



computational proteomics

Laboratory for Computational Proteomics

www.FenyoLab.org

E-mail: Info@FenyoLab.org

Facebook: [*NYUMC Computational Proteomics Laboratory*](#)

Twitter: [*@CompProteomics*](#)

Determinants for *Escherichia coli* RNA Polymerase Assembly within the β Subunit

Yi Wang¹, Konstantin Severinov¹, Nick Loizos¹, David Fenyö²
Ewa Heyduk³, Tomasz Heyduk³, Brian T. Chait² and Seth A. Darst^{1*}

¹Laboratory of Molecular Biophysics, The Rockefeller University, 1230 York Avenue New York, NY 10021, USA

²Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, 1230 York Avenue New York, NY 10021, USA

³Department of Biochemistry and Molecular Biology St. Louis University Medical School, 1402 S. Grand Boulevard, St. Louis MO 63104, USA

*Corresponding author

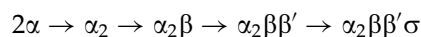
We used binding assays and other approaches to identify fragments of the *Escherichia coli* RNAP β subunit involved in the obligatory interaction with the α subunit to form the stable assembly intermediate $\alpha_2\beta$ as well as in the interaction to recruit the β' subunit into the $\alpha_2\beta$ sub-assembly. We show that two regions of evolutionarily conserved sequence near the C terminus of β (conserved regions H and I) are central to the assembly of RNAP and likely make subunit-subunit contacts with both α and β' .

© 1997 Academic Press Limited

Keywords: *E. coli* RNA polymerase assembly; transcription; limited proteolysis; protein-protein footprinting

Introduction

The *Escherichia coli* DNA-dependent RNA polymerase (RNAP) comprises an essential catalytic core of two α -subunits (each 36.5 kDa), one β (150.6 kDa), and one β' (155.2 kDa) subunit. The holoenzyme has an additional regulatory subunit, normally σ^{70} (70.2 kDa). The RNAP assembles according to the following pathway (reviewed by Zillig *et al.*, 1976; Ishihama, 1981):



The RNAP α subunit serves as the initiator for RNAP assembly. Igarashi & Ishihama (1991) showed that truncated α derivatives lacking about 100 C-terminal residues are completely competent for dimerization, RNAP assembly, and basal transcription. Five regions of the α subunit (designated A to E), all within the N-terminal assembly domain, are conserved in sequence among α homologs of prokaryotic, eukaryotic, archaeobacterial,

and chloroplast RNAPs (reviewed by Ebright & Busby, 1995; Heyduk *et al.*, 1996). A number of mutagenesis studies have localized determinants for α dimerization and interaction with the other RNAP subunits to within these conserved regions (Igarashi *et al.*, 1990, 1991; Hayward *et al.*, 1991; Igarashi & Ishihama, 1991; Kimura *et al.*, 1994; Kimura & Ishihama, 1995a,b). Hydroxyl-radical protein footprinting was used to confirm these results and strengthen the conclusion that α conserved regions A and B are involved in direct interactions with the β subunit while α conserved regions C and D are involved in direct interactions with the β' subunit (Heyduk *et al.*, 1996). The sequence conservation within these regions in all organisms suggests that α homologs in archaeobacterial and eukaryotic RNAPs perform analogous functions (Kolodziej & Young, 1991; Azuma *et al.*, 1993; Lalo *et al.*, 1993; Pati, 1994; Ulmasov *et al.*, 1996).

In contrast to the α subunit, very little is known about regions of the β and β' subunits involved in RNAP assembly and subunit-subunit interactions. Sequence comparisons among eubacterial, archaeobacterial, and eukaryotic β subunit homologs reveal regions of high sequence similarity separated by regions that are poorly conserved (Falkenburg

Abbreviations used: MALDI-MS, Matrix-assisted laser desorption/ionization mass spectrometry; NTD, N-terminal domain; RNAP, RNA polymerase; SDS/PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; wt, wild type.

et al., 1987; Sweetser *et al.*, 1987; Berghofer *et al.*, 1988; Pühler *et al.*, 1989; see Figure 4). We have recently shown that active *E. coli* RNAP can be assembled *in vitro* using β subunit fragments along with the other RNAP subunits (Severinov *et al.*, 1995a, 1996). The N and C-terminal boundaries of these fragments roughly delineate independent structural domains of the β subunit that do not have to be covalently linked for assembly and activity of the RNAP.

Because of the apparently analogous role of the α subunit and its regions of conserved primary structure in *E. coli* RNAP assembly to the role of α homologs in RNAP assembly in other organisms including man, it seems reasonable to assume that regions of conserved sequence in the β subunit are also involved in assembly. Which conserved region(s) of β ? In this study, we use a combination of approaches to identify the conserved regions of the *E. coli* RNAP β subunit involved in the obligatory interaction with α to form the stable assembly intermediate $\alpha_2\beta$ as well as in the interaction to recruit the β' subunit into the $\alpha_2\beta$ sub-assembly. We show that the two C-terminal most conserved regions of β (conserved regions H and I according to the notation of Sweetser *et al.*, 1987) are central to the assembly of RNAP and likely make subunit-subunit contacts with both α and β' .

Results

The C-terminal portion of the β subunit interacts with α

We have shown previously that the β subunit can be split nearly in half, into two separate peptide fragments, β residues 1 to 643 and 643 to 1342 (denoted β_{ABCD} and β_{EFGHI} , respectively, to reflect the locations of the conserved regions; see Table 1 for a description of all the β subunit fragments used in this study), but still assemble *in vitro* with the other RNAP subunits into active RNAP (Severinov *et al.*, 1996). We used RNAP α subunit engineered with an N-terminal hexahistidine tag ($\text{His}_6\text{-}\alpha$; Tang *et al.*, 1995) in a Ni^{2+} -NTA agarose co-immobilization assay to investigate the binding of these β subunit fragments to α (Figure 1). RNAP immobilized by virtue of $\text{His}_6\text{-}\alpha$ subunits was transcriptionally active in the immobilized state (Tang *et al.*, 1995), indicating that the N terminus of at least one of the α subunits must be solvent exposed since immobilization of the RNAP in this way did not interfere with any necessary subunit-subunit interactions. Since overexpressed β and its fragments were generally found in inclusion bodies, the binding experiments were done by mixing $\text{His}_6\text{-}\alpha$ and β (or β fragments) together in the denatured state (in buffer containing 6 M guanidinium-HCl) and then removing the denaturant by dialysis (see Materials and Methods). When a mixture of $\text{His}_6\text{-}\alpha$ and β was reconstituted and then loaded onto Ni^{2+} -NTA agarose beads (Figure 1a, lane 1), the unbound material removed (Figure 1a, lane 2),

Table 1. Characteristics and nomenclature of β subunit fragments used in α binding experiments

Protein	β Conserved regions	β Residues
β	ABCDEFGHI	1-1342
β_{ABCDEFG}	ABCDEFG	1-989
β_{ABCD}	ABCD	1-643
β_{EFGHI}	EFGHI	643-1342
β_{FGHI}	FGHI	711-1342
β_{FGHI_N}	FGHI_N	711-1246
β_{FGH}	FGH	711-1114
β_{HI}	HI	907-1342
β_{HI_N}	HI_N	907-1246
$\beta_{\text{[950-1342]}}$	HI	950-1342
β_{EFGH}	EFGH	643-1153
β_{I}	I	1153-1342
$\beta_{\text{FGHI}_N(\Delta[937-983])}$	FGHI_N	711-1246($\Delta[937-983]$)
$\beta_{\text{FGHI}_N(\Delta[946-1062])}$	FGHI_N	711-1246($\Delta[946-1062]$)

the beads washed with 25 mM imidazole buffer (Figure 1a, lane 3), then eluted with 100 mM imidazole buffer, β was found in the eluted fraction along with $\text{His}_6\text{-}\alpha$ (Figure 1a, lane 4). Since renatured β in the absence of $\text{His}_6\text{-}\alpha$ did not interact with the beads (Figure 1a, lane 8), we conclude that β was retained in the presence of $\text{His}_6\text{-}\alpha$ by virtue of a direct interaction with immobilized $\text{His}_6\text{-}\alpha$, as expected. When a mixture of $\text{His}_6\text{-}\alpha$ and β_{ABCD} was loaded onto Ni^{2+} -NTA agarose beads (Figure 1a, lane 9), all of the β_{ABCD} appeared in the unbound fraction (Figure 1a, lane 10), whereas an analogous experiment with $\text{His}_6\text{-}\alpha$ and β_{EFGHI} indicated an apparently stoichiometric interaction of β_{EFGHI} with the immobilized $\text{His}_6\text{-}\alpha$. In the absence of $\text{His}_6\text{-}\alpha$, β_{EFGHI} did not interact with the Ni^{2+} -NTA agarose beads (Figure 1a, lane 20) and so we conclude that β_{EFGHI} , by itself, directly interacts with the α subunit whereas the remainder of the β polypeptide, β_{ABCD} , does not. Moreover, when a mixture of $\text{His}_6\text{-}\alpha$, β_{ABCD} , and β_{EFGHI} were loaded together onto Ni^{2+} -NTA agarose beads, only β_{EFGHI} was retained with $\text{His}_6\text{-}\alpha$ in the eluted fraction (Figure 1b). Thus, β_{ABCD} also does not interact with the complex between α and β_{EFGHI} and so all of the determinants for the interaction between α and β appear to be contained within β_{EFGHI} .

