Structure-Function Relationship of the Alpha Subunit of *Escherichia coli* RNA Polymerase

Tomofumi Negishi

Doctor of Philosophy

Department of Genetics, School of Life Science
The Graduate University for Advanced Studies

1996

Contents

Abbreviations

aa amino acid(s)

bp base pair(s)

C- carboxy-

cAMP cyclic adenosine monophosphate

CAPS cyclohexylamino-1-propanesulphonic acid

CBB Coomassie brilliant blue

CRP cAMP receptor protein

DEAE diethylaminoethyl

DFP diisopropylfluoro phosphate

DTT dithiothreitol

EDTA diethylaminoethyl

HPLC high performance liquid chromatography

IPTG isopropyl-β-D-thiogalactopyranoshide

LB Luria Broth

N- amino-

PAGE polyacrylamide gel electrophoreisis

PCR polymerase chain reaction

PMSF phenylmethanesulfonyl fluoride

PITC phenylisothiocyanate

PTC phenylthiocarbamyl

PTH phenylthiohydantoin

SDS sodium dodecyl sulfate

TFA trifluoroacetic acid

TFMSA trifluoromethanesulphonic acid

TGE Tris-Glycerol-EDTA

TGED Tris-Glycerol-EDTA-DTT

1. Summary

The α subunit of *Escherichia coli* RNA polymerase consists of 329 amino acids and plays an essential role in protein-protein contacts not only for RNA polymerase assembly but also for transcription regulation by class-I factors and DNA enhancer elements. To reveal the structure-function relationship of α subunit, I studied the organization of structural domains by analysis of the pattern of limited proteolysis with two endoproteases, V8 protease and trypsin. Results indicated that one region, Arg235–Glu244, was highly accessible to endoproteases. I propose that the α subunit consists of two major structural domains, an amino (N) terminal-proximal domain upstream from Arg235 and a carboxy (C) terminal-proximal domain downstream from Glu245, both being connected by an inter-domain linker formed by the spacer between these two amino acid residues. The structural organization obtained was in good agreement with its functional map. The N-terminal domain corresponds to the assembly domain of RNA polymerase while the C-terminal domain corresponds to the transcription regulatory domain including the contact sites with class-I transcription factors and DNA UP elements. Upon prolonged proteolysis, the N-terminal domain was cleaved into two subdomain fragments (Na and Nb) between Arg45 and Glu68. On the other hand, the C-terminal proximal domain of 85 amino acids in length (amino acid (aa) residues 245-329) was highly resistant to endoproteases. This domain was assumed to form a compact structure.

Next, I determined the conformation of this C-terminal domain (αCTD) by NMR analysis in collabolation with Prof. Y. Kyogoku and colleagues (Institute for Protein Research, Osaka University). NMR analysis revealed that αCTD forms a compact and rigid structure consisting of four helices and two long loops at both termini. The helix1

(Val264 to Leu273) is perpendicular to the largest helix 4 (Lys297 to Ser309) and the N-termini of helix 1 and helix 4 are very close to each other on the tertiary structure. The location of the contact sites for class-I transcription factors, mapped based on mutant studies, is discussed in relation to the tertiary structure of αCTD.

To identify the residue(s) involved in interaction with UP (DNA enhancer) elements, chemical shift perturbation experiments of NMR analysis were performed. Results indicated that the helix 1 and the N-terminal region of helix 4 interact with the *rrnB*P1 promoter UP element. It seems that the configuration formed between helix 1 and helix 4 is appearently similar to helix-turn-helix motif which is present in DNA-binding proteins.

Finally, I tried to confirm the contact site of CRP on helix 1 using a newly developed transcription inhibition assay by α-related peptides. For this purpose, I prepared a set of synthetic peptides containing the CRP contact site sequence of α subunit. This transcription inhibition experiment showed that a peptide including the whole sequence of helix 1 significantly inhibited CRP-RNA polymerase interaction, suggesting that the intact conformation of helix 1 is necessary for effective interaction with CRP. The inhibition test of transcription by synthetic peptides can be used for mapping contact sites with other class-I factors.

2. Introduction

Escherichia coli cells have evolved global regulation mechanisms of transcription for growth and survival under various conditions. The RNA polymerase plays a essential role in transcription of all species of cellular RNA. Control of the activity and specificity of RNA polymerase is now recognized as a major mechanism of the global regulation of transcription.

(Val264 to Leu273) is perpendicular to the largest helix 4 (Lys297 to Ser309) and the N-termini of helix 1 and helix 4 are very close to each other on the tertiary structure. The location of the contact sites for class-I transcription factors, mapped based on mutant studies, is discussed in relation to the tertiary structure of αCTD.

To identify the residue(s) involved in interaction with UP (DNA enhancer) elements, chemical shift perturbation experiments of NMR analysis were performed. Results indicated that the helix 1 and the N-terminal region of helix 4 interact with the *rrnB*P1 promoter UP element. It seems that the configuration formed between helix 1 and helix 4 is appearently similar to helix-turn-helix motif which is present in DNA-binding proteins.

Finally, I tried to confirm the contact site of CRP on helix 1 using a newly developed transcription inhibition assay by α-related peptides. For this purpose, I prepared a set of synthetic peptides containing the CRP contact site sequence of α subunit. This transcription inhibition experiment showed that a peptide including the whole sequence of helix 1 significantly inhibited CRP-RNA polymerase interaction, suggesting that the intact conformation of helix 1 is necessary for effective interaction with CRP. The inhibition test of transcription by synthetic peptides can be used for mapping contact sites with other class-I factors.

2. Introduction

Escherichia coli cells have evolved global regulation mechanisms of transcription for growth and survival under various conditions. The RNA polymerase plays a essential role in transcription of all species of cellular RNA. Control of the activity and specificity of RNA polymerase is now recognized as a major mechanism of the global regulation of transcription.

The RNA polymerase of *Escherichia coli* consists of four core subunits (2α, β and β′) and one of the multiple molecular species of σ subunit (Fig. 1). On the basis of the known molecular weights and relative contents of these subunits, the RNA polymerase is known to exist in two forms, core enzyme ($\alpha_2\beta\beta'$) and holoenzyme ($\alpha_2\beta\beta'\sigma$) (Burgess, 1969). The genes encoding for these subunits have been located on the *E. coli* chromosome (Hayward and Scaife, 1976). The core subunit genes are organized into operons together with the genes for ribosomal proteins. The genes, *rpoB* (β) and *rpoC* (β′), are included in a single large operon with the genes encoding ribosomal proteins L11 (*rplK*), L1 (*rplA*), L10 (*rplJ*) and L12 (*rplL*) (Bendiak *et al.*, 1977; Lindahl *et al.*, 1977b; Yamamoto and Nomura, 1978; An and Friesen, 1980). The gene *rpoA* (α) forms an operon with the genes for ribosomal protein S13 (*rpsM*), S11 (*rpsK*), S4 (*rpsD*) and L17 (*rplQ*) (Jaskunas *et al.*, 1975; 1977; Lindahl *et al.*, 1977a). The *rpoD* gene for σ^{70} subunit froms an operon with the genes for ribosomal protein S21 (*rpsU*) and the primase (*dnaG*) (Harris *et al.*, 1977; 1978; Nakamura, 1978 Travers *et al.*, 1978; Gross *et al.*, 1978). From the sequence of *rpoA* (α), *rpoB* (β), *rpoC* (β′) and *rpoD* (σ^{70}), the molecular weights of the α, β, β′ and σ has been calculated to be 36,512 (Ovchinnikov *et al.*, 1977), 150,618 (Ovchinnikov *et al.*, 1981), 155,163 (Ovchinnikov *et al.*, 1982), and 70,263 (Burton *et al.*, 1981), respectively. The assembly of RNA polymerase proceeds sequentially under the order: $2\alpha \rightarrow \alpha_2 \rightarrow \alpha_2\beta \rightarrow \alpha_2\beta\beta'$ (core enzyme) $\rightarrow E\sigma$ (holoenzyme)(Saitoh and Ishihama, 1976; reviewed in Ishihama, 1981).

The rate of synthesis of core enzyme subunits in exponentially growing cells increases together with those of ribosomal proteins (Iwakura *et al.*, 1974). The total number of core enzyme is maintained in the range of 1,000 to 3,000 moleculers per cell at exponential growth phase (Ishihama, 1991). The total number of genes on the *E. coli* chromosome is, however,

more than 4,000 as estimated from its DNA sequence. The limited number of core RNA polymerase is interconvertible between different holoenzyme formed by replacement of σ subunit. The promoter selectivity of holoenzyme is further modulated by interaction with one of many transcription factors (reviewed in Ishihama, 1988).

Our knowledge of the functional role(s) of each subunit was established by the genetic and biochemical studies. Functional analyses of β subunit indicated that: i) affinity labeling with substrate analogues demonstrated that the substrate binding site is located in the C-terminal region of the β subunit (Frishauf *et al.*, 1973; Armstrong *et al.*, 1976; Grachv *et al.*, 1987; 1989; Mustaev *et al.*, 1991); ii) mutations conferring resistance to rifampicin are mapped in three clusters in the central region of β subunit (Jin and Gross, 1988). A rifampicin resistant mutant (Arg529Cys) (Arg was substituted for Cys at position 529), showing reduced affinity for UTP and decreased promoter clearance, implied close proximity of this position to the catalytic center of RNA polymerase (Jin *et al.*, 1994). Similarly, another rifampicin resistant mutant carrying Gln513 to Pro substitution showed decreased affinity for purine nucleotides during elongation (Jin and Gross, 1991). The close relation between substrate binding site and rifampicin binding site was also suggested by affinity labeling (Mustaev *et al.*, 1994; Severinov *et al.*, 1995) and resonance energy transfer experiments (Kumar and Chatterji, 1990); iii) RNA polymerase containing β mutant with four Ala substitutions between codons 1065-1073 showed defect in promoter clearance and decreased elongation rate, suggesting that this region is related to translocation of RNA products and is close to the catalytic center of RNA polymerase (Sagitov *et al.*, 1993); iv) a small internal deletion mutant of β subunit suggested that a region between codons 965 and 1083 is involved in promoter selectivity (Glass *et al.*, 1986b); v) RNA polymerase mutants

carrying aa substitution at approximate positions of aa residue 736 and 906 in β subunit show relaxed phenotypes insensitive to stringent control (Glass *et al.*, 1986a). Moreover, ppGpp analogue was cross-linked to β subunit (Owens, *et al.*, 1987; Dipankar *et al.*, submitted for publication); vi) deletion mutants lacking longer than 20% from C-terminus of β subunit can not form complexes with σ subunit (Glass *et al.*, 1986c), and mutants lacking 40% from the C-terminus of β subunit show defect in the core enzyme assembly (Glass *et al.*, 1988).

Functional analyses of β′ indicated that: i) comparison of aa sequence of largest subunit (RNA polymerase II) of yeast (Allison et al., 1985; Memet et al., 1988), *Drosophila* (Jokerst *et al.*, 1989), mouse (Ahearn *et al.*,) with *E. coli* (Squires *et al.*, 1981;Ovchinnikov *et al.*, 1982), indicated that there are eight highly conserved segments (A-H) in β′ subunit. ii) photochemical cross-link experiment of the complex formed by *E. coli* RNA polymerase and T7 DNA, demonstrated that β′ subunit was cross-linked to T7 DNA. β′ subunit can bind to DNA. (Wu and Hillel, 1978; Harrison *et al.*, 1982). iii) mutations (Met747Ile, Arg780His and Ser793Phe) conferring resistance to streptolydigin, a potent inhibitor of elongation, were located in one of the highly conserved regions (Segment F) of β′ subunit and these mutants showed a change in elongation rate (Severinov *et al.*, 1995); iv) the *rpoC* mutation (Tyr75Asp) reduced readthrough at a rho-independent transcription terminator placed downstream of the phage HK022 P_L promoter (Clerget *et al.*, 1995). This mutation is located in the putative Zinc-binding motif [one kind of the DNA-binding motifs, which include in addition, helix-loop-helix (HLH) motif (observed in MyoD, n-Myc) (Murre *et al.*, 1989), lucine zipper motif (GCN4, Jun, Fos, Myc) (Landschulz *et al.*, 1988; Gentz *et al.*, 1989), and helix-turn-helix (HTH) motif (λ Cro repressor, CRP, λ repressor) (Anderson *et al.*, 1981; Ohlendorf *et al.*, 1982; Matsuo *et al.*, 1995; Mckay

et al., 1981; Pabo *et al.*, 1982)] of β′ subunit, suggesting that this domain is related to termination (Clerget *et al.*, 1995); v) the *rpoC* mutant (Glu402Lys) is defective in contact between β′ subunit and regulatory factors required for transcription termination and antitermination including NusA (Ito and Nakayama, 1993).

σ^{70} subunit is essential for the recognition of specific promoters by RNA polymerase. Functional analyses of the σ^{70} subunit indicated that: i) comparison of nucleotide sequence of *E.coli* (Burton *et al.*, 1981), *Salmonella typhimurium* (Erickson *et al.*, 1985) and *Bacillus subtilis* (Stragier *et al.*, 1985; 1986; Gribskav *et al.*, 1986) indicated that conserved regions (1, 2, 3and 4) exist among these σ^{70} subunits. ii) photochemical cross-link experiment showed that σ subunit was cross-linked to *lac*UV5 (Simpson, 1979) and T7 DNA (Wu and Hillel, 1978; Harrison *et al.*, 1982). σ subunit recognizes promoter through direct interaction with DNA. iii) The five substitution mutants affecting promoter recognition were isolated. Substitution mutants, Gln437His and Tyr440Ile in region 2, affected -10 promoter recognition, and Arg584Cys, Glu585Gln and Arg588His in region 4 affected -35 promoter recognition (Gardella *et al.*, 1989; Siegele *et al.*, 1989; Waldburger *et al.*, 1990; Keener and Nomura, 1993). Therefore, these regions seemed to be concerned in direct interaction of the σ subunit with -10 and -35 signals of the promoter, respectively; iv) deletion mutation analysis exhibited that the region between aa residues 530 to 539 of σ^{70} is essential for the interaction with CRP in *gal*P1 transcription and region between aa residues 557 to 575 contains PhoB contact site (Kumar *et al.*, 1994). They proposed that the contact sites for a group of proteins are located in the C-terminal region of σ^{70} subunit and called this region the contact site II. Substitution mutants (Asp570Gly, Tyr571Ala, Tyr572Leu, Glu575Lys, Val576Tyr, Lys578Glu and Phe580Val) reduced transcription from PhoB-dependent promoters including the *pstS* promoter

(Makino *et al.*, 1993; Kim *et al.*, 1995). This result indicated that the first helix of region 4.2 plays an important role in the direct interaction with PhoB. The substitution of Arg 596 to His in σ^{70} specifically suppressed the activation defect caused by the pc2 mutation (Arg38Asn) of λ cI protein. This result indicated that the C-terminal region of σ^{70} subunit interacts with CI protein to activate transcription from P_{PM} promoter (Li *et al.*, 1994). v) The cross link experiment of substrate analogue indicated that cross-link takes place not only to β subunit but also σ^{70} subunit within Glu508-Met561 containing the C-terminal protein of region 3 (Severinov *et al.*, 1994), suggesting close location of the σ^{70} region 3 to the substrate-binding site of β.

The α subunit is composed of 329 aa residues. Since an N-terminal fragment of α lacking the C-terminal segment downstream from aa 236 can be assembled into functional core enzyme, the C-terminal portion is not required for the enzyme assembly (Igarashi and Ishihama, 1991). In a temperature–sensitive *E. coli* strain carrying *rpoA*112 mutation (Ishihama et al., 1971), which encodes α subunit with Arg to Cys substitution at aa 45 (Igarashi *et al.*, 1990), the subunit assembly is blocked in either α dimerization or subsequent $\alpha_2\beta$ complex formation step at a non-permissive temperature (Kawakami and Ishihama, 1980). These observations altogether indicate that the N-terminal two-thirds of the α subunit carries the protein-protein contact sites for subunit assembly (Igarashi *et al.*, 1991a; Hayward *et al.*, 1991). Recently, Kimura *et al.* (1994; 1995a; 1995b; 1996) carried out fine mapping of this N-terminal domain (subunit assembly domain) of α by making series of deletion, insertion and aa substitution mutants. This result indicated that: the N-terminal region between aa 20-235 is sufficient and necessary for core enzyme assembly (Kimura *et al.*, 1994); the β-binding sites are located in two regions, one around aa 45 and the other around 80; and β′-binding

sites are located on two regions, one between aa 80-86 and the other between 173-200 (Kimura and Ishihama,1995a; 1995b; 1996).

