THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 11, Issue of March 14, pp. 7201–7210, 1997

Printed in U.S.A.

Domain Organization of *Escherichia coli* Transcript Cleavage Factors GreA and GreB*

(Received for publication, September 18, 1996, and in revised form, December 10, 1996)

Dmitry Koulich‡§¶, Marianna Orlova¶‖, Arun Malhotra, Andrej Sali**, Seth A. Darst**, and Sergei Borukhov‡ ‡‡**

*From the ‡Department of Microbiology and Immunology, State University of New York, Health Science Center at Brooklyn, Brooklyn, New York 11203, the ‖Public Health Research Institute, New York, New York 10016, and **The Rockefeller University, New York, New York 10021*

The GreA and GreB proteins of *Escherichia coli* induce cleavage of the nascent transcript in ternary elongation complexes of RNA polymerase. Gre factors are presumed to have two biologically important and evolutionarily conserved functions: the suppression of elongation arrest and the enhancement of transcription fidelity. A three-dimensional structure of GreB was generated by homology modeling on the basis of the known crystal structure of GreA. Both factors display similar overall architecture and surface charge distribution, with characteristic C-terminal globular and N-terminal coiled-coil domains. One major difference between the two factors is the "basic patch" on the surface of the coiled-coil domain, which is much larger in GreB than in GreA. In both proteins, a site near the basic patch cross-links to the 3′ terminus of RNA in the ternary transcription complex. GreA/GreB hybrid molecules were constructed by genetic engineering in which the N-terminal domain of one protein was fused to the C-terminal domain of the other. In the hybrid molecules, both the coiled-coil and the globular domains contribute to specific binding of Gre factors to RNA polymerase, whereas the antiarrest activity and the GreA or GreB specificity of transcript cleavage is determined by the N-terminal domain. These results implicate the basic patch of the N-terminal coiled-coil domain as an important functional element responsible for the interactions with nascent transcript and determining the size of the RNA fragment to be excised during the course of the cleavage reaction.

Two closely related *Escherichia coli* proteins, GreA and GreB, participate in RNA polymerase (RNAP)[1] transcription elongation by preventing and/or suppressing the condition of elongation arrest (1, 2). In addition, both factors have been shown to facilitate the transition of RNAP from the stage of abortive initiation to productive elongation (3). GreA and GreB may also have a proofreading role in transcription (4). It is thought that the Gre activity is accomplished by endonucleolytic cleavage of RNAs within the ternary elongation complexes (TECs) (2). The cleavage is followed by dissociation of the 3′-terminal fragment and restart of elongation from a newly generated 3′-OH terminus (2). Similar reactions are induced in eukaryotic TECs of RNA polymerase II by transcription elongation factor TFIIS (5–8), which performs the same functions as GreA and GreB but lacks any sequence similarity.

GreA and GreB have almost the same molecular mass and share substantial amino acid sequence homology (2). However, there are several differences in their functional and biochemical properties *in vitro*. First, in all studied TECs that are susceptible for cleavage reactions, GreA induces hydrolysis of short RNAs 2 or 3 nucleotides long from the RNA 3′ terminus ("type A" cleavage activity), whereas GreB stimulates hydrolysis of RNAs that are 2–18 nucleotides long, depending on the stage of transcription elongation ("type B" cleavage activity) (2, 9). Second, GreA can only prevent the formation of arrested TEC (read-through activity), whereas GreB, besides displaying read-through activity, can convert the preformed arrested TEC to a productive complex by a cleavage and restart mechanism (2). Finally, we have recently shown that Gre factors form reversible complexes with RNAP core or holoenzyme with apparent K_d values that differ by at least 2 orders of magnitude: 2–3 × 10^{-5} M (GreA) and 1–2 × 10^{-7} M (GreB). These values were estimated by size exclusion high pressure liquid chromatography of RNAP complexes with ^{35}S-labeled Gre proteins and were further supported by affinity chromatography of ^{35}S-labeled Gre factors on immobilized His-tagged RNAP (10). Similar K_d values were obtained for the complexes of Gre proteins with purified TECs (prepared on a ribosomal *rrnB* P1 promoter DNA) carrying 6-mer and 12-mer transcripts (10).

TFIIS and Gre factors do not cleave free RNA or RNA·DNA duplexes (2, 10) and require the formation of a ternary complex between RNAP, DNA, and RNA transcript for their activity (10) or a specific binary complex between RNAP and RNA (11). On the other hand, both *E. coli* RNAP and eukaryotic RNAP II show a weak ability to cleave RNA in the absence of factors, and this activity is stimulated by mild alkaline pH (12) or pyrophosphate (13). These data suggest that the transcript cleavage factors are not nucleases *per se* but cofactors that stimulate an intrinsic nucleolytic activity of RNAP. A working model that emerged from these data envisages the RNAP catalytic center as the performer of the cleavage reaction (12, 13). In this model, the catalytic center disengages from the RNA 3′ terminus, slips back along the transcript, hydrolyzes or pyro-

* This work was supported, in part, by National Institutes of Health (NIH) Grant GM54098 (to S. B.), NIH Grant GM49242 (to A. Goldfarb), and Russian Foundation for Fundamental Research Grant 96-04-49-019 (to V. Nikiforov). 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: Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, 123182 Russia.

¶ These two authors contributed equally to this work.

‡‡ To whom correspondence should be addressed: Dept. of Microbiology and Immunology, State University of New York, Health Science Center at Brooklyn, 450 Clarkson Ave., Brooklyn, NY 11203. Tel.: 718-270-3752; Fax: 718-270-2656; E-mail: serbor@aol.com.

[1] The abbreviations used are: RNAP, RNA polymerase; 8-N_3-ATP, 8-azidoadenosine triphosphate; TEC, ternary elongation complex; BSA, bovine serum albumin; NTCB, 2-nitro-5-thiocyanobenzoic acid; WT, wild type; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

phosphorolizes an internal phosphodiester bond and reengages with the newly generated 3′ terminus. The model is consistent with the notion of conformational flexibility of the TEC that is central to the discontinuous mechanism of RNA polymerization (14, 15).

Understanding the mechanism by which Gre proteins induce these perturbations within the TEC and stimulate cleavage would help unravel the mechanism of elongation. In a previous paper, we reported the 2.2-Å resolution crystal structure of GreA (16), which suggested a way this protein might interact with the TEC. The GreA structure consists of the C-terminal globular and N-terminal extended coiled-coil domains, resembling a fist with an outstretched index finger. One surface of the molecule is uniformly negatively charged, while the opposite surface is essentially hydrophobic with a small basic patch near the tip of the finger. We suggested that during interaction with the TEC, GreA's acidic side is oriented away from the acidic surface of RNAP, while the basic patch contacts RNA in the complex.

In the present work, this model was further explored through comparative analysis of GreA and GreB domain organization and function. To this end, we sought to relate the similarities and differences between biochemical functions of the two factors to the common and divergent features of their structures.

EXPERIMENTAL PROCEDURES

Materials—Nucleoside triphosphates, RNase-free BSA, sequencing grade endoproteinases ArgC, AspN, and thrombin were from Boehringer Mannheim. [α-^{32}P]CTP and [α-^{32}P]GTP (3000 Ci/mmol) were purchased from ICN. 8-N_3ATP, 2-nitro-5-thiocyanobenzoic acid (NTCB), and BNPS-skatole were from Sigma. Prestained protein molecular weight markers were from Amersham Corp.

Strains and Plasmids—E. coli greA$^-$greB$^-$ strain AD8571 (12) was used for overexpression of wild type (WT) and mutant Gre factors and for preparation of GreA/GreB-free RNAP. *E. coli* strain XL1-Blue (Stratagene) was used for selection of all plasmids after ligation. pTRC99A (Pharmacia Biotech Inc.) was used as an expression vector for construction of plasmids overproducing WT and mutant Gre proteins. Plasmid pCF3 is a derivative of pUC19 containing an insertion of the cloned *greB* gene (2). All sequences of WT and mutant *gre* genes described in this work were verified by dideoxy chain termination sequencing of the double-stranded plasmid DNA.

Oligonucleotides—The following oligonucleotides were used for construction of overexpressing plasmids and PCR mutagenesis (italicized nucleotides correspond to *greA* and *greB* genes, as indicated in parentheses; underlined nucleotides define the restriction sites of the enzymes, shown in parentheses; boldface nucleotides correspond to the region coding the His_6 tag): 1) 5′-*pCAAGCTATTCCGATGACCTTA*-3′ (*GreA 5′ terminus*); 2) 5′-CGCGC<u>GGATCC</u>*TTACAGGTATTCCACCTTAATTAC*-3′ (*GreA 3′ terminus*, <u>*Bam*HI</u>); 3) 5′-GCGCG<u>TCATGA</u>*AAACGCCCCTGGTTACC*-3′ (*GreB 5′ terminus*, <u>*Bsp*HI</u>); 4) 5′-CGCGC<u>GGATCC</u>*TTACGGTTTCACGTACTCGATAGC*-3′ (*GreB 3′ terminus*, <u>*Bam*HI</u>); 5) 5′-CGCGC<u>GGATCC</u>TTA**GTGATGGTGATGGTGATG***CAGGTATTCCACCTTAATTAC*-3′ (*GreA 3′ terminus*, <u>*Bam*HI</u>); 6) 5′-CGCGC<u>GGATCC</u>TTA**GTGATGGTGATGGTGATG***CGGTTTCACGTACTCGATAGC*-3′ (*GreB 3′-terminus*, <u>*Bam*HI</u>).