Identification of β subunit domains in complex with α by limited proteolysis

Limited proteolysis has often been used to define the domain organization of proteins (Wilson, 1991). When β alone was treated with limiting amounts of a variety of proteases, the C-terminal half was rapidly degraded (K.S., D.F., E. Severinova, B.T.C., and S.A.D., unpublished results). We designed an assay to identify, if present, the proteolytically-resistant core of β_{EFGHI} while in a complex with α . A battery of five proteases was used; trypsin, chymotrypsin, proteinase K, Pronase, and endoproteinase Glu-C. Rather than full length α , which itself is susceptible to proteolysis (Blatter *et al.*, 1994), we used the assembly-competent N-

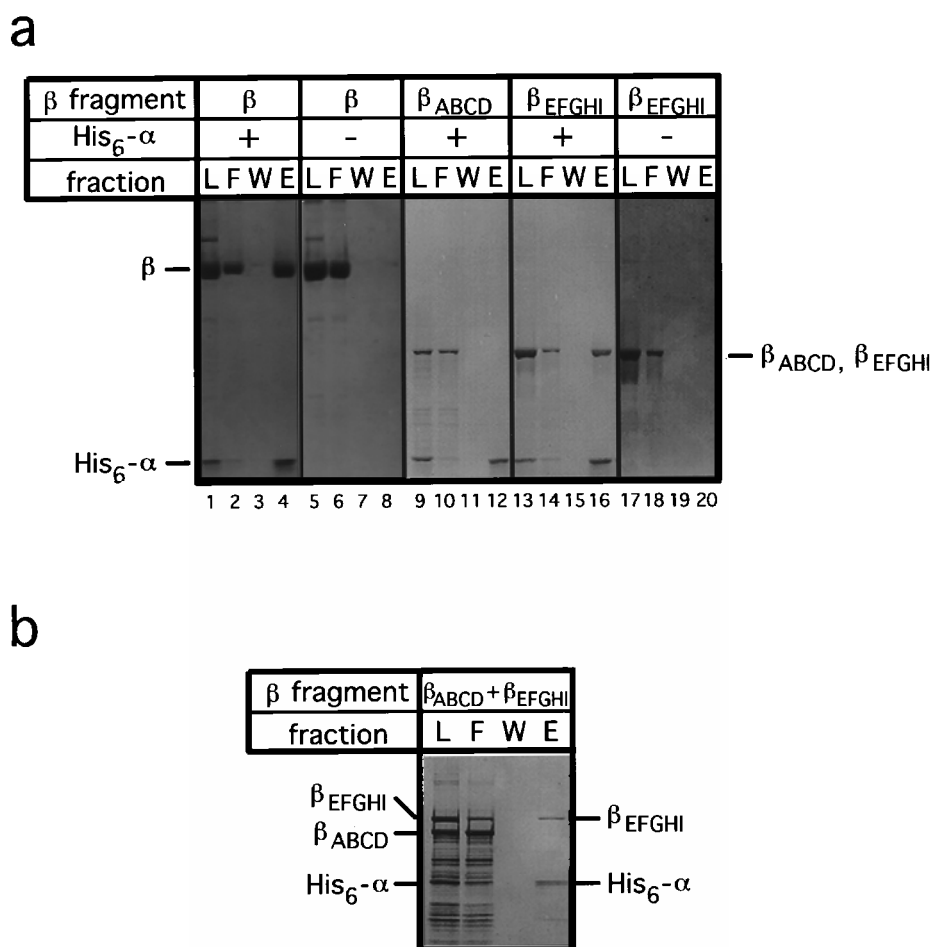


Figure 1. Binding of full length β and β_{EFGHI} , but not β_{ABCD} , to immobilized His₆- α by the Ni²⁺-co-immobilization assay. **a**, His₆- α was renatured with a molar excess of full length β (lanes 1 to 4), β_{ABCD} (lanes 9 to 12), or β_{EFGHI} (lanes 13 to 16), loaded onto Ni²⁺-NTA-agarose beads (L) and the unbound protein (or flow-through) removed (F). The beads were then washed with buffer containing 25 mM imidazole (W), then the bound proteins were eluted with buffer containing 100 mM imidazole (E). The protein fractions were then analyzed by SDS-PAGE on an 8% gel (NOVEX). His₆- α was omitted from the experiments shown in lanes 5 to 8 and lanes 17 to 20 as a control for non-specific binding of β or its fragments to the Ni²⁺-NTA-agarose beads. **b**, Ni²⁺-co-immobilization assay with His₆- α renatured with molar excesses of β_{ABCD} and β_{EFGHI} [$\Omega T988$] simultaneously. Since β_{ABCD} and β_{EFGHI} are nearly the same size and co-migrate on the SDS-PAGE system used, β_{EFGHI} [$\Omega T988$], with a 127 residue insertion after amino acid 988 (Kashlev *et al.*, 1989; Borukhov *et al.*, 1991), and thus with a lower mobility, was used. In this assay, β_{EFGHI} [$\Omega T988$] was added to the renaturation mix from a crude inclusion body preparation, explaining the presence of numerous contaminating bands in the load and flow-through fractions.

terminal domain (NTD) of α engineered with an N-terminal hexahistidine tag (His₆- α [1-235], hereafter referred to as His₆- α NTD; Tang *et al.*, 1995). Control experiments showed that His₆- α NTD, when immobilized on Ni²⁺-NTA agarose, was completely resistant to cleavage even by the highest concentrations of each of the five proteases used (data not shown). The assay, then, went as follows. (1) (His₆- α NTD)₂ β_{EFGHI} complex was reconstituted by renaturation from the denatured state. (2) The complex was immobilized on Ni²⁺-NTA agarose beads. The beads were washed to remove unbound material (mainly excess β_{EFGHI}). (3) The immobilized complexes were treated with limiting amounts of protease. In separate reactions, a wide concentration range of each of the five proteases was tested. (4) The beads were washed

again to remove protease and unbound proteins. (5) The bound proteins were eluted from the beads with buffer containing 100 mM imidazole and analyzed by SDS-PAGE.

Limited proteolysis of the complexes with chymotrypsin, proteinase K, and trypsin revealed proteolytically-resistant domains of β_{EFGHI} in that discrete bands were generated that were stable over a wide range of protease concentrations (Figure 2). Pronase and endoproteinase Glu-C did not yield stable products that could be analyzed further, probably due to a preponderance of accessible sites for these two enzymes.

Limited proteolysis with chymotrypsin, and analysis of the bound peptides by SDS-PAGE, yielded two major products derived from β_{EFGHI} (Figure 2a). Treatment with low concentrations of

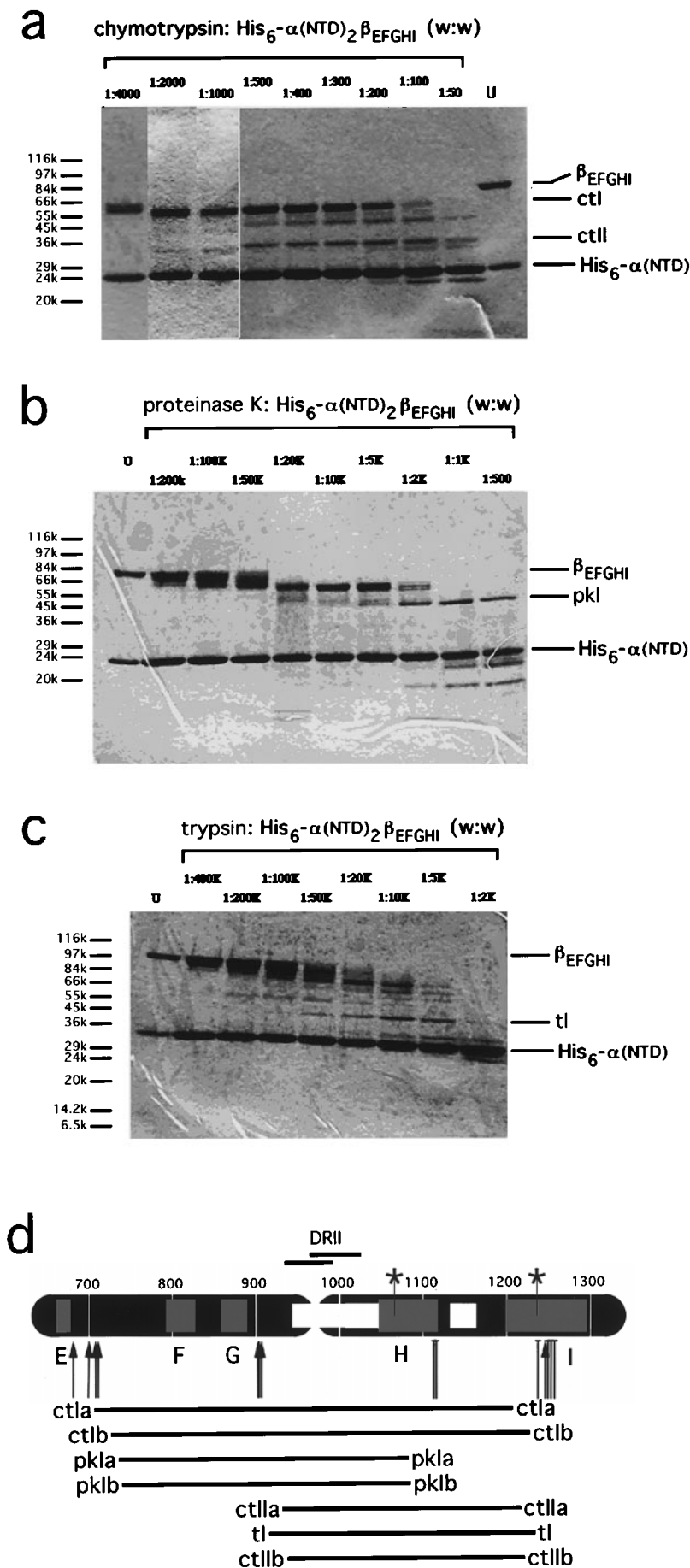


Figure 2. Limited proteolysis of Ni²⁺-NTA-agarose-immobilized (His₆-α(NTD)₂β_{EFGHI}) analyzed by SDS-PAGE on 4% to 20% gradient gels (Novex). In each gel, the lane labeled U is untreated with protease. The ratios across the top of the other lanes indicate the mass ratio of protease to (His₆-α(NTD)₂β_{EFGHI}). The positions of molecular mass markers (Sigma) are indicated to the left of each gel. a, Limited proteolysis using chymotrypsin. b, Limited proteolysis using proteinase K. c, Limited proteolysis using trypsin. d, Summary of proteolysis results and mapping of β fragments using N-terminal sequencing and MALDI-MS. The horizontal black bar represents the primary sequence of β_{EFGHI} with amino acid numbering corresponding to full length β shown above the bar. Evolutionarily conserved regions are shaded grey and labeled E to I below the bar according to Sweetser *et al.* (1987). The open box represents a region containing large deletions found in several chloroplast RNAP β homologs (Hudson *et al.*, 1988). Deletions encompassing dispensable region II (DRII; Borukhov *et al.*, 1991) are represented as bars above the primary structure. The locations of Lys1065 and His1237, which cross-link to initiating nucleotide analogs (Mustaev *et al.*, 1991), are indicated (*). Below the primary structure are the results of mapping the N and C termini of the major protease-resistant fragments of β resulting from limited proteolysis of (His₆-α(NTD)₂β_{EFGHI}). Cleavage sites that were mapped to within one residue are indicated by vertical arrows, sites that were mapped to within a range of residues are indicated by a line spanning that range. The labeled horizontal lines below illustrate the identified fragments in the context of the primary structure.

