Tethering of the Large Subunits of *Escherichia coli* RNA Polymerase*

(Received for publication, July 3, 1997)

Konstantin Severinov§, Rachel Mooney¶, Seth A. Darst†, and Robert Landick‡

From §The Rockefeller University, New York, New York 10021 and the †Department of Bacteriology, University of Wisconsin, Madison, Wisconsin 53706

The rpoB and rpoC genes of eubacteria and archaea, coding, respectively, for the β and β′-like subunits of DNA-dependent RNA polymerase, are organized in an operon with rpoB always preceding rpoC. Here, we show that in *Escherichia coli* the two genes can be fused and that the resulting 2751-amino acid β-β′ fusion polypeptide assembles into functional RNA polymerase in vivo and in vitro. The results establish that the C terminus of the β subunit and the N terminus of the β′ subunit are in close proximity to each other on the surface of the assembled RNA polymerase during all phases of the transcription cycle and also suggest that RNA polymerase assembly in vivo may occur co-translationally.

DNA-dependent RNA polymerase (RNAP) is the central enzyme of gene expression and a major target for regulation. Cellular RNAPs are large, multisubunit protein complexes. Core RNAP of *Escherichia coli* (~380 kDa) contains four polypeptides: β′ (155 kDa), β (150 kDa), and a dimer of α (37 kDa). Core RNAPs from eukaryotes contain 12 or more subunits with a total mass in excess of 600 kDa (1). The subunit composition of archaeal, chloroplast, and some eubacterial RNAPs is even more complex due to splits in the genes coding for the largest subunits (2, 3). Despite these differences, all cellular RNAPs exhibit striking and co-linear sequence similarities with the bacterial subunits. The two largest subunits comprise 60% of the RNAP mass and appear principally responsible for most of the enzyme’s functions.

The synthesis of RNAP subunits is coordinately regulated (4), but the exact mechanisms are unknown. In eukaryotes, many genes coding for RNAP subunits have similar regulatory similarities with the bacterial subunits. The two largest subunits (rpoB and rpoC) genes are described as pKS1000, pKS1010 (a derivative of pRL709) that expresses the β′-β′ fusion protein with a C-terminal His8 tag from an IPTG-inducible trc promoter, was constructed by a two-step procedure. In the first step, a DNA fragment containing codons 1–941 of rpoC was PCR-amplified from pT7β (8) using the upstream primer AACTCGGACGGAGCGCTGAG-GTGAAGATTTATTAAG, which places an XhoI site (underlined) encoding Leu-Glu directly before the rpoC GTG initiation codon (in bold). The amplified rpoC fragment was treated with XhoI and SacI (SacI cuts at rpoC codon 877 and is compatible with XhoI) and ligated into the XhoI site of the rpoB expression plasmid pRL706 to yield pKS1000 (Table I). pRL706 contains a modified BamHI (filled in) to SacI fragment from pRL385 (9) between the NcoI (filled in) and SacI sites of the lac promoter. pRL709 (10). The modification places a unique XhoI site encoding Leu-Glu directly after the last codon of rpoB and before a His8 tag, so that insertion of the XhoI-SacI fragment in pKS1000 fused the C-terminal GAG(Glu) codon of rpoB to the N-terminal GTG codon of rpoC (encodes Val in fusion), separated only by the XhoI site encoding Leu-Glu. Plasmids corresponding to pRL706 without the His8 tag or XhoI site, but containing an SF531(Rif18) substitution (pRW229) or an SF531(Rif18) substitution and a C-terminally His8 tagged rpoC gene from pRL663 (11) inserted between XhoI and HindIII downstream of rpoB (pRL717), also were used in our experiments (Table I).

In the second step of pRL719 construction, a BamI-Sse8387 fragment of pKS1000 was inserted between the corresponding sites in the rpoB expression plasmid pRL717. The intact rpoB and rpoC genes as described for pKS1000, pKS1010 (a derivative of pRL719) that expresses the β′-β′ fusion without a His8 tag was constructed by replacement of the BamHI-HindIII fragment of pRL719 with the corresponding fragment of pRW308 (11). pKS1020 (which expresses the β-β′ fusion polypeptide with an internal His8 tag at the β′-junction site) was constructed by treating pKS1010 with XhoI and ligating with a compatible double-stranded linker containing six histidine codons.