Construction of Plasmids Overproducing WT GreA, GreB, and His-tagged Gre Proteins—-Oligonucleotides 1 and 2, synthesized according to the DNA sequence of the *greA* gene (17), were used for PCR amplification of *greA* from *E. coli* DNA. The PCR product was purified by electrophoresis and digested with *Bam*HI. The expression vector pTRC99A was digested with *Nco*I, filled in by Klenow fragment of DNA polymerase I, and then cleaved by *Bam*HI. The *greA* fragment was inserted into the expression vector pTRC99A between the blunt end and *Bam*HI sites. Ligation of these fragments yielded plasmid pMO1.1, which contained the *greA* gene under the control of isopropyl-1-thio-β-D-galactopyranoside-inducible *trc* promoter. For engineering of the plasmid overproducing GreB, oligonucleotides 3 and 4 were used for PCR amplification of the *greB* gene from the plasmid pCF3. The purified PCR product was digested with *Bsp*HI and *Bam*HI and inserted into linearized pTRC99A between the *Nco*I and *Bam*HI sites. Ligation of these fragments yielded pMO1.4, which contained the *greB* gene under the control of the *trc* promoter. Plasmids overproducing His_6-tagged GreA (pMO1.1His) and GreB (pMO1.4His) were obtained similarly, except that oligonucleotides 5 and 6 carrying His_6 codons at the end of each *gre* gene were used for PCR.

Construction of GreA/GreB Hybrids—We engineered a set of 10 chimeric GreA/GreB molecules that carry N-terminal polypeptide fragments of one Gre protein fused to the C-terminal part of the other at amino acid positions 6, 12, 31, 81, and 107 of GreA. Ten hybrid primers were synthesized: the 5′-proximal half of each primer corresponded to the DNA fragment of one *gre* gene immediately upstream of the fusion and the 3′-proximal half corresponded to the fragment of the other *gre* gene immediately downstream of the fusion. All *greA/greB* hybrids were obtained by a conventional two-step PCR procedure.

Protein Purification—-WT GreA and GreB were purified to apparent homogeneity as described (18). For isolation of His_6-tagged GreA, *greA$^-$greB$^-$ E. coli* (AD8571) cells transformed with pMO1.1His were grown and induced with isopropyl-1-thio-β-D-galactopyranoside as described (18). The cells were pelleted by low speed centrifugation, homogenized in 5 ml of buffer A (7 M guanidine HCl, 40 mM Tris-HCl, pH 7.5, 0.8 M NaCl, 1 mM EDTA, and 1 mM dithiothreitol) and centrifuged at 25,000 × *g* for 20 min. The supernatant was added to 0.8 ml of Ni^{2+}-chelating NTA-agarose beads (Qiagen) preequilibrated in buffer A. The suspension was incubated for 20 min at 25 °C on a rotary shaker, and the beads were pelleted by low speed centrifugation and washed with 3 × 10 ml of buffer A. The adsorbed GreA was renatured by five subsequent washings of the beads with 5 ml of buffer B (the same as buffer A but without guanidine HCl) and eluted with 2 ml of buffer B, containing 0.6 M imidazole. The eluate was desalted and concentrated in Centricon-10 (Amicon). The resulting electrophoretically homogeneous His_6-tagged GreA (1.2 mg in 200 μl) was stored at −20 °C in buffer B containing 50% glycerol. His_6-tagged GreB and GreA/GreB hybrids were purified essentially as described above.

GreA/GreB-free RNAP holoenzyme ($\alpha,\beta\beta'\sigma$) was isolated from *E. coli* AD8571 strain, and RNAP carrying an insertion of 130 amino acids in β-subunit (β_{ins}RNAP) was isolated from *E. coli* MKDC747 strain as described (12).

In Vitro Transcription Assays—The read-through, antiarrest, and transcript cleavage assays were performed using a 202-base pair *E. coli rrnB* P1 DNA fragment as described previously (1, 2, 18, 30) (see "Results" and legends to Figs. 3 and 7 for details). The NTP concentrations for chain extension reactions were 5 μM except for 8-N_3ATP, which was 100 μM. The RNA products were analyzed, autoradiographed, and quantified by PhosphorImager (Molecular Dynamics) as described (12, 18).

Gre-RNAP Binding Assay—15 μg of each His_6-tagged Gre factor were incubated for 10 min at 4 °C with 7 μl of Ni^{2+}-chelating NTA-agarose beads in 20 μl of buffer C (40 mM Tris-HCl, pH 7.5, 30 mM KCl, 0.1 mM EDTA, and 0.1 mM dithiothreitol) containing 0.8 M NaCl. The beads were then washed in 3 × 500 μl of the same buffer. The yield of immobilized Gre factors was ~95% for all proteins. 5 μg of RNAP holoenzyme in 20 μl of buffer C containing 0.1 M NaCl, 10 mM $MgCl_2$, and 0.5 mg/ml of BSA were added to the beads and incubated with gentle shaking for 15 min. Free RNAP was removed by centrifugation, and the beads were washed twice with 100 μl of the same buffer. The beads carrying adsorbed RNAP were suspended in 20 μl of transcription buffer (buffer C, supplemented with 0.5 mg/ml BSA and 10 mM $MgCl_2$) containing 0.5 μg of *rrnB* P1 DNA template, 0.5 mM CpA, 2 μM ATP, and 1 μM [α-^{32}P]CTP. After incubation at 37 °C for 10 min, the beads were washed with 4 × 200 μl of buffer C containing 0.2 mg/ml of BSA and 0.2 M NaCl. The adsorbed proteins and TEC carrying radiolabeled hexameric transcript CpA**p**C**p**CpA**p**C (6C-TEC) (here and elsewhere, boldface type symbolizes radioactive phosphates) were eluted from the beads with 50 μl of buffer C containing 20 mM EDTA and 0.8 M NaCl. An aliquot of each eluate (5 μl for all samples except 0.5 μl for WT GreB) was analyzed by denaturing 23% polyacrylamide gel electrophoresis (PAGE) followed by autoradiography and quantitation by PhosphorImager. In the control experiment, 6C-TEC was prepared as above but in the absence of beads, and His_6-tagged GreA or GreB (100 μg/ml) were mixed with 2 μg of RNAP in 20 μl of transcription buffer. The resulting 6C-TEC was purified by gel filtration using Quick-Spin G-50, and an aliquot (1/200 of the total material) from each reaction was analyzed by PAGE.

RNA-Protein Photocross-linking—Photocross-linking experiments were performed using radioactively labeled 9A*-TEC carrying 8-N_3AMP at the RNA 3′ terminus essentially as described previously for the cross-linking of GreA (16).

Proteolytic Degradation of the Cross-linked Proteins—The cross-linked β′-subunit (~3 × 10^4 cpm) carrying radioactive RNA was purified by SDS-4% PAGE as described (19) and mixed with 5 μg of purified unlabeled β′-protein. The resulting material was precipitated by cold

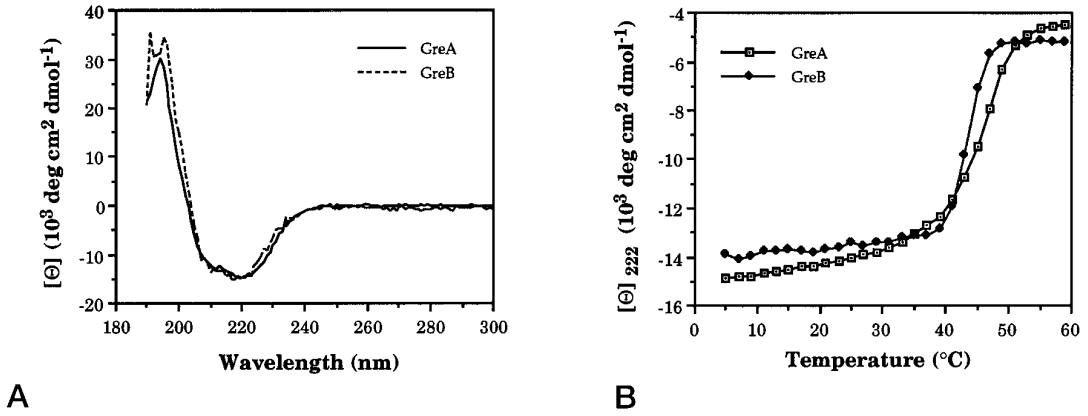

FIG. 1. *A*, circular dichroism spectra of GreA and GreB at 4 °C. *B*, thermal unfolding transitions of GreA and GreB monitored from the CD signals at 222 nm.

