexpression systems, the levels of overproduced proteins ranged from 1 to 50% of the total cell protein (16). In our case, the 40% yield of σ^A was probably due to the high transcription rate by T7 RNA polymerase, the very good ribosome-binding sequence for the *sigA* gene, the formation of protease-resistant inclusion bodies, and the strong inhibition of *E. coli* RNA polymerase by rifampin.

FIG. 8. Peptide mapping of σ^A protein. (A) Tryptic and chymotryptic fragments which were N-terminally sequenced (the molecular mass of each fragment is labeled as in Fig. 6 and 7). The numbers at the left indicate the molecular weights of the peptide fragments as determined by SDS-PAGE mobility. The C terminus of each fragment was not determined, and therefore the exact size of each fragment is not known. The length of each fragment in the figure was calculated by using the following equation: length of $\sigma^A \times$ (apparent molecular mass [in kDa] of each fragment/55 kDa). a.a., Amino acid number. (B) , Hydrophobic regions in σ protein. The left bar has 54% and the right bar has 50% hydrophobic amino acid residues. (C) Charge distribution of σ protein. The net charge is indicated below each region.

FIG. 9. Hydrophilicity (A) and hydrophobicity (B) profiles of σ^A protein. (A) The area above the zero line is the more hydrophilic region of σ^A protein; (B) the area above the zero line is the more hydrophobic region of σ^A protein. The 2.4 region contains amino acids 193 to 212 of the σ^A protein sequence.

The purification of σ^A was made relatively simple by the formation of σ^A inclusion bodies, which are characteristic for many overproduced proteins (16, 23). Since proteins in inclusion bodies are not active, it was necessary to extract the σ^A from the inclusion bodies and cell debris by the use of denaturing and renaturing conditions (12, 23). The denaturing agent guanidine hydrochloride was used successfully in extracting *E. coli* σ^A from the inclusion bodies (11) and in renaturing σ^A to an active form (12). The use of anion exchange chromatography and a molecular sieve column removed the major contaminant (EF-Tu) and other, minor protein contaminants from σ^A .

Several criteria were used to prove the authentic nature of *E. coli*-produced σ^A . By Western blot analysis, the purified σ^A was found to have the same mobility and antigenicity as native *B. subtilis* σ^A . Minor bands in the σ^A preparation that cross-reacted with anti- σ^A probably resulted from the proteolytic degradation of the overproduced protein in the cell (6, 14) or from chemical degradation during SDS-PAGE. The most convincing evidence of the authentic nature of σ^A was found during RNA polymerase holoenzyme reconstitution studies, in which the addition of purified σ^A to core enzyme resulted in the expected transcription pattern of known *B. subtilis* σ^A promoters. There was also a direct relationship between increasing transcription and increasing levels of σ^A . Thus, the specificity of overproduced σ^A in promoter recognition was identical to that of native σ^A .

Partial protease digestion can be used to study protein conformation (7, 26, 27) and to detect peptide chains exposed on the surface of a protein molecule (22). Peptide sequences in tightly folded domains of a protein can be detected by using larger amounts of protease or longer digestion periods, which resulted in proteolytic fragmentation (22). Peptide mapping experiments with *E. coli* σ^{70} showed that a 40-kDa trypsin-resistant peptide fragment was capable of binding to core enzyme (1, 11, 19). In another study of *E. coli* σ^{70} , it was found that the abnormal mobility of σ^{70} was due to the very acidic nature of the N-terminal half of the protein. In addition, the very positive or the very negative charge of proteolytic fragments may affect their mobilities on SDS-PAGE (29).