RNAP Purification—Cells overproducing the β′-β′ fusion from plasmids were grown to mid-log phase in 1 liter of LB medium plus 1 mM IPTG and 50 μg of ampicillin/ml. After recovery by centrifugation, cells were resuspended in 40 ml of 50 mM Tris-HCl, 300 mM KCl, 1 mM EDTA, pH 7.9, and lysed by passage through a French press. The lysate was cleared by low speed centrifugation, and RNAP was recovered by polyethyleneimine (PEI) precipitation, high salt extraction, and ammonium sulfate precipitation essentially as described by Burgess and Jendrisak (12). The ammonium sulfate pellet containing RNAP was redissolved in 20 ml of 20 mM Tris-HCl, 1 mM EDTA, 5% glycerol, pH 7.9 (TGE buffer), loaded on a 1-ml heparin HiTrap cartridge (Pharmacia Biotech AB) equilibrated in the same buffer, washed with TGE + 1 mM NaCl, and eluted with TGE + 0.6 mM NaCl. The eluted protein was precipitated with ammonium sulfate, dissolved in 0.5 ml of TGE + 0.3 mM NaCl, and passed through a Superose 6 column (Pharmacia) equilibrated in the same buffer. Fractions containing RNAP (monitored by
SDS-PAGE, see Fig. 2A) were pooled, dialyzed into 50 mM phosphate buffer, 0.5 mM NaCl, pH 8.0 (P buffer), bound to a 1-mL MC-20 column (Perspective Biosystems) charged with Ni²⁺, and equilibrated in P buffer, washed with P buffer + 5 mM imidazole, and the RNAP eluted in 10 mL of P buffer + 100 mM imidazole. RNAP was concentrated on a C-100 concentrator (Amicon, Inc., Beverly, MA) to ~1 mg/mL, diluted 2-fold with glycerol, and stored at -20 °C.

Affinity Labeling and in Vivo Transcription—Lys⁰¹⁰⁰⁵ affinity labeling reactions on T7 A2 promoter containing DNA fragment (13) were performed essentially as described for the T7 A1 promoter-dependent labeling (13) with the following modifications. A derivatized GMP analog, specified by position +1 of the T7 A2 promoter, was used to cross-link Lys⁰¹⁰⁰⁵ at the first stage of the reaction. The synthesis of the GMP derivative was analogous to the AMP derivative synthesis described previously (14). To radioactively label the cross-linked subunits, [α-³²P]CTP, specified by position +2 of the T7 A2 promoter, was used. Reactions were terminated by the addition of an equal volume of Laemmli loading buffer, and proteins were resolved by SDS-PAGE in 8% precast Tris-glycine gels (Novex, San Diego, CA) followed by autoradiography.

Im mobilized transcription on a T7 A1 promoter-driven DNA template was performed as described (15). Elongation complexes stalled at position +20 were prepared in 50-µL reactions containing 25 µL of N7-⁰-nitrotriacetic acid-agarose (Qiagen, Inc.), 20 µL of the 302-base pair T7 A1 promoter DNA fragment (template 3 of Nudler et al. (16)), 40 mM RNA, 0.5 mM Apl, 50 mM CTP and GTP, 2.5 mM [α-³²P]ATP (3000 Ci/nmol), 40 mM Tris-HCl, pH 7.9, 40 mM KCl, and 10 mM MgCl₂. Reactions proceeded for 15 min at 23 °C and then were washed three times with 1.5 µL of buffer (40 mM Tris-HCl, pH 7.9, 40 mM KCl, 10 mM MgCl₂) as described (15). To desorb transcription complexes from Ni²⁺ beads, 2 mM imidazole, pH 8.0, was added to a final concentration of 100 mM, and reactions were incubated with occasional shaking for 1 h at 4 °C. The Ni²⁺ beads were pelleted by centrifugation, and the supernatant, containing desorbed transcription complexes, was transferred into a fresh Eppendorf tube. Transcription by the +20 elongation complexes was synchronously initiated by adding rNTPs to a final concentration of 10 µM or 1 mM. Reactions proceeded for 5 min at 23 °C and were then terminated by addition of an equal volume of loading buffer containing 6 M urea. Transcription products were analyzed by urea-polyacrylamide gel electrophoresis (7 M urea, 6% polyacrylamide), followed by autoradiography.