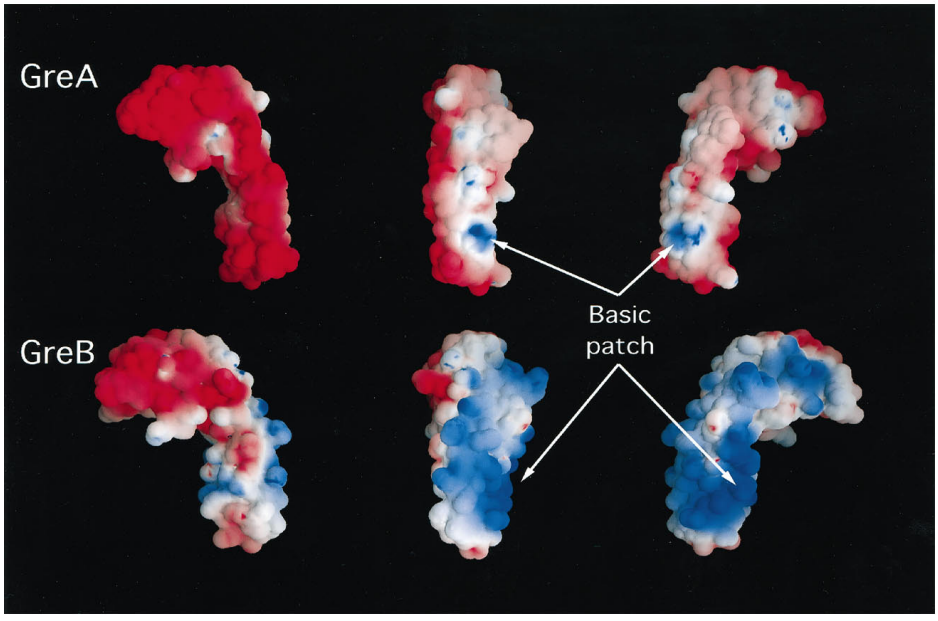

FIG. 2. *Top*, water-accessible surface and charge distribution of GreA, generated using GRASP (41). Three orthogonal views are shown. The *center* and *left* views were obtained by rotating the rightmost view by 90° or 180° about the vertical axis, respectively. The surface is colored by the electrostatic potential. *White*, uncharged; *red*, negative (Asp, Glu); *blue*, positive (Arg, Lys). *Bottom*, the same views showing the water-accessible surface and charge distribution of the GreB model.

acetone (1:1 (v/v)), redissolved in 4 μl of buffer D (40 mM Tris-HCl, pH 8.2, 40 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA) containing 1% SDS, and incubated for 5 min at 37 °C. The mixture was diluted with 200 μl of buffer D containing 0.15 μg of thrombin and incubated at 37 °C for 5 h. The reaction was terminated by the addition of 10 μl of 10 mM diisopropylfluorophosphate and analyzed by Tris-glycine SDS-14% PAGE followed by autoradiography. For proteolytic cleavage of GreB, the reaction mixture containing cross-linked GreB (1×10^5 cpm) was mixed with 4 μg of unlabeled GreB and incubated with 0.1 μg of AspN protease in 50 μl of buffer D containing 5 mM $CaCl_2$ for 1 h at 30 °C. Under these conditions, the cross-linked β′ was resistant to proteolysis by AspN (data not shown). The reaction was terminated by the addition of EDTA (20 mM final concentration) and analyzed by Tris-Tricine SDS-14% PAGE and autoradiography. The 16-kDa proteolytic fragment carrying cross-linked RNA was purified as described above for β′ and digested with 1 μg of endoproteinase ArgC in buffer D containing 10 mM $CaCl_2$ and 10 mM dithiothreitol for 4 h at 37 °C. The reaction was terminated and analyzed as described above for AspN.

Chemical Degradation of the Cross-linked Proteins—The purified cross-linked species were mixed with 5 μg of the same purified unlabeled protein prior to chemical degradation. For cleavage at Cys, the proteins were treated with NTCB for 4 h at 37 °C as described (20). For cleavage at Trp, the proteins were incubated with BNPS-skatole (10 mg/ml) in 70% acetic acid for 2 h, at 37 °C according to Ref. 21.

CD Spectrometry—CD spectra were obtained using an AVIV circular dichroism DS-62 spectrophotometer at 4 °C. The buffers were 50 mM sodium phosphate, pH 7.5, containing 50 mM NaCl for GreA and 250 mM NaCl for GreB. Protein concentrations used for experiments were in the range of 15–40 μM. The path length of the cuvette was 0.1 cm. Data were collected every 1 nm with a bandwidth of 1.5 nm, and the CD signal was averaged over 2 s. Each spectrum was averaged over four scans from 190 to 300 nm. The temperature dependence of the CD profiles was studied at 222 nm using a cuvette with the path length of 0.1 cm, and the temperature was varied from 5 to 60 °C in 2 °C increments. The samples were equilibrated for 2 min at each temperature point, and the CD signal was averaged over 30 s.

Protein concentration was determined using Bradford's protein assay (22). Tris-glycine and Tris-Tricine SDS-PAGE was performed according to Laemmli (23) and Schagger *et al.* (24), respectively, and gels were stained with Coomassie Brilliant Blue R-250.

RESULTS

The Model Three-dimensional Structure of GreB—Diffraction quality crystals of GreB have not been obtained yet. However, the high degree of functional (2) and sequence homology (35% sequence identity) between *E. coli* GreA and GreB suggests that they have similar structures (25). Moreover, examination of the sequence alignment reveals that key features of the GreA structure are conserved in GreB (16). These features include hydrophobic residues that participate in forming the hydrophobic core of the coiled-coil domain; charged residues that form interhelical salt bridges that stabilize the coiled-coil structure; all of the residues involved in interactions between the N- and C-terminal domains; and the helical bulge at position 24, causing a skip in the coiled-coil heptad repeat pattern.

Ultraviolet CD spectroscopy provides a rough indicator of the

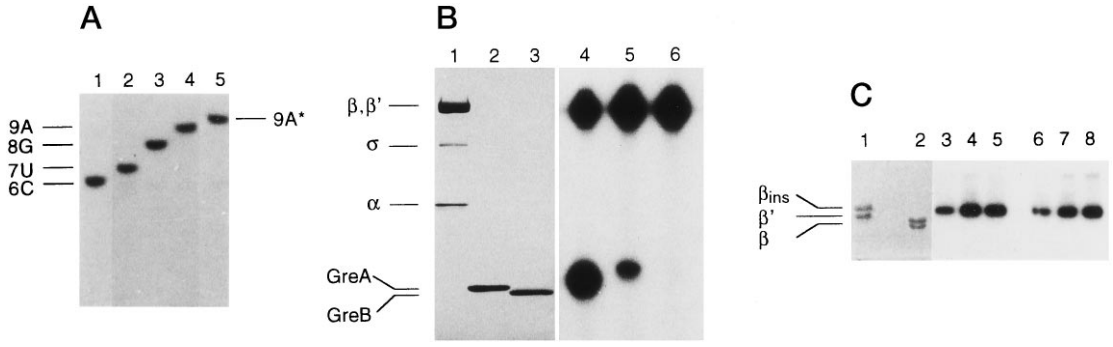

FIG. 3. *A*, generation of TECs with 3′-terminal 8-N_3AMP probe on a ribosomal *rrnB* P1 DNA. An autoradiograph of the urea-23% PAGE shows RNA chain extension of the initial radiolabeled 6C-TEC (~10 fmol) (*lane 1*) after incubation with UTP (7U-TEC, *lane 2*), UTP plus GTP (8G-TEC, *lane 3*), UTP, GTP, and ATP (9A-TEC, *lane 4*), and UTP, GTP, and 8-N_3ATP (9A*-TEC, *lane 5*). *B*, RNA-protein photocross-linking. *Lanes 1*, *2*, and *3*, Coomassie staining after SDS-14% PAGE of the purified WT RNAP holoenzyme (3 μg), GreA (2 μg), and GreB (2 μg), respectively. *Lanes 4–6*, autoradiograph after SDS-PAGE of 9A*-TEC (~0.5 *pmol*) UV-irradiated in the absence (*lane 6*) and the presence of 0.5 μg of GreB (*lane 4*) or 2 μg of GreA (*lane 5*). *C*, control experiments using WT RNAP (*lanes 3–5*) and β_{ins}-RNAP (*lanes 6–8*). *Lanes 1* and *2*, Coomassie staining after SDS-4% PAGE of the purified WT RNAP (2 μg), and β_{ins}-RNAP (3 μg), respectively. *Lanes 3–8*, autoradiograph after SDS-4% PAGE of the cross-linking reactions performed with 9A*-TEC (~0.5 pmol) in the absence (*lanes 5* and *8*) and in the presence of 2 μg of GreA (*lanes 4* and *7*) or 0.5 μg of GreB (*lanes 3* and *6*).

secondary structure of globular proteins. The measured CD spectra of GreA and GreB are nearly identical (Fig. 1*A*). The minima at 208 and 222 nm are indicative of helical secondary structure (26), as expected from the GreA structure. The $[\phi]_{222}/[\phi]_{208}$ ratio has previously been used to assess the number of helical strands within a molecule (27). A value for $[\phi]_{222}/[\phi]_{208}$ of about 0.8 is associated with single-stranded α-helix, whereas a value of about 1.0 or more is suggestive of a two-stranded coiled-coil. The value of $[\phi]_{222}/[\phi]_{208}$ for GreA and GreB (1.3 for both) is suggestive of double-stranded coiled-coils, as observed in the GreA x-ray crystal structure. In addition, the CD signals at 222 nm were monitored as a function of temperature to investigate the thermal unfolding of the proteins (Fig. 1*B*). Both proteins behave similarly, with cooperative thermal unfolding transitions at about 46 and 43 °C for GreA and GreB, respectively.

Taken together, these observations strongly support the conclusion that the structure of GreB is very similar to that of GreA. A three-dimensional model of GreB was calculated using comparative protein modeling by satisfaction of spatial restraints as implemented in the program MODELLER[2] (28). First, the relatively high sequence similarity between GreA and GreB resulted in reliable alignment with only one single residue deletion. Second, the alignment was used to derive many distance and dihedral angle restraints on the GreB sequence. Finally, the GreB model was calculated by minimizing the violations of these restraints. The main chain atoms of the model are generally within 0.4 Å of the equivalent GreA atoms. Assuming that the relative orientation of the coiled-coil and the globular domain is conserved, the main chain root mean square error in the model is expected to be approximately 1.5 Å, and about 70% of the side chain rotamers are expected to be modeled correctly (29).