B. subtilis σ^A factor has a molecular weight of 43,000 (10); however, it moves to a 55-kDa position in SDS-PAGE. This

abnormal mobility is probably due to a highly negative charge in the N-terminal part of the σ^A protein (Fig. 8). Since there are positive- and negative-charge clusters in the σ^A protein (Fig. 8), it is also very possible that the proteolytic fragments have abnormal mobilities on SDS-PAGE, as seen for *E. coli* σ^{70} proteolytic fragments (29). In the partial proteolysis of σ^A protein, a 47-kDa tryptic fragment and a 42-kDa chymotryptic fragment were found after treatment with low amounts of protease; both had the same N-terminal sequence as native σ^A (Table 2 and Fig. 8). The 47-kDa fragment did not have sigma activity, as tested by runoff transcription assays (data not shown). These results indicate that there is a very protease-sensitive surface region in the C-terminal part of σ^A and that removal of this C-terminal region affects the transcription activity of RNA polymerase.

Subsequently, a 20-kDa tryptic fragment and an 18-kDa chymotryptic fragment were generated. The two fragments were very resistant to protease digestion and had their N-terminal sequence in the proposed promoter -10 binding region of σ^A . In addition, they contained one of the most hydrophobic regions of σ^A , region 2.4 and the connecting region between regions 2 and 3 (15) (Table 2 and Fig. 8 and 9). This sequential proteolysis of σ^A suggests that the -10 binding region (region 2.4) is probably in a tightly folded hydrophobic domain of σ^A and that its accessibility to protease probably needs modification of σ^A protein conformation.

Since the 2.4 region of σ^A is involved in promoter -10 region recognition (3), its low protease accessibility seems to indicate a change of σ^A protein conformation during σ^A and core enzyme association or during holoenzyme and promoter recognition. This question is currently under investigation in our laboratory.

The absence of aggregation of σ^A at relatively high concentrations (data not shown) under denaturing and nondenaturing conditions and the analysis of partial proteolysis of σ^A have now provided information and a rationale for attempts to crystallize the σ^A protein.

ACKNOWLEDGMENTS

We thank W. Studier for providing us with the T7 expression system.

This research was supported in part by Public Health Service grant GM19673 from the National Institutes of Health.

LITERATURE CITED

- Burgess, R. R., B. Erickson, D. Gentry, M. Gribskov, D. Hager, S. Lesley, M. Strickland, and N. Thompson. 1987. Bacterial RNA polymerase subunits and genes, p. 13-16. In W. S. Reznikoff, R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr., and M. P. Wickens (ed.), RNA polymerase and the regulation of transcription. Elsevier Science Publishing, Inc., New York.
- Burgess, R. R., A. A. Travers, J. J. Dunn, and E. K. F. Bautz. 1969. Factor stimulating transcription by RNA polymerase. *Nature (London)* 221:43-46.
- Chamberlin, M. J. 1974. The selectivity of transcription. *Annu. Rev. Biochem.* 43:721-775.
- Davison, B. L., T. Leighton, and J. Rabinowitz. 1979. Purification of *Bacillus subtilis* RNA polymerase with heparin agarose. *J. Biol. Chem.* 254:9220-9226.
- Doi, R. H., and L.-F. Wang. 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. *Microbiol. Rev.* 50:227-243.
- Fallon, A. M., C. S. Jinks, M. Yamamoto, and M. Nomura. 1979. Expression of ribosomal protein genes cloned in a hybrid plasma in *Escherichia coli*: gene dosage effects on synthesis of ribosomal proteins and ribosomal protein messenger ribonucleic

