Laboratory for Computational Proteomics

www.FenyoLab.org

E-mail: Info@FenyoLab.org

Facebook: *NYUMC Computational Proteomics Laboratory*

Twitter: *@CompProteomics*

Communication

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 33, Issue of August 19, pp. 20826–20828, 1994

The σ Subunit Conserved Region 3 Is Part of "5′-Face" of Active Center of *Escherichia coli* RNA Polymerase*

(Received for publication, May 31, 1994)

Konstantin Severinov‡§, David Fenyö§, Elena Severinova§¶, Arkady Mustaev‡, Brian T. Chait§, Alex Goldfarb‡, and Seth A. Darst§‖

From the ‡Public Health Research Institute, New York, New York 10016 and §The Rockefeller University, New York, New York 10021

Ribonucleotide analogs bound in the initiating site of *Escherichia coli* RNA polymerase holoenzyme in open promoter complexes were cross-linked to the β and σ^{70} subunits. Using limited proteolysis and chemical degradation, the cross-link site in σ^{70} was mapped to a segment between amino acids Glu^{508} and Met^{561} containing the C-terminal part of conserved region 3. This result, when reconciled with genetic data on the interaction of σ^{70} conserved regions 2 and 4 with the −10 and −35 promoter regions, respectively, allows us to model the orientation of the σ^{70} subunit domains within the open promoter complex.

In prokaryotes, specific initiation of transcription by the catalytically competent core RNAP[1] (subunit composition, $\alpha_2\beta\beta'$) is dependent upon binding of specificity σ factors. Different σ factors allow transcription from different sets of promoters (reviewed in Ref. 1). Sequence analysis reveals four colinear regions of extensive homology among σ factors (1, 2). The high degree of evolutionary conservation between σ factors suggests that the conserved regions participate in functions common to all σ factors, *e.g.* binding to core enzyme, promoter recognition, and DNA melting.

The main σ factor of *Escherichia coli*, σ^{70}, is also the best studied. Genetic studies indicate that part of conserved region 2 of σ^{70} (amino acids 435–443) interacts with the −10 box of the promoter (3, 4), while part of the C-terminal conserved region 4 (amino acids 584–588) recognizes the −35 promoter box (4, 5). Data from different laboratories demonstrate that σ can be cross-linked to the 5′ end of nascent RNA oligomers in early ternary complexes (6–8). Therefore, a segment of σ in the open promoter complex is close to the +1 position of the DNA template. Unfortunately, no data on localization of the nascent RNA cross-link sites in σ are available. This is partially due to the low yields of cross-linking (7). Mapping of the adduct sites in σ is further complicated by the anomalous mobility of the σ polypeptide and its fragments on SDS gels (9). In this work we map the site in σ cross-linked to initiating ATP analogs derivatized with broadly specific, cross-linkable groups. To identify σ fragments harboring the cross-link site we used a combination of chemical and proteolytic degradation of native and denatured σ subunit-RNA adducts and matrix-assisted laser desorption mass spectroscopy.

We modified RNAP in open promoter complexes with cross-linkable nucleotide derivatives (Fig. 1*A*) as the priming (+1) substrates (10, 11). The enzyme was then allowed to form the first phosphodiester bond with [α-^{32}P]UTP as specified by the +2 position of the template. As a result, the radioactive dinucleotide reporter group was covalently attached to amino acids in the vicinity of the initiating center of the enzyme (10). Denaturation and SDS-polyacrylamide gel electrophoresis of the reaction products, followed by autoradiography, revealed the derivatized RNAP subunits (Fig. 1*B*).

The alkylating groups positioned at the γ-, β-, or α-phosphates of the initiating adenine nucleotide all resulted in highly effective labeling of the β subunit[2] (10). The ATP analogue also modified appreciable amounts of the σ subunit (Fig. 1*B*). Modification of σ with the ADP* reagent was much less efficient and was only evident when the gel was overexposed. No modification of σ occurred when the AMP* reagent was used. Similar results were obtained when RNAP in open complexes formed on the *lac*UV5 and the phage λ P_L promoters were labeled (data not shown).