The most striking result from the homology model of the GreB structure is revealed by examining the charge distribution around the water-accessible surface (Fig. 2). As noted earlier (16), the charge distribution around the surface of GreA exhibits a remarkable asymmetry. One face of the molecule is strongly acidic (Fig. 2, *top left*), whereas the opposite face is neutral except for a small basic region formed by Arg^{52} and Arg^{37}. The asymmetry in the GreB charge distribution is even more dramatic. As with GreA, one face of GreB is acidic (Fig. 2, *bottom left*); however, the opposite face is strongly basic and has a much larger basic patch than GreA. These structural features may be related to the different functional properties of the two factors.

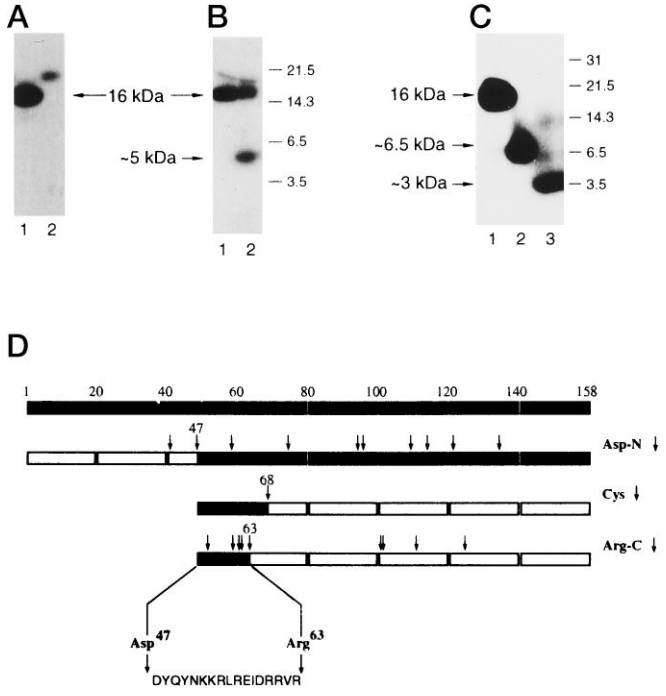

FIG. 4. **Mapping of the cross-linking site in GreB.** *A*, an autoradiograph of SDS-12% PAGE shows the initial cross-linked GreB (*lane 2*) and the cleavage product generated by AspN (*lane 1*). *B*, an autoradiograph of Tris-Tricine SDS-16% PAGE shows the purified 16-kDa fragment (*lane 1*) and the products of its chemical degradation by NTCB (*lane 2*). *C*, an autoradiograph of Tris-Tricine gradient (10–20%) SDS-PAGE shows the initial 16-kDa fragment (*lane 1*), the products of chemical degradation by BNPS-skatole (*lane 2*), and enzymatic digestion by ArgC (*lane 3*). The positions and the molecular masses (kDa) of prestained markers are indicated on the *right* in all autoradiographs. *D*, schematic illustration of the mapping procedure. The *bar* at the *top* represents 158-amino acid-long GreB polypeptide. Three successive steps of enzymatic and chemical degradation of the cross-linked GreB corresponding to the experimental data of *A*, *B*, and *C* are shown schematized *below*. The Asp, Cys, and Arg cleavage sites of GreB are indicated by *arrows*. The *numbers above* the *arrows* designate cleavages that separate fragments carrying cross-linked radioactive adduct (*black areas*).

Cross-linking of RNA 3′ Terminus to GreA, GreB, and RNAP in

[2] MODELLER is available on the Internet by anonymous ftp from guitar.rockefeller.edu and also as part of QUANTA and INSIGHTII (MSI, San Diego, CA; E-mail address, blp@msi.com).

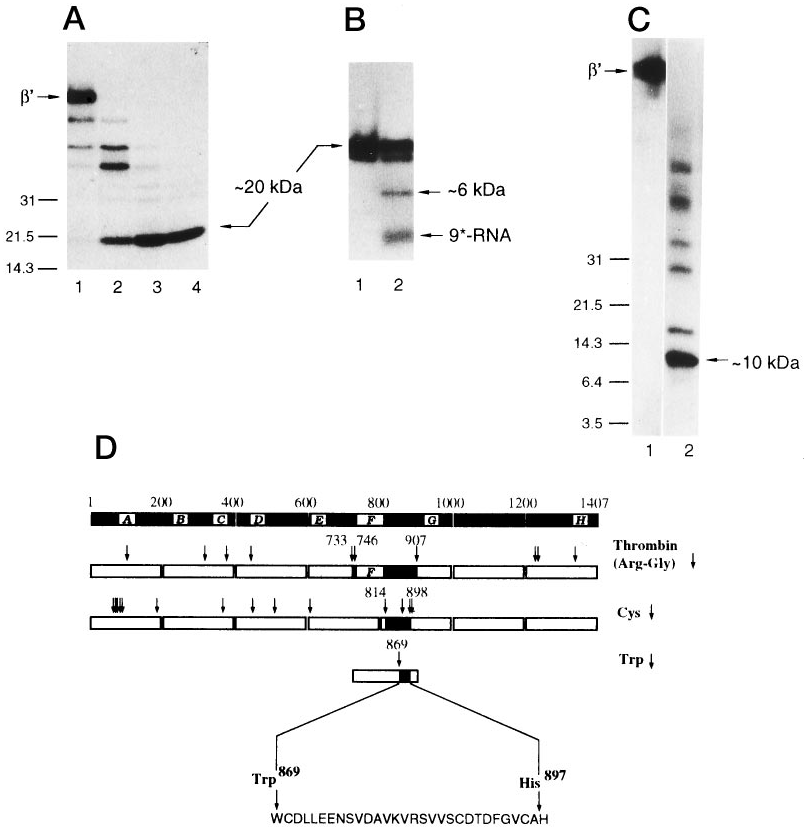

FIG. 5. **Mapping of the cross-linking site in β′.** *A*, an autoradiograph of a gradient (4–20%) SDS-PAGE shows the time course of enzymatic cleavage of the initial cross-linked β′ by thrombin for 15, 40, 120, and 300 min. (*lanes 1*, *2*, *3*, and *4*, respectively). *B*, an autoradiograph of Tris-Tricine gradient (10–20%) SDS-PAGE shows the initial purified 19- and 20-kDa fragments (*lane 1*) and the products of chemical degradation by BNPS-skatole (*lane 2*). *C*, an autoradiograph after SDS-PAGE (same as in *B*) shows the initial cross-linked β′ (*lane 1*) and the products of chemical degradation by NTCB (*lane 2*). *D*, schematic presentation of the mapping procedure (for explanation, see Fig. 4*D*). The *bar* at the *top* represents 1407-amino acid-long β′ polypeptide, with *lettered boxes* symbolizing highly conserved regions (42). The Arg-Gly, Cys, and Trp cleavage sites that follow from the sequence of β′ are indicated by *arrows*.

TECs—To investigate the orientation of GreA and GreB relative to RNA in the TECs we employed RNA-protein cross-linking using a photoreactive analog of AMP (8-N_3-AMP) incorporated into the 3′ terminus of radioactively labeled nascent RNA. On the ribosomal *rrnB* P1 promoter (starting sequence: CACCACUGACACGG . . .) in the presence of the dinucleotide CpA, ATP, and [α-^{32}P]CTP, RNA polymerase forms a stable TEC carrying the hexameric transcript CpA**p**C**p**CpA**p**C (6C) (30). The 6C-TEC was purified by gel filtration (Fig. 3*A*, *lane 1*) and used for stepwise extension of the transcript to 7U, 8G, and 9A (*lanes 2*, *3*, and *4*, respectively). The addition of a mixture of UTP, GTP, and 8-N_3-ATP to radiolabeled 6C-TEC led to the formation of 9A*-TEC carrying transcript CpA**p**C**p**CpA**p**CpUpGpA* (*lane 5*) with photoactive azido probe at the RNA 3′ end.

In the absence of Gre proteins, irradiation of the 9A*-TEC generated a single radioactive band visible on an autoradiograph of SDS-PAGE (Fig. 3*B*, *lane 6*), which corresponds to the cross-linked β′-subunit of RNAP. A control experiment was performed with a mutant RNAP carrying a 127-residue insertion in a nonessential region of the β-polypeptide (β_{ins}) (31). The β_{ins}-subunit migrates in SDS-PAGE much slower than the WT β (32) (Fig. 3*C*, *lanes 1* and *2*). However, cross-linking experiments conducted with both WT and β_{ins}RNAP produced a single radioactive band of the same mobility (*lanes 5* and *8*, respectively), eliminating β as the possible cross-linked species. The conclusion that the 3′ end of RNA cross-links to β′ is in agreement with our earlier result obtained for the 22A*-TEC formed on T7AI promoter (19).

The addition of GreB to the 9A*-TEC followed by UV-irradiation resulted in a 3-fold decrease of β′ cross-linking and the appearance of a second radioactive band, which migrates slightly above GreB with an apparent molecular mass of 19 kDa (Fig. 3*B*, *lane 4*). The yield of GreB cross-linking was approximately 18% of the initial 9A*-TEC. Under the same conditions, GreA cross-linked with the yield of about 5% and did not cause decrease in the cross-linking of β′ (Fig. 3*B*, *lane 5*). Fig. 3*C* demonstrates that neither GreA (*lanes 4* and *7*) nor GreB (*lanes 3* and *6*) induced the cross-linking of the β-subunit. The cross-linked sites on GreB and β′ were mapped by specific chemical and enzymatic degradation as shown in Figs. 4 and 5.