RESULTS

Generation of rpoB::rpoC Fusion Gene and Analysis of Its Function in Vivo—The plasmid pRL719, which overproduces the β′:β′ fusion protein from the IPTG-inducible trc promoter, was generated as described under “Experimental Procedures.” The rpoB portion of the fused gene harbored the dominant rifampicin resistance mutation rif⁰¹⁴¹⁸ and allowed us to test for β function of the fused protein in vivo. The growth on plates of rifampicin-sensitive E. coli RL602 cells harboring either pRL719 (pβ'rfl14¹⁸), rif²⁰⁰⁰ rpoB expression plasmid pRW225 (pβ'rif²⁰⁰⁰), rif²⁰⁰⁰ rpoB expression plasmid pRW308 (pβ'), or control plasmid pBR322 was investigated (Fig. 1). After streaking on plates and overnight growth, only cells expressing rpoB::rpoC from pRL719 or rif²⁰⁰⁰ rpoB from pRW225 grew in the presence of rifampicin; cells expressing rpoC from plasmid pRW308 as well as cells harboring pBR322 completely failed to grow in the presence of rifampicin. Growth in the presence of rifampicin was dependent on rpoB::rpoC expression, since cells did not grow in the absence of IPTG, which is needed to derepress the trc promoter of the structural gene of pRL719. Since the expression of rif²⁰⁰⁰::β′ confers rifampicin resistance to the cells, we conclude that the rpoB::rpoC fusion gene is able to function as a source of the β subunit in vivo.

To test for function of the β′:β′ fusion in place of β′, E. coli RL602 harboring pRL719 was plated at an elevated (42 °C) temperature. RL602 cells harbor an amber mutation in the chromosomal copy of rpoC that is suppressed by a temperature-sensitive suppressor (17). RL602 grows at 30 °C, but not at 42 °C. The experiment shown in Fig. 1C demonstrates that cells expressing either the β′:β′ fusion or β′, but not β, grew at 42 °C. In the absence of IPTG, RL602 cells harboring pRL719 completely failed to form colonies at 42 °C. Since the expression of rif²⁰⁰⁰::β′ confers temperature resistance to the cells, we conclude that the rpoB::rpoC fusion gene is able to function as a source of the β′ subunit in vivo. RL602 cells were also able to grow at 42 °C and in the presence of rifampicin in an IPTG-dependent manner (Fig. 1D), demonstrating that the rpoB::rpoC fusion gene simultaneously functioned as a source of both β and β′ in vivo.

Purification of RNA Polymerase Containing the β′:β′ Fusion—The results of the in vivo experiments presented above could be explained by the presence in the RL602 cells harboring pRL719 of (i) functional RNA polymerase comprising aβ′:β′, (ii) functional RNA polymerase comprising αβ′β′:β′, and (iii) β and β′ resulting from proteolytic cleavage of the β′:β′ fusion protein at the subunit junction site followed by assembly of “end-type” β and β′ subunits into RNA polymerase comprising αβ′β′. To distinguish between these possibilities, RNA from RL602 cells overproducing β′:β′ was purified using our standard purification procedure (18) involving PEI precipitation and extraction of RNA from the PEI pellet with 1 M NaCl, followed by heparin affinity chromatography (see “Experimental Procedures”). Heparin affinity chromatography fractions containing RNA were loaded on a
Superose 6 gel-filtration column attached to the fast protein liquid chromatography system, and the eluted proteins were analyzed by SDS-PAGE (Fig. 2). The β'-β' fusion protein co-eluted exactly with the normal β', and α polypeptides and with RNAP activity (data not shown) during the gel-filtration step. From this experiment we conclude that the fusion protein assembled in an RNAP molecule comprising α2β'-β'.