Trypsin and CNBr cleavage were used to fragment the σ polypeptide in order to localize the cross-linking site of the ATP* derivative. Preliminary experiments showed that the radioactive band of cross-linked σ could be excised from an SDS gel and renatured to yield active protein. The renatured σ subunit-dinucleotide adduct bound RNAP core enzyme (data not shown). Moreover, the resulting holoenzyme formed an open complex on the T7 A1 promoter, and the dinucleotide adduct could be extended with [α-^{32}P]CTP, the nucleotide specified by position +3 of the promoter (data not shown). We therefore reasoned that cleavage of the renatured σ subunit-pppApU adduct would proceed just as native σ. To this end, renatured and derivatized σ subunit was mixed with unlabeled carrier σ protein purified from superproducing cells (12) for the cleavage reactions. In each case, the products of the reactions were resolved by SDS-polyacrylamide gel electrophoresis and Coomassie staining, and polypeptides containing the cross-linked adduct were visualized by autoradiography.

Trypsin degradation of the σ polypeptide proceeds in a highly ordered manner[3] (13) and could therefore be used for cross-link mapping (Fig. 2). At low trypsin concentrations, σ was cleaved into two fragments with apparent mobilities of 65.0 and 22.5 kDa (Fig. 2*A*, *lane 2*). The 22.5-kDa fragment contained radioactivity, while the 65.0-kDa fragment did not (Fig. 2*B*, *lanes 2–5*). At higher trypsin concentrations, this band disappeared and radioactivity accumulated in a band with apparent mobil-

* This work was supported in part by grants from the Lucille P. Markey Charitable Trust, the Irma T. Hirschl Trust, and the Human Frontiers in Sciences program (to S. D.) and by National Institutes of Health Grants GM30717 (to A. G.) and RR00862 (to B. T. C.). 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.

¶ On leave from the Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia.

‖ Lucille P. Markey Scholar. To whom correspondence should be sent. Tel. 212-327-7479; Fax: 212-327-7477.

[1] The abbreviations used are: RNAP, RNA polymerase; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine.

[2] K. Severinov, A. Mustaev, E. Severinova, M. Kozlov, S. A. Darst, and A. Goldfarb, manuscript in preparation.

[3] E. Severinova, D. Fenyö, K. Severinov, B. Berkowitz, B. T. Chait, and S. A. Darst, manuscript in preparation.

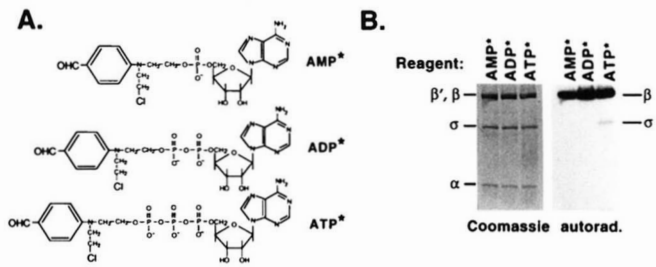

FIG. 1. **Affinity labeling of RNA polymerase.** *A*, cross-linkable derivatives of adenosine nucleotides used in this work. *B*, RNA polymerase in open complexes at the phage T7 A1 promoter was modified with the reagents shown in *panel A*. The reaction products were separated on an 8% Tris/glycine SDS-polyacrylamide gel and stained with Coomassie Blue (*left panel*). The modified subunits were visualized by autoradiography (*right panel*). Reactive derivatives of ATP, ADP, and AMP were synthesized as described (10, 11). 10-μl modification reactions contained ~0.5 μg of reconstituted wild type RNAP holoenzyme (14) and 0.5–1.0 mM derivatized initiating substrate in the transcription buffer (14). Reactions were supplemented with 10 mM sodium borohydride and incubated at 37 °C for 1 h. 100 ng of 137-base pair polymerase chain reaction-generated DNA fragment carrying the A1 promoter of phage T7 (14) was added together with 10 μCi of [α-^{32}P]UTP (3000 Ci/mmol), and incubation was continued for 30 min. The resulting ^{32}P labeling of RNAP depended on the addition of the template DNA.