Mapping of the RNA-cross-linking Site in GreB—First, the cross-linked GreB was subjected to limited proteolysis with endoproteinase AspN, which specifically hydrolyzes peptide bonds at the N-terminal side of Asp (33). Under nondenaturing conditions, this enzyme quantitatively cleaves free GreB only at Asp^{47} generating two polypeptide fragments, Met^{1}–Ala^{46} and Asp^{47}–Pro^{158}, as revealed by N-terminal amino acid sequencing (data not shown). The reaction yielded a radioactive product with an apparent molecular mass of 16 kDa (Fig. 4*A*, *lane 1*), which is the expected mass of the C-terminal fragment Asp^{47}–Pro^{158} plus the RNA probe. Next, the 16-kDa fragment was chemically cleaved at Cys residues by NTCB (20). Since GreB contains only one Cys at position 68, the only expected cleavage products are the 2.5-kDa N-terminal and 10.5-kDa C-terminal fragments. Two radioactive bands were detected by autoradiography after SDS-PAGE analysis: the uncleaved GreB and a peptide with an apparent molecular mass of ~5.0 kDa (Fig. 4*B*, *lane 2*), which we interpreted as the N-terminal fragment carrying an additional mass of 2.5 kDa contributed by the cross-linked 9A-RNA. These results map the cross-link within the fragment Asp^{47}–Lys^{67}. This conclusion was confirmed by a chemical cleavage of the 16-kDa proteolytic fragment at Trp residues with BNPS-skatole (21). This cleavage resulted in a single radioactive product with an apparent mass of ~6.5 kDa (Fig. 4*C*, *lane 2*), which could only correspond to Asp^{47}–Trp^{91}. For further mapping, the 16-kDa fragment was digested by endoproteinase ArgC, which specifically cleaves peptide bonds at the C-proximal side of Arg (34). The major radioactive product observed near the 3.5-kDa marker (Fig. 4*C*, *lane 3*) corresponds to Asp^{47}–Arg^{63} or shorter fragments within. These results localize the site of the photocross-link to the 17-amino acid-long Arg-rich segment between Asp^{47} and

Hybrid #	GreA/GreB Hybrids	Type of cleavage activity toward 9A-TEC	Specific cleavage activity toward TEC (U/μg)	
			7U-	9A-
	Gre A (1–158)	A	1000	100
	Gre B	B	300	1000
1	GreB $^{1-6}$ /GreA $^{7-158}$	A	140	0.1
2	GreB $^{1-12}$ /GreA $^{13-158}$	A	2000	20
3	GreB $^{1-31}$ /GreA $^{32-158}$	-	0	0
4	GreB $^{1-81}$ /GreA $^{82-158}$	B	0.2	2
5	GreB $^{1-105}$/GreA $^{106-158}$	B	30	30
6	GreA $^{1-6}$ /GreB $^{7-158}$	B	20	50
7	GreA $^{1-12}$ /GreB $^{13-158}$	-	0	0
8	GreA $^{1-31}$ /GreB $^{32-158}$	-	0	0
9	GreA $^{1-81}$ /GreB $^{82-158}$	A	20	30
10	GreA $^{1-107}$/GreB $^{108-158}$	A	20	1

FIG. 6. **Cleavage activities of His_6-tagged WT and hybrid Gre factors toward 7U- and 9A-TECs.** Specific activities are expressed as relative units/μg of Gre protein. The type of activity and the values of cleavage activity were determined after PAGE analysis of the transcription reactions performed as described under "Experimental Procedures" and in Ref. 18. Each value represents the average of at least three experiments.

Arg^{63}, which coincides with the location of the large basic patch of GreB.

In our previous paper (16), we described the localization of the cross-linked site on GreA in 9A* TEC between Asp^{41} and Phe^{57}. Thus, the cross-linking of the RNA 3′ terminus to both Gre factors occurs within an overlapping 22-amino acid-long fragment between positions 41 and 63, suggesting that this region is responsible for the interactions with the nascent RNA 3′ terminus.

Mapping of the RNA-cross-linking Site in the β′-subunit of RNAP—For localization of the cross-link in β′, the initial radioactive material was isolated from SDS gel and subjected to exhaustive degradation by thrombin. This enzyme recognizes the consensus sequence Leu-Val-Pro-Arg-Gly-Ser. However, it is also highly specific for the sequence Arg-Gly and cleaves peptide bonds at the C-proximal side of Arg (35). After a prolonged incubation in the presence of SDS, the reaction yielded a single prominent radioactive band visible on an autoradiogram after SDS-PAGE (Fig. 5*A*, *lanes 1–4*) with apparent molecular mass of ~20 kDa. Further analysis of the isolated 20-kDa fragment by Tris-Tricine SDS-PAGE revealed two closely migrating bands with masses of ~20 and ~19 kDa (Fig. 5*B*, *lane 1*). Taking into account the additional mass of RNA these could only correspond to fragments Gly^{733}–Arg^{907} and Gly^{746}–Arg^{907}. Next, the mixture of the two Gly–Arg fragments was subjected to chemical cleavage at Trp residues by BNPS-skatole. The reaction yielded two radioactive cleavage products: a band corresponding to the detached RNA and a band with a mass of ~6 kDa (Fig. 5*B*, *lane 2*), which we interpreted as the C-terminal fragment Trp^{869}–Arg^{907}. Finally, the initial cross-linked β′ was chemically cleaved at Cys residues by NTCB. The major radioactive product with a mass of ~10 kDa, representing about 70% of the total radioactivity (Fig. 5*C*, *lane 2*) was observed, which was also the smallest fragment visible on the autoradiograph. According to the distribution of Cys residues in β′ (see Fig. 5*D*), the 10-kDa product corresponds to a fragment flanked by Cys^{814} and Cys^{898}. The other radioactive bands above the 10-kDa product visible on the autoradiograph presumably correspond to the products of incomplete digestion, since the yield of cleavage reactions with NTCB may vary from 90 to 30% (20), depending on the susceptibility of Cys residues. However, the possibility of minor cross-linking sites (representing less than 10% of the total radioactivity in cross-linked β′) outside of the localized Cys^{814}–Cys^{898} fragment cannot be excluded. Thus, the major cross-link site is mapped to the 29-amino acid-long segment between Trp^{869} and Cys^{898}. This segment is located near the conserved "region G" of β′, where we have previously mapped the cross-link site for the RNA 3′ terminus in the 22A*-TEC formed on a T7A1 DNA (19).

Since the azido probe is directly attached to the adenine at the RNA 3′ end and has a reactivity radius of less than 5 Å, the simultaneous cross-linking of β′ and Gre proteins strongly suggests that they are situated in close proximity to each other.

Swapping of Gre Domains in GreA/GreB Hybrids—To delineate functional domains in Gre proteins, we constructed a set of chimeric GreA/GreB molecules carrying the N-terminal fragment of one Gre protein fused to the C-terminal fragment of the other. Each GreA/GreB hybrid carried six His residues at the C terminus in order to facilitate their purification. His-tagged GreA, GreB, and GreA/GreB hybrids (purified as described under "Experimental Procedures") were compared in their ability to activate the transcript cleavage reaction in TECs formed on a ribosomal *E. coli rrnB* P1 DNA fragment (2, 18).

The transcript cleavage assays used in these experiments are based on two properties of Gre proteins: first, the "catalytic" nature of their activity and, second, the type specificity of their respective activities. Gre proteins are able to induce complete transcript cleavage in TECs in substoichiometric amounts even when they are present at a Gre/TEC molar ratio of 0.01:1 (1). This catalytic property of Gre factors enables us to quantitate their specific activity in enzymological terms. Thus, in quantitative assays, one unit of cleavage activity is defined as the