chymotrypsin yielded a band with a mobility corresponding to about 60 kDa (Figure 2a, ct-I). As the chymotrypsin concentration was increased, a second stable product accumulated with a mobility corresponding to about 35 kDa (Figure 2a, ct-II). N-terminal sequencing of these two products revealed that each contained a major and a minor component. The band labeled ct-I contained proteins with N termini at β residue 700 (major component) and 681 (minor). The band labeled ctII contained proteins with N termini corresponding to 907 (major) and 903 (minor). All of these N termini are consistent with the cleavage specificity of chymotrypsin. Next, the following experiment was performed. (1) $(\text{His}_6\text{-}\alpha\text{NTD})_2\beta_{\text{EFGHI}}$ complex, immobilized on Ni^{2+} -NTA agarose, was subjected to chymotrypsin cleavage at a ratio of chymotrypsin:complex of 1:300 (w/w). At this concentration of chymotrypsin, significant amounts of both ct-I and ct-II were generated (Figure 2a). (2) The beads were washed to remove unbound proteins. (3) The beads were then suspended in buffer containing 6 M guanidine-HCl to denature and elute the remaining β fragments from the immobilized $\text{His}_6\text{-}\alpha\text{NTD}$. (4) The eluted β fragments were then subjected to matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS; Hillenkamp *et al.*, 1991) to accurately measure their masses.

From the N-terminal sequences, the measured masses, and with consideration of the cleavage specificity of chymotrypsin, the C termini of the proteolytic fragments could be roughly inferred. The N-terminal sequencing and MALDI-MS data that led to the identification of the β fragments are tabulated in Table 2.

Limited proteolysis with proteinase K, and analysis of the bound peptides by SDS-PAGE, yielded one major, stable product from β_{EFGHI} (pk-I, Figure 2b) with a mobility corresponding to about 46 kDa. N-terminal sequencing of the product labeled pk-I revealed that it contained two components with N termini at β residues 708 and 711, consistent with the broad specificity of proteinase K for peptide bonds C-terminal of hydrophobic amino acids. The products (predominantly pk-I) resulting from limited proteolysis at a ratio of proteinase K:complex of 1:500 (w/w) were analyzed by MALDI-MS, allowing the C termini to be roughly determined (Table 2).

Limited proteolysis with trypsin, and analysis of the bound peptides by SDS-PAGE, yielded one major, stable product from β_{EFGHI} (t-I, Figure 2c) with a mobility corresponding to about 37 kDa. N-terminal sequencing of the product labeled t-I revealed that it contained one component with N terminus at β residue 904, consistent with the cleavage specificity of trypsin and also with an earlier study of RNAP trypsinolysis (Borukhov *et al.*, 1991). The products (predominantly t-I) resulting from limited proteolysis at a ratio of trypsin:complex of 1:5000 (w/w) were analyzed by MALDI-MS, allowing the C terminus to be determined (Table 2). The results of this analysis of the domain organization of β_{EFGHI} in complex with α are schematically summarized in Figure 2d.

Binding of β subunit fragments to αNTD

The experiments described in the previous section were designed to identify proteolytically resistant β subunit fragments that remained specifically bound (and thus immobilized on the Ni^{2+} -NTA agarose) to $\text{His}_6\text{-}\alpha\text{NTD}$. However, control experiments were performed where the β subunit fragments, generated by proteolysis, were eluted from the Ni^{2+} -NTA agarose beads and immobilized $\text{His}_6\text{-}\alpha\text{NTD}$ by buffer containing 6 M guanidine-HCl, renatured by dialysis in the absence of α , then reloaded on Ni^{2+} -NTA agarose beads. Many of the β subunit fragments were immobilized even though $\text{His}_6\text{-}\alpha$ was not present, indicating non-specific binding of the β subunit fragments to the Ni^{2+} -NTA agarose beads. The β subunit fragments could be eluted from the beads as the imidazole concentration was increased, but the complicated nature of the proteolysis reactions combined with the differential elution of the β subunit fragments from the beads over a wide range of imidazole concentrations made unambiguous determination of specific α binding *versus* non-specific immobilization difficult. In order to simplify the analyses, we used PCR methods to subclone five β subunit fragments (β_{FGHI} , β_{FGHI_N} , β_{FGH} , β_{HI} , and β_{HI_N} ; Table 1), incorporating representative N and C-terminal cleavage sites identified in the proteolysis experiments, into a T7-RNAP based overexpression system (Studier *et al.*, 1990).

Table 2. Identification of β subunit fragments generated by limited proteolysis

Protein	β Conserved regions	N-terminal sequence	Observed mass (kDa)	β Residues		Calculated mass (kDa)
				N	C	
β_{ctIb}	FGHI_N	XXANMQRQAVP	61.5	681	1234-1241	61.2-61.9
β_{ctIIb}	FGHI_N	VGTMERAV	61.5	700	1253-1261	61.1-61.9
β_{pkIa}	FGH	VAVDSGVTAV	44.7	708	1110-1116	44.3-45.0
β_{pkIb}	FGH	DSGVTAVAKR	44.7	711	1113-1119	44.3-45.0
β_{ctIIa}	HI_N	XAIFGEKAS	38.9	903	1245-1252	38.5-39.3
β_{tI}	HI_N	AIFGEKASDV	38.8	904	1246	38.5
β_{ctIIb}	HI_N	XEKASDVKDD	38.9	907	1249-1256	38.5-39.2

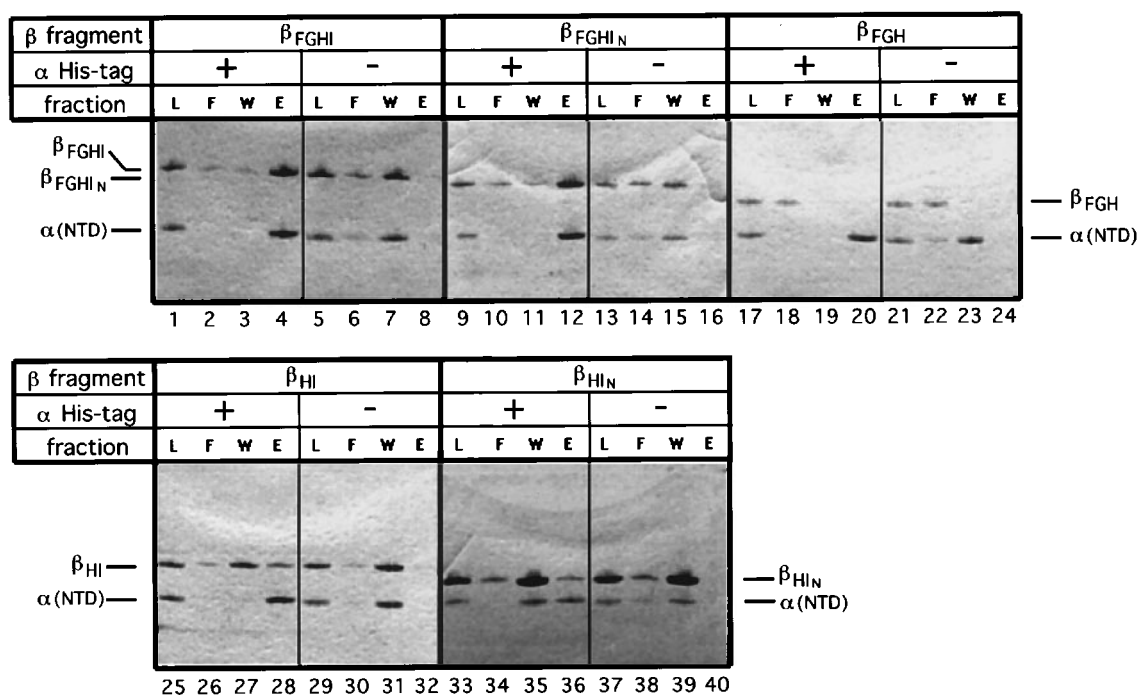


Figure 3. Binding of β fragments to α NTD. Ni^{2+} -co-immobilization assays of β fragment binding to either His₆- α NTD (lanes 1 to 4, 9 to 12, 17 to 20, 25 to 28, and 33 to 36) or α NTD (lanes 5 to 8, 13 to 16, 21 to 24, 29 to 32, and 37 to 40), as a control for non-specific binding of the β fragment to the Ni^{2+} -NTA-agarose beads. Load (L), flow-through (F), wash (W), and elution (E) fractions were analyzed by SDS-PAGE on 8 to 25% gradient PhastGels (Pharmacia).