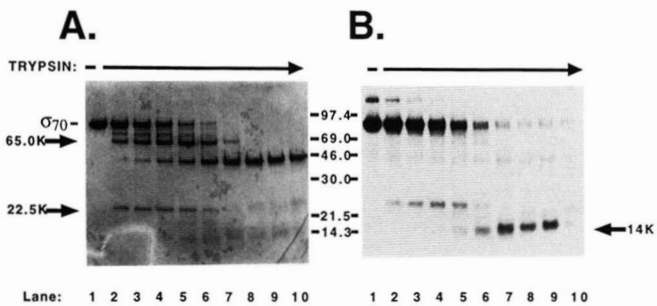

FIG. 2. **Limited proteolysis of derivatized σ subunit.** σ subunit affinity-labeled with ATP* was excised from the gel shown in Fig. 1*B* and renatured in the presence of unlabeled carrier σ subunit. Aliquots of renatured σ subunit were then treated with increasing amounts of trypsin, and reaction products were separated on a 10% Tris-Tricine SDS-polyacrylamide gel that was stained with Coomassie Blue (*A*) and autoradiographed (*B*). Affinity labeling reaction products were resolved by electrophoresis in precast Tris/glycine SDS-polyacrylamide gels (Novex). Gels were stained with 0.2 M cold KCl and autoradiographed. The σ polypeptide band was extracted from the gel and renatured according to Ref. 24. The renatured protein was mixed with unlabeled carrier σ and subjected to trypsin digestion for 30 min at room temperature. The molar ratios of trypsin to σ subunit ranged from 1:400 to 1:25. Reactions were stopped by the addition of SDS-containing electrophoresis loading buffer and immediate boiling. The low mobility radioactive band in *B* is the cross-linked β subunit, which invariably was present in trace amounts in our cross-linked σ preparations.

ity of about 14 kDa (*lanes 6–9*), which disappeared at still higher trypsin concentrations (*lane 10*).

The pattern of trypsin degradation of native σ was identical to that of cross-linked σ (data not shown). Native σ was digested under conditions that yielded the 22.5- and 65.0-kDa mobility fragments (Fig. 2*A*, *lane 2*). The N-terminal sequence of the 65.0-kDa fragment, determined using standard methods (14), was identical to the N-terminal sequence of the intact σ subunit. The N-terminal sequence of the 22.5-kDa fragment indicated that the fragment was generated by cleavage at Arg^{448}, localizing the cross-link site between Thr^{449} and the C terminus of σ at amino acid 613.

Next, σ was digested under conditions that yielded the 14-kDa mobility fragments. The products of the tryptic reaction were purified by fast protein liquid chromatography gel filtration on a Superose-6 column (Pharmacia Biotech Inc.) in the presence of 6 M guanidine hydrochloride (14). Well resolved peaks were collected, and guanidine hydrochloride was removed by membrane filtration on a Centricon-10 concentrator (Amicon). The purified polypeptides were then subjected to matrix-assisted laser desorption/ionization mass spectrometry (15) using a time-of-flight mass spectrometer constructed at the Rockefeller University (16, 17). Because of the low yield of the cross-linked σ, the mass spectrometric analysis was carried out on native σ (unmodified with affinity reagent). The mass spectrometric analysis revealed the presence of two polypeptides in the 14-kDa fraction, one of 10,933 ± 4 Da (fragment a) and the other of 10,263 ± 4 Da (fragment b). Based on the known specificity of trypsin cleavage, fragment a must correspond to either amino acids 1–99 or 466–562 of σ, while fragment b must correspond to either 1–93 or 487–578. Because we have already concluded that the cross-link site must be between Thr^{449} and the C terminus of σ the label must be within amino acids 466–562 and/or 487–578.

We next performed CNBr cleavage of the purified, cross-linked σ subunit. Upon exhaustive treatment with CNBr, radioactivity accumulated in a single band with an apparent mobility of 8 kDa (Fig. 3*A*). Taking into account the presence of radioactive dinucleotide in the adduct, which adds about 1 kDa to the mass of the polypeptide itself, and also the aberrant mobility of the σ subunit and its fragments (9), the only possible fragments with such a mobility (between Thr^{449} and the C terminus) are Glu^{508}–Met^{561} (expected M_r of 5.8 kDa) or Asn^{568}–Asp^{613} (expected M_r of 5.6 kDa). Under *single-hit* CNBr cleavage conditions (11), the smallest labeled band has a mobility of about 14 kDa (Fig. 3*B*). If the cross-link site was within Asn^{568}–Asp^{613}, the smallest labeled band would have an M_r of 5.6 kDa (as seen in the Coomassie Blue-stained gel). Thus, we finally conclude that the cross-link site is within Glu^{508}–Met^{561}.

In addition to the experiments described here, we used Trp-specific BNPS-skatole (3-bromo-3-methyl-2-(2-nitrophenylmercapto)-3*H*-indole) and Cys-specific 2-nitro-5-thiocyanobenzoic acid treatment to fragment the σ polypeptide (18). The results of these experiments were all consistent with and confirmed our conclusions presented above (data not shown). Thus, at least one σ subunit amino acid between Glu^{508} and Met^{561} is located within ~5 Å of the γ-phosphate of the priming nucleotide, which is the effective range of the probe used in this study. It should be noted that although this region of the σ subunit is close to the priming substrate, the cross-linked amino acid in this region cannot be involved in catalysis since the cross-linked adducts are still able to form at least two phosphodiester bonds (not shown). Elsewhere, we show that the β subunit Rif-region^2, as well as Lys^{1065} and His^{1237}, are also cross-linked with the reagents used in this work (19). Hence, in the open complex the distance between these sites in the β subunit and the cross-link site in σ should not exceed ~10 Å.