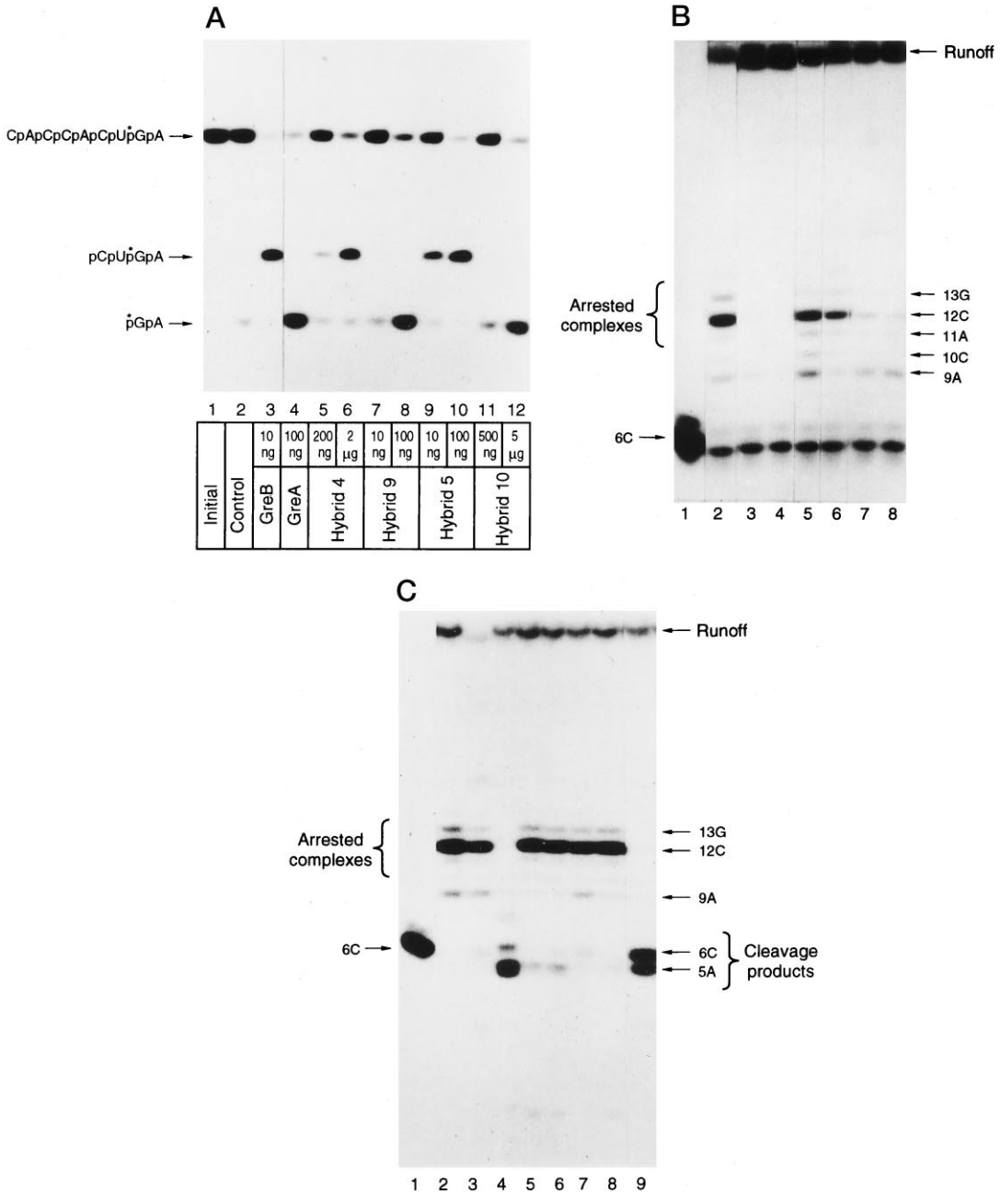

FIG. 7. ***In vitro* transcription activities of His$_6$-tagged WT and hybrid Gre factors.** *A*, *B*, and *C* are the autoradiographs after urea-23% PAGE. *A*, transcript cleavage of 9A-TEC. The starting 9A-TEC (~1 nM) carrying radiolabeled nonamer CpApCpCpApCpU**p**GpA (*lane 1*) was incubated in transcription buffer (buffer C containing 10 mM $MgCl_2$) for 10 min at 37 °C alone (*lane 2*) or in the presence of GreA (*lane 3*), GreB (*lane 4*) and GreA/GreB hybrids (*lanes 5–12*). The *asterisk* symbolizes radioactive phosphates. *B*, read-through assay. The starting 6C-TEC (~4 nM) carrying radiolabeled hexamer CpA**p**C**p**CpA**p**CpUpGpA (*lane 1*) was incubated with 30 µM NTPs in transcription buffer for 10 min at 37 °C alone (*lane 2*), or in the presence of 5 µg/ml GreA (*lane 3*), 1 µg/ml GreB (*lane 4*), and 200 µg/ml hybrids 4, 10, 9, and 5 (*lanes 5*, *6*, *7*, and *8*, respectively). *C*, antiarrest assay. Approximately 2 nM of arrested TECs (see Fig. 7*B*) were purified by gel filtration (*lane 2*) and incubated for 15 min at 37 °C in transcription buffer alone (*lane 3*) or in the presence of 5 µg/ml GreB (*lane 4*), 200 µg/ml GreA (*lane 5*), and 300 mg/ml hybrids 4, 10, 9, and 5 (*lanes 6*, *7*, *8*, and *9*, respectively).

amount of Gre protein required for the hydrolysis of 50% of the RNA in 7U- or 9A-TECs under standard reaction conditions (18). The specific cleavage activity is expressed as units/µg of Gre protein. In 9A-TEC carrying the radiolabeled transcript CpApCpCpApCpU**p**GpA, GreA stimulates cleavage and release of dinucleotides **p**GpA and pCpU (type A cleavage activity), whereas GreB induces the cleavage of the tetranucleotide pCpU**p**GpA (type B cleavage activity) (2). Thus, in the qualitative assay performed on the 9A-TEC, the type of products generated define the GreA- or GreB-type activity. The results obtained from comparative analyses of hybrid and WT Gre proteins using these criteria are summarized in Fig. 6.

For hybrids 1 and 6, in which only the first six N-terminal residues were exchanged, the type of cleavage activity was determined by the C-terminal part of the resulting polypeptide. Thus, hybrid 1 retained the GreA-type activity, while hybrid 6 retained the GreB-type activity. Both hybrids exhibited decreased cleavage activity toward 7U- and 9A-TECs compared with the corresponding WT Gre factors. The extent of the decrease varied from 7- to 1000-fold, depending on the combination of N- and C-terminal fragments and the type of TECs. Hybrid 1 and GreA displayed, for example, 140 and 1000 units/µg cleavage

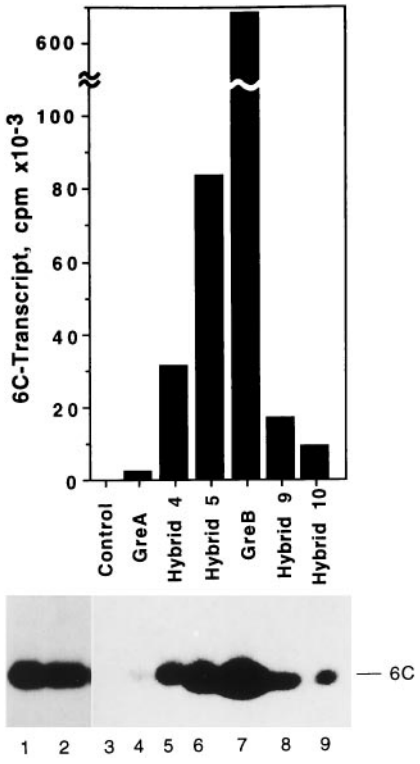

FIG. 8. **Binding affinities of His_6-tagged WT and hybrid Gre factors toward RNAP.** *Bottom*, an autoradiograph after urea-23% PAGE shows the formation of radiolabeled 6C-TEC by RNAP in free solution in the presence of GreA (*lane 1*) or GreB (*lane 2*), by RNAP adsorbed on Ni^{2+}-NTA-agarose matrix (*lane 3*), and by RNAP adsorbed on immobilized WT GreA, GreB, and hybrids 4, 5, 9, and 10 (*lanes 4*, *7*, *5*, *6*, *8*, and *9*, respectively) (see "Experimental Procedures" for details). *Top*, graphic presentation of the data obtained after quantification of the gel shown at the *bottom* by PhosphorImager. Each *bar* corresponds to the amount of radioactive 6C transcript generated by RNAP adsorbed on immobilized factors.

activity toward 7U-TEC, or 0.1 and 100 units/μg toward 9A-TEC, respectively. Hybrid 6 and GreB displayed 20 and 300 units/μg activity toward 7U-TEC or 50 and 1000 units/μg toward 9A-TEC, respectively. Thus, the changes within the short N-terminal segment of GreA and GreB affected the specific activity (units/μg) of these proteins and their selectivity toward TECs but did not affect the type of cleavage reaction.

Increasing the N-terminal portion of GreB in GreA to 12 residues (hybrid 2) restored the original cleavage activity of GreA. However, similar substitution in GreB (hybrid 7) led to a total loss of cleavage activity toward both TECs. Apparently, this effect was due to incorrect protein folding, since hybrid 7, when expressed in *E. coli*, was found in the inclusion bodies, less than 1% of which was recovered in soluble form after renaturation from 7 M guanidine HCl. The same results were obtained when the first 12 N-terminal residues in GreA and GreB were deleted (data not shown). According to the three-dimensional structure of GreA and GreB, four nonconserved residues at positions 4, 5, 8, and 9 and conserved Thr^6 are buried inside the Gre molecule and, presumably, participate in forming the hydrophobic core of the N-terminal coiled-coil domain (16). Our results together with crystallographic data suggest that the N-terminal 12 residues form an essential structural element for initiating the proper folding of the coiled-coil domain of Gre molecules.

The exchange between GreA and GreB of the N-terminal 31 residues (hybrids 3 and 8) as well as 44 residues (data not shown) yielded insoluble proteins that refolded poorly and displayed no detectable cleavage activity. The N-terminal residues 1–44 comprise the first α-helix of the coiled-coil domain, and most of the residues in this region that are critical for the formation of correct coiled-coil structure are highly conserved in both Gre molecules (2, 16). The only exceptions are those that comprise two of the four interhelical bonds that stabilize the overall structure of N-terminal domain: salt bridge Lys^{22}–Glu^{66} and hydrogen bond Arg^{25}–Asn^{54} in GreA and salt bridge Glu^{25}–Arg^{54} and hydrophobic bond Trp^{22}–Val^{62} in GreB (2, 16). Presumably, the disruption of these two interhelical bonds in GreA/GreB hybrids led to destabilization of coiled-coil structure and resulted in proteins that failed to assume their native conformation.

The Type of Gre-specific Cleavage Activity Is Determined by the N-terminal Domain—Exchanging the N-terminal 81 residues of GreA with the corresponding residues of GreB (hybrid 4) resulted in the switch of cleavage activity from type A to type B (Fig. 6, *hybrid 4*; Fig. 7*A*, *lanes 5* and *6*). The specific transcript cleavage activity of hybrid 4 toward 7U- and 9A-TECs decreased 1500- and 500-fold, respectively, in comparison with GreB. (Because the cleavage activity of hybrid 4 switched to type B, its specific activity can only be compared with GreB.) Similarly, the hybrid 9 displayed type A cleavage activity (Fig. 6, hybrid 9; Fig. 7*A*, *lanes 7* and *8*); however, the specific activity of this hybrid toward 7U- and 9A-TECs decreased only moderately (50- and 3-fold, respectively) compared with GreA. Hybrids 5 (Fig. 7*A*, *lanes 9* and *10*) and 10 (Fig. 7*A*, *lanes 11* and *12*) displayed the same types of cleavage activity as hybrids 4 and 9, respectively. However, the specific cleavage activity of hybrid 5 showed substantial improvement over that of hybrid 4, whereas the activity of hybrid 10 toward 9A-TEC was lower than that of hybrid 9 (Fig. 6). These data demonstrate that the N-terminal coiled-coil domain of Gre protein dictates the type of cleavage activity characteristic for each factor. These results also suggest that the C-terminal globular domain does not affect the type of cleavage reaction but is required for full transcript cleavage activity toward TECs.