After overexpression, about 50% of β_{FGHI_N} was found in the cytosol, the other 50% in inclusion bodies. All of the other fragments were found almost totally in inclusion bodies and so all of the reconstitutions were done from inclusion body preparations. The binding of each fragment to α was investigated using Ni^{2+} -co-immobilization assays with His₆- α NTD, and in side-by-side reactions with α NTD as a control for non-specific binding of the β fragments to the beads (Figure 3). After binding to the beads and removal of the unbound material, a wash at an intermediate concentration of 20 to 30 mM imidazole was sufficient in each case to remove non-specifically bound β fragments. For example, in parallel reactions, co-immobilization of β_{FGHI} by either His₆- α NTD (Figure 3, lanes 1 to 4) or α NTD (Figure 3, lanes 5 to 8) was tested. Renatured material was loaded onto the beads at 2.5 mM imidazole (lanes 1 and 5) and unbound material removed (lanes 2 and 6). Non-specifically bound proteins were then removed by washing at 25 mM imidazole (lanes 3 and 7). Finally, specifically immobilized proteins were eluted at 100 mM imidazole (lanes 4 and 8). Roughly stoichiometric amounts of β_{FGHI} eluted from the beads with His₆- α NTD (assuming two copies of His₆- α NTD per complex; Figure 3, lane 4), while under identical conditions, no β_{FGHI} remained on the beads when α NTD was used (Figure 3, lane 8), indicating strong, specific binding of β_{FGHI} to α NTD. Strong, specific binding of β_{FGHI_N} (Figure 3, lanes 12 and 16) indicates that β residues C-terminal of 1246, including approximately the C-terminal half of con-

served region I, do not participate either directly or indirectly in α binding. In contrast, β_{FGH} was not found in the elution fractions whether His₆- α NTD or α NTD were used (Figure 3, lanes 20 and 24, respectively), indicating a requirement for the N-terminal half of conserved region I for α binding. The fragments β_{HI} and β_{HI_N} both bound specifically to α (Figure 3, lanes 28, 32, 36, and 40), confirming that the C-terminal half of conserved region I is not involved in α binding and also indicating that regions F and G are not required for α binding. The binding of β_{HI} and β_{HI_N} to α was weak in comparison with the α binding of β_{FGHI} and β_{FGHI_N} however, indicating that regions F and G may participate somehow in stabilizing the α - β interaction. The α binding results described thus far (Figures 1 and 3) are summarized in Figure 4. Also summarized in Figure 4 are the α binding results for three additional β fragments that were available in the laboratory (Table 1), $\beta[950-1342]$, containing conserved regions H and I but with 43 amino acids truncated from the N-terminus compared with β_{HI} (Severinov *et al.*, 1995a), and two β fragments corresponding to β_{FGHI_N} but with deletions in dispensable region II between conserved regions G and H, $\beta_{FGHI_N}(\Delta[937-983])$ (Borukhov *et al.*, 1991) and $\beta_{FGHI_N}(\Delta[946-1062])$. From the α binding results of all these β subunit fragments derived from β_{FGHI} (Figure 4), we conclude that β conserved region H and the N-terminal half of conserved region I (I_N) are necessary and sufficient, and that I_N is required, for binding of β to α .

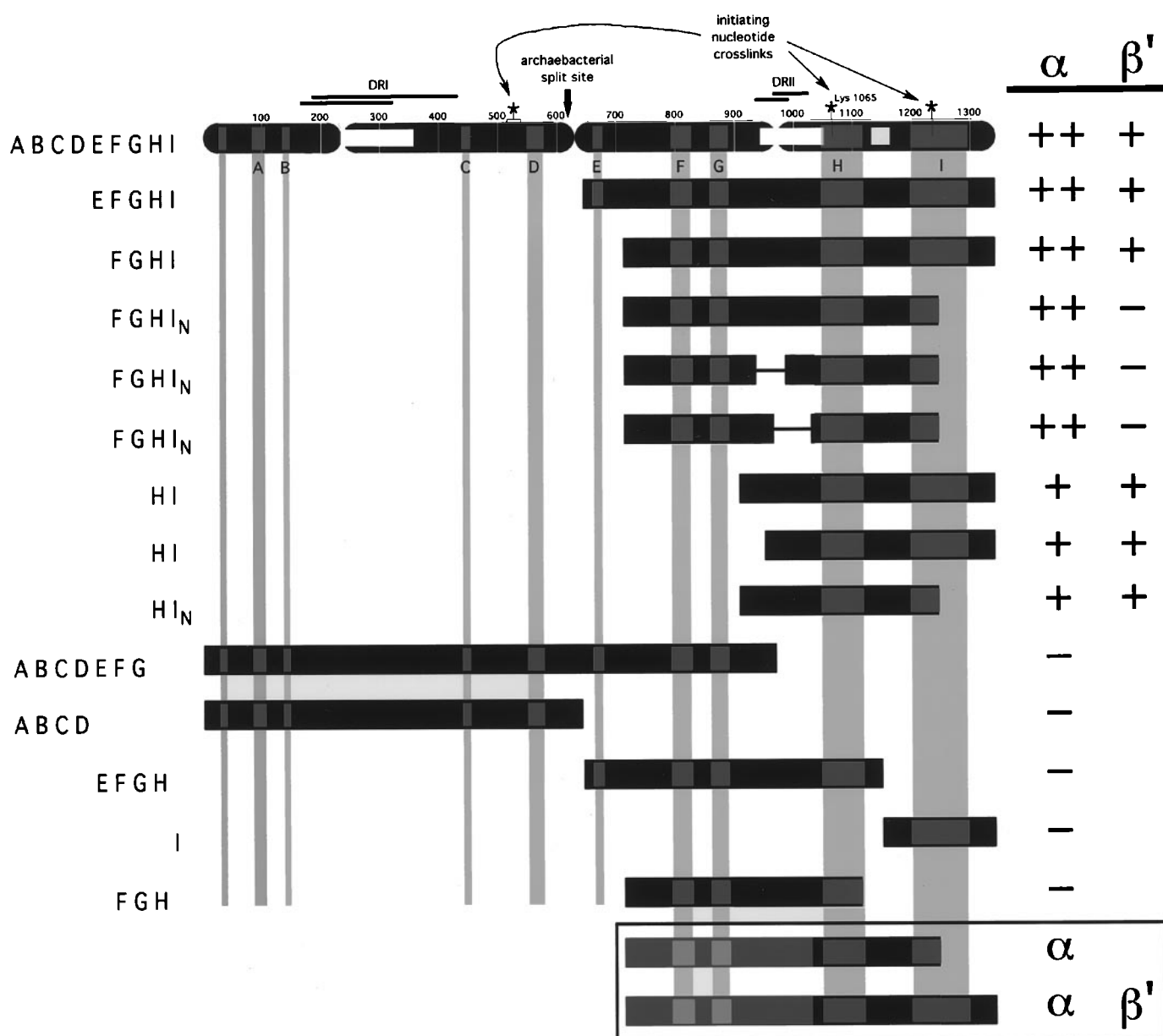


Figure 4. Summary of assembly results. The horizontal black bar at the top represents the primary sequence of β with amino acid numbering shown above the bar. Evolutionarily conserved regions are shaded grey and labeled A to I below the bar according to (Sweetser *et al.*, 1987). The open boxes represent regions containing large deletions found in several chloroplast RNAP β homologs (Hudson *et al.*, 1988). Dispensable regions are represented as bars above the primary structure and labeled dispensable region (DR) I or II. The locations of initiating nucleotide crosslinks (Mustaev *et al.*, 1991; Severinov *et al.*, 1995a,b) are indicated (*). Below the primary structure are illustrated the β subunit fragments investigated in this study and the conserved regions contained within them are listed on the left. The results of α binding are shown qualitatively in the first column to the right of the fragments (++ indicates approximately stoichiometric binding to α , + indicates weaker binding, - indicates no binding observed by the Ni^{2+} -co-immobilization assay). For the β fragments that bound α (++ or +), the ability to recruit β' into the α - β complex was investigated and the results are shown in the rightmost column. In the rectangular box at the bottom are shown the β regions required (darkest shading) and helpful (lighter shading) for binding α (top) and for recruiting β' into the α - β complex (bottom).

To address whether β conserved region H is required for α interaction, we investigated the binding of two additional β fragments, β_{EFGH} and β_{I} (Table 1 and Figure 4). Neither fragment bound α in the Ni^{2+} -co-immobilization assay (data not shown). Thus, we conclude that amino acid resi-

dues in both β conserved regions H and I_N are required for α binding.

Specificity of β fragment interactions with α

We used three additional tests to investigate the specificity and authenticity of the observed inter-

actions between the α subunit and β fragments; protein footprinting, binding to an α point mutant defective in β binding, and competition between full length β and its fragments.

Protein footprinting

Hydroxyl-radical protein footprinting was recently used to show that residues within α conserved regions A and B (specifically α residues 30 to 55 and 65 to 75, respectively) are involved in direct interactions with the β subunit in the $\alpha_2\beta$ sub-assembly, while residues within α conserved regions C and D (specifically α residues 175 to 185 and 195 to 210, respectively) are involved in direct interactions with the β' subunit (Heyduk *et al.*, 1996). We utilized the same footprinting method to ask if the interactions observed between α and β fragments containing conserved regions H and I_N were identical to the interactions between α and full length β .

For this work, we used an α derivative with an N-terminal heart muscle protein kinase recognition site (Li *et al.*, 1989). With the purified, ^{33}P -end-labelled α derivative (α^*), complexes were formed with excess β or β fragments and purified to homogeneity. For each complex, we then performed hydroxyl-radical-mediated cleavage and compared the cleavage pattern quantitatively (Heyduk *et al.*, 1996) with α_2^* alone (Figure 5). Separate, independent cleavage experiments for α_2^* alone were compared to determine the level of background noise (Figure 5a). In addition to the analysis of Heyduk *et al.* (1996), we tested the statistical significance of the differences between the (α_2^* - α_2^*) control experiments and the ($\alpha_2^*\beta$ - α_2^*), ($\alpha_2^*\beta_{\text{FGHI}}$ - α_2^*), and ($\alpha_2^*\beta_{\text{FGHI}_N}$ - α_2^*) experiments using a Student's *t*-test for small samples with unequal variances (Mendenhall & Sincich, 1988). Regions where statistically significant differences from the (α_2^* - α_2^*) control were observed at a confidence level of 0.99999% are shown as black bars just above the *x*-axis in each plot of Figure 5. There were no statistically significant differences between two independent data sets of (α_2^* - α_2^*).

Figure 5b compares the cleavage pattern of the $\alpha_2^*\beta$ complex to that of α_2^* alone. Negative values of the curve represent protection of α from cleavage by the binding of β . The characteristics of the difference plot resemble qualitatively the difference plot of Heyduk *et al.* (1996), demonstrating the reproducibility of the method. Figure 5c and d compares cleavage patterns of the $\alpha_2^*\beta_{\text{FGHI}}$ and $\alpha_2^*\beta_{\text{FGHI}_N}$ complexes, respectively, to that of α_2^* alone. All three plots (Figure 5 to c) show qualitatively and quantitatively (in terms of statistically significant differences) the same protection pattern. For all three experiments, statistically significant differences occurred within two broad peaks of protection centered around α amino acids 43 to 52 and 61 to 76, corresponding with the earlier work of Heyduk *et al.* (1996). In addition, weaker but statistically significant protection was also observed in

a region around residues 159 to 176. These results demonstrate that all of the contacts made by the β subunit on α are also made identically by β_{FGHI} and β_{FGHI_N} .

We attempted to determine footprints of β_{HI} and β_{HI_N} on α but were unable to obtain reproducible difference plots. We believe this is due to the relatively weak binding of these shorter β derivatives to α (Figure 3). After the formation of complexes in the presence of a large excess of $\beta_{\text{HI(N)}}$, several steps of purification were required to obtain homogeneous complexes free of excess $\beta_{\text{HI(N)}}$, $\alpha^*\beta_{\text{HI(N)}}$ aggregates, and free α_2^* , resulting in disassociation and very low yields of complex.