The C terminus of the β' portion of the fusion protein encoded by pRL719 is extended with a His6 tag (see “Experimental Procedures”). As is shown elsewhere (19), the His6 tag allows facile separation of RNAP containing the plasmid-borne, His6-tagged β' subunit from RNAP containing untagged β' of chromosomal origin by ion metal affinity chromatography (IMAC). Superose 6 fractions containing RNAP were pooled, fractionated by IMAC, and the flow-through and eluant were analyzed by SDS-PAGE (Fig. 2, lanes E and F). The flow-through fractions contained wild-type RNAP (α2β'α), while fractions retained on the column and eluted with 100 mM imidazole contained β':β', α, and σ and did not contain the normal β and β' polypeptides. Thus, the fusion polypeptide assembled into RNAP did not undergo substantial proteolytic degradation into individual β and β' polypeptides in the cell or during purification.

The rpoB::rpoC Fusion Gene Can Be the Only Source of the β' Subunit in vivo—The principal conclusion of this work is that the C terminus of β and the N terminus of β' are positioned close to each other on the surface of E. coli RNAP. Since the β:β' fusion polypeptide can support RNAP function in vivo and in vitro with only a two-amino acid linker inserted between the β and β' sequences, the C terminus of β and the N terminus of β' must be positioned within about 5 angstroms of each other. In addition, since RNAP harboring the β:β' fusion polypeptide appears to be fully functional in vivo and in vitro, the C terminus of β and the N terminus of β' must remain in close proximity throughout all phases of the transcription cycle. Independent support for this conclusion comes from both genetic and biochemical studies. Yano and Nomura (20) demonstrated that a temperature-sensitive substitution in the N-terminal-most conserved segment A of the β' homolog of Saccharomyces cerevisiae RNAP I can be suppressed by a substitution near the C terminus of the β homolog, suggesting that
these two segments of the largest subunits interact with each other. Nucler et al. (21) investigated protein-DNA contacts at the leading edge of E. coli RNAP elongation complexes and found that the N-terminal conserved segment A of β and the C-terminal segment I of β are both cross-linked to a nucleotide derivative positioned 5 nucleotides downstream of the catalytic center in the template strand.

**Evolutionary Implications**—In eubacteria and archaea, the genes coding for the largest RNAP subunits (β and β') are organized similarly. The eubacterial rpoBC gene, coding for the β subunit, always precedes rpoC, which codes for β' (6). The two genes are co-transcribed and are separated by a short, untranslated linker. Similarly, archaeal genes coding for the β and β' homologs are also co-transcribed as part of an operon, and their relative position is the same as in prokaryotes (7). The linker separating the genes is either very short or the two genes overlap. The functional significance of such an organization of the rpo genes is unknown. In contrast to prokaryotes and archaea, genes coding for the RNAP largest subunits in eukaryotes are located on different chromosomes (1). It is tempting to speculate that the organization of rpoB and rpoC in archaea and prokaryotes is due to transcription/translation coupling that occurs in these organisms in the absence of nuclei and allows RNAP assembly to occur co-translationally. Several lines of evidence support this hypothesis. The pathway of E. coli RNAP assembly in vivo and in vitro is α·β·α·β (24). Our recent data demonstrate that evolutionarily conserved segments H and the N-terminal half of segment I of the E. coli β subunit are necessary and sufficient for the specific and obligatory interaction with the α subunit (25). In addition, the C-terminal-most half of conserved segment I is necessary for recruitment of β' into the α·β complex (25). Data of Luo et al. (26) suggest that the N-terminal portion of β' is required for interactions with α·β. Thus, ordered translation of the rpoBC mRNA is compatible with co-translational assembly of RNAP. Experiments directly testing this hypothesis are currently in progress.

Interestingly, recent genomic sequencing of Helicobacter pylori revealed that in this organism, the rpoB and rpoC genes form a continuous open reading frame (28). In addition, a cytoplasmic killer plasmid of the yeast Kluyveromyces lactis contains an open reading frame corresponding to a fusion of rpoB to the N-terminal portion of rpoC (28). While RNAP has not been purified from these organisms, our results show that such an arrangement is compatible with RNAP function, since the protein product resulting from the E. coli rpoB:rpoC fusion can assemble into functional RNAP.