The Antiarrest Activity of Gre Factors Is Determined by Their N-terminal Coiled-coil Domain—The hybrids 4, 5, 9, and 10 were further compared with the WT factors for their ability to suppress the formation of arrested TECs during transcription elongation on a ribosomal *rrnB* P1 promoter (read-through assay) and to induce the transcript cleavage in preformed arrested TECs (antiarrest assay) (2, 18).

For the read-through assay (Fig. 7*B*), the radiolabeled 6C-TEC (Fig. 7*B*, *lane 1*) was incubated with four NTPs alone or in the presence of GreA, GreB, or hybrid proteins. In the absence of Gre factors, about 50% of the TEC was arrested at positions +12 and +13 (*lane 2*). The addition of GreA, GreB, or hybrids 5 and 9 to the initial 6C-TEC, prior to the addition of NTPs, reduced the formation of arrested TECs and increased the total amount of the full-length run-off transcription product (*lanes 3*, *4*, *7*, and *8*, respectively). The hybrids 4 and 10 had little or no effect on the formation of arrested TECs (*lanes 5* and *6*, respectively).

For the antiarrest assay (Fig. 7*C*), the arrested 12C- and 13G-TECs obtained by chasing the initial 6C-TEC with four NTPs were further purified by gel filtration (*lane 2*), and then were exposed to either hybrid or WT factors. Of all the proteins tested, only GreB and hybrid 5 induced the cleavage of arrested TECs (*lanes 4* and *9*, respectively), yielding TECs carrying 5A and 6C transcripts that could be converted into full-length run-off products upon the addition of NTPs (data not shown). Under the same conditions, GreA and hybrids 9, 10, and 4 did not cleave any arrested TECs (*lanes 5*, *6*, *7*, and *8*, respectively). The fact that GreA was inactive toward arrested 12C- and 13G-TECs is consistent with our earlier observation that, in

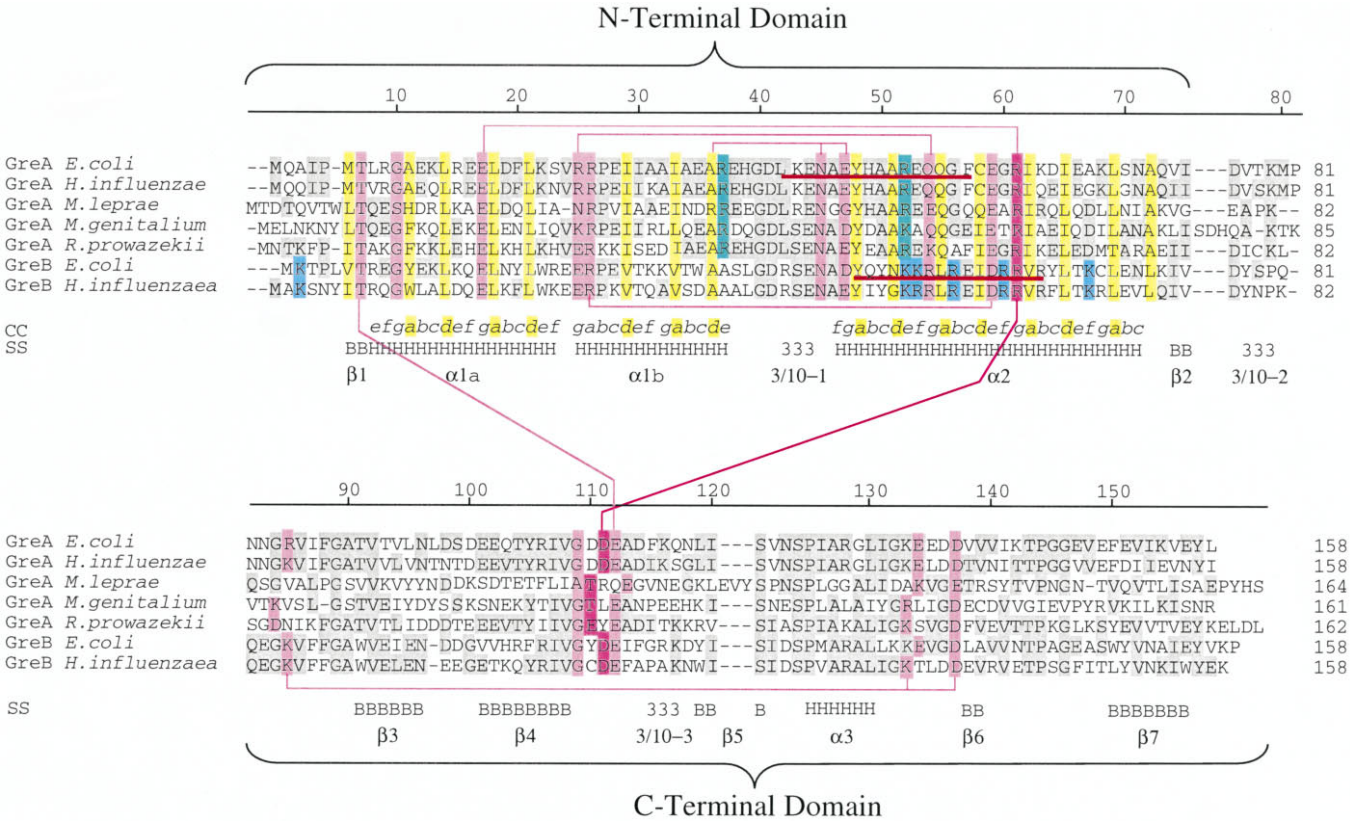

FIG. 9. **Sequence alignment of Gre proteins from five different organisms (see "Discussion" for details) and schematic illustration of *E. coli* GreA structural features.** The *numbering above* the *sequences* is with reference to *E. coli* GreA. The sequences, *underlined* in *red*, denote the fragments of *E. coli* GreA and GreB, containing the cross-link with RNA 3′ terminus. *Below* are shown the following features of the *E. coli* GreA structure: positions within the coiled-coil heptad repeat pattern (*CC*); secondary structure (*SS*) (β-strand (*B*), α-helix (*H*), 3/10 helix (*3*)); and labels for the elements of secondary structure. The *color shading* represents the following: amino acid identities (*gray*); generally conserved hydrophobic residues occupying *a* and *d* positions of the coiled-coil heptad repeat and positions 6 and 72 (*yellow*); residues that participate in interhelical and interdomain interactions (*magenta*); conserved residues at positions 37 and 52 forming small basic patch of GreA (*green*); residues forming large basic patch of GreB (*blue*).

contrast to GreB, GreA can act only before the arrest took place and not after the fact (2).

Thus, the functional activity of GreA/GreB hybrids in the read-through and the antiarrest assays was determined by their N-terminal domain and correlated with their specific transcript cleavage activity on 9A-TEC (see Figs. 6 and 7*A*). Indeed, hybrid 9 carrying the coiled-coil domain of GreA and globular domain of GreB acted as GreA displaying only the read-through activity, and hybrid 5 composed of the coiled-coil domain of GreB and two-thirds of the globular domain of GreA behaved as GreB, exhibiting both the read-through and the antiarrest activities. At the same time, hybrids 4 and 10 were inactive in both assays apparently due to their low specific transcript cleavage activity (see Fig. 6).

Both the Coiled-coil and the Globular Domains Contribute to Specific Binding of Gre Factors to RNAP—The functional roles of Gre domains were further explored by comparing the binding affinities of GreA/GreB hybrids and WT Gre proteins with RNAP. The addition of six His residues to the C terminus of Gre proteins does not alter the binding properties of GreB but causes a ~10-fold decrease in the binding affinity of GreA (10). When Gre factors are immobilized on Ni^{2+}-NTA-agarose through His tags, they are still able to bind RNAP from solution in accordance with their binding affinities. Moreover, RNAP molecules bound to immobilized Gre factors are able to form a stable 6C-TEC on *rrnB* P1 promoter, and neither GreA nor GreB affects the formation and the stability of 6C-TEC.[3] Therefore, the amount of 6C transcript generated by RNAP on the beads is proportional to the amount of RNAP molecules adsorbed to immobilized Gre proteins.