The effect of an α point mutant defective in β binding

Kimura & Ishihama (1995a,b) described a mutant α subunit with a single amino acid substitution, Arg45 to Ala (α_{R45A}), that was properly folded and dimerized but was defective in binding β to form the assembly intermediate $\alpha_2\beta$. If the interactions of the β fragments with α are authentic, the β fragment interactions with the α_{R45A} mutant should also be negatively affected. This was tested using Ni^{2+} -co-immobilization assays with His₆- α_{R45A} NTD and β or β fragments containing conserved regions H and I_N (Figure 6).

Full length β bound to His₆- α_{R45A} NTD in the Ni^{2+} -co-immobilization assay (Figure 6, lane 2), but the binding was weak compared to wild-type (wt) α NTD (Figure 1, lane 4). Estimating from the Coomassie stained bands, there was about a five-fold reduction in the amount of β retained (relative to α) when His₆- α_{R45A} NTD was used rather than wt His₆- α NTD.

The binding of the β fragments β_{EFGHI} , β_{FGHI} , and β_{FGHI_N} to the α_{R45A} mutant was undetectable by the Ni^{2+} -co-immobilization assay (Figure 6, lanes 4, 6, and 8). In contrast, the shorter β fragments β_{HI} (Figure 6, lane 10), $\beta_{[950-1342]}$ (Figure 6, lane 12), and β_{HI_N} (not shown) were co-immobilized similarly by His₆- α_{R45A} NTD and wt His₆- α NTD. We attempted to titrate down the molar ratio of β fragment: α in the renaturation mixture (normally 4:1) but this did not reveal any obvious difference in binding affinity between the smallest β fragments and the α_{R45A} or wt α .

Competition between full length β and its fragments for α binding

Only one β subunit interacts with an α dimer in the RNAP assembly. We tested the ability of full length β to bind competitively with its fragments to α . In all cases tested, β fragments containing conserved regions H and I_N could not bind α simultaneously with full length β . For example, equimolar amounts of His₆- α NTD and β_{HI} were renatured in the presence of increasing amounts of either full length β or β_{FGHI} , and binding to the His₆- α NTD was then tested by the Ni^{2+} -co-immobilization assay.

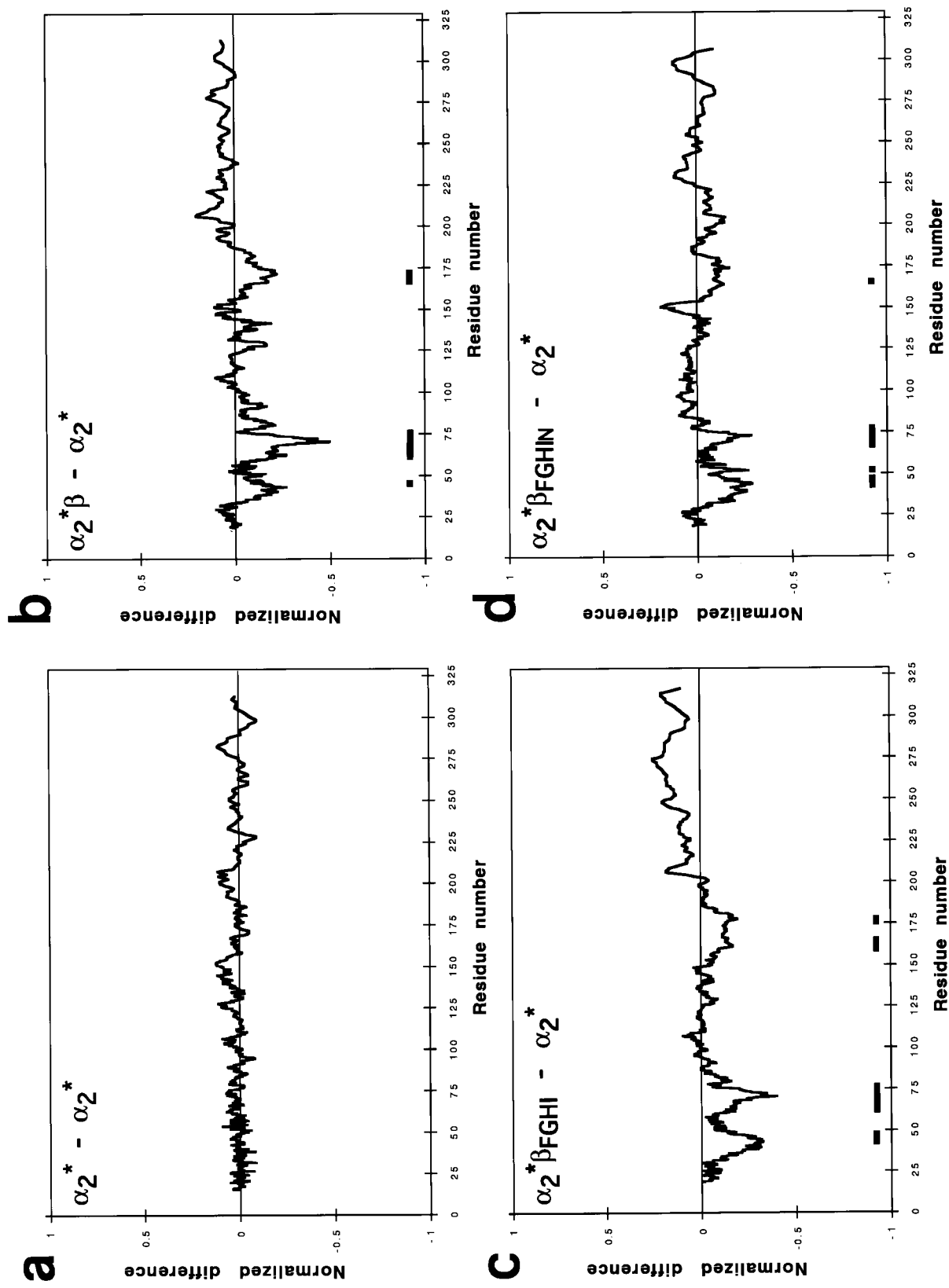


Figure 5. Difference plots showing hydroxyl-radical protein footprinting results. Normalized difference intensity, $(I - I_{\alpha_2^*})/I$, is plotted against α residue number, where I is the corrected intensity for the complex under study, and $I_{\alpha_2^*}$ is the corrected intensity for α_2^* . Statistically significant differences according to a Student's t -test (confidence level of 0.99999) are denoted by the thick black bars just above the x-axis. a, $\alpha_2^* - \alpha_2^*$ (averaged data for 17 lanes); b, $\alpha_2^* \beta_{FGHI} - \alpha_2^*$ (averaged data for 11 lanes); c, $\alpha_2^* \beta_{FGHI} - \alpha_2^*$ (averaged data for 9 lanes); d, $\alpha_2^* \beta_{FGHIN} - \alpha_2^*$ (averaged data for 17 lanes).

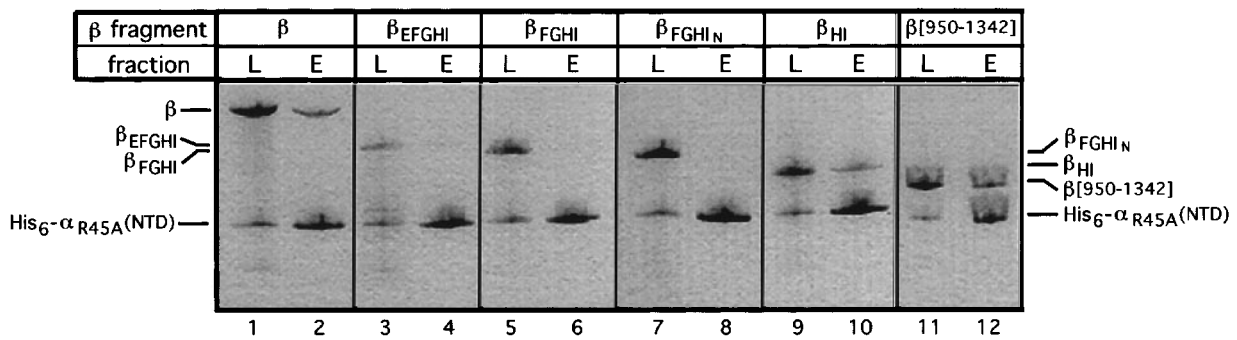


Figure 6. Effect of the α_{R45A} mutant on β fragment binding. Ni^{2+} -co-immobilization assays of binding of full length β or β fragments to His₆- α_{R45A} NTD. The load (L) and elution (E) fractions were analyzed by SDS-PAGE on 8 to 25% gradient PhastGels (Pharmacia).

bilization assay (Figure 7). Addition of an equimolar amount of full length β or β_{FGHI} reduced the binding of β_{HI} to His₆- α NTD (Figure 7, lanes 1 and 4), while a tenfold molar excess of β or β_{FGHI} was sufficient to eliminate detectable β_{HI} binding (Figure 7, lanes 2 and 5). Excess amounts of β_{HI} were also able to reduce the binding of β or β_{FGHI} or α (data not shown), but we were unable to add a molar excess of β_{HI} sufficient to completely compete β or β_{FGHI} binding due to the solubility limit of β_{HI} in the renaturation assay.