The C-terminal portion of σ conserved region 3 (σ^{70} amino acids 475–520) is contained within the σ^{70} fragment we have identified as containing the ATP* cross-link site. Region 3 is found in only a subset of σ factors and when present, the sequence is weakly conserved relative to conserved regions 2 and 4 (1, 2). Nevertheless, the sequence conservation observed is highly suggestive of structural and functional conservation. Also in contrast to regions 2 and 4, genetic studies to elucidate the function of region 3 have been few. A mutant *E. coli* σ^{32} with a small deletion in region 3 (corresponding to σ^{70} amino acids 503–520 in the aligned sequences) has been investigated (20). The protein, which otherwise appears to function normally, exhibits a reduced affinity for core RNAP. This is not inconsistent with our results as we have shown this region of σ^{70} is close to specific sites on the β subunit.

The principal conclusion from our experiments pertinent to

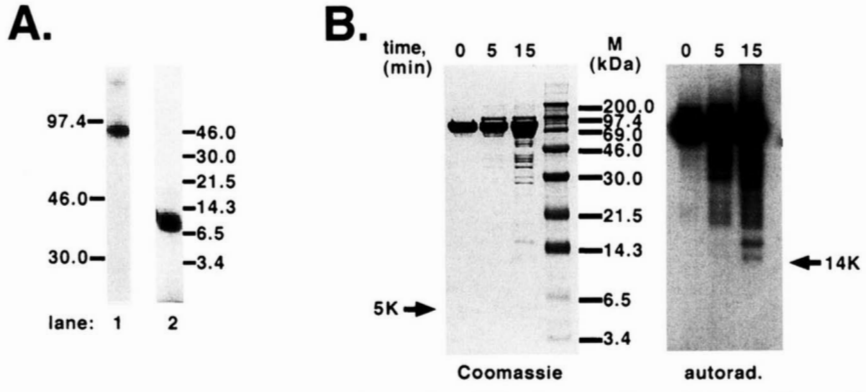

FIG. 3. **CNBr degradation of the cross-linked σ subunit.** *A*, complete CNBr cleavage of cross-linked σ. The purified radioactive σ shown in *lane 1* (autoradiograph of a 8% Tris/glycine polyacrylamide gel) was treated with CNBr as described (25), and reaction products were separated on a 16% Tris-Tricine SDS-polyacrylamide gel (*lane 2*) and visualized by autoradiography. *B*, incomplete CNBr cleavage of cross-linked σ. The reaction products were resolved on a 16% Tris-Tricine polyacrylamide gel, stained with Coomassie Blue (*left panel*), and autoradiographed (*right panel*) (see Ref. 11 for the conditions of incomplete CNBr cleavage and interpretation of the results).

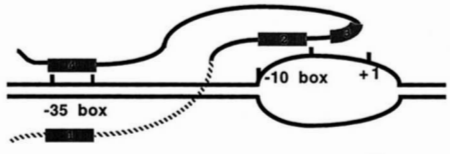

FIG. 4. **A model of the orientation of the σ^{70} domains in the open complex relative to T7 A1 promoter DNA.** See the text for discussion. The N-terminal part of σ containing conserved region 1 is drawn to present the conserved regions 1 and 4 interaction proposed to occur in free σ (21). This interaction may not occur in the open complex (see Ref. 21 for discussion).

the structure of RNA polymerase is that the results of our cross-link mapping, taken together with available genetic evidence on the interaction of σ conserved regions 2 and 4 with the −10 and −35 promoter boxes, respectively, allow us to model the orientation of the σ conserved domains in the open complex relative to promoter DNA as shown on Fig. 4. The main feature of this model is that σ domains 2 and 3 are aligned "parallel" with respect to promoter DNA, but then the protein must flip back somewhere near the +1 position of the template DNA and proceed in an antiparallel orientation to the −35 promoter box. We note that such an orientation would allow an interaction between conserved regions 1 and 4 (Fig. 4), which agrees with recent findings of Dombroski *et al.* (21).