**Acknowledgments**—We are grateful to Arkady Mustasev for guidance during the synthesis of derivatized GMP and to J.-F. Tomb for communicating his data prior to publication.

**REFERENCES**

1. Buhler, J.-M., Riva, M., Mann, C., Thoriaux, P., Memet, S., Micozin, J. Y., Treich, I., Mariotte, S., and Sentenac, A. (1987) in RNA Polymerase and the Regulation of Transcription (Kornblatt, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, M. T., Jr., and Wikens, M. P., eds) pp. 25–36, Elsevier Science Publishing Co., New York

2. Puhler, G., Leeff, H., Grupp, F., Palm, P., Klenk, H.-P., Lottepecht, F., Garrett, R. A., and Zillig, W. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 4569–4573

3. Leffers, H., Grupp, F., Lottepecht, F., Zillig, W., and Garrett, R. A. (1989) J. Mol. Biol. 206, 1–17

4. Dykhxhoorn, D. M., St. Pierre, R., and Linn, T. (1996) Mol. Microbiol. 19, 485–496

5. Jansma, D. B., Archambault, J., Mostachfi, O., Freisen, J. D. (1996) Nucleic Acids Res. 24, 4543–4551

6. 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., et al. (1995) Science 260, 496–512

7. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., Fritzgerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. M. S., and Venter, J. C. (1996) Science 273, 1058–1073

8. Zalenskaya, K., Lee, J., Gujulova, C. N., Shiu, Y. K., Slutskey, M., and Goldfarb, A. (1996) Gene (Amst.) 85, 7–12

9. Landick, R., Colwell, A., and Stewart, J. (1996) J. Bacteriol. 178, 2844–2854

10. Amann, E., Ochs, B., and Abel, K.-J. (1989) Gene (Amst.) 80, 301–315

11. Wang, D., Meier, T., Chan, C., Feng, G., Lee, D., and Landick, R. (1995) Cell 81, 341–350

12. Burgess, R. R., and Jendrisak, J. J. (1975) Biochemistry 14, 4634–4638

13. Godovikova, T. S., Grachev, M. A., Vatyan, I. V., Toore, I. G., Zarytova, V. F., and Zaychikov, E. F. (1987) Eur. J. Biochem. 166, 611–616

14. Grachev, M. A., Kolocheva, T. I., Lukhanov, E. A., and Mustave, A. (1987) Eur. J. Biochem. 163, 113–121

15. Kashlev, M., Nucler, E., Severyn, K., Burukov, S., Komissarova, N., and Goldfarb, A. (1996) Methods Enzymol. 274, 326–334

16. Nucler, E., Kashlev, M., Nikolforov, V., and Goldfarb, A. (1995) Cell 81, 353–357

17. Ridley, S. P., and Oeschger, M. P. (1982) J. Bacteriol. 152, 736–746

18. Polyakov, A., Severinova, E., and Darst, S. A. (1995) Cell 83, 365–373

19. Kashlev, M., Martin, E., Poljakov, A., Severinov, K., Nikolforov, V., and Goldfarb, A. (1996) Gene (Amst.) 130, 9–14

20. Yano, R., and Nomura, M. (1991) Mol. Cell. Biol. 11, 754–764

21. Nucler, E., Avetissova, E., Markovtsev, V., and Goldfarb, A. (1996) Science 273, 211–217

22. Deleted in proof

23. Deleted in proof

24. Ishihama, A. (1981) J. Mol. Biol. 149, 707–727

25. Wang, Y., Severinov, K., Loizos, N., Fenoy, D., Heyduk, K., Heyduk, T., Chait, B. T., and Darst, S. A. (1997) J. Mol. Biol., 270, 648–662

26. Luo, J., Sharif, K. A., Jin, R., Fujita, N., Ishiihama, A., and Krakow, J. S. (1996) Genes Cells 1, 819–827

27. Weilbaecher, R., Hebran, C., Feng, G., and Landick, R. (1994) Genes Dev., 8, 2513–2517

28. Tomch, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., et al. (1997) Nature 388, 539–547

29. Wilson, D. W., and Meacock, P. A. (1988) Nucleic Acids Res. 16, 8097–8111