Recruitment of β' to complexes between β fragments and α

Mutational analysis and hydroxyl-radical protein footprinting have shown that α conserved regions C and D make direct interactions with the β' subunit that are critical for core RNAP assembly (Kimura *et al.*, 1994; Kimura & Ishihama, 1995a,b; Heyduk *et al.*, 1996). Nevertheless, in the absence of β , β' does not interact with α , as shown by the Ni^{2+} -co-immobilization assay (Figure 8, lane 4). Thus, both the α and β subunits make critical interactions with β' to form core RNAP (Figure 8, lane 8). Interestingly, we found that the complex of His₆- α NTD with β_{EFGHI} was sufficient to recruit β' (Figure 8, lane 12). We used the Ni^{2+} -co-immobilization assay with His₆- α NTD and the various β fragments to determine the conserved regions of β required for recruitment of β' into their complex with α . Since the formation of the α - β complex is obligatory, we only investigated the β fragments containing, at a minimum, conserved regions H and I_N since, as determined above, only these β fragments could interact with α . While all of the β' binding results are summarized schematically in Figure 4, the most revealing experiments are shown in Figure 8, lanes 13 to 20. When a mixture of His₆- α NTD, β' , and β_{FGHI_N} was renatured and subjected to the Ni^{2+} -co-immobilization assay (Figure 5, lanes 13 to 16), the complex between His₆- α NTD and β_{FGHI_N} formed, as expected, but this complex was unable to recruit β' (Figure 5, lane 16). However, renaturation of a mixture of His₆-NTD, β' , and β_{FGHI} resulted in the formation

of the ternary complex (Figure 5, lane 20), indicating that the C-terminal half of β conserved region I was required for recruitment of β' . We also tested the ability of β_{HI} and β_{HI_N} to recruit β' to the complexes with α (data not shown). Surprisingly, both β fragments were equally able to recruit β' into their complexes with α .

Discussion

The impetus for this study came from the initial observation using the Ni^{2+} -co-immobilization assay that the C-terminal half of β (β_{EFGHI}) alone interacts with α while the N-terminal half (β_{ABCD}) does not (Figure 1), suggesting that this assay could be conveniently used to parse the β subunit into functional domains that interact with α to initiate RNAP assembly. To further divide β into smaller fragments that were functionally relevant, we did not want to rely completely on the locations of the evolutionarily-conserved regions since structural domains may not necessarily correspond to the regions of conserved sequence (Severinova *et al.*, 1996). Therefore, we used limited

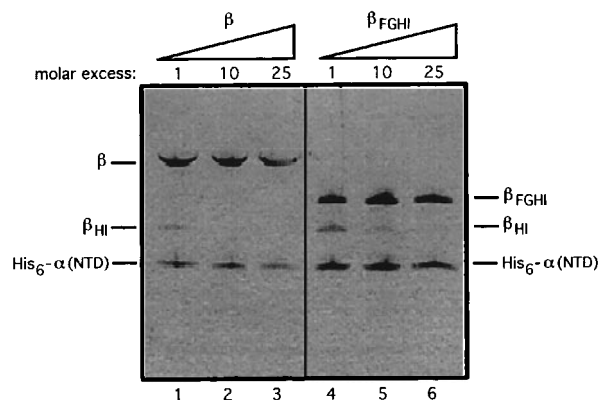


Figure 7. Competitive binding of β or β fragments to α NTD. Full length β (lanes 1 to 3) or β_{FGHI} (lanes 4 to 6) were renatured with His₆- α NTD and β_{HI} at the indicated molar ratios over β_{HI} . The proteins found in the elution fraction from the Ni^{2+} -co-immobilization assay were then analyzed by SDS-PAGE on an 8 to 25% PhastGel (Pharmacia).

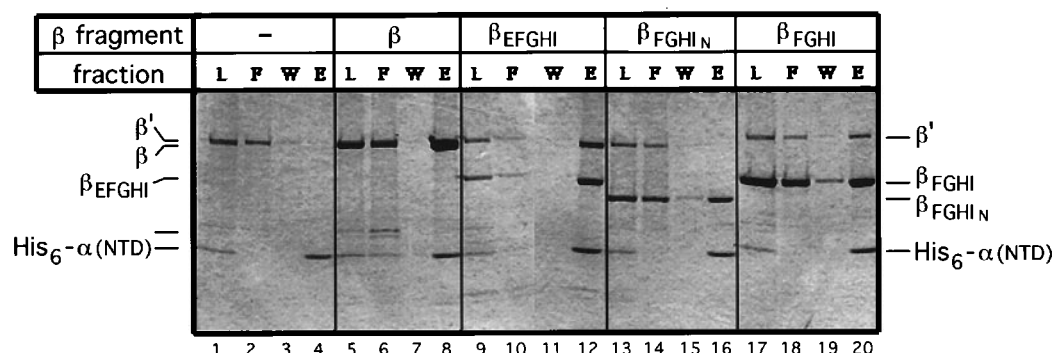


Figure 8. Binding of β' to β or β fragments complexed with α NTD. The indicated β fragments were renatured with β' and $\text{His}_6\text{-}\alpha$ NTD and complex formation was analyzed by Ni^{2+} -co-immobilization assays and SDS-PAGE on 8 to 25% gradient PhastGels (Pharmacia).

proteolysis to define domains of β_{EFGHI} complexed with α . Of most significance from the proteolysis studies was the high degree of clustering of all but one of the N and C-terminal cleavage sites (Figure 2d and Table 2) for three different proteases, indicating that the locations of these sites contain information about the domain architecture of β rather than the properties of the individual proteases. Two clusters of N-terminal cleavage sites were observed. The first N-terminal cluster contained four cleavage sites (681, 700, 708, and 711) from two different proteases (chymotrypsin and proteinase K), all within 30 residues of each other, just C-terminal of β conserved region E. The second N-terminal cluster contained three cleavage sites (903, 904, and 907) from two different proteases (chymotrypsin and trypsin), all within four residues of each other, just C-terminal of conserved region G. A cluster of C-terminal cleavage sites occurred with two different proteases (chymotrypsin and trypsin) within a region of 28 residues (1234-1261) in the middle of conserved region I, again indicating that structural domains as defined by limited proteolysis may not necessarily correspond to domains defined by sequence analysis.

β Conserved regions H and I_N are required for the interaction with α

Using domains of β defined by the proteolysis study as well as additional β fragments (Table 1), we used the Ni^{2+} -co-immobilization assay to investigate α binding. The results are summarized in Figure 4. The different β fragments grouped qualitatively into three classes; those that bound α approximately stoichiometrically (based on the Coomassie stained gels, denoted ++ in Figure 4), those that bound α but obviously less tightly (+), and those that did not detectably bind α (-). Only β fragments containing both conserved regions H and the N-terminal half of region I (I_N) interacted with α and the results in total allow us to conclude that β conserved regions H and I_N are required for the interaction of β with α . This is the main conclusion from this study. While conserved regions F and G are not required for α binding, they appear

to play some role in stabilizing the interaction since the only difference between the strong binding β fragments (++) and those that bound weakly (+) was the presence of regions F and G.

While we have shown that active RNAP can be reconstituted from peptides comprising the β subunit split into at least four separate fragments (Severinov *et al.*, 1995a, 1996), the β subunit split into two separate peptides at a site between conserved regions H and I will not support the assembly of functional enzyme (K. S., A. Vasenko, I. Bass, S. A. D., unpublished results). The finding that both regions H and I_N of β are required for α interaction explains the inability of the β polypeptide split between the two regions to assemble into functional enzyme since the assembly process would not be able to initiate.

Because almost all of the β fragments used in this study were relatively insoluble and tended to aggregate and precipitate on their own in solution, the α - β fragment complexes were renatured by dialysis from the denatured state, leading to the possibility of artifacts due to non-specific aggregation effects. Therefore, we performed a number of experiments, in addition to Ni^{2+} -co-immobilization, to test the specificity and authenticity of the observed interactions between β fragments containing conserved regions H and I_N and α . The most stringent test of specific and authentic binding was protein footprinting, which was used to show that the fragments β_{FGHI} and β_{FGHI_N} bound to α and protected the same α residues from hydroxyl-radical mediated cleavage as full length β (Figure 5). This unequivocally verifies that the negative results obtained in the Ni^{2+} -co-immobilization assay for some of the β fragments are not artifacts of the assay and demonstrate that these regions of β do not interact with α , at least not in a way that is detectable by two very different assays, the Ni^{2+} -co-immobilization or by hydroxyl-radical protein footprinting. In addition, these two β fragments (β_{FGHI} and β_{FGHI_N}); (1) did not bind an α mutant with a point substitution (R45A) known to be specifically compromised in its ability to bind β (Figure 6), and (2) bound α competitively with full

length β . Thus, we conclude that the fragments β_{FGHI} and β_{FGHIN} bound α in a highly specific manner and using the same interactions as full length β in forming the assembly intermediate $\alpha_2\beta$.

We were unable to observe a reproducible protein footprint on α from the binding of the shorter β fragments β_{HI} and β_{HIN} , probably due to their weaker binding affinity. Also, the shorter β fragments bound the α_{R45A} mutant similarly to wt α . It should be noted, however, that in the Ni^{2+} -co-immobilization assay, the α_{R45A} mutant is not completely defective in binding full length β (Figure 6), and under conditions that drive *in vitro* RNAP assembly (Tang *et al.*, 1995), the α_{R45A} mutant is equally effective as wt α in reconstituting active RNAP as measured by an affinity-labeling assay (Grachev *et al.*, 1987; Severinov *et al.*, 1995a; Y.W., K.S., A. Mustaev, & S.A.D., unpublished results), suggesting that the α_{R45A} mutant may not be an appropriate test of α - β interactions *in vitro*. These results are not inconsistent with the data presented by Kimura & Ishihama (1995a,b). In favor of the specificity of β_{HI} binding to α , β_{HI} bound α competitively with full length β , indicating that they interact with the same site on α . In addition, the β subunit, when split into two separate peptides between conserved regions G and H, remains competent to assemble active RNAP (Severinov *et al.*, 1995a), indicating that β fragments containing only regions H and I are competent for initiating RNAP assembly.

Functions for β conserved regions H and I_{N} are not known but of note are two sites, Lys1065 within region H, and His1237 within I_{N} , that cross-link to initiating nucleotide analogs (Mustaev *et al.*, 1991), placing these residues within about 3 Å of the initiating nucleotide α -phosphate. Furthermore, protein-DNA cross-linking has implicated a region of β between Met1230 and Met1273 within region I as participating in critical protein-DNA interactions near the transcript 3'-end (Nudler *et al.*, 1996). Thus, residues within β conserved regions H and I_{N} are positioned very close to, and appear to contribute essential components to the enzymes active center. Since there is no evidence that αNTD is located near the DNA template or RNA transcript nor that it plays a role in the RNAP catalytic function, it is surprising to find that RNAP assembly involves interactions between αNTD and β conserved regions H and I_{N} .