- acid. *J. Bacteriol.* **138**:383–396.
7. Fisher, R., and T. Blumenthal. 1980. Analysis of RNA polymerase by trypsin cleavage. *J. Biol. Chem.* **255**:11056–11062.
 8. Fukuda, R., and R. H. Doi. 1977. Two polypeptides associated with the ribonucleic acid polymerase core of *Bacillus subtilis* during sporulation. *J. Bacteriol.* **129**:422–432.
 9. Gardella, T., H. Moyle, and M. M. Susskind. 1989. A mutant *Escherichia coli* σ^{70} subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.* **206**:579–590.
 10. Gitt, M. A., L.-F. Wang, and R. H. Doi. 1985. A strong sequence homology exists between the major RNA polymerase sigma factors of *Bacillus subtilis* and *Escherichia coli*. *J. Biol. Chem.* **260**:7178–7185.
 11. Gribskov, M., and R. R. Burgess. 1983. Overexpression and purification of the sigma subunit of *E. coli* RNA polymerase. *Gene* **26**:109–118.
 12. Hager, D. A., and R. R. Burgess. 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzyme activity: result with sigma subunit of *Escherichia coli* RNA polymerase, wheat germ topoisomerase, and other enzymes. *Anal. Biochem.* **109**:76–86.
 13. Halling, S. M., F. J. Sanchez-Anzaldo, R. Fukuda, R. H. Doi, and C. F. Meares. 1977. Zinc is associated with the β subunit of DNA-dependent RNA polymerase of *Bacillus subtilis*. *Biochemistry* **16**:2880–2884.
 14. Hayward, R. S., S. J. Austin, and J. G. Scaife. 1974. The effect of gene dosage on the synthesis and stability of RNA polymerase subunits in *Escherichia coli*. *Mol. Gen. Genet.* **131**:173–180.
 15. Helman, J. D., and M. J. Chamberlin. 1988. Structure and function of bacteria sigma factors. *Annu. Rev. Biochem.* **57**:839–872.
 16. Kane, J. F., and D. L. Hartley. 1988. Formation of recombinant protein inclusion bodies in *Escherichia coli*. *Trends Biotechnol.* **6**:95–101.
 17. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
 18. Lesley, S. A., and R. R. Burgess. 1989. Characterization of the *Escherichia coli* transcription factor σ^{70} : localization of a region involved in the interaction with core polymerase. *Biochemistry* **28**:7728–7734.
 19. Lowe, P. A., D. A. Hager, and R. R. Burgess. 1979. Purification and properties of the sigma subunit of *Escherichia coli* DNA-dependent RNA polymerase. *Biochemistry* **18**:1344–1352.
 20. Maniatis, T. E., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
 21. Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membrane. *J. Biol. Chem.* **262**:10035–10038.
 22. Mihalyi, E. 1978. Application of proteolytic enzymes to protein structure study, 2nd ed., vol. 1. CRC Press, Inc., Boca Raton, Fla.
 23. Mitraki, A., and J. King. 1989. Protein folding intermediates and inclusion body formation. *Trends Biotechnol.* **7**:690–697.
 24. Reznikoff, W. S., D. A. Siegele, D. W. Cowing, and C. A. Gross. 1985. The regulation of transcription initiation in bacteria. *Annu. Rev. Genet.* **19**:355–387.
 25. Rosenberg, A. H., B. N. Lade, D. S. Chui, S.-W. Lin, J. J. Dunn, and F. W. Studier. 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**:125–135.
 26. Rupley, J. A. 1967. Susceptibility to attack by proteolytic enzymes. *Methods Enzymol.* **11**:905–917.
 27. Scheinker, V. S., S. F. Beresten, S. K. Degtyaev, and L. L. Kisselev. 1979. The effect of tRNA and tryptophanyl adenylate on limited proteolysis of beef pancreas tryptophanyl-tRNA synthetase. *Nucleic Acids Res.* **7**:625–637.
 28. Siegele, D. A., J. C. Hu, W. A. Walter, and C. A. Gross. 1989. Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **206**:591–603.
 29. Strickland, M. S., N. E. Thompson, and R. R. Burgess. 1988. Structure and function of the sigma-70 subunit of *Escherichia coli* RNA polymerase. Monoclonal antibodies: localization of epitopes by peptide mapping and effect on transcription. *Biochemistry* **27**:5755–5762.
 30. Studier, F. W., and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113–130.
 31. Zuber, P., J. Healy, H. L. Carter, S. Cutting, C. P. Moran, and R. Losick. 1989. Mutation changing the specificity of a RNA polymerase sigma factor. *J. Mol. Biol.* **206**:605–614.