On the other hand, the C-terminal region of α downstream from aa 236 is required for transcription regulation by providing interaction with class-I transcription factors (Igarashi and Ishihama, 1991; Igarashi *et al.*, 1991b) (Fig. 2). Genetic analyses of the sites interacting with various transcription factors, CRP (Zou *et al.*, 1992; Murakami *et al.*, 1996), OxyR (Tao *et al.*, 1993; 1995), CysB, MelR and AraC (Giffard *et al.*, 1988; Thomas and Glass, 1991), Ogr (P2) and δ (P4) (Sacer *et al.*, 1975; 1982; Ayers *et al.*, 1994), OmpR (Tao *et al.*, 1993; 1995) and Fnr (Lambard *et al.*, 1991) revealed that the contact site I region is composed of several subsites, each subsite being composed of about 5-10 aa residues (reviewed in Ishihama, 1992; 1993). In addition, the contact site I region of α subunit also participates in specific binding activity of DNA elements with transcription enhancement activity, located upstream of the *rrnB*P1 promoter (Ross *et al.*, 1993). Upstream sequences of the core promoter, which increase promoter strength without accessory factors, were also observed in *tyrT* promoter (Lamond *et al.*, 1983), *tufB* promoter (Delft *et al.*, 1987), *levV* promoter (Bauer *et al.*, 1988), *rrnA*P1 (Nachaliel *et al.*, 1989), *argT* promoter (Hsu *et al.*, 1991), and *gal*P1 promoter (Lavigene *et al.*, 1992). These observations altogether indicate that the N-terminal and C-terminal regions form two distinct functional domains.

In order to reveal the domain organization of α subunit on the structural basis, I employed in this study the limited proteolysis method under non-denaturing conditions. Results using two endoproteases, V8 protease and trypsin, indicate that the α subunit is composed of two major structural domains linked by an intra-domain spacer between aa residues 235-244 (Negishi *et al.*, 1995). To determine the tertiary structure of αCTD, I isolated both unlabeled and isotope-labeled αCTD [C-terminal

fragment of α subunit consisting of a 98 amino acids between amino acid residues 233-329 (plus methionine at the NH_2-terminus)]. NMR anlysis of this αCTD fragment was carried out in colaboration with Dr. Y. Kyogoku, Institute for Protein Research, Osaka University. From the structural information with atomic resolution of αCTD in solution, it was found that αCTD has four amphipathic helices and two long loops enclosing the hydrophobic core to fold the compact structure.

Fine mapping of the contact site for CRP was carried out in this laboratory by systematic mutagenesis of αCTD (Zou *et al.*, 1992; Murakami *et al.*, 1996). These studies revealed that Arg265 is the most important residue for contact between αCTD and CRP, and in addition, Asp268 Cys269 and Leu270 are also involved in αCTD-CRP interaction. These residues are all located on the surface of helix 1. In order to get a direct evidence for protein-protein interaction, I tested possible inhibitiory effect of short peptides with the sequences corresponding to the CRP contact site on α (hereafter reffered to "alpha peptide") on CRP-dependent *lac* transcription. Results will show that a peptide covering the entire helix 1 has an inhibitory effect on CRP-dependent transcription.

3. Materials and Methods

3-1 *Purification of α subunit*

α subunit was purified from *E. coli* BL21(λDE3) transformed with the expression plasmid, pGEMAX185, essentially according to Igarashi and Ishihama (1991). All steps of the purification were carried out at 4 °C. Frozen cells at - 80 °C were suspended in 5 volumes of 50 mM Tris-HCl buffer (pH 8.0 at 4 °C) containing 1 mM EDTA [lysis buffer]. After adding phenylmethylsulfonylfluoride (PMSF)(Boehringer Mannheim) and

fragment of α subunit consisting of a 98 amino acids between amino acid residues 233-329 (plus methionine at the NH_2-terminus)]. NMR anlysis of this αCTD fragment was carried out in colaboration with Dr. Y. Kyogoku, Institute for Protein Research, Osaka University. From the structural information with atomic resolution of αCTD in solution, it was found that αCTD has four amphipathic helices and two long loops enclosing the hydrophobic core to fold the compact structure.

Fine mapping of the contact site for CRP was carried out in this laboratory by systematic mutagenesis of αCTD (Zou *et al.*, 1992; Murakami *et al.*, 1996). These studies revealed that Arg265 is the most important residue for contact between αCTD and CRP, and in addition, Asp268 Cys269 and Leu270 are also involved in αCTD-CRP interaction. These residues are all located on the surface of helix 1. In order to get a direct evidence for protein-protein interaction, I tested possible inhibitiory effect of short peptides with the sequences corresponding to the CRP contact site on α (hereafter reffered to "alpha peptide") on CRP-dependent *lac* transcription. Results will show that a peptide covering the entire helix 1 has an inhibitory effect on CRP-dependent transcription.

3. Materials and Methods

3-1 *Purification of α subunit*

α subunit was purified from *E. coli* BL21(λDE3) transformed with the expression plasmid, pGEMAX185, essentially according to Igarashi and Ishihama (1991). All steps of the purification were carried out at 4 °C. Frozen cells at - 80 °C were suspended in 5 volumes of 50 mM Tris-HCl buffer (pH 8.0 at 4 °C) containing 1 mM EDTA [lysis buffer]. After adding phenylmethylsulfonylfluoride (PMSF)(Boehringer Mannheim) and

lysozyme (SEIKAGAKU, 6x cryst) to make the final concentration of 0.27 mM and 0.3 mg/ml, respectively, the cell suspension was incubated on ice for 20 min. Cells were lysed by sonication (5 sec x 6 times) (SONICATOR, HEAT SYSTEMS-ULTRASONICS, INC.) and then centrifuged at 10,000 xg for 10 min at 4 °C. α subunit recovered in the precipitate was extracted with 0.5 M NaCl in the lysis buffer, and then centrifuged at 10,000 xg for 10 min at 4 °C. Supernatant was brought to 60% saturation of ammonium sulfate, and stored at 4 °C overnight. α subunit was recovered in the precipitate after centrifugation at 10,000 x g for 20 min, dissolved in TGED buffer [10 mM Tris-HCl (pH 7.6 at 4 °C), 5% (v/v) glycerol, 0.1 mM EDTA and 0.1 mM dithiothreitol (DTT)], and then dialyzed against 0.1 M NaCl in TGED buffer. After centrifugation at 10,000 x g for 10 min at 4 °C, the supernatant was applied on to a DEAE-TOYOPEARL (TOSOH) column (3.0 x 62.5 cm) equilibrated with TGED buffer containing 0.1 M NaCl. After washing the column with two column volumes of TGED buffer, proteins were eluted with a linear gradient (1,600 ml) of 0.1–0.5 M NaCl in TGED buffer. Fractions of 10 ml were collected at a flow rate of 2 ml/min, and aliquots were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) according to the method of Laemmli (1970). Fractions containing α subunit were pooled, and proteins were salted out with 60% saturation ammonium sulfate. After 2 hr on ice, the precipitate was collected by centrifugation at 10,000 x g for 20 min at 4 °C, and dissolved in TGED containing 0.1 M NaCl. After overnight dialysis against the same buffer, the partially purified α subunit was applied on to a TOYOPEARL HW55F (TOSOH) column (1.6 x 94 cm) and eluted with TGED buffer. Fractions of 2 ml were collected at a flow rate of 1 ml/min. The pooled fractions containing α subunit were dialyzed against a storage buffer [10 mM Tris-HCl (pH 7.6 at 4 °C), 10 mM $MgCl_2$, 0.1 mM EDTA, 200 mM KCl, 50% glycerol and 1 mM DTT]

and stored at -80 °C. Starting from 15 g of induced cells, 79 mg of α subunit was purified. The purity of α subunit used in this study was more than 98% as judged by SDS-PAGE.

3-2 *Proteolytic digestion*

For trypsin digestion, the stock preparation of α subunit was diluted to 5 μg/μl with 40 mM Tris-HCl (pH 8.0 at 4 °C) containing 40 mM KCl and 5% (v/v) glycerol, and incubated with 50 ng/μl of trypsin (SIGMA) for various times at 37 °C [the ratio of α subunit to trypsin was 100 : 1 (w/w)]. The reaction was terminated by adding PMSF (final concentration, 5 mM) and stored at -30 °C. Cleavage products were analyzed by SDS-PAGE.

In the case of *S. aureus* V8 protease (Pierce) digestion, α subunit was diluted to 5 μg/μl with 125 mM phosphate buffer (pH 7.8) containing 1 mM EDTA and incubated for various times at 37 °C with 50 ng/μl of *S. aureus* V8 protease (Pierce) [100 : 1 (w/w) ratio]. The reaction was terminated by adding diisopropylfluoro-phosphate (DFP) at a final concentration of 5 mM.

3-3 *N-terminal sequence analysis*

Proteins on gels were blotted onto polyvinylidene difluoride-type supports (ProBlott, Applied Biosystems) according to the manufacturer's protocols with minor modifications. In brief, proteins were transferred electrophoretically (Bio-Rad, Transblott Cell) to ProBlott membranes in 10% cyclohexylamino-1-propanesulphonic acid (CAPS) containing 10% methanol. The membranes were stained briefly with 0.5% Coomassie blue R-250 in 10% acetic acid and 25% isopropanol. Stained bands were

excised from the membranes and subjected to Edman degradation analysis using an Applied Biosystems Model 477A Protein/Peptide Sequencer equipped with an Applied Biosystems Model 120 phenylthiohydantoin (PTH) Analyzer for on-line analysis of PTH derivatives. Twenty reaction cycles were performed for each proteolysis fragment.

3-4 *C-terminal sequence analysis*

After SDS-PAGE of proteolytic fragments, each gel band was excised into small gel slices. The gel slices were homogenized in an ice-cold elution buffer (total volume, less than 1 ml) composed of 50 mM Tris-HCl (pH 8 at 25 °C), 0.1 mM EDTA, 5 mM DTT, 0.15 M NaCl and 0.1% SDS. Elution was performed by mixing gently the gel suspension overnight at 4 °C. After centrifugation at 15,000 rpm for 20 min at 4 °C, the supernatant was concentrated to less than 100 μl using a vacuum centrifuge (Speedvac concentrator SVC 200H, SAVANT) and then applied to Quick Spin column (Boehringer Mannheim) to remove remaining polyacrylamide gel particles. Proteolytic fragments (100 to 200 pmol each) were hydrolyzed with 10 ng carboxypeptidase Y (Takara) at 37 °C in 100 mM pyridine-acetate buffer (pH 5). At 1 hr intervals, aliquots were heated at 95 °C for 5 min and dried up by vacuum centrifugation.

Phenylthiocarbamyl (PTC) coupling of liberated amino acids was carried out as follows: Dried sample was dissolved in 100 μl of a freshly prepared coupling buffer (acetonitorile : triethylamine : distilled water = 7 : 2 : 1), dried again by evaporating the coupling buffer with a vacuum centrifuge, and then redissolved in 100 μl of the coupling buffer containing 5% phenylisothiocyanate (PITC). After incubation at room temperature

for 20 min, reaction products were dried up by vacuum centrifugation, and stored under argon at -30 °C until analysis.

For PTC-amino acid analysis, the sample was dissolved in 50 mM sodium acetate (pH 5) and applied to Applied Biosystems Model 120 PTH Analyzer (C18, PTC-column). Elution conditions were: column temperature, 38 °C; wave length, 254 nm; flow rate, 0.3 ml/min; and gradient system [93% solvent A (50 mM sodium acetate, pH 5.4) and 7% solvent B (70% (v/v) acetonitrile / water) at 0 min; 68% A and 32% B at 10 min; 45% A and 55% B at 20 min; and 100 % B at 25 min and thereafter].

3-5 *Expression and purification of αCTD*

A 3'-terminal fragment of the *rpoA* gene from nucleotide 697 (counting from the first base of the initiation codon) to 993 was PCR-amplified using a pair of primers. The 5'-primer was attached with *Nde*I site while the 3'-primer with *Bam*HI site. PCR was carried out according to the operater manual (PERKIN ELMAR CETUS). In brief, reaction mixture contained in 100μl: template DNA, 1ng; 10x reaction buffer, 10 μl; dNTP mixture (1.25mM each of dATP,dCTP,dGTP and dTTP); primer, 0.6μg each, and Amplitaq, 0.5μl (2.5 unit/tube). Reaction program was: melting (94 °C, 1 min), annelling (37 °C, 1 min) and polymerization (72 °C, 2 min) for 25 cycles. The PCR-amplified *rpoA* fragment was inserted into an expression vector pET-3a between *Nde*I and *Bam*HI site to make pETαCTD. The resulting plasmid expresses a carboxy-terminal α fragment of 98 amino acids in length [97 amino acids from α carboxy-terminus plus methionine encoded by the initiaton codon within the *Nde*I site].

Transformation and expression of αCTD was carried out essentially according to Igarashi and Ishihama (1991). In brief, *E. coli* BL21(λDE3) was

transformed with pETαCTD, and the transformant was grown in LB medium counting 100 μg/ml of ampicillin. When the culture reached to 30 Klett units, 30 units as mesured with a Klett-Summerson photometer (filter, 600nm) expression of αCTD was induced by adding IPTG (WAKO) at 1 mM. After incubation at 37 °C for 2 hrs, cells were harvested and stored at -80 °C.

For purification of αCTD, frozen cells were suspended in lysis buffer, and lysozyme (SEIKAGAKU, 6x Cryst) was added to the cell suspension to make a final concentration of 3.0 mg/ml. After incubaton for 20 min on ice, the cell suspention was sonicated (5 sec, 6 times, at 4 °C) (SONICATOR™ HEAT SYSTEMS-ULTRASONICS, INC) and then centrifuged at 1500 x g for 15 min at 4 °C. αCTD was recoverd in the supernatant fraction [intact α is recoverd in the precipitate fraction (Igarashi and Ishihama, 1991)]. The supernatant was applied to a DEAE-TOYOPEAL (TOSOH) column (1.5 x 11 cm) and the column was developed with a 0.05-0.5 M linear gradient of NaCl in TGE buffer [10 mM Tris-HCl (pH 7.6 at 4 °C), 0.1 mM EDTA and 5% glycerol]. The peak fractions of αCTD as monitored by SDS-PAGE were pooled and applied onto TOYOPEAL-HW55F (TOSOH) gel filtration column (1.6 x 90 cm) and proteins were eluted with the TGE buffer containing 0.1 M NaCl. The peak fraction of αCTD at this step was more than 99% pure as juged dy SDS-PAGE followed by stainning with Comassie brilliant blue (see Fig. 7).

For preparation of the NMR sample of higher than 1 mM concentration, pooled αCTD fractions were applied onto a Protein PAK G-DEAE (Waters) column (8.2 x 75 mm) equilibrated with 50 mM phosphate buffer (pH 6.0) and eluted with the same buffer containing 0.5 M NaCl. The purified αCTD was diluted with 50 mM phosphated buffer (pH 6.0) to make 0.3 M NaCl solution. Starting from 4.95 g cells, the yield of αCTD at the final step was 6.9 mg. ^{15}N-labeled and ^{13}C/^{15}N-double labeled αCTD were expressed in M9 medium containing $^{15}NH_4Cl$ (0.05%) and ^{13}C-D-glucose (0.1%) as nitrogen and carbon sourses. Transformant was grown in the same medium to OD_{600} = 0.6.

Expression of αCTD was induced by adding IPTG at 1mM. After incubation at 37 °C for 2 hrs, cells were harvested. The labeled αCTD was purified by the same procedure as employed for purification of the unlabelled sample for NMR analysis.

3-6 *NMR measurment*

The purified αCTD was dissolved to make a final concentration of 1 mM in H_2O/D_2O (90%/10%, v/v). For ^1H NMR analysis, pH was adjusted to 5.5 with 5 mM phosphate buffer containing 20 mM KCl and 1 mM dithiothreitol. For NMR analysis of the ladeled αCTD, the protein concentration was adjusted to 2 mM solution in 20 mM phosphate buffer (pH 6) containing 30 mM KCl and 1 mM DTT.

The structure of αCTD was determined by multidimensional heteronuclear magnetic resonance spectroscopy (Jeon *et al.*, 1995). Resonance assignments of ^1H, ^{15}N and ^{13}C nuclei were given by means of the 3D NMR experiments. For the backbone nuclei resonances, ^{15}N-edited total correlation spectroscopy (TOCSY)-HSQC was performed (Grzesiek *et al.*, 1992). For the side chain nuclei resonance, HEHOHEHAHA was performed (Grzesiek *et al.*, 1993). Distance information was collected using two-dimensional homonuclear nuclear Overhauser effect spectroscopy (NOESY) with a 100-ms mixing time. NMR measurements were carried out at 500 MHz with Bruker AMX500 and DMX 500 spectrometers by Prof. Y. Kyogoku and colleagues (Institute for Protein Research, Osaka University).

3-7 *Synthesis of Alpha peptides*

α peptides were synthesized with an Applied Biosystems 430A automated peptide synthesizer by the standard tert-butyloxycarbonyl (t-Boc)

method using pheny-lacetamidomethyl (PAM) resin. Couplings were mediated by symmetrical anhydrides, dicyclohexylcarbodiimide (DCC) and 1-hydroxybenzotriazole (HOBT) active esters. Cys269 was protected with 4-methoxy benzyle, while Arg255 and Arg265 were with mesitylene-2-sulfonyl. Other amino acids were protected according t-Boc method (other aa was protected by t-butoxycarbonil group). Cleavage of the peptides from the resin and removal of the protecting groups were achived by treatment with TFMSA/TFA and a mixture of scavengers, ethandithiothiol and anisole. The peptides were precipitated by the addition of diethyl ether at 4 °C, dried up by lyophilization, and purified by reverse phase HPLC using μ BONDASPHERE column (Waters) at a flow rate of 5 ml/min, and a linear gradient of acetonitrile formed from between 0.05% (v/v) TFA solution and acetonitrile containing 0.05% (v/v) TFA. The purified peptides were dried up by vaccum centrifugation. The sequences were checked using an Applied Biosystems model 477A Protein/Peptide Sequencer equipped with an Applied Biosystems model 120 phenylthiohydantoin (PTH) Analyzer for on-line analysis of PTH derivatives. The purified alpha peptides were dissolved in water for use. The peptide concentration was determined by using the fluoromeric assay of proteins based on the content of free amino groups essentially according to Bolen *et al.* (1973).

3-8 *In vitro transcription*

In vitro transcription was carried out using wild-type *lac* and *lac*UV5 promoter fragments as templates. The following templates were used: a 205 bp *Eco*RI fragment of plasmid pKB252 carrying the *lac*UV5 promoter (Kajitani and Ishihama, 1983), a 205 bp *Pvu*II-*Xba*I fragment of pUC19 carrying the wild-type *lac* promoter. The standard assay system was as follows (total volume, 50 μl): template DNA (*lac*P1, 0.1 pmol; *lac*UV5, 0.02

pmol), CRP (2.5 pmol) and cAMP (10 μM) were incubated for 3 min at 37 °C; $E\sigma^{70}$ holoenzyme (1 pmol) was then added and preincubation was carried out for 5 min at 37°C; and finally, 15 μl of substrate mixture (4 mM ATP, 4 mM GTP, 4 mM CTP, 1.25 mM UTP and 2μCi [α-^{32}P] UTP) containing heparin (final concentration, 200 μg/ml) was added to initiate RNA synthesis (5 min, 37 °C). Labeled transcripts were ethanol-precipitated and analyzed by electrophoresis on polyacrylamide gels containing 8 M urea. The gels were exposed to imaging plates and analyzed with a BAS-2000 image analyzer (Fuji).

For the competitive inhibition assays, α subunit, αCTD or synthetic peptides were mixed with template DNA for 3 min at 37 °C before the addition of CRP and cAMP.

4. Results

4-1 Domain organization of α subunit

4-1-1 *Tryptic cleavage of α subunit*

There are 39 possible sites of trypsin cleavage in α subunit (Fig. 3). Limited digestion of α subunit (5 mg/ml) with 50 μg/ml trypsin led to produce two large fragments with the apparent molecular mass of 31 and 28 kDa (Fig. 4A, upper duplex bands), and two small fragments of 11 and 9 kDa (Fig. 4A, lower duplex bands). After prolonged digestion, the large fragments disappeared and the cleavage products migrated on SDS-PAGE as a smeared band with the averaged molecular mass of 11 kDa. The 11 kDa fragments were, however, resistant to overnight digestion with the same concentration of trypsin. A half of the input α subunit was cleaved within 30 min, and the complete conversion of α subunit to the 11 kDa fragments took place after 2 hr incubation under the conditions employed.

pmol), CRP (2.5 pmol) and cAMP (10 μM) were incubated for 3 min at 37 °C; $E\sigma^{70}$ holoenzyme (1 pmol) was then added and preincubation was carried out for 5 min at 37°C; and finally, 15 μl of substrate mixture (4 mM ATP, 4 mM GTP, 4 mM CTP, 1.25 mM UTP and 2μCi [α-^{32}P] UTP) containing heparin (final concentration, 200 μg/ml) was added to initiate RNA synthesis (5 min, 37 °C). Labeled transcripts were ethanol-precipitated and analyzed by electrophoresis on polyacrylamide gels containing 8 M urea. The gels were exposed to imaging plates and analyzed with a BAS-2000 image analyzer (Fuji).

For the competitive inhibition assays, α subunit, αCTD or synthetic peptides were mixed with template DNA for 3 min at 37 °C before the addition of CRP and cAMP.

4. Results

4-1 Domain organization of α subunit

4-1-1 *Tryptic cleavage of α subunit*

There are 39 possible sites of trypsin cleavage in α subunit (Fig. 3). Limited digestion of α subunit (5 mg/ml) with 50 μg/ml trypsin led to produce two large fragments with the apparent molecular mass of 31 and 28 kDa (Fig. 4A, upper duplex bands), and two small fragments of 11 and 9 kDa (Fig. 4A, lower duplex bands). After prolonged digestion, the large fragments disappeared and the cleavage products migrated on SDS-PAGE as a smeared band with the averaged molecular mass of 11 kDa. The 11 kDa fragments were, however, resistant to overnight digestion with the same concentration of trypsin. A half of the input α subunit was cleaved within 30 min, and the complete conversion of α subunit to the 11 kDa fragments took place after 2 hr incubation under the conditions employed.

4-1-2 *V8 proteolytic cleavage of α subunit*

There are 56 possible cleavage sites of V8 protease in α subunit (Fig. 3). By a mild treatment of α subunit (5 mg/ml) with V8 protease (50 μg/ml), intact α subunit disappeared after more than 3 hr of digestion. In parallel, three proteolytic fragments with the apparent molecular mass of 28, 24, and 9 kDa appeared within a short period (about 15 min) and subsequently, 14.4 and 5.8 kDa fragments were generated at 1 hr (Fig. 4B). In addition, a 4.2 kDa fragment was detected in the 1 hr digestion sample when more than 100 μg of the sample was analyzed. All these five cleavage fragments were resistant to further digestion with the same concentration of V8 protease at least for 3 hr. After overnight digestion, however, a small amount of 28 kDa fragment remained but all other four fragments were digested to smaller fragments.

4-1-3 *Determination of the N-terminal sequences of tryptic fragments*

Tryptic digestion generated four major fragments: 31, 28, 11 and 9 kDa polypeptides. These fragments (100 – 200 pmol) were fractionated by SDS-PAGE and electroblotted from the gels on to PVDF membranes. Proteins were detected by staining with Coomassie blue R-250, and each protein band was directly subjected to determination of the N-terminal amino acid sequence using a protein sequencer. Both the 31 and 28 kDa fragments were found to have the same N-terminal sequence as intact α subunit (Fig. 5A). Judging from the fragment sizes and the distribution of potential tryptic cleavage sites, these two fragments were estimated to correspond to N-terminal fragments, T(N-26) and T(N-24 or N-25) [T represents a tryptic fragment and the numbers in parentheses represent the tryptic cleavage sites shown in Fig. 3 (N and C represent N- or C-terminus, respectively)] (see Fig. 6).

The 11 kDa band was a mixture of two major and two minor fragments. The two major components, 11 kDa-A and 11 kDa-B, constituted 51.4% and 42.0%, respectively [the molar ratio of 11 kDa-A to 11 kDa-B fragment was 1 : 0.8] and their N- terminal residues were determined as Ile46 and Asp236, respectively. Thus, the 11 kDa-B was identified to be the C-terminal proximal fragment generated by the initial cleavage. The contents of two minor fragments, 11 kDa-C and 11 kDa-D, were only 5.4 and 1.2%, and their the N-terminal amino acid residues were determined as Ala220 and Gly34, respectively. The N-terminal amino acid residue of 9 kDa fragment corresponded to Glu244, indicating that this is a secondary cleavage product derived from the 11 kDa-B fragment (see Fig. 6).

4-1-4 *Determination of the C-terminal sequences of tryptic fragments*

The N-terminal sequence analysis indicated that the initial cleavage by trypsin takes place at Arg235 (trypsin site 26). The N-terminal proximal fragment, T(N-26), was further cleaved into more than three fragments including 11 kDa-A fragment. The one cleavage site was identified to be between Arg44 (trypsin site 5) and Arg45 (site 6), but the other was not detected. In order to identify the second cleavage site, the C-terminal amino acid sequence was determined for the 11 kDa-A and 11 kDa-B fragments. The C-terminal sequences of tryptic fragments were determined after step-wise digestion with carboxypeptidase Y (TAKARA) followed by analysis of amino acids liberated. The C-terminal sequence of the 11 kDa-A and 11 kDa-B fragments were determined to be -R-G-R and -X-D-E, respectively (Fig. 5A). Thus, we concluded that the 11 kDa-A and 11 kDa-B frgments were T(6-16) and T(26-C), respectively (see Fig. 6).

4-1-5 *Determination of the N-terminal sequences of V8 fragments*

V8 digestion generated five major fragments, 28, 24, 14.4, 9 and 5.8 kDa polypeptides. Sequence determination revealed that the 28 kDa fragment carries Phe8 at its N-terminus. The 28 kDa preparation contained another fragment (31 kDa fragment), which retained the intact N-terminal sequence of α, in a molar ratio of 3 : 1 [these two fragments were estimated to be V(1-41) and V(N-41)]. By N-terminal sequence analysis, the 9 kDa fragment was found to correspond to the C-terminal proximal half of the initial cleavage products [V(41-C)]. The N-terminal aa residue of 24 kDa fragments was found to correspond to Phe8, indicating that this fragment must be generated from the 28 kDa N-terminal fragment [V(1-41)] by a secondary cleavage near its C-terminus. The 24 kDa fragment was further digested by V8 protease to generate 14.4 kDa fragment, which carried Tyr68 at its N-terminus (Fig. 5).

The 5.8 kDa preparation was a mixture of two fragments, each carrying the N- terminal residue of Val242 and Glu245, respectively. The molar ratio of the two fragments was 3.3 : 1. Thus, we concluded that the first cleavage by V8 protease took place at Glu241 (V8 site 41) [and Glu7 (V8 site 1)]. The N-terminal fragments [V(N-41) and V(1-41)] were then cleaved at Glu67 (V8 site 9) to split into two sub-fragments, 4.2 kDa-A and 14.4 kDa-A (see Fig. 6). When 100 μg of the digestion mixture was applied on SDS-PAGE, a minor fragment of 4.2 kDa-B was also observed. The N-terminal amino acid residue of these 4.2 kDa fragments were determined to be Phe8 and Asn137, indicating that this fraction contained two fragments, 4.2 kDa-A and 4.2 kDa-B (Fig. 5).

4-1-6 *Determination of the C-terminal sequences of V8 fragments*

In order to confirm the cleavage sites in further details, the C-terminal sequences of 9 and 5.8 kDa fragments were determined. From the C-terminal sequences -X-D-E (9 kDa) and -E-V-E (5.8 kDa), we concluded that the 9 kDa fragment was the initial cleavage product [V(41-C)] but the two 5.8 kDa components were secondary cleavage products, derived from this 9 kDa fragment, carrying the same C-terminal sequence of -E-V-E, corresponding from Val242 to Glu288 [V(41-52)] and from Glu245 to Glu288 [V(42-52)] respectively (see Fig. 6). This indicates that, although the C-terminal domain with transcription activation function forms a rigid and compact structural domain, the region including aa 288 can be cleaved by proteases.

4-2 Structure of αCTD

4-2-1 *Expression and purification of αCTD*

PCR-amplified *rpoA* fragment including the coding capacity of αCTD was inserted into an expression vector pET-3a to make pETαCTD. pETαCTD was transformed into *E. coli* BL21(λDE3). The transformant was grown in LB medium containing 100 μg/ml of ampicillin to 30 Klett unit. After adding IPTG at 1mM, the expression level of αCTD reached to maximum at 2 hrs (37°C). The expressed αCTD was recoverd in the soluble fraction (intact α subunit forms inclusion bodies). In DEAE ion exchange column chromatography, αCTD was eluted in 0.14 M NaCl fraction and in HW55F gel filtration column chromatography, αCTD was eluted in the dimer position [but it stays as monomer (see Discussion)]. αCTD was purified to homogeneity by column works (Fig. 7). For NMR anlysis, αCTD was concentrated to more than 1mM by adsorbing the αCTD fractions to a small column (Protein PAK-DEAE (HPLC)) and then

eluting with a small volume of elution buffer. This purified sample gave high resolution of signals in ^{1}H NMR (Fig. 8).

The purification of isotope-labeled sample was performed by the same procedure and the purity was as high as the unlabeled sample.

4-2-2 *Structure of αCTD*

The structure of αCTD was determined by multidimentional heteronuclear magnetic resonance spectroscopy. NMR measurments were carried out at 500 MHz. The structure of αCTD consists of four amphipathic helices and two long loops including the domain termini and altogether forming a compact and rigid structure (Fig. 10). These four helices formed the hydrophobic core in αCTD. Helix 1 contains residues Val264 to Leu273, helix 2 Ile278 to Gln283, helix 3 Glu286 to Thr292 and helix 4 Lys297 to Ser309. The N-terminal loop contains two α-helical turns, Ile252 to Arg255 and Pro256 to Asp259 (Fig. 9). The helix 1 is perpendicular to the longest helix 4. The N-terminus of helix 1 and that of helix 4 are very close to each other on the tertiary structure. The N-terminal loop upstream from Thr263 and the C-terminal loop downstream from Gly311 close to each other between Phe249 (N-terminal loop) and Trp321 and Ile326 (C-terminal loop). The root-means-square deviation (RMSD) for the back bone heavy atoms of 50 structures from Phe249 to Ile326 was 0.67 Å and the region from Val264 to Ser309 was 0.57Å.

αCTD is able to bind to not only transcription factors but also the *rrnB*P1 promoter UP element (Blatter *et al.*, 1994). In order to determine the residue of αCTD involved in interaction with *rrnB*P1 promoter UP element, ^{15}N-labeled αCTD was subjected to chemical shift perturbation experiments by using the *rrnB*P1 promoter UP element duplex DNA with a sequence [d(TCAGAAAATTATTTTAAATTTCCTC)] (Jeon *et al.*, 1994).

It is based on the idea that NH and NH_2 groups, which are involved in protein-DNA interactions, are likely to undergo relatively large chemical shift changes on complex formation. The selective signal loss is observed due to either line broadening or intensity reduction in (^{15}N, ^{1}H) hetero nuclear signal quantum correlation (HSQC) spectra of ^{15}N enriched protein in the presence of small amount of DNA. The substantial selective signal losses were observed for the resonance of Arg265, Asn268, Leu270, Thr292, Asn294, Leu295 and Gly296 (Joen *et al.*, 1995), indicating that these residues interact with *rrnB*P1 promoter UP element. The intermediate broading effects were observed for Thr263, Val264, Ala274, Ser299 and Glu302, suggesting that these residues are in weak interaction with the DNA UP element, or the observed signal loses are due to indirect effects. Val264, Arg265, Asn268, Leu270 are located in the helix 1, and Ser299 and Glu302 are in the N-terminal region of helix 4. Asn294 and Leu295 are located in the loop region between helix 3 and helix 4. Thr263 and Ala274 are close-neighboring residues near N-terminal and C-terminal end of helix 1. These results altogether indicated that the entire helix 1 and the N-terminal region of helix 4 are involved in the interaction with *rrnB*P1 promoter UP element.

4-3 Inhibition of *lac* transcription by alpha peptides

4-3-1 *Theoretical considerations*

Complex formtion between DNA and proteins is inhibited by addition of oligonucleotides with the sequence of protein-binding sites. Likewise, the formation of protein-protein complexes should be competed by adding oligopeptides with the aa sequence corresponding to the contact sites. Such attempts have not been made for analysis of RNA polymerase-transcription

factor contacts.

In order to confirm that helix 1 in αCTD is indeed involved in direct contact with CRP, I tried to test the inhibition of CRP-α subunit contact by synthetic peptides with the sequence of αCTD helix 1 region. Four kinds of alpha peptides were synthesized as shown in Fig. 11. Peptides A, B and C correspond to parts of α sequence between Asp250 and Asp280. Peptide A is a tetrapeptide including an important residue Arg265 for CRP contact. Peptide B is an undecapeptide including the N-terminal proximal half of helix 1. Peptide C contains the whole helix 1. Peptide D is a mutant of peptide C, which differs by one residue at Arg265 (the most important residue for CRP-dependent transcription) substituted to Ala. To monitor the effective inhibition of CRP-α subunit contact, I measured transcription inhibition of CRP-dependent *lac* transcription *in vitro*.

4-3-2 *Inhibition of lac transcription*

First I carried out inhibition assay of *lac* transcription by α subunit and αCTD. Both α subunit and αCTD inhibited CRP-dependent transcription from *lac*P1 promoter even though CRP-independent transcription from *lac*UV5 promoter was also inhibited but to lesser extent (Fig. 12). The inhibition of *lac*UV5 transcription was considered to be non-specific, and the ratio of *lac*P1/*lac*UV5 transcription was taken as an indication of specific inhibition. Upon increase in the concentration of α subunit or αCTD to 2500 pmol, transcription level decreased to 40 and 80%, respectively (Fig. 13). The results indicated that this transcription inhibition assay can be used for specification of the CRP contact site on α subunit.

Next, I analyzed effect of the synthetic peptides. Peptide A, B and D were added in the reaction mixture from 0 to 30,000 pmol and peptide C was added from 0 to 10,000 pmol. The gel patterns of transcripts are shown

in Fig. 14 and the quantitative data are shown in Fig. 15. Peptide A did not show any inhibitory activity at least up to the highest concentration examined. Peptide B inhibited transcription level to approximetry 70% at 30,000 pmol. Peptide C inhibited transcription level to 50% at 10,000 pmol. Peptide D, a mutant derivative of peptide C, inhibited to about 80% at 10,000 pmol. Thus, peptide C is the most effective inhibitor but peptide D has a half inhibitory activity of the level of peptide C. These results are consistent with the idea that peptide C carries the CRP contact site.

5. Discussion

5-1 *Domain organization of α subunit*

Domain structure is defined as a stable unit of protein structure, which can fold autonomously, and generally consists of a continuous single stretch of polypeptide chain from 100-200 amino acids in length (for examples see Gõ, 1983; Janin, J. and Chothia, C., 1985). The domain structure is often a basic unit of a function (Janin *et al.*, 1985). For examples, DNA polymerase I from *E. coli* consists of two fragments, large fragment and small fragment. These two fragments can be separated by limited proteolysis with subtilisin. The large fragment retains the polymerase activity and 3′→5′ exonuclease activity and the small fragment retains only 5′→3′ exonuclease activity (Klenow *et al.*, 1971; Jacobsen *et al.*, 1974). CRP contains two functional domains, DNA-binding domain and dimerization domain (cAMP binds to this domain) (Mckay *et al.*, 1981). Nitrate reductase from spinach has a multidomain structure, consisting of molybdenum domain, cytochrome b_{557} domain and FAD

in Fig. 14 and the quantitative data are shown in Fig. 15. Peptide A did not show any inhibitory activity at least up to the highest concentration examined. Peptide B inhibited transcription level to approximetry 70% at 30,000 pmol. Peptide C inhibited transcription level to 50% at 10,000 pmol. Peptide D, a mutant derivative of peptide C, inhibited to about 80% at 10,000 pmol. Thus, peptide C is the most effective inhibitor but peptide D has a half inhibitory activity of the level of peptide C. These results are consistent with the idea that peptide C carries the CRP contact site.

5. Discussion

5-1 *Domain organization of α subunit*

Domain structure is defined as a stable unit of protein structure, which can fold autonomously, and generally consists of a continuous single stretch of polypeptide chain from 100-200 amino acids in length (for examples see Gõ, 1983; Janin, J. and Chothia, C., 1985). The domain structure is often a basic unit of a function (Janin *et al.*, 1985). For examples, DNA polymerase I from *E. coli* consists of two fragments, large fragment and small fragment. These two fragments can be separated by limited proteolysis with subtilisin. The large fragment retains the polymerase activity and 3′→5′ exonuclease activity and the small fragment retains only 5′→3′ exonuclease activity (Klenow *et al.*, 1971; Jacobsen *et al.*, 1974). CRP contains two functional domains, DNA-binding domain and dimerization domain (cAMP binds to this domain) (Mckay *et al.*, 1981). Nitrate reductase from spinach has a multidomain structure, consisting of molybdenum domain, cytochrome b_{557} domain and FAD

domain. This domain organization was revealed by limited proteolysis with trypsin and V8 protease (Kubo *et al.*,1988).

The domain structure can be further separated to subdomains. The subdomain is a unit of folded structure larger than an isolated helix or sheet, but smaller than an entire domain (Rose, G. D., 1979). Oct-1 and Oct-2, transcription factors with the recognition activity of octamer sequence (5′-ATTTGCAT-3′), contain a POU domain (DNA-binding domain) consisting of two subdomains, POU-specific subdomain and POU-type homeo subdomain. POU-specific subdomain plays a role in sequence recognition while POU-type homeo subdomain carries DNA binding avtivity (Assa-Munt *et al.*, 1993).

Limited proteolysis under non-denaturing conditions is one of the experimental methods frequently used to investigate the domain structure of multidomain proteins. Under mild reaction conditions, endoproteases cleave proteins preferentially at the region exposed into the solvent reflecting the domain organization of the proteins. Domain linkers are often exposed into the solvent. Conformational change of a protein under different conditions can also be detected by following the change of proteolytic cleavage pattern. For an example, TyrR protein shows different patterns of trypsin cleavage in the presence or absence of ATP (Cui *et al.*, 1993).

The proteolytic cleavage was applied for analysis of subunit-subunit interactions within RNA polymerase (Lill *et al.*, 1975; Ishihama *et al.*, 1979; 1987). Proteolytic cleavage pattern is different depending on the assembly states such as between native and premature core enzymes (Ishihama *et al.*, 1979). In order to get insight into the domain organization within a single subunit, I analyzed in this study the pattern of limited proteolysis of isolated α subunit by endoproteases, V8 protease and trypsin. The results indicate that the region between Arg235 to Glu244 are highly

accessible for cleavage by endoproteases, and that the initial cleavage takes place within this region (see Fig. 6). These regions were assumed to be exposed on protein surface, presumably forming an interdomain linker. I carried out surface probability analysis by using Gene Works (TEIJIN Co.) according to Janin *et al.* (1978) and Emini *et al.* (1985). This analysis indicated that the region between Arg235 to Glu248 has a high possibility to be exposed on protein surface. This surface prediction is in agreement with the result of proteolytic cleavage. I thus propose two major structural domains for α subunit, *i.e.*, N-terminal domain (aa residues 8 – 235), and C-terminal domain (aa residues 245 – 329). This two-structural domain model is in good agreement with the functional map of α subunit (Igarashi and Ishihama, 1991; Kimura *et al.*, 1994; also reviewed in Ishihama, 1992; 1993). The N-terminal region between aa 21 to 235 is involved in the subunit assembly (Kimura *et al.*, 1994; Kimura and Ishihama, 1995a; 1995b; 1996), while the C-terminal region up to aa 235 carries the class-I factor contact sites (Igarashi and Ishihama, 1991; Igarashi *et al.*, 1991b; for reviews see Ishihama, 1992; 1993).

Upon prolonged proteolysis, the N-terminal domain was cleaved into two subdomain fragments (Na and Nb) between Arg45 and Glu68. The proximal N-terminal Na subdomain was rapidly digested upon further proteolysis with trypsin, suggesting that the subdomain Na undergoes rapid unfolding after the secondary cleavage. In this small domain of 38 amino acids in length, there are 4 possible sites for trypsin cleavage. Analyses of insertion mutants of two amino acids, Ala and Ser, and Ala-substitution mutants at various positions of the N-terminal domain indicates that α dimer contains more than one contact including the site of around Gly40 (Kimura and Ishihama, 1995a; 1995b; 1996). Thus, the subdomain Na may be involved in inter-subunit contact between two α subunits. Subdomain Na alone may not be able to form dimer, and as a result, it must

be degraded rapidly. The subdomain Nb plays essential roles not only in α dimerization but also in linking large two subunits, β and β′, all together forming an assembly core of RNA polymerase. The subdomain Nb of 168 amino acids in length was cleaved between Arg150 and Gly151 into two smaller fragments. The proximal N-terminal fragment containing the β subunit-binding site was resistant to cleavage by trypsin, but the rest carrying the β′ subunit-binding site was digested rapidly upon continued proteolysis. It seems that β′ subunit-binding site tends to unfold after second cleavage.

The C-terminal proximal domain of 85 amino acids in length (aa residues 245–329) was highly resistant to endoproteinases. In spite of the fact that this domain has a number of contact sites with class I transcription factors (reviewed by Ishihama, 1992; 1993), it seems that this region maintains a compact and rigid structure at least under the conditions employed. Independently, Blatter *et al.* (1994) carried out limited digestion of α subunit with trypsin and found that it is composed of two domains, N-terminal proximal domain comprised of aa 2 to 241 and C-terminal domain of aa 249 to 329.

5-2 *Structure of αCTD*

Since the finding that the C-terminal domain is involved in protein-protein and protein-DNA contacts for transcription regulation, it is supposed that this domain changes its conformation upon interaction with transcription factors or DNA UP elements and alters promoter recognition properties of the RNA polymerase. It is therefore important to analyze the conformational change in the C-terminal domain upon transcription activation.

NMR spectroscopy has become a powerful technique to get structural and dynamic information of proteins in solution at the atomic level. At present, however, detailed structures can be obtained only for nonaggregating proteins whose molecular weights are less than 20 kDa. Since the molecular weight of αCTD is about 11 kDa, it was suitable for NMR analysis. The spectrum of ^{1}H NMR provided high resolution signals from individual side chain groups. Using analytical size-exclusion chromatography, Blatter *et al.* (1994) showed that αCTD forms a dimer at 8 μM. At a higher concentration (1 mM) used in the present experiment, however, the profile of ^{1}H NMR indicated that αCTD stayed as a monomer. And a sedimentation equilibrium experiment on a 0.1 mM αCTD solution showed a molecular weight of 11,800 as a monomer (Jeon *et al.*, 1995). αCTD seems to have affinity for certain resins of size-exclusion chromatograpy and thus it is eluted at a position of dimer (Kimura *et al.*, 1994).

Detailed mapping of the contact site for CRP has been carried out by systematic mutagenesis within or near the contact site I of αCTD (Zou *et al.*,1992; Murakami *et al.*, 1996). Zou *et al.* (1992) prepared a set of *rpoA* mutant library by using PCR mutagenesis and isolated α mutants defective in response to CRP. The *rpoA*129 (Arg265Cys) mutant showed the most decreased *lac* transcription; *rpoA*127 (Leu270Pro) mutant also showed decreased expression of *lac*; but the level of *lac* expression in other mutant *rpoA*123 (Asn268Asp) is higher than those of *rpoA*129 and *rpoA*127. Murakami *et al.* (1996) carried out a systematic mutagenesis in aa residues 258-275 and aa 297-298 by substitution of alanine or tryptophan and examined *lac* transcription by reconstituted mutant RNA polymerases carrying these mutant α subunits *in vitro*. The results indicated that the mutant enzymes carrying [260A]α, [262A]α, [265A]α, [268A]α, [297A]α and [270A]α showed a decreased level of *lac* transcription, among which

[265A]α mutant enzyme exhibited the most decreased activity. The results were essentially the same as those obtained with the enzymes carrying Trp-scanning α subunits. Therefore mutations affecting the CRP response have been mapped in two regions of αCTD, aa 265-270, and 296 to 299. The one region, aa 265-270, was on helix 1 and the other region, aa 296 and 299, was on helix 4.

The two OxyR contact regions aa 265-269 and aa 293-300 are also located on helix 1 and helix 4, respectively, and very close to those of CRP (Tao *et al.*, 1993; 1995). The contact sites for CysB, MelR and AraC were suggested to include Lys271 by analysis of *rpoA*341, which shows pleiotropic defects in CysB, MelR and AraC-dependent gene expression (Giffard *et al.*, 1988; Thomas and Glass, 1991). This site is also located on helix 1. The contact sites for bacteriophage P2 ogr protein and bacteriophage P4 δ protein were suggested to include Leu290 by analysis of *rpoA*109 (Sauer *et al.*, 1975; 1982). This site is located near the N-terminal region of helix 3 and near V8 protease cleavage site Glu288 (see Fig. 9). This region is exposed on the protein surface and is able to interact with transcription factors. The contact site for Fnr, a transcription factor for the genes induced under anaerobic condition, was suggested to locate within a stretch of seven amino acid between Gly311 to Arg317 (Lambardo *et al.*, 1991). The contact site for OmpR was indicated to be mapped at Pro322 and Pro323 on the extream C-terminal region of αCTD. OmpR controls transcription activation in a positive or negative fashion at the *ompC* and *ompF* promoters in response to medium osmolarity (Slauch *et al.*, 1989; 1991). These amino acid residues of OmpR contact site were included within the C-terminal loop downstream from Gly311. This OmpR contact site may be exposed on the protein surface and will be able to interact with OmpR.

A set of monoclonal antibodies specific for α subunit have been isolated by Krakow and colleagues (Venezia and Krakow, 1990). Epitope mapping experiments indicated that α subunit contained three major antigenic regions, aa residues 1-23, 190-210 and 310-320 (Sharif *et al.*, 1994). The third region, amino acide residues 310-320, is located within the C-terminal loop of α.

The contact sites on α with transcription factors consist of segments of about 8-10 amino acid residues, which are close to antigenic epitopes. Thus the type of protein-protein interaction between α subunit and transcription factors may be simillar to that of epitope-paratope interaction between antigens and antibodies, but it may be lower than epitope-paratope interaction in strength of binding.

The genetic studies in this laboratory revealed that Leu260, Leu262, Arg265 Asn268 and Lys297 were important for interaction with both *rrnB*P1 promoter UP element and CRP, while Leu270, Ile275 and Lys298 were only involved in interaction with CRP and Cys269 was only responsible for interaction with *rrnB*P1 promoter UP element (Zou *et al.*,1992; Murakami *et al.*, 1996). Both genetic studies and NMR analysis showed that both Arg265 and Asn268 interact with *rrnB*P1 promoter UP element. However selective signal loss in NMR analysis was not observed for Leu260, Leu262, Cys269 and Lys297 predicted by genetic studies. These residues might affect α-UP element interaction indirectly. NMR analysis provides direct evidence about protein-DNA interaction by using intact ^{15}N-labeled αCTD and *rrnB*P1 promoter UP element duplex DNA, while genetic studies can not rule out indirect effect(s).

The tertiary structure of αCTD shows that both N-terminal region of helix 1 and helix 4 are close to each other and that helix 1 is perpendicular to helix 4. The relative orientation between helix1 and helix 4 is similar to helix-turn helix (HTH) motif. HTH motif is one of the DNA-binding

motifs, which include in addition, helix-loop-helix (HLH) motif (observed in MyoD, n-Myc) (Murre *et al.*, 1989), Zn finger motif (Sp1,GAL4) (Pavletich *et al.*, 1991; Kraulis *et al.*, 1992; Marmorstein *et al.*, 1992) and lucine zipper motif (GCN4, Jun, Fos, Myc) (Landschulz *et al.*, 1988; Gentz *et al.*, 1989). All these DNA-binding motifs contain helix as a part of structure.

The HTH structure was the first discoverd DNA-recognition motif that exists in a large family of prokaryotic DNA-binding proteins such as λ Cro repressor (Anderson *et al.*, 1981; Ohlendorf *et al.*, 1982; Matsuo *et al.*, 1995), CRP (Mckay *et al.*, 1981) and λ repressor (Pabo *et al.*, 1982). The basic structure of HTH motif consists of an α helix, a turn, and a second α helix (Steiz *et al.*, 1982 ; Ohlendorf *et al.*, 1983). For instance, HTH motif of λ Cro repressor is a segment of 20 amino acid residues in length, consisting of the first helix (amino acid residues 1-7) and the second helix (residues 12-20). This second helix is a DNA recognition helix, and fits into the major groove of DNA. Side chains from this helix are considered to make a site-specific contacts with bases in the major groove.

αCTD recognized *rrnB*P1 promoter UP element by using two helices, the entire helix 1 and the N-terminal region of helix 4. However, helix 1 and helix 4 are not directly connected, and the orientation of both helices are opposite to each other. This αCTD motif can interact with not only DNA UP element but also transcription factors, CRP and OxyR. Such type of motif have never been identified.

5-3 *Inhibition of transcription activation by alpha peptides*

To determine contact sites with the transcription factor, we usually employ genetic methods by making a library of *rpoA* point mutation. Mutation always accompany indirect effect due to conformational changes

of a protein. Therefore, in order to confirm the contact site of α for CRP, I synthesized a set of peptides containing the sequence of the CRP contact site and these peptides were subjected to inhibition assay of *lac* transcription. In the previous studies, anti-α monoclonal antibodies have been used for inhibition studies of the RNA polymerase (Venezia and Krakow, 1990; Sharif *et al.*, 1994). Krakow and colleagues (1990) isolated four monoclonal antibodies against the purified α subunit and these four monoclonal antibodies strongly inhibited CRP-dependent initiation from *lac*P1 promoter and partially inhibited initiation directed by *lac*UV5 promoter (Venezia *et al.*, 1990). Interaction between CRP and RNA polymerase is affected by binding of one of the anti-α monoclonal antibodies to α (Riftina *et al.*, 1990). The epitope of this anti-α monoclonal antibody may be proximal to the contact site between CRP and RNA polymerase. However, antibodies are huge proteins, and an indirect effect(s) of the monoclonal antibobies can not be ruled out.

Recently, Joseph *et al.*, (1995) carried out peptide inhibition assay for mapping of functional domains on GTP-binding protein, Rac1, and proposed to call this method as "peptide walking". They succeeded to map five functional sites in the primary structure of Rac1 by inhibition test of NADPH oxidase activation. Four of the five domains are exposed on the surface of Rac1 and were not identified previously by mutational analyssis. Peptide inhibition assay can supplement to mutational analysis. Coen *et al.* (1995) identified a class of specific peptide inhibitors of herpes simplex virus (HSV) DNA polymerase to disrupt subunit-subunit interaction. The HSV DNA polymerase consists of a catalytic subunit (Pol) and a smaller accssesory subunit UL42 that increases the processivity of the enzyme. They synthesized peptides corresponding to the C terminus of Pol and identified some peptides with blocking activity of the protein-protein interaction between Pol and UL42. These peptides and their structure gave

an information for the synthesis of clinically effective drugs, because these peptides were found to block virus replication *in vivo* (Digared *et al.*, 1995).

I tried "peptide walking" to cofirm the CRP contact site on α subunit by using four peptides (see Fig. 11). First I examined the inhibition assay by using α subunit and αCTD. α subunit was two fold more active in inhibition than αCTD. Since α subunit exists as dimer in solution (Saitoh and Ishihama, 1976) and binds to UP elements template more stably than αCTD (Blatter *et al.*, 1994), it is reasonable that α subunit is more active than αCTD. Next I tried the inhibition assay by peptides. Peptide A is very short and unable to fold any stable secondary structure even though it contains the most important residue, Arg265. This peptide did not inhibit *lac* transcription. On the other hand, the lengths of peptide B, C and D are 11, 16 and 16 aa residues, enough to form secondary structure (1 helix turn is 3.5 aa residues). Peptide B cotains major aa residues involved in the CRP contact sites but only half of the helix 1. Peptide B lacks two Ala residues, Ala272 and Ala274, located on the C-terminal proximal region of helix 1. Since Ala has the highest helical potential than other amino acids (Marqusee *et al.*, 1989), peptide B may form less stable helix structure than peptides C and D. As expected, peptide B inhibited *lac* transcription in about half the level of peptide C. Peptide C contains both the CRP contact site and the entire helix 1. In fact, peptide C was found to be the most active inhibitor among the test peptides. Peptide D is a mutant derivative of peptide C, carrying a substitution of Ala for Arg265. The level of transcription inhibition by peptide D was half of that by peptide C and close to that by peptide B. The replacement of Arg265 by Ala provided the most drastic damage to CRP interaction (Murakami *et al.*, 1996). NMR or CD spectroscopy will be necessary to reveal the relationship between the observed inhibition activity and the tertiary structure.

The effective concentration of peptide C for inhibition of activation by CRP is 200 μM. In the case of Rac1-GTP interaction, the effective concentration of peptides is 20 μM, while inhibition of HSV Pol-UV42 interaction is observed in the peptide concentration of 10-100 μM. Thus, peptide C was required about 2-10 fold higher concentrations than these examples (Rac 1 and Pol). Molecular interaction between CRP and RNA polymerase takes place on the same DNA molecule and may be stronger than the other two cases (Rac1-GTP and HSV Pol-UV42) which take place in free solution. It is reasonable that inhibition of CRP-RNA polymerase contact requires higher concentrations of inhibitory peptides than those required for other two cases. The difference may also be due to differences in peptide length and tertiary structure.

This new approach confirmed the contact site on α with CRP and suggested that the whole secondary structure of helix 1 is needed for interaction with CRP. The inhibition test for transcription by alpha peptides will be a useful method for mapping of the contact sites of other class-I factors.

6. References

Ahearn, J. M., Bartolomei, M. S., West, M. L., Cisek,L. J. & Corden, J. L. (1987). Cloning and sequence analysis of the mouse genomic locus encoding the largest subunit of RNA polymerase II. *J. Biol. Chem.* **262**, 10695-10705.

Allison, L. A., Moyle, M., Shales, M. & Ingles, J. (1985). Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* **42**, 599-610.

An, G. & Friesen, J. D. (1980). Charactarization of promoter-cloning plasmids: analysis of operon structure in the *rif* regulon of *Escherichia coli* and isolation of an enhanced internal promoter mutant. *J. Bacteriol.* **144**, 904-916.

Anderson, W. F., Takeda, Y. & Matthews, B. W. (1981). Structure of the *cro* repressor from bacteriophage λ and its interaction with DNA. *Nature* **30**, 754-758.

Armstrong,V. W., Sternbach, H. & Eckstein, F. (1976). Affinity labeling of *Escherichia coli* RNA polymerase with 5-formyl-1-(α-D-ribofuranosyl) uracil 5'-triphosphate. *Biochemistry* **15**, 2086-2091.

Assa-Munt, N., Mortishire-Smith, R. J., Aurora, R., Herr, W. & Wright, P. E. (1993). The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage λ repressor DNA-binding domain. *Cell* **73**, 193-205.

Ayers, DJ., Sunshine, M. G., Six, E. W. & Christie, G. E. (1994). Mutations affecting two adjacent amino acid residues in the alpha subunit of RNA polymerase block transcriptional activation by the bacteriophage P2 Ogr protein. *J. Bacteriol.* **176**, 7430-7438.

Bauer, B. F., Kar, E. G., Elford, R. M. & Holmes, M. (1988). Sequence determinants for promoter strength in the *leuV* operon of *Escherichia coli*. *Gene* **63**, 123-134.

Bendiak, D. S., Parker, J. & Friesen, J. D. (1977). Fine-structure mapping of the *rts*, *rplK*, *rplL*, and *rpoB* genes of *Escherichia coli*. *J. Bacteriol.* **129**, 536-539.

Blatter, E. E., Ross, W., Tang, H., Gourse, L. R. & Ebright, R. H. (1994). Domain organization of RNA polymerase α subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**, 889-896.

Bohlen, P., Stein, S., Dairman, W. & Udenfield, S. (1973). Fluorometric assay of proteins in the nanogram range. *Arch. Biochem. Biophys.* **155**, 213-220.

Burgess, R. R. (1969). Separation and charactarization of the subunit of ribonucleic acid polymerase. *J. Biol. Chem.* **244**, 6168-6176.

Burtou, Z., Burgess, R. R., Lin, J., Moor, D., Holder, S. & Gross, C. A. (1981). The nucleotide sequence of the cloned *rpoD* gene for the RNA polymerase sigma subunit from *E. coli* K-12. *Nucleic. Acid. Res.* **9**, 2889-2903.

Clerget, M., Jin, J. D. & Weisberg, R. A. (1995). A zinc-binding region in the β′ subunit of RNA polymerase is involved in antitermination of early transcription of phage HK022. *J. Mol. Biol.* **248**, 768-780.

Cui, J. & Somerville, R. (1993). The TyrR Protein of *Escherichia coli*: Analysis by limited proteolysis of domain structure and ligand-mediated conformational changes. *J. Biol. Chem.* **268**, 5040- 5047.

Dekker, N., Cox, M., Boelens R., Verrijzer C. P., Vliet, P. C. V. & Kaptein, R. (1993). Solution structure of the POU-specific DNA-binding domain of Oct-1. *Nature* **362**, 852-855.

Delft, J. H. M., Marinon, B., Schmidt, D. S. & Bosch, L. (1987). Transcription of the tRNA-*tufB* operon of *Escherichia coli*: activation, termination and antitermination. *Nucl. Acids. Res.* **15**, 9515-9530.

Digard, P., Williams, K. P., Hensley, P., Brooks, I. S., Dahl, C. E. & Coen, D. M. (1995). Specific inhibition of herpes simplex virus DNA polymerase by helical peptides coresponding to the subunit interface. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1456-1460.

Emini, E., A., Hughes, J., V., Perlow, D., S. & Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J. Virol.* **55,** 836-839.

Erickson, B. D., Burton, Z. F., Watanabe, K. K. & Burgess, R. R. (1985). Nucleotide sequence of the *rpsU-dnaG-rpoD* operon from *Salmonella typhimurium* and a comparison of this sequence with the homologous operon of *Escherichia coli*. *Gene* **40**, 67-78.

Frischauf, A. M., Scheit, K. H. (1973). Affinity labeling of *E. coli* RNA polymerase with substrate and template analogues. *Biochem. Biophys. Res. Commun.* **53**, 1227-1233.

Gardella, T., Moyle, H. & Susskind, M. M. (1989) A mutant *Eschrichia coli* σ^{70} subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.* **206**, 579-590.

Gentz, R., Rauscher, F. J., Abate, C. & Curran, T. (1989). Parallel association of Fos and Jun leucine zippers juxtaposes DNA binding domains. *Science* **243**, 1695-1699.

Giffard, P. M. & Booth, I. R. (1988). The *rpoA*341 allele of *Escherichia coli* specifically impairs the transcription of a group of positively-regulated operons. *Mol. Gen. Genet.* **214**, 148-152.

Glass, R. E., Jones, S. T. & Ishihama, A. (1986a). Genetic studies on the β subunit of *Escherichia coli* RNA polymerase, VII : RNA polymerase is a target for ppGpp. *Mol. Gen. Genet.* **203**, 265-268.

Glass, R. E., Jones, S. T., Nene, V., Nomura, T., Fujita, N. & Ishihama, A. (1986b). Genetic studies on the β subunit of *Escherichia coli* RNA polymerase, VIII : localization of a region involved in promoter selectivity. *Mol. Gen. Genet.* **203**, 487-491.

Glass, R. E., Honda, A. & Ishihama, A. (1986c). Genetic studies on the β subunit of *Escherichia coli* RNA polymerase, IX : the role of the carboxy-terminus in enzyme assembly. *Mol. Gen. Genet.* **203**, 492-495.

Glass, R.E., Ralphs, N. T., Fujita, N. & Ishihama, A. (1988). Assembly of amber fragments of the beta subunit of *Escherichia coli* RNA polymerase. *Eur. J. Biochem.* **176**, 403-407.

Gõ, M. (1983) Modular structure units, exons and function of chicken lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 1964-1968.

Grachev, M. A., Kolocheva, T. I., Lukhtanov, E. A. & Musyaev, A. 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.

Grachev, M. A., Lukhtanov, E. A., Mustaev, A. A., Zaychikov, E. F., Abdukayumov, M. N., Rabinov, I. V., Richter, V. I., Skoblov, Y. S. & Chistyakov, P. G. (1989). Studies of the functional topography of *Escherichia coli* RNA polymerase. A method for localization of the sites of affinity labelling. *J. Biochem.* **180**, 577-585.

Grzesiek, S., Dobeli, H., Gent, R., Garotta, G., Labhardt, A. M. & Bax, A. (1992). 1H, 13C, and 15N NMR backbone assigments and secondary structure of human interferon-gamma. *Biochemistry* **31**, 8180-8190.

Grzesiek, S. & Bax, A. (1993). Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. *J. Biomol. NMR* **3**, 185-204.

Gribskav, M. & Burgess, R. R. (1986). Sigma factors from *E. coli*, *B. subtilis*, Pharge SPO1, and Phage T4 are homologous proteins. *Nucleic. Acid Res.* **14**, 6745-6763.

Gross, C., Hoffman, J., Ward, C., Hager, D., Burdick, G., Berger, H. & Burgess, R. (1978). Mutation affecting thermostability of sigma subunit of *Escherichia coli* RNA polymerase lies near the *dnaG* locus at about 66 min on the *E. coli* genetic map. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 427-431.

Gross, C. A., Blattner, F. R., Taylor, W. E., Lowe, P. A. & Burgess, R. R. (1979). Isolation and charactarization of transducing phage coding for σ subunit of *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5789-5793.

Harris, J. D., Martinez, I. I. & Calendar, R. (1977). A gene from *Escherichia coli* affecting the sigma subunit of RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1836-1840.

Harris, J. D., Heilig, J. S., Martinez, I. I., Calendar, R. & Isaksson, L. A. (1978). Temprature-sensitive *Escherichia coli* mutant producing a temperature-sensitive σ subunit of DNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 6177-6181.

Hsu, L. M., Giannini, J. K., Leung, T. C. & Crosthwaite, J. C. (1991). Upstream sequence activation of *Escherichia coli argT* promoter in vivo and in vitro. *Biochemistry*. **30**, 813-822.

Hayward, R. S. & Scaife, J. G. (1976). Systematic nomenclature for the RNA polymerase genes of prokaryotes. *Nature* **260**, 646-647.

Hayward, R. S., Igarashi, K. & Ishihama, A. (1991). Functional specialization within the α-subunit of *Eschericia coli* RNA polymerase. *J. Mol. Biol.* **221**, 23-29.

Igarashi, K. & Ishihama, A. (1991). Bipartite function map of the *E. coli* RNA polymerase α 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 (*rpo*A) of *Escherichia coli* RNA polymerase. *Nucleic. Acids Res.* **18,** 5945-5948.

Igarashi, K., Fujita, N. & Ishihama, A. (1991a). Identification of a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **218**, 1-6.

Igarashi, K., Hanamura, A., Makino, K., Aiba, H., Aiba, H., Mizuno, T., Nakata, A. & Ishihama, A. (1991b). Bipartitie functional organization of the α subunit of *Escherichia coli* RNA polymerase: two modes of transcription activation by positive factors. *Proc. Natl. Acad. Sci., U. S. A.* **88**, 8958-8962.

Ishihama, A., Aiba, H., Saitoh, T. & Takahashi, S. (1979). Subunits of RNA polymerase in function and structure, 7. Structure of premature core enzyme. *Biochemistry* **18**, 972-978.

Ishihama, A. (1981). Subunit assembly of *Escherichiae coli* RNA polymerase. *Adv. Biophys.* **14** , 1- 35.

Ishihama, A. (1988). Promoter selectivity of prokaryotic RNA polymerases. *Trends Genet.*, **4**, 282-286.

Ishihama, A. (1991). Global control of gene expression in bacteria. *In* "*Control of Cell Growth and Division*," ed. by A. Ishihama and H. Yoshikawa, pp. 121-140, Springer-Verlag, Berlin.

Ishihama, A. (1992). Role of the RNA polymerase α subunit in transcription activation. *Mol. Miclobiol.* **6**, 3283-3288.

Ishihama, A. (1993). Protein-protein communication within the transcription apparatus. *J. Bacteriol.* **175**, 2483-2489.

Ishihama, A., Fujita, N. & Glass, R. E. (1987). Subunit assembly and metabolic stability of *E. coli* RNA polymerase. PROTEINS: Structure, Function, and Genetics **2**, 42-53.

Ito, K. & Nakayama, Y. (1993). Pleiotropic effects of the *rpoC*10 mutation affecting the RNA polymerase β′ subunit of *Escherichia coli* on factor-dependent transcription termination and antitermination. *Mol. Microbiol.* **9**, 285-293.

Iwakura, Y., Ito, K & Ishihama, A. (1974). Biosynthesis of RNA polymerase in *Escherichia coli*, I. Control of RNA polymerase content at various growth rate. *Mol. Gen. Genet.*, **133**, 1-23.

Jacobson, H., Klenow, H. & Hansen, K. O. (1974). The N-terminal amino acid sequences of DNA polymerase I from *Escherichia coli* and of the large and the small fragments obtained by a limited proteolysis. *Eur. J. Biochem.* **45**, 623-627.

Janin, J. & Chothia, C. (1985). Domain in proteins: Definitions, location, and structural principles. *Meth. Enzymol.* **115**, 420-430.

Janin, J., Wodak, S., Levitt, M. & Maigret, B. (1978). Conformation of amino acid side-chains in proteins. *J. Mol. Biol.* **125**, 357-386.

Jaskunas, S. R., Burgess, R. R. & Nomura, M., (1975). Identification of a gene for the α-subunit of RNA polymerase at the *str-spc* region of the *Eschericha coli* chromosome. *Proc. Nat. Acad. Sci. U. S. A.* **72**, 5036-5040.

Jakunas, S. R., Fallon, A. M. & Nomura, M. (1977). Identification and organization of ribosomal protein genes of *Escherichia coli* carried by *λfus*2 transducing phage. *J. Biol. Chem.* **252**, 7323-7336.

Jeon, Y. H., Negishi, T., Shirakawa, M., Yamazaki, T., Fujita, N., Ishihama, A., & Kyogoku, Y. (1995). Solution structure of the

activator contact domain of the RNA polymerase α subunit. *Science* **270** 1495-1497.

Jin, D. J. & Gross, C. A. (1988). Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J. Mol. Biol.* **202**, 45-58.

Jin. D. J. & Turnbough, C. L. Jr. (1994). An *Escherichia coli* RNA polymerase defective in transcription due to its overproduction of abortive initiation products. *J. Mol. Biol.* **236**, 72-80.

Jokerst, R. S., Weeks, J. R., Zehring, W. A. & Greenleaf, A. L. (1989). Analysis of the gene encoding the largest subunit of RNA polymerase II in *Drosophila*. *Mol. Gen. Genet.* **215**, 266-275.

Joseph, G. & Pick, E. "Peptide Walking" is a novel method for mapping functional domains in proteins. (1995). *J. Biol. Chem.* **270**, 290079-29082.

Kawakami, K. & Ishihama, A. (1980). Defective assembly of ribonucleic acid polymerase subunit in a temperature-sensitive α-subunit mutant of *Escherichia coli*. *Biochemistry* **19**, 3491- 3495.

Kajitani, M. & Ishihama, A. (1983). Determination of the promoter strength in the mixed transcription system: promoters of lactose, tryptophan and ribosomal protein L10 operons from *Escherichia coli*. *Nucl. Acids. Res.* **11**, 671-686.

Kay, H. K., Hansen, K. O. & Patkar S. A. (1971). Proteolytic cleavage of native DNA polymerase into two different catalytic fragments. (1971). *Eur. J. Biochem.* **22**, 371-381.

Keener, J. & Nomura, M. (1993). Dominant lethal phenotype of a mutation in the -35 recognition region of *Escherichia coli* σ^{70}. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 1751-1755.

Kim, S., Makino, K., Amemura, M., Nakata, A. & Shinagawa, H. (1995). Mutational analysis of the role of the first helix of region 4.2 of the σ^{70}

subunit of *Escherichia coli* RNA polymerase in transcriptional activation by activator protein PhoB. *Mol. Gen. Genet.* **248**, 1-8.

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.

Kimura, M. & Ishihama, A. (1995a). 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. & Ishihama, A. (1995b). 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. (1996). Subunit assembly *in vitro* of *Escherichia coli* RNA polymerase: role of the amino-terminal assembly domain of alpha subunit. Genes to Cells, in press

Kraullis, P. J., Raine, A. R. C., Gadhavi, P. L. & Laue, E. D. (1992). Structure of the DNA-binding domain of zinc GAL4. *Nature* **356**, 448-450.

Kubo, Y., Ogura, N. & Nakagawa, H. (1988). Limited proteolysis of the nitrate reductase from spinach leaves. *J. Biol. Chem.* **263**, 19684-19689.

Kumar, A., Grimes, B., Fujita, N., Makino, K., Malloch, R.A., Hayward, R.S. & Ishihama, A. (1994). Role of the sigma^{70} subunit of *Escherichia coli* RNA polymerase in transcription activation. *J. Mol. Biol.* **235**, 405-413.

Kumar, K. P. & Chatterji, D. (1990). Resonance energy transfer study on the proximity relationship between the GTP binding site and the rifampicin binding site of *Escherichia coli* RNA polymerase. *Biochemistry* **29**, 317-322.

Kyogoku, Y., Kojima, C., Lee, S. J., Tochio, H., Suzuki, N., Matsuo, H.& Shirakawa, M. (1994). Induced structural change in protein-DNA complexes. *Meth. Enzymol.* **261**, 524-541.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Lambardo, M. S., Bagga, D. & Miller, C. G. (1991) Mutation in *rpoA* affect expression of anaerobically regurated gene in *Salmonella typhimurium*. *J. Bacteriol.* **173**, 7511-7518.

Lamond, A. I. & Travers, A. A. (1983). Requirement for an upstream element for optimal transcription of a bacterial tRNA gene. *Nature* **305**, 248-250.

Landshulz, W. H., Johnson, P. F. & Mcknight S. L. (1988). The leucine zippar: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.

Lavigne, M., Herbert, M., Kolb, A. & Buc, H. (1992). Upstream curved sequences influence the initiation of transcription at the *Escherichia coli* galactose operon. *J. Mol. Biol.* **224**, 293-306.

Lie, M., Moyle, H. & Susskind. (1994). Target of transcriptional activation function of phage λ cI protein. *Science* **263**, 75-77.

Lill, H. & Hartmann, G. (1975). Digestion with matrix-bound proteases as a possible probe for the topography of the DNA-dependent RNA polymerase from *Escherichia coli*. *Eur. J. Biochem.* **54**, 45-53.

Lindahl, L., Post, L., Zengel, J., Gilbert, S. F., Strycharz, W. A. & Nomura, M. (1977a) Mapping of ribosomal protein gene *in vitro* protein synthesis using DNA fragments of λ*fus*3 transducing phage DNA as templates. *J. Biol.Chem.* **252**, 7365-7383.

Lindahl, L., Yamamoto, M., Nomura, M., Kirschbaum, J. B., Allet, B., Rochaix, J. D. (1977b). Mapping of a cluster of genes for

components of transcriptional machineries of *Escherichia coli*. *J. Mol. Biol.* **109**, 23-47.

Makino, K., Amemura, M., Kim, S.K., Nakata, A. & Shinagawa, H. (1993). Role of the σ^{70} subunit of RNA polymerase in transcriptional activation by activation protein PhoB in *Escherichia coli*. *Genes Devel.* **7**, 149-160.

Matsuo, H., Shirakawa, M., & Kyogoku, Y., (1995). Three-dimensional dimer structure of the λ-cro repressor in solution as determined by heteronuclear multidimensional NMR. *J. Mol. Biol.* **254**, 668-680.

Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S. C. (1992). DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356**, 408-424.

Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989). Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5286-5290.

Maxwell, A. & Reece, R. (1989). Tryptic fragments of the *Escherichia coli* DNA gyrase A protein. *J. Biol. Chem.* **264**, 19648-19653.

Mckay, D. B. & Steitz, T. A. (1981). Structure of catabolite gene activatior protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**, 744-749.

Memet, S., Gouy, M., Mark, C., Sentenac, A. & Buhler, J. M. (1988). *RPA190*, the gene coding for the largest subunit of yeast RNA polymerase A. *J. Biol. Chem.* **263**, 2830-2839.

Mosteller, R. D. & Yanofsky, C. (1970). Transcription of the tryptophan operon in *Escherichia coli*: rifampicin as an inhibition of initiation. *J. Mol. Biol.* **48**, 525-531.

Murakami, K., Fujita, N. & Ishihama, A. (1996). Transcription factor recognition surface on the RNA polymerase alpha subunit is involved in contact with the DNA enhancer element. *EMBO J.*, in press.

Murre, C., McCaw, P S. & Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD,* and *myc* proteins. *Cell* **56**, 777-783.

Mustaev, A., Kashlev, M., Lee, J., Polyakov, A., Lebedev, A., Zalenskaya, K., Grachev, M., Goldfarb, A. & Nikiforov, V. (1991). Mapping of priming substrate contacts in the active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **266**, 23927-23931.

Mustaev, A., Zaychikov, E., Severinov, K., Kashlev, M., Polyakov, A., Nikiforov, V. & Goldfarb, A. (1994). Topology of the RNA polymerase active center probed by chimeric rifampicin-nucleotide compounds. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12036-12040.

Nachaliel,N., Melnick, J., Gafny, R. and Glaser, G. (1989). Ribosome associated protein(s) specifically bind(s) to the upstream activator sequence of the *E. coli rrnA*P1 promoter. *Nucl. Acids. Res.* **17**, 9811-9822.

Nakamura, Y. (1978). RNA polymerase mutant with altered sigma factor in *Escherichia coli.* *Molec. Gen. Genet.* **165**, 1-6.

Nakamura, Y., Osawa, T. & Yura, T. (1977). Chromosomal location of a structural gene for the RNA polymerase σ factor in *Escherichia coli.* *Proc. Natl. Acad. Sci. U. S. A.* **74**, 1831-1835.

Negishi, T., Fujita, N. & Ishihama A. (1995). Structural map of the alpha subunit of *Escherichia coli* RNA polymerase: structural domains identified by proteolytic cleavage. *J. Mol. Biol.* **248**, 723-728.

Nene, V. & Glass, R. E. (1982). Genetic studies on the β subunit of *Escherichia coli* RNA polymerase. *Mol. Gen. Genet.* **188**, 399-404.

Ohlendorf, D. H., Anderson, W. F., Fisher, R. G., Takada, Y. & Matthews, B. W. (1982). The molecular basis of DNA-protein regulation inferred from the structure of cro repressor. *Nature* **298**, 718-723.

Ohlendorf, D. H., Anderson, W. F., Lewis, M., Pabo, C. O. & Matthews,

B. W. (1983). Comparison of the structure of cro and λ repressor proteins from bacteriophage λ. *J. Mol. Biol.* **169**, 757-769.

Ovchinnikov, Y. A., Lipkin, V. M., Modyanov, N. N. & Sverdlov, Y. V. (1977). Primary structure of α subunit of DNA-dependent RNA polymerase from *Escherichia coli*. *FEBS Let.* **76**, 108-111.

Ovchinnikov, Y. A., Monastyrskaya, G. S., Gubanov, V. V., Guryev, S. O., Chertov, O. Y., Modyanov, N. N., Grinkevich, V. A., Makarova, I. A., Marchenko, T. V., Polovnikova, I. N., Lipkin, V. M., & Sverdlov, E. D. (1981). The primary structure of *Escherichia coli* RNA polymerase Nucleotide sequence of the *rpoB* gene and amino-acid sequence of the β-subunit. *Eur. J. Biochem.* **116**, 621-629.

Ovchinnikov, Y. A., Monastyrskaya, G. S., Gubanov, V. V., Guryev, S. O., Salomatina, I. S., Shuvaeva, T. M., Lipkin, V. M. & Sverdlov, E. D. (1982). The primary structure of *Escherichia coli* RNA polymerase: Nucleotid sequence of the *rpoC* gene and amino acid sequence of the β ′ subunit. *Nucreic. Acid. Res.* **10**, 4036-4043.

Owens, J. R., Woody, A. M. & Haley, B. E. (1987). Characterization of the guanosine-3′-diphosphate-5′-diphosphate binding site on *E. coli* RNA polymerase using a photoprobe, 8-azidoguanosine-3′-5′-bisphosphate. *Biochem. Biophys. Res. Commun.* **142**, 964-971.

Pabo, C. O. & Lewis, M. (1982) The operator-binding domain of λ repressor: structure and DNA recognition. *Nature*, **298**, 443-447.

Pavletich, N. P. & Pabo, C. O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809-817.

Riffina, F., DeFalco, E. & Krakow, J. S. Effects of an monoclonal antibody on interation of *Escherichia coli* RNA polymerase with *lac* promoters. *Biochemistry* **29**, 4440-4446.

Rose, G. D. (1979). Hierarchic organization of domains in globular proteins. *J. Mol. Biol.* **134**, 447-470.

Ross, W., Gosink, K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severionv, K. & Gourse, R. (1993). A third recognition element of bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262**, 1407-1413.

Sagitov, V., Nikiforov, V. and Goldfarb A. (1993). Dominant lethal mutation near the 5′ substrate binding site affects RNA polymerase propagation. *J. Biol. Chem.* **268**, 2195-2202.

Saitoh, T. & Ishihama, A. (1976). Subunits of RNA polymerase in function and structure. VI. Sequence of the assembly *in vitro* of *Escherichia coli* RNA polymerase. *J. Mol. Biol.*, **104**, 621-635.

Sauer, B., Calendar, R., Ljungquist, E., Six, E. & Sunshine, M. G. (1982). Interaction of satellite phage P4 with phage 186 helper. *Virology* **116**, 523-534.

Severinov, K., Fenyo, D., Severinov, E., Mustaev, A., Chait, B. T., Goldfarb, A. & Darst, S. A. (1994). The σ subunit conserved region 3 is part of "5′-face" of active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **269**, 20826-20828.

Severinov, K., Markov, D., Sevarinova, E., Nikiforov, V., Landick, R., Darst, S. A. & Goldfarb, A. (1995). Streptolydigin-resistanct mutants in an evolutionarily conserved region of the β′ subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **270**, 23926-23929.

Severinov, K., Markov, D., Sevarinova, E., Kozlov, M., Darst, S. A. & Goldfarb, A. (1995). 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.

Sharif, K. A., Fujita, N., Jin, R., Igarashi, K., Ishihama, A. & Krakow, J. S. (1994). Epitope mapping and functional characterization of monoclonal antibodies specific for the α subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **269**, 23655-23660.

Siegele, D. A., Hu, J. C., Walter, W. A. & Gross, C. A. (1989). Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **206**, 591-603.

Slauch, J. M. & Silhavy, T. J. (1989). Genetic analysis of the switch that controls porin gene expression in *Escherichia coli* K-12. *J. Mol. Biol.* **210**, 281-292.

Slauch, J. M., Russo, F. D. & Silhavy, T. J. (1991). Suppressor mutation in *rpo*A suggests that OmpR controls transcription by direct intraction with the α subunit of RNA polymerase. *J. Bacteriol.* **173**, 7501-7510.

Squires, C., Krainer, A., Barry, G., Shen, W. F. & Squires, C. L. (1981). Nucleotide sequence at the end of the gene for the RNA polymerase β′ subunit (*rpoC*). *Nucl. Acids Res.* **9**, 6827-6840.

Steiz, T. A., Ohlendorf, D. H., Mckay, D. B., Anderson, W. F. & Matthews, B. W. (1982). Structural similarity in the DNA binding domains of catabolite gene activatior and cro repressor proteins. *Proc. Natl. Acad. Sci., U. S. A.* **79**, 3097-3100.

Stragier, P., Parsort, C. & Bouvier, J. (1985). Two functional domains conserved in major and alternate bacterial sigma factors. *FEBS Lett.* **187**, 11-15.

Stragier, P. (1986). Comment on ' duplicated sporulation genes in bacteria' by J. Errington, P. Fort and J. Mandelstam (*FEBS Letters* **188** (1985) 184-188). *FEBS Lett.* **195**, 9-11.

Sunshin, M. G. & Sauer, B. A bacterial mutation blocking P2 phage late gene expression. (1975). *Proc. Nat. Acad. Sci., U.S.A.* **72**, 2770-2774.

Tao, K., Zou, C., Fujita, N. & Ishihama, A. (1993). Involvement of the RNA polymerase α subunit C-terminal region in co-operative interaction and transcriptional activation with OxyR protein. *Mol. Microbiol.* **7**, 859-864.

Tao, K., Zou, C., Fujita, N. & Ishihama, A. (1995). Mapping of the OxyR protein contact site in the C-terminal region of RNA polymerase α subunit. *J. Bacteriol.* **177**, 6740-6744.

Thomas M. S. & Glass, R. E. (1991). *Escherichia coli rpoA* mutation of positively regulated systems. *Mol. Miclobiol.* **5**, 2719-2725.

Travers, A. A., Buckland, R., Goman, M., LeGrice, S. S. G. & Scaife, J. G. (1978). A mutation affecting the σ subunit of RNA polymerase changes transcription specificity. *Nature* **273**, 354-358.

Venezia, N. D. & Krakow, J. S. (1990). Effects of anti-α monoclonal antibodies on initiation and elongation by the *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **265**, 8122-8126.

Waldburger, C., Gardella, T., Wong, R. & Susskind, M. M. (1990). Change in conserved region 2 of *Escherichia coli* σ^{70} affecting promoter recognition. *J. Mol. Biol.* **215**, 267-276.

Yamamoto, M. & Nomura, M. (1978). Cotranscription of genes for RNA polymerase subunit β and β′ with genes for ribosomal proteins in *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.* **75**, 3891-3895.

Yura, T., &. Ishihama, A. (1979). Genetics of bacterial RNA polymerase. *Ann. Rev. Genet.* **13**, 59-97.

Zou, C., Fujita, N., Igarashi, K. & Ishihama, A. (1992). Mapping the cAMP receptor protein contact site on the α subunit of *Escherichia coli* RNA polymerase. *Mol. Miclobiol.* **6**, 2599-2605.

Table 1.

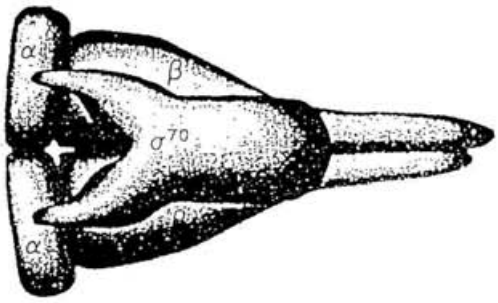
Buffer solutions

1) Lysis buffer	50 mM Tris-HCl (pH 8.0 at 4 °C) 1 mM EDTA (pH 8.0) 500 mM NaCl
2) TGE buffer	10 mM Tris-HCl (pH 7.6 at 4 °C) 50% (v/v) glycerol 1 mM EDTA 0.1 mM DTT
3) Strage buffer	10 mM Tris-HCl (pH 7.6 at 4 °C) 10 mM $MgCl_2$ 0.1 mM EDTA 200 mM KCl 50% glycerol 1 mM DTT

Table 2.

Mediums

LB medium	Trypton	10 g
	Yeast ext.	5 g
	NaCl	5 g
	Glucose	1 g
	H_2O (pH 7.2)	1 liter
Medium 1		
(^{15}N-labeled sample)	Na_2HPO_4	6.2 g
	KH_2PO_4	2.9 g
	NaCl	19 g
	CaCl2	0.01 g
	N-NH_4Cl	0.475 g
	thiamine, thymine, biothin, adenosine, guanosine and cytidine	2 mg each
	Ampicillin	50 μg/ml
	H_2O	
Medium 2		
(^{13}C-^{15}N-labeled sample)	Na_2HPO_4	7 g
	KH_2PO_4	3 g
	NaCl	20 g
	CaCl2	0.01 g
	^{5}N-NH_4Cl	0.5 g
	thiamine, thymine, biothin, adenosine, guanosine and cytidine	2 mg each
	Ampicillin	50 μg/ml
	H_2O	



Subunit	Function
α subunit	core assembly (α dimerization, β binding, β′binding)
	transcription activation (contact site I)
β subunit	binding of substrates
	binding of RNA products
	binding of rifampicin and streptolydigin
	binding of σ subunit
	stringent control (ppGpp binding)
	core assembly (α binding)
β′ subunit	binding of template DNA
	binding of σ subunit
	core assembly (α binding)
σ subunit	promoter recognition
	transcription activation (contact site II)

Figure 1. Structural model of *E. coli* RNA polymerase holoenzyme ($E\sigma^{70}$) and function of each subunit.

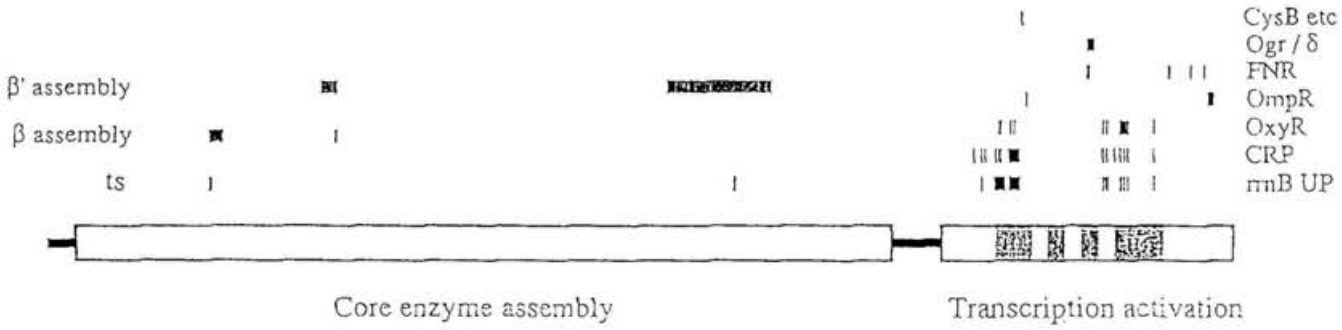


Figure 2. Functional map of the RNA polymerase α subunit.

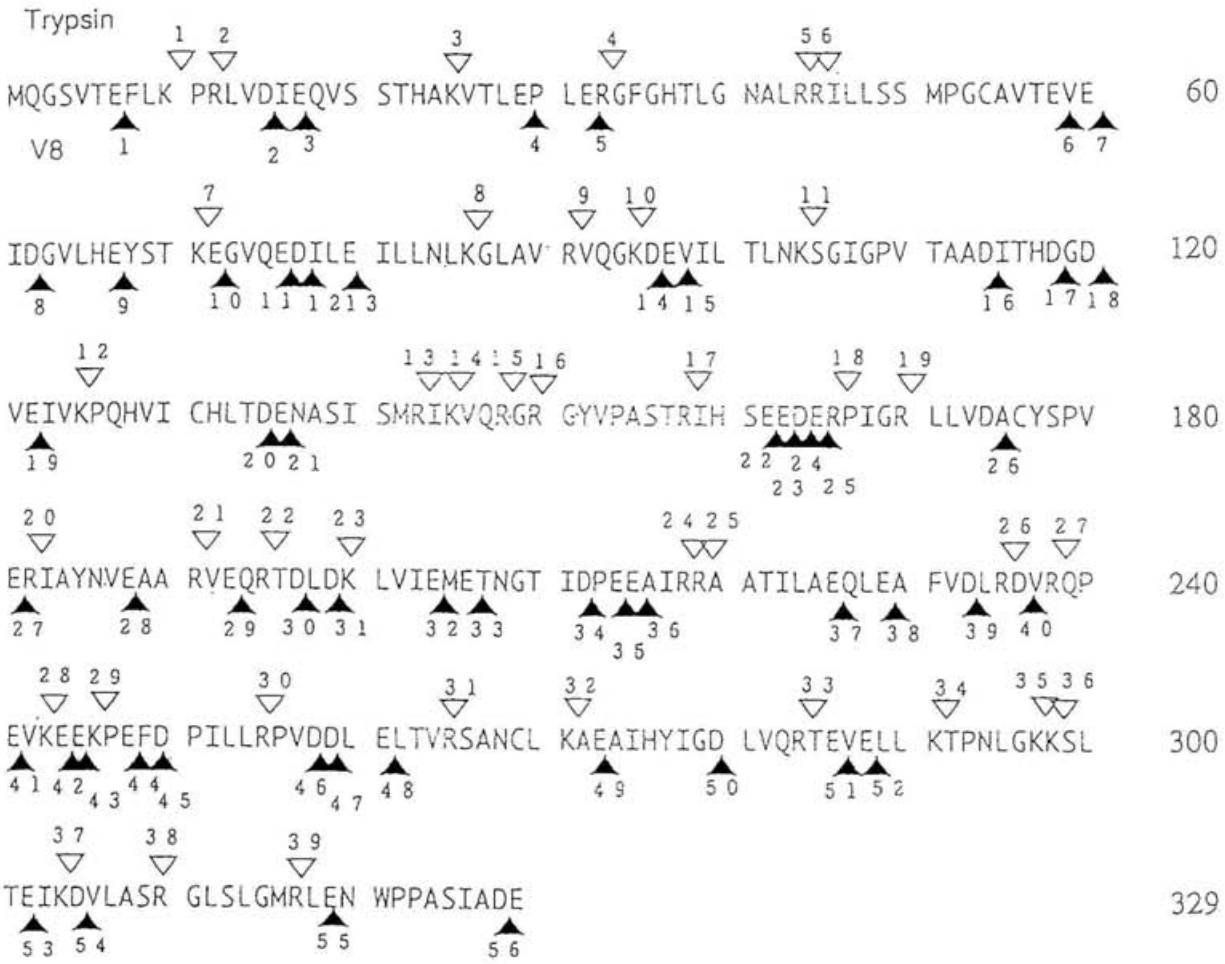


Figure 3. Potential proteolytic cleavage sites of RNA polymerase α subunit. A total of 39 cleavage sites for trypsin and 56 sites for V8 proteinase are indicated by open and closed triangle, respectively.

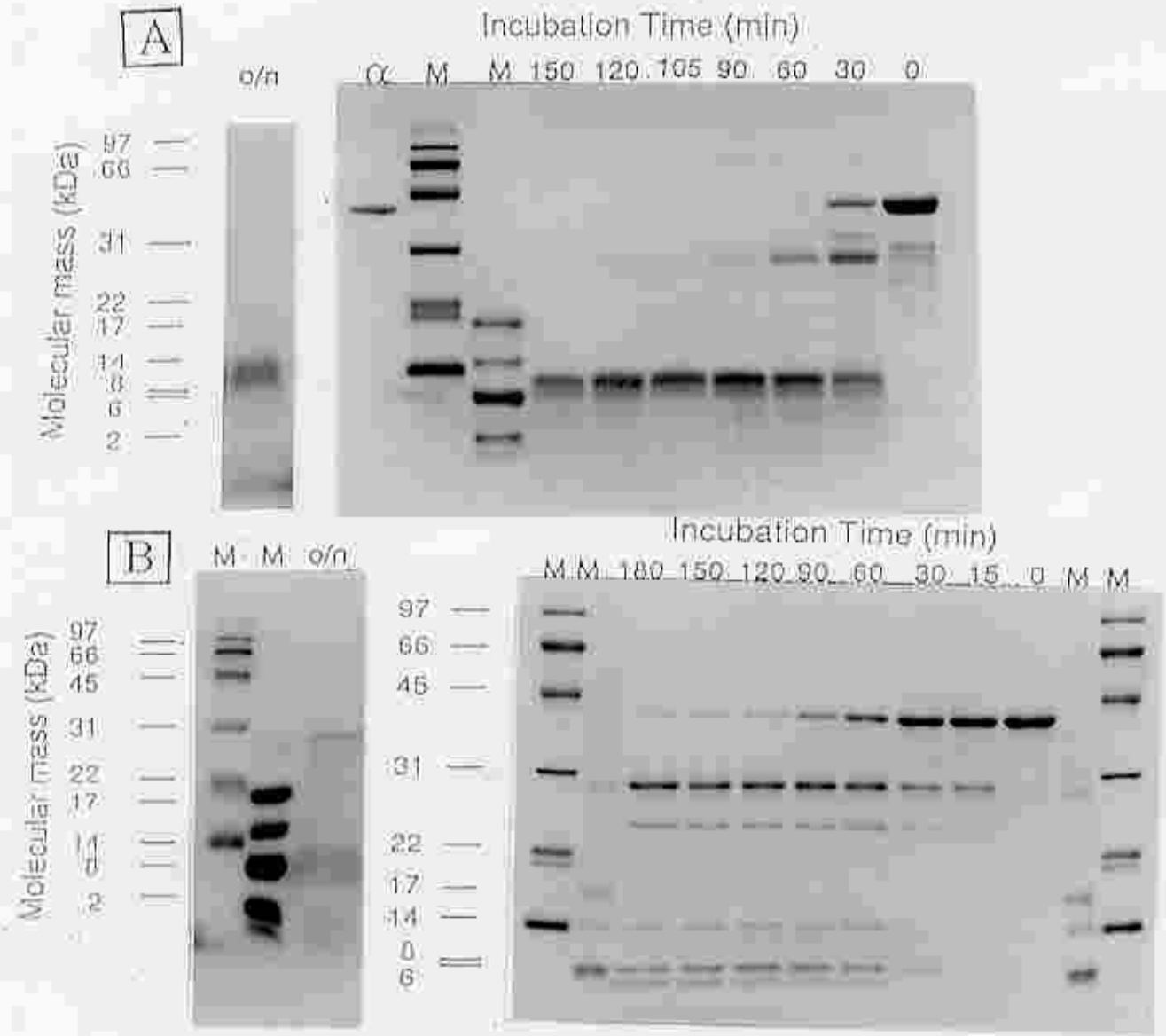


Figure 4. Proteolytic cleavage of RNA polymerase α subunit. [A] α subunit (5 μg/μl) was incubated with trypsin (50 ng/μl) at 37 °C and at the times indicated, analyzed by SDS-PAGE. Gels were stained with CBB. O/N shows the gel pattern of overnight incubation. [B] α subunit (5 μg/μl) was incubated with V8 protease (50 ng/μl) at 37 °C.

A Trypsin cleavage

Sample		Content	N-terminal Sequence	C-terminal Sequence	Fragment
31kDa		100%	^{1}M-Q-G-S-V-T-E-F-		T(N-26)
28kDa		100%	^{1}M-Q-G-S-V-T-E-F-		T(N-24/25)
11kDa					
	A	51.4%	^{46}I-L-L-S-S-M-P-G-	-V-X-R-G-R^{150}	T(6-16)
	B	42.0%	^{236}D-V-R-Q-P-E-V-K-	-P-P-X-S-I-X-D-E^{329}	T(26-C)
	C	5.4%	^{220}A-A-T-I-X-A-X-Q-		T(25-?)
	D	1.2%	^{34}G-F-G-H-T-L-G-N-		T(4-?)
9kDa		100%	^{244}E-E-K-P-E-F-D-P-		T(28-C)

B V8 protease cleavage

Sample		Content	N-terminal Sequence	C-terminal Sequence	Fragment
28kDa					
	A	75%	^{8}F-L-K-P-R-L-V-D-I-		V(1-41)
	B	25%	^{1}M-Q-G-S-V-T-E-F-L-		V(N-41)
24kDa		100%	^{8}F-L-K-P-R-L-V-D-I-		V(1-?)
14.4kDa		100%	^{68}Y-S-T-K-E-G-X-Q-X-		V(9-?)
9kDa		100%	^{242}V-K-E-E-K-P-E-F-D-	-I-X-D-E^{329}	V(41-C)
5.8kDa					
	A	77%	^{242}V-K-E-E-K-P-E-F-D-	-R-X-E-V-E^{288}	V(41-52)
	B	23%	^{245}E-K-P-E-F-D-P-L-X-	-R-X-E-V-E^{288}	V(42-52)
4.2kDa					
	A	45%	^{8}F-L-K-P-R-L-V-D-I-		V(1-?)
	B	55%	^{137}N-A-S-I-S-M-R-I-K-		V(21-?)

Figure 5. Amino acid sequences of proteolytic cleavage fragments of RNA polymerase α subunit. [A] The amino-terminal and carboxy-terminal amino acid sequences were determined for tryptic fragments using 670 pmol (31 kDa), 710 pmol (28 kDa), 1000 pmol (11 kDa) and 100 pmol (9 kDa) fragments. Unidentified amino acid residues are denoted by X. The number at the left and right side of each sequence indicates the amino acid residue in the intact α subunit. Fragments are designated by showing the sites of potential cleavage, shown in Fig. 3 (N and C represents the amino and carboxy terminus, respectively). When an SDS-gel band contains more than one fragments, the percentage of each fragment content in the total mixture is indicated at the left side of the sequence. [B] The N-terminal and C-terminal amino acid sequences were determined for V8 proteolytic fragments using 536 pmol (28 kDa), 67 pmol (24 kDa), 56 pmol (14.4 kDa), 138 pmol (9 kDa), 69 pmol (5.8 kDa), and 50 pmol (4.2 kDa) samples.

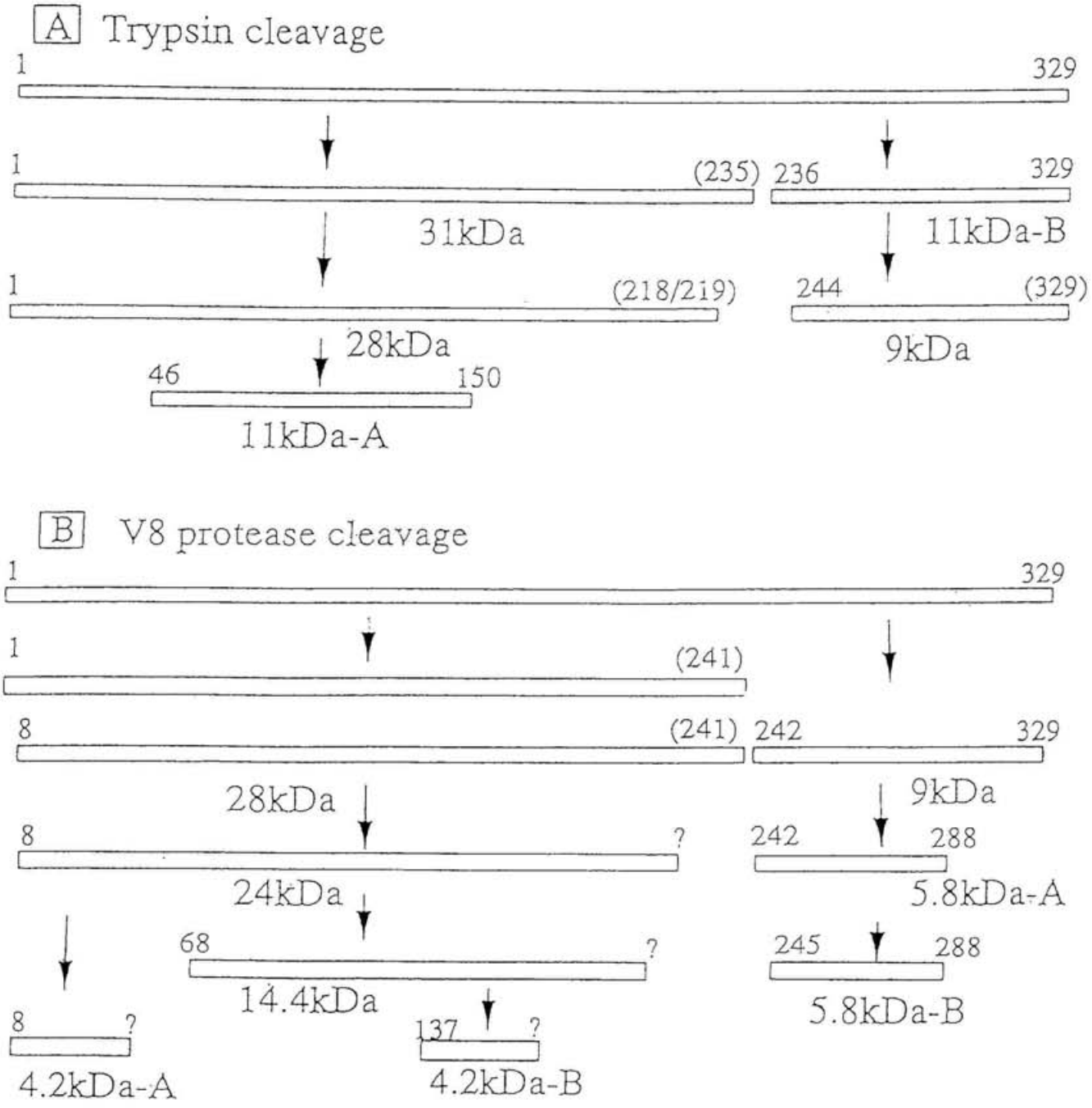


Figure 6. The pathway of proteolytic cleavage of RNA polymerase α subunit. The cleavage sites of α subunit by trypsin [A] and V8 proteinase [B] are shown. The numbers in parentheses represent the cleavage sites estimated from the fragment size and the distribution of potential cleavage sites.

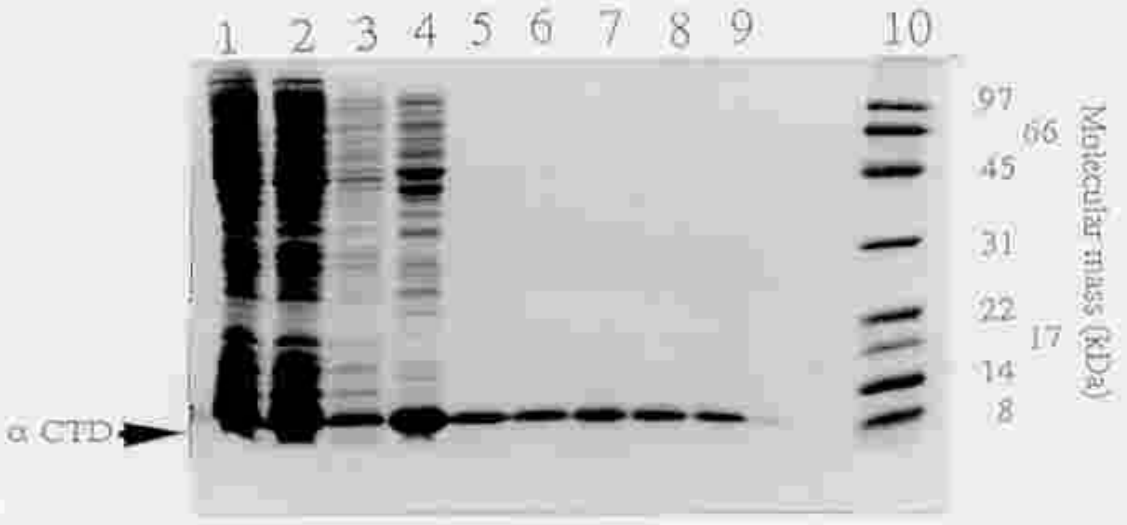


Figure 7. SDS-polyacrylamide gel electrophoresis of fraction from purification of αCTD for NMR analysis. Lane 1, crude extract after sonication; lane 2, supernatant fraction after centrifugation; lane 3, inclusion body fraction after centrifugation; lane 4, DEAE-TOYOPEAL (1.5 x 11cm); lane 5, TOYOPEAL-HW55F ; lane 6-8, Protein PAK G-DEAE (8.2 x 75 mm); lane 9, αCTD (sample for NMR analysis). Each column condition is shown in the text.

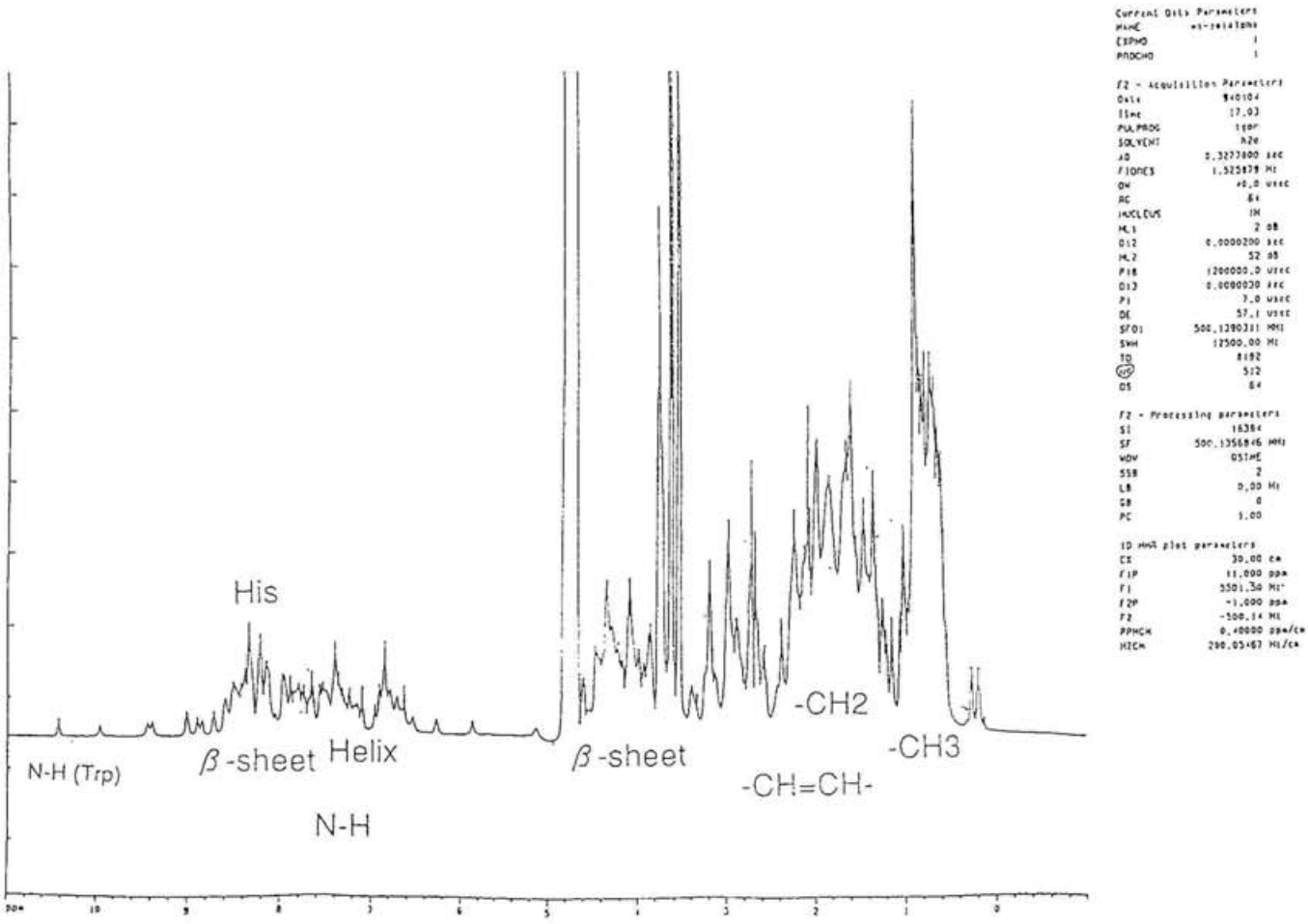


Figure 8. ^{1}H NMR spectra (500 MHz) of α CTD in H_2O/D_2O (90%/10%, v/v: pH 5.5) at 30 °C. PH was adjusted to 5.5 with 5mM phosphate buffer containing 20 mM KCl and 1 mM DTT. Sample preparation method is shown in the text.

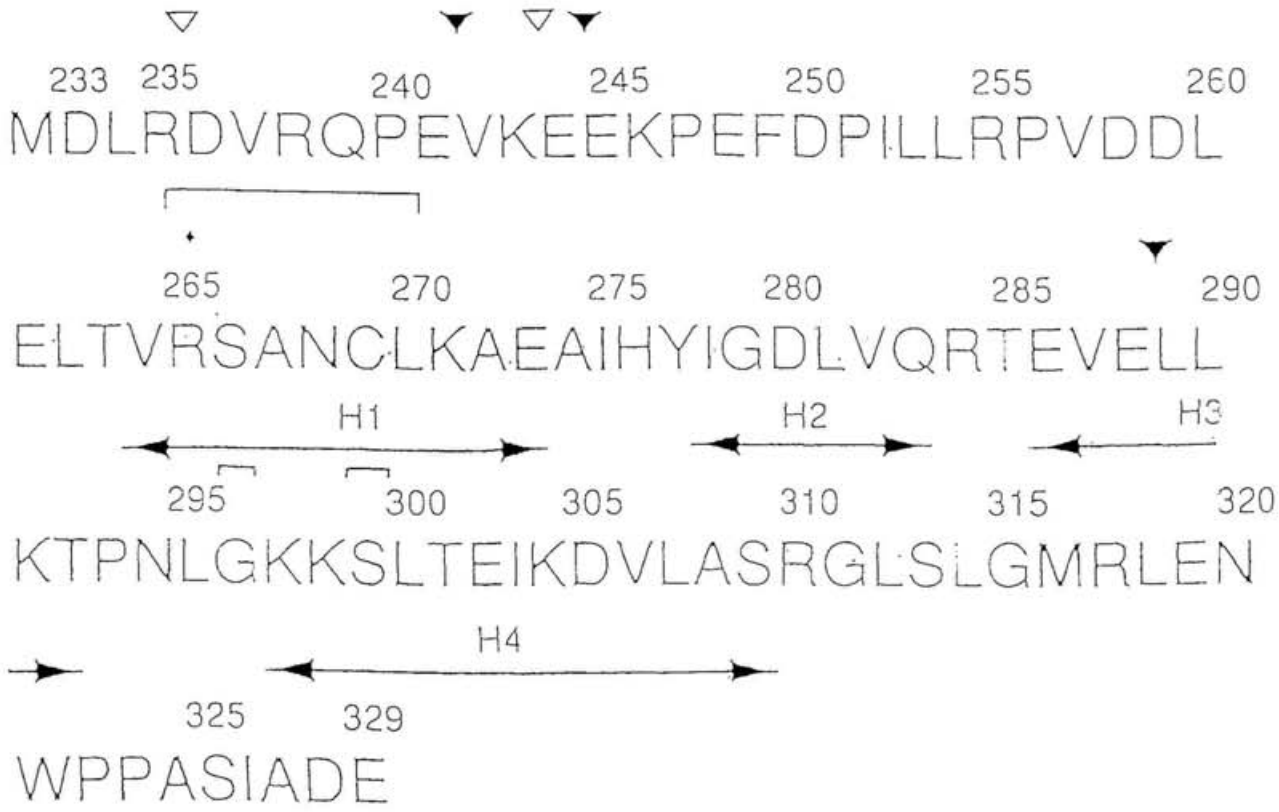


Figure 9. Secondary structure of the αCTD. The sequence of the α CTD is shown in one-letter code. The location of the helices is indicated by arrows. The cleavage sites of the αCTD with trypsin and V8 protease are shown by open and filled triangles, respectively. The bracket indicates the CRP contact site, 265-270, 296 and 299. Astarisk indicates the most important residues, Arg265.