β Conserved region I_{C} appears to be involved in recruiting β' into the α - β complex

An additional observation that β_{FGHI} was sufficient to recruit the β' subunit into its complex with α (Figure 8) led us to take the Ni^{2+} -co-immobilization assay one step further to investigate which β fragments, once bound to α , were able to recruit β' . The results with some of the fragments clearly indicated that, in addition to the β regions previously identified as being required for α binding, the C-terminal half of β conserved region I (I_{C}) was re-

quired to bind β' . So, for instance, β_{FGHI} and β_{FGHIN} both bound α quite strongly, but only β_{FGHI} was able to recruit β' into the complex (Figure 8). However, β_{HI} and β_{HIN} both bound α and both appeared able to recruit β' into the complex, clouding this interpretation. Indirect support for the conclusion that a direct interaction occurs between β conserved region I_{C} and β' comes from genetic suppressor studies. Experiments designed to find interacting regions within the β subunit failed to find intragenic suppressor mutations of three different alleles within conserved region I_{C} (substitutions at β positions 1249, 1266, and 1272; Tavormina *et al.*, 1996). In *S. cerevisiae* RNA polymerase II, a conditional mutation resulting from substitution of a highly conserved Gly residue (corresponding to position 1282 of *E. coli* β) within conserved region I_{C} of the β homolog was suppressed by mutations in the β' homolog (Martin *et al.*, 1990; Scafe *et al.*, 1990). These results together indirectly support the possibility that region I_{C} of β may interact with regions of β' .

An emerging view of RNA polymerase assembly

The results of this study, combined with results from many other studies (Igarashi *et al.*, 1990; Hayward *et al.*, 1991; Igarashi *et al.*, 1991; Igarashi & Ishihama, 1991; Blatter *et al.*, 1994; Kimura *et al.*, 1994; Kimura & Ishihama, 1995a,b; Severinov *et al.*, 1995a,b; Heyduk *et al.*, 1996) provide support for the following sequence of events along the RNAP assembly pathway (Zillig *et al.*, 1976; Ishihama, 1981). (1) Assembly is initiated by the dimerization of the α subunits, which is mediated by the αNTD . (2) The assembly intermediate $\alpha_2\beta$ is formed by interactions between α conserved regions A and B and β conserved regions H and I_{N} . The complex is stabilized by regions F and G of β , probably through interactions with regions H and I. (3) β' is recruited into the $\alpha_2\beta$ complex through interactions with conserved regions in both α (regions C and D) and β (including region I_{C}). (4) The condensation of further interactions between conserved regions in β and β' (or between intrasubunit regions), which are not currently known, results in the formation of core RNAP. (5) RNAP holoenzyme subsequently forms by the addition of the σ subunit through unknown interactions but which do not involve the α subunit.

The β subunit homolog in archaeobacteria is encoded by two separate genes (Berghofer *et al.*, 1988; Pühler *et al.*, 1989). Conserved regions A to I are arranged co-linearly with respect to *E. coli* β except the split site (corresponding to approximately codon 650 of *E. coli* *rpoB*) separates conserved regions A to D (β_{N}) from E to I (β_{C}) on two separate peptides. Our results also suggest that in archaeobacteria, assembly of RNAP involves the initial interaction of β_{C} with α_2 prior to the recruitment of β' . Subsequently, β_{N} would enter the complex to form the core RNAP.

Further studies investigating the interactions of other large subunit domains (Severinov *et al.*, 1995a,b, 1996) with each other and with sub-assemblies of other RNAP subunit domains will help elucidate additional subunit-subunit interactions that occur in the RNAP assembly.

Materials and Methods

Proteins

α Derivatives

Plasmids expressing His₆- α , α NTD, His₆- α NTD (Tang *et al.*, 1995), and His₆- α NTD with an N-terminal calf heart protein kinase site (Heyduk *et al.*, 1996) were a generous gift from R. Ebricht (Waksman Institute, Rutgers University). The proteins were expressed and purified as described (Tang *et al.*, 1995).

β Derivatives

Full length β and β' were obtained from expression plasmids pMKSe2 (Severinov *et al.*, 1993) and T7 β' (Zalenskaya *et al.*, 1990), respectively. Construction of expression plasmids for β_{ABCDEFG} and $\beta_{\text{[950-1342]}}$ was described (Severinov *et al.*, 1995a), as were expression plasmids for β_{ABCD} and β_{EFGHI} (Severinov *et al.*, 1996). $\beta_{\text{EFGHI}}[\Omega\text{T988}]$ was cloned from *rpoB* containing the ΩT988 insertion (Kashlev *et al.*, 1989; Borukhov *et al.*, 1991) into the plasmid expressing β_{EFGHI} . The β fragments β_{FGHI} , β_{FGHI_N} , β_{FGH} , β_{HL} , β_{HL_N} and β_{EFGH} were obtained from expression plasmids constructed by PCR cloning from pMKSe2 into *NcoI* and *BamHI* sites of the T7-based pET15b vector (Novagen). Use of the 3'-*NcoI* site results in loss of the vector-based N-terminal His₆-tag; the resulting protein products contain only native β sequences. β_1 was cloned from pMKSe2. $\beta_{\text{FGHI}_N}(\Delta[910-982])$ and $\beta_{\text{FGHI}_N}(\Delta[946-1062])$ were cloned from *rpoB* mutants harboring the deletions (constructed as described by Borukhov *et al.*, 1991) into the plasmid expressing β_{EFGHI_N} .

For all of the β fragments, overexpression was induced with 1 mM isopropyl β -D-thiogalactoside at 37°C for three hours. The proteins in inclusion bodies were then prepared according to Tang *et al.* (1995).

Reconstitution of α - β complexes

Purified α or α NTD derivatives were mixed with β or its fragments (from washed inclusion bodies) at a molar ratio of α : β of 1:4 and a total protein concentration of not more than 0.5 mg/ml in reconstitution buffer (40 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM EDTA, 20 mM MgCl₂, 10 μ M ZnCl₂, 20% (v:v) glycerol, 7 mM β -mercaptoethanol, 6 M guanidine-HCl) and dialyzed overnight at 4°C against 2 \times 2 l of the same buffer without guanidine-HCl.

Limited proteolytic digestion

Complexes of His₆- α NTD with β fragments (about 10 μ g total protein) were immobilized on Ni²⁺-NTA-agarose beads (Qiagen) by incubation in 50 μ l of buffer A (20 mM Tris-HCl, pH 7.9, 100 mM NaCl, 1 mM EDTA, 1 mM β -mercaptoethanol, 5% (v:v) glycerol) + 2.5 mM imidazole and incubated for 30 minutes at 4°C with gentle mixing. Protease was then added and digestion

reactions were carried out at 25°C with gentle mixing. After 30 minutes, phenylmethanesulfonyl fluoride was added to a final concentration of 1 mM, the beads were pelleted by centrifugation, washed four times with 500 μ l of buffer A, and eluted with 10 μ l of buffer A + 100 mM imidazole. The eluted products were then analyzed by SDS-PAGE using 4 to 20% gradient gels (Novex).

Matrix-assisted laser desorption mass spectrometry

Matrix-assisted laser desorption mass spectra (Hillenkamp *et al.*, 1991) of the proteolytic fragments of β were collected using a time-of-flight mass spectrometer constructed at The Rockefeller University (Beavis & Chait, 1989). The fragments were mixed with α -cyano-4-hydroxycinnamic acid (10 g/l in formic acid/water/isopropanol (1:3:2, by vol.)) to obtain a final protein concentration of about 2 pmol/ μ l. An aliquot (0.5 μ l) was placed on the mass spectrometer probe tip and air dried. The sample was irradiated with 10 ns duration laser pulses (355 nm wavelength) from a Nd(YAG) laser. The resulting ions were accelerated in an electrostatic field and their time-of-flight was measured with a LeCroy 9350 M oscilloscope. The observed masses are listed in Table 2.

Ni²⁺-NTA agarose co-immobilization binding assays

Protein complexes (about 10 μ g) were mixed with pre-equilibrated Ni²⁺-NTA-agarose beads (Qiagen) in 50 μ l buffer A and incubated for 30 minutes at 4°C with gentle mixing. The beads were pelleted by centrifugation and washed twice with 500 μ l of buffer A, 2.5 mM imidazole, then twice with 500 μ l of buffer A, 25 mM imidazole. The protein samples were then eluted from the beads with buffer A, 100 mM imidazole and analyzed by SDS-PAGE using 8 to 25% gradient PhastGels (Pharmacia).

Hydroxyl-radical protein footprinting

Purified α derivative was ³³P-end-labelled in a 230 μ l reaction containing 90 μ M α derivative, 90 units bovine heart protein kinase (Sigma), 0.1 μ M [γ -³³P]ATP (New England Nuclear, 7-110 TBq/mmol), 10 mM Tris-HCl (pH 7.9), 50 mM NaCl, 0.5 mM EDTA, 25% (v/v) glycerol, and 30 mM DTT for one hour at 37°C. Complexes ($\alpha_2^*\beta$, $\alpha_2^*\beta_{\text{FGHI}}$, and $\alpha_2^*\beta_{\text{FGHI}_N}$) were prepared essentially as described (Tang *et al.*, 1995). Briefly, labelled α derivative and excess β (or β fragment) was mixed under denaturing conditions. Following dialysis into non-denaturing conditions, complexes were purified by batch-mode metal ion-affinity chromatography, then gel-filtration chromatography on a Superose-6 FPLC column (Pharmacia) as in Borukhov & Goldfarb (1993) except that DTT and glycerol were omitted from the running buffer. Footprinting reactions were performed at room temperature in a buffer containing 10 mM Mops-NaOH (pH 7.2), 100 mM KCl, 10 mM MgCl₂, 10 μ M α_2^* or complex, 1 mM (NH₄)Fe(SO₄)₂, 2 mM EDTA, 20 mM ascorbate and 1 mM H₂O₂ for 30 minutes. Reactions were initiated and terminated as described (Heyduk *et al.*, 1996) and the products were analyzed by tricine SDS/PAGE (Schagger & von Jagow, 1987; Heyduk & Heyduk, 1994) followed by phosphorimager analysis (Molecular Dynamics Storm). The gels were quantified and the difference plots were generated essentially as described (Heyduk *et al.*, 1996), as was the calibration of amino acid positions

using standards generated by CNBr, endoproteinase Lys-C, or endoproteinase Glu-C cleavage.