The −35 and −10 promoter boxes are separated by a distance of 25 base pairs or about 85 Å, assuming straight, B-form DNA. If σ were spherical and had a protein density of 1.3 g/cm^3, its diameter would be about 56 Å, much less than the distance it must span to interact simultaneously with the −35 and −10 promoter boxes. Our model in Fig. 4 now suggests that the distance from +1 to −35 (about 119 Å for straight, B-form DNA) must be spanned by only domains 3 and 4 (consisting of only about 18-kDa protein mass). Bending of the promoter DNA in the open complex so that the −35 promoter region is closer to +1 seems highly likely given these circumstances. Bent DNA in RNA polymerase-promoter binary complexes has previously been proposed based on gel electrophoretic experiments (22) and structural considerations (23).

REFERENCES

1. Helmann, J. D., and Chamberlin, M. J. (1988) *Annu. Rev. Biochem.* **57,** 839–872
2. Lonetto, M., Gribskov, M., and Gross, C. A. (1992) *J. Bacteriol.* **174,** 3843–3849
3. Siegele, D. A., Hu, J. C., Walter, W. A., and Gross, C. A. (1989) *J. Mol. Biol.* **206,** 591–603
4. Waldburger, C., Gardella, T., Wong, R., and Susskind, M. M. (1990) *J. Mol. Biol.* **215,** 267–276
5. Gardella, T., Moyle, T., and Susskind, M. M. (1989) *J. Mol. Biol.* **206,** 579–590
6. Grachev, M. A., and Zaychikov, E. F. (1981) *FEBS Lett.* **130,** 841–847
7. Bernhard, S. L., and Meares, C. F. (1986) *Biochemistry* **25,** 5914–5919
8. Osumi-Davis, P. A., Woody, A.-Y. M., and Woody, R. W. (1987) *Biochim. Biophys. Acta* **910,** 130–141
9. Strickland, M. S., Thompson, N. E., and Burgess, R. R. (1988) *Biochemistry* **27,** 5755–5762
10. Grachev, M. A., Kolocheva, T. I., Lukhtanov, E. A., and Mustaev, A. A. (1987) *Eur. J. Biochem.* **163,** 113–121
11. Grachev, M. A., Lukhtamov, E. A., Mustaev, A. A., Zaychikov, E. F., Abdukayumov, M. N., Rabinov, I. V., Richter, V. I., Skoblov, Y. S., and Chistyakov, P. G. (1989) *Eur. J. Biochem.* **180,** 577–585
12. Gribskov, M., and Burgess, R. R. (1983) *Gene (Amst.)* **25,** 167–178
13. Burgess, R. R., Erickson, B., Gentry, D., Gribskov, M., Hager, D., Lesley, S., Strickland, M., and Thompson, N. (1987) in *RNA Polymerase and the Regulation of Transcription* (Reznikoff, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, M. T., Jr., and Wickens, M. P., eds) pp. 3–15, Elsevier Science Publishing Corp., New York
14. Severinov, K., Mustaev, A., Kashlev, M., Borukhov, S., Nikiforov, V., and Goldfarb, A. (1992) *J. Biol. Chem.* **267,** 12813–12819
15. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T. (1991) *Anal. Chem.* **63,** 1193A-1203A
16. Beavis, R. C., and Chait, B. T. (1989) *Rapid Commun. Mass Spectrom.* **3,** 233–237
17. Beavis, R. C., and Chait, B. T. (1990) *Anal. Chem.* **62,** 1836–1840
18. Borukhov, S., Lee, J., and Goldfarb, A. (1991) *J. Biol. Chem.* **266,** 23932–23935
19. Mustaev, A., Kashlev, M., Lee, J., Polyakov, A., Lebedev, A., Zalenskaya, K., Grachev, M., Goldfarb, A., and Nikiforov, V. (1991) *J. Biol. Chem.* **266,** 23927–23931
20. Zhou, Y. N., Walter, W. A., and Gross, C. A. (1992) *J. Bacteriol.* **174,** 5005–5012
21. Dombroski, A. J., Walter, W. A., and Gross, C. A. (1993) *Genes & Dev.* **7,** 2446–2455
22. Kuhnke, G., Fritz, H., and Ehring, R. (1987) *EMBO J.* **6,** 507–513
23. Darst, S. A., Kubalek, E. W., and Kornberg, R. D. (1989) *Nature* **340,** 730–732
24. Hager, D. A., and Burgess, R. R. (1980) *Anal. Biochem.* **109,** 76–86
25. Price, N., Redpath, N., Severinov, K., Campbell, D., Russel, R., and Proud, C. (1991)*FEBS Lett.* **282,** 253–258