To assess the contribution of each Gre domain to the binding of Gre molecules to RNAP, Ni^{2+}-NTA-agarose beads carrying WT or hybrid factors were incubated with RNAP holoenzyme at a molar ratio of 10:1 in the presence of BSA as a carrier protein. After removal of free RNAP, the immobilized Gre·RNAP complex was incubated with *rrnB* P1 DNA, CpA, ATP, and [α-^{32}P]CTP to allow the formation of 6C-TEC. The free NTPs were removed, and the radioactivity in 6C-TEC formed on the beads was quantified following urea-PAGE analysis (Fig. 8). According to this semiquantitative assay, the radioactivity of 6C transcript associated with immobilized GreB was ~300 times higher than that of GreA (*lanes 7* and *4*, respectively). In solution, the amount of 6C-TEC formed was the same in the presence of GreA and GreB (Fig. 8, *lanes 1* and *2*, respectively), indicating that the low recovery of 6C transcript from immobilized GreA (*lane 4*) was due to a low amount of RNAP bound to GreA. These results are in good agreement with our earlier observation that the apparent K_d values for GreB·RNAP and GreA·RNAP complexes differ by approximately 2 orders of magnitude (10). When Ni^{2+}-NTA-agarose was used for the assay without immobilized Gre factors, no radioactive 6C transcript was detected (*lane 3*). In comparison with GreA, hybrids 4, 5, 9, and 10 displayed a 15-, 50-, 8-, and 4-fold increase in the apparent binding affinity toward RNAP (Fig. 8, *lanes 5*, *6*, *8*, and *9*, respectively), although none of them reached the binding efficiency of GreB. Both hybrids 5 and 9, carrying, respectively,

[3] S. Borukhov, unpublished observation.

the N- and C-terminal domains of GreB, have higher binding affinity than GreA, suggesting that the sites responsible for specific binding to RNAP are equally distributed between the two domains of the Gre molecule. These results also suggest that the binding of Gre factors to RNAP requires cooperative action of their coiled-coil and globular domains.

DISCUSSION

The transcript cleavage reaction appears to be a ubiquitous and evolutionarily conserved function among multisubunit RNA polymerases. It has been observed in TECs formed by vaccinia virus RNA polymerase (36), prokaryotic RNAP of *E. coli* (2), and eukaryotic RNA polymerases II (5–8) and III (37). The elongation factor TFIIS facilitates the transcript cleavage and the read-through/antiarrest activities of RNA polymerase II in such divergent organisms as yeast (8), insects (6), and mammals (5). In prokaryotes, the genes encoding Gre factors have been identified and sequenced from five different organisms including *E. coli* (2, 17), *Hemophilus influenzae* (38), *Rickettsia prowazekii* (39), *Mycobacterium leprae*,[4] and *Mycoplasma genitalium* (40). In addition, GreA- and GreB-like activities were detected in *Pseudomonas sp.*, *Acinetobacter sp.*, and *Bacillus subtilis*.[3] The biological importance of Gre factors is also underscored by the fact that the *greA* gene is present in the smallest known genome of any free-living organism, *M. genitalium* (40).

The amino acid sequence alignment of seven homologous GreA and GreB proteins is shown in Fig. 9. Although the percentage of identical residues among the members of the Gre family varies, most of the key structural elements of the *E. coli* GreA (16) are preserved with few exceptions. The conserved structural elements include hydrophobic residues at the expected *a* and *d* positions of the characteristic heptad repeat, $(abcdefg)_2$, and at positions 6 and 72 that form the hydrophobic core of the N-terminal domain (Fig. 9, *yellow shading*) and most of the residues participating in interhelical and interdomain salt bridges and/or hydrogen bonds stabilizing the overall Gre structure (*magenta* and *orchid shading*). In addition, all five members of the GreA family contain two highly conserved basic residues at positions 37 and 52 (*turquoise shading*), comprising a small basic patch on the water-accessible surface of GreA (see Fig. 2, *top*). Similarly, two representatives of the GreB family contain six conserved basic residues at positions 2, 52, 53, 56, 60, and 67 (Fig. 9, *blue shading*) that form a large basic patch on the surface of GreB (see Fig. 2, *bottom*). These observations suggest that all Gre proteins have similar structural organization and spatial arrangement of their N- and C-terminal domains. Comparison of the CD spectra of *E. coli* GreA and GreB and homology modeling of the GreB three-dimensional structure strongly support this conclusion.

Our results of RNA-protein cross-linking demonstrate that both GreA and GreB interact with the 3′ end of RNA and presumably β′-subunit of RNAP in TEC through a 16–17-residue-long peptide segment that overlaps with basic patches (Fig. 9, *red underline*). This supports our earlier hypothesis that the surface of Gre molecules harboring the basic patch contacts RNAP and RNA in TEC (16). The analysis of the functional roles of Gre domains using Gre hybrids revealed that the N-terminal domain is responsible for specific binding of Gre factors to RNAP and carries the structural determinants conferring the GreA or GreB type of cleavage activity and, presumably, the antiarrest and read-through activities. The C-terminal domain does not have a direct role in transcript cleavage or antiarrest function but participates in binding to RNAP and is required for full activity of Gre factors.

[4] GenBank™ accession number U15183.

Acknowledgments—This work was initiated while D. K., M. O., and S. B. were at A. Goldfarb's lab at the Public Health Research Institute. We thank A. Goldfarb and V. Nikiforov for generous support and advice and J. Lee for critical reading of the manuscript. We are also grateful to W. McAllister, A. Das, and R. Landick for comments and D. Cowburn and S. Cahill for help with CD measurements.

REFERENCES

1. Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 8899–8902
2. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993) *Cell* **72,** 459–466
3. Hsu, M. H., Vo, N. V., and Chamberlin, M. J. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 11588–11592
4. Erie, D. A., Hajiseyedjavadi, O., Young, M. C., and von Hippel P. H. (1993) *Science* **262,** 867–873
5. Reines, D., Chamberlin, M. J., and Kane, C. M. (1989) *J. Biol. Chem.* **264,** 10799–10809
6. Sluder, A. E., Greenleaf, A. L., and Price, D. H. (1989) *J. Biol. Chem.* **264,** 8963–8969
7. Izban, M. G., and Luse, D. S. (1992) *Genes & Dev.* **6,** 1342–1356
8. Reines, D. (1992) *J. Biol. Chem.* **267,** 37995–3800
9. Lee, D. N., Feng, G., and Landick, R. (1994) *J. Biol. Chem.* **269,** 22295–22303
10. Orlova, M. (1995) *Intrinsic Transcript Cleavage Activity of RNA Polymerase.* Ph.D. thesis, Institute of Molecular Genetics, Russian Academy of Science, Moscow
11. Altmann, C. R., Solow-Cordero, D. E., and Chamberlin, M. J. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 3784–3788
12. Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 4596–4600
13. Rudd, M. D., Izban, M. G., and Luse, D. S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 8057–8061
14. Chamberlin, M. J.(1994) *Harvey Lect.* **88,** 1–21
15. Nudler, E., Goldfarb, A., and Kashlev, M. (1994) *Science* **265,** 793–796
16. Stebbins, C. E., Borukhov, S., Orlova, M., Goldfarb, A., and Darst, S. A. (1995) *Nature* **373,** 636–640
17. Sparkowski, J., and Das, A. (1990) *Nucl. Acids Res.* **18,** 6443
18. Borukhov, S., and Goldfarb, A. (1996) *Methods Enzymol.* **274B,** 315–326
19. Borukhov, S., Lee, J., and Goldfarb, A. (1991) *J. Biol. Chem.* **266,** 23932–23935
20. Jacobson, G. R., Schaffer, M. H., Stark, G. R., and Vanaman, T. C. (1973) *J. Biol. Chem.* **248,** 6583–6591
21. Hunziker, P. E., Hughes, G. J., and Wilson, K. J. (1980) *Biochem. J.* **187,** 575–579
22. Bradford, M. (1976) *Anal. Biochem.* **72,** 248–254
23. Laemmli, U. K. (1970) *Nature* **227,** 680–685
24. Schagger, H., Link, T. A., Engel, W. D., and von Jagow, G. (1986) *Methods Enzymol.* **126,** 224–237
25. Sander, C., and Schneider, R. (1991) *Proteins* **9,** 56–68
26. Holzwarth, G., and Doty, P. (1965) *J. Am. Chem. Soc.* **87,** 218–228
27. Lau, S. Y. M., Taneja A. K., and Hodges, R. S. (1984) *J. Biol. Chem* **259,** 13253–13261
28. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol* **234,** 779–815
29. Sali, A., Pollerton, L., Yuan, F., van-Vlijmen, H., and Karplus, M. (1995) *Proteins* **23,** 318–326
30. Borukhov, S., Sagitov, V., Josaitis, C. A., Gourse, R. L., and Goldfarb, A. (1993) *J. Biol. Chem.* **268,** 23477–23482
31. Kashlev, M., Bass, I., Lebedev, A., Kalyaeva, E., and Nikiforov, V. (1989) *Genetika* **25,** 396–405
32. Borukhov, S., Severinov, K., Kashlev, M., Lebedev, A., Bass, I., Rowland, G. C., Lim, P. P., Glass, R. E., Nikiforov, V., and Goldfarb, A. (1991) *J. Biol. Chem.* **266,** 23921–23926
33. Drapeau, G. R. (1980) *J. Biol. Chem.* **255,** 839–840
34. Mitchell, W. M., and Harrington, W. F. (1968) *J. Biol. Chem.* **243,** 4683–4692
35. Nishikawa, S., Yanase, K., Tokunaga-Doi, T., Kodama, K., Gomi, H., Uesugi, S., Ohtsuka, E., Kato, Y., Suzuki, F., and Ikehara, M. (1987) *Protein Eng.* **1,** 487–492
36. Hagler, J., and Shuman, S. (1993) *J. Biol. Chem.* **268,** 867–873
37. Whitehall, S. K., Bardeleben, C., and Kassavetis, G. A. (1994) *J. Biol. Chem.* **269,** 2299–2306
38. Fleischmann R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., Craig Venter, J. (1995) *Science* **269,** 496–511
39. Marks, G. L., and Wood, D. O. (1992) *Nucleic Acids Res.* **20,** 3785
40. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A., III, Craig Venter, J. (1995) *Science* **270,** 397–403
41. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296
42. Iwabe, N., Kuma, K.-K., Kishino, H., Hasegawa, M., and Miyata, T. (1991) *J. Mol. Evol.* **32,** 70–78