Acknowledgments

We thank R. Ebricht (Waksman Institute, Rutgers University) for plasmids expressing α NTD derivatives and for helpful discussions, G. Meredith (Stanford University) for assistance with the *t*-test analysis of the protein footprinting results, and E. Campbell for helpful discussions. K.S. is a recipient of The Burroughs Wellcome Career Award In the Biomedical Sciences. N.L. was supported in part by funds granted by the Norman and Rosita Winston Foundation. S.A.D. is a Lucille P. Markey Scholar and a Pew Scholar in the Biomedical Sciences. This work was supported in part by grants from the Lucille P. Markey Charitable Trust, the Irma T. Hirschl Trust, the Pew Foundation, and NIH GM53759 to S.A.D., NIH GM50514 to T.H., and NIH RR00762 to B.T.C.

References

- Azuma, Y., Yamagishi, M. & Ishihama, A. (1993). Subunits of the *Schizosaccharomyces pombe* RNA polymerase II: enzyme purification and structure of the subunit 3 gene. *Nucl. Acids Res.* **21**, 3749–3754.
- Beavis, R. C. & Chait, B. T. (1989). Factors affecting the ultraviolet laser desorption of proteins. *Rapid Commun. Mass Spectrom.* **3**, 233–237.
- Berghofer, B., Krockel, L., Kortner, C., Truss, M., Schallenberg, J. & Klein, A. (1988). Relatedness of archaeobacterial RNA polymerase core subunits to their eubacterial and eukaryotic equivalents. *Nucl. Acids Res.* **16**, 813–8128.
- Blatter, E., Ross, W., Tang, H., Gourse, R. & Ebricht, R. (1994). Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell*, **78**, 889–896.
- Borukhov, S. & Goldfarb, A. (1993). Recombinant *Escherichia coli* RNA polymerase: purification of individually overexpressed subunits and *in vitro* assembly. *Protein Expt. Purif.* **4**, 503–511.
- Borukhov, S., Severinov, K., Kashlev, M., Lebedev, A., Bass, I., Rowland, G. C., Lim, P.-P., Glass, R. E., Nikiforov, V. & Goldfarb, A. (1991). Mapping of trypsin cleavage and antibody-binding sites and delineation of a dispensable domain in the β subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **266**, 23921–23926.
- Ebricht, R. H. & Busby, S. (1995). *Escherichia coli* RNA polymerase α subunit: Structure and function. *Curr. Opin. Genet. Dev.* **5**, 197–203.
- Falkenburg, D., Dworniczak, B., Faust, D. M. & Bautz, E. K. F. (1987). RNA polymerase II of *Drosophila*. Relation of its 140,000 M_r subunit to the β subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **195**, 929–937.
- Grachev, M. A., Kolocheva, T. I., Lukhtanov, E. A. & Mustaev, A. (1987). Studies on the functional topography of *Escherichia coli* RNA polymerase: highly selective affinity labelling by analogues of initiating substrates. *Eur. J. Biochem.* **163**, 113–121.
- Hayward, R. S., Igarashi, K. & Ishihama, A. (1991). Functional specialization within the alpha-subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **221**, 23–29.
- Heyduk, E. & Heyduk, T. (1994). Mapping protein domains involved in macromolecular interactions: A novel protein footprinting approach. *Biochemistry*, **33**, 9643–9650.
- Heyduk, T., Heyduk, E., Severinov, K., Tang, H. & Ebricht, R. H. (1996). Rapid epitope mapping by hydroxyl-radical protein footprinting: Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits. *Proc. Natl Acad. Sci. USA*, **93**, 10162–10166.
- Hillenkamp, F., Karas, M., Beavis, R. C. & Chait, B. T. (1991). Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* **63**, 1193A–1203A.
- Hudson, G. S., Holton, T. A., Whitefield, R. P. & Bottomley, W. (1988). Spinach chloroplast *rpoBC* genes encode three subunits of the chloroplast RNA polymerase. *J. Mol. Biol.* **200**, 639–654.
- Igarashi, K. & Ishihama, A. (1991). Bipartite functional map of the *E. coli* RNA polymerase alpha subunit: Involvement of the C-terminal region in transcription activation by cAMP-CRP. *Cell*, **65**, 1015–1022.
- Igarashi, K., Fujita, N. & Ishihama, A. (1990). Sequence analysis of two temperature-sensitive mutations in the alpha subunit gene (*rpoA*) of *Escherichia coli* RNA polymerase. *Nucl. Acids Res.* **18**, 5945–5948.
- Igarashi, K., Fujita, N. & Ishihama, A. (1991). Identification of a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **218**, 1–6.
- Ishihama, A. (1981). Subunit assembly of *Escherichia coli* RNA polymerase. *Advan. Biophys.* **14**, 1–35.
- Kashlev, M. V., Bass, I. A., Lebedev, A. N., Kalyaeva, E. S. & Nikiforov, V. G. (1989). Deletсионно-интерсионное картирование области, не существенной для функционирования β -суб'единицы RNK-полимеразы *Escherichia coli*. *Genetika*, **25**, 396–405.
- Kimura, M. & Ishihama, A. (1995a). Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: amino acid substitution within the amino-terminal assembly domain. *J. Mol. Biol.* **254**, 342–349.
- Kimura, M. & Ishihama, A. (1995b). Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: insertion analysis of the amino-terminal assembly domain. *J. Mol. Biol.* **248**, 756–767.
- Kimura, M., Fujita, N. & Ishihama, A. (1994). Functional map of the alpha subunit of *Escherichia coli* RNA polymerase. Deletion analysis of the amino-terminal assembly domain. *J. Mol. Biol.* **242**, 107–115.
- Kolodziej, P. A. & Young, R. A. (1991). Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. *Mol. Cell. Biol.* **11**, 4669–4678.
- Lalo, D., Carles, C., Sentenac, A. & Thuriaux, P. (1993). Interactions between three common subunits of yeast RNA polymerases I and III. *Proc. Natl Acad. Sci. USA*, **90**, 5524–5528.
- Li, B.-L., Langer, J., Schwartz, B. & Pestka, S. (1989). Creation of phosphorylation sites in proteins: construction of a phosphorylatable human interferon alpha. *Proc. Natl Acad. Sci. USA*, **86**, 558–562.
- Martin, C., Okamura, S. & Young, R. (1990). Genetic exploration of interactive domains in RNA polymerase II subunits. *Mol. Cell. Biol.* **10**, 1908–1914.
- Mendenhall, W. & Sincich, T. (1988). *Statistics for the Engineering and Computer Sciences. 2nd edit.*, pp. 378, Dellen/MacMillan, San Francisco.

- Mustaev, A., Kashlev, M., Lee, J., Polyakov, A., Lebedev, A., Zalenskaya, K., Grachev, M., Goldfarb, A. & Nikiforov, V. (1991). Mapping of the priming substrate contacts in the active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **266**, 23927–23931.
- Nudler, E., Avetisova, E., Markovtsov, V. & Goldfarb, A. (1996). Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science*, **273**, 211–217.
- Pati, U. K. (1994). Human RNA polymerase II subunit hRPB14 is homologous to yeast RNA polymerase I, II, and III subunits (AC19 and RPB11) and is similar to a portion of the bacterial RNA polymerase alpha subunit. *Gene*, **145**, 289–292.
- Pühler, G., Leffers, H., Groppe, F., Palm, P., Klenk, H.-P., Lottspeich, F., Garrett, R. A. & Zillig, W. (1989). Archaeobacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proc. Natl Acad. Sci. USA*, **86**, 4569–4573.
- Scafe, C., Nonet, M. & Young, R. A. (1990). RNA polymerase II mutants defective in transcription of a subset of genes. *Mol. Cell. Biol.* **10**, 1010–1016.
- Schagger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379.
- Severinov, K., Soushko, M., Goldfarb, A. & Nikiforov, A. (1993). Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **268**, 14820–14825.
- Severinov, K., Mustaev, A., Severinova, E., Bass, I., Landick, R., Nikiforov, V., Goldfarb, A. & Darst, S. A. (1995a). Assembly of functional *Escherichia coli* RNA polymerase using beta subunit fragments. *Proc. Natl Acad. Sci. USA*, **92**, 4591–4595.
- Severinov, K., Mustaev, A., Severinova, E., Kozlov, M., Darst, S. A. & Goldfarb, A. (1995b). The β subunit Rif-cluster I is only Angstroms away from the active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **270**, 29428–29432.
- Severinov, K., Mustaev, A., Kukarin, A., Muzzin, O., Bass, A., Darst, S. A. & Goldfarb, A. (1996). Structural modules of the large subunits of RNA polymerase: introducing archaeobacterial and chloroplast split sites in the β and β' subunits of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **271**, 27969–27974.
- Severinova, E., Severinov, K., Fenyö, D., Marr, M., Brody, E. N., Roberts, J. W., Chait, B. T. & Darst, S. A. (1996). Domain organization of the *Escherichia coli* RNA polymerase σ^{70} subunit. *J. Mol. Biol.* **263**, 637–647.
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
- Sweetser, D., Nonet, M. & Young, R. A. (1987). Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc. Natl Acad. Sci. USA*, **84**, 1192–1196.
- Tang, H., Severinov, K., Goldfarb, A. & Ebright, R. H. (1995). Rapid RNA polymerase genetics: one-day, no-column preparation of reconstituted recombinant *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 4902–4906.
- Tavormina, P. L., Reznikoff, W. S. & Gross, C. A. (1996). Identifying interacting regions the β subunit of *Escherichia coli* polymerase. *J. Mol. Biol.* **258**, 213–223.
- Ulmasov, T., Larkin, R. M. & Guilfoyle, T. J. (1996). Association between 36- and 13.6-kDa α -like subunits of *Arabidopsis thaliana* RNA polymerase II. *J. Biol. Chem.* **271**, 5085–5094.
- Wilson, J. (1991). The use of monoclonal antibodies and limited proteolysis in elucidation of structure-function relationships in proteins. *Methods Biochem. Anal.* **35**, 207–250.
- Zalenskaya, K., Lee, J., Gujuluva, C. N., Shin, Y. K., Slutsky, M. & Goldfarb, A. (1990). Recombinant RNA polymerase: inducible overexpression, purification and assembly of *Escherichia coli* rpo gene products. *Gene*, **89**, 7–12.
- Zillig, W., Palm, P. & Heil, A. (1976). Function and reassembly of subunits of DNA-dependent RNA polymerase. In *RNA Polymerase* (Losick, R. & Chamberlin, M., eds), pp. 101–125, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Edited by R. Enright

(Received 25 February 1997; received in revised form 25 April 1997; accepted 25 April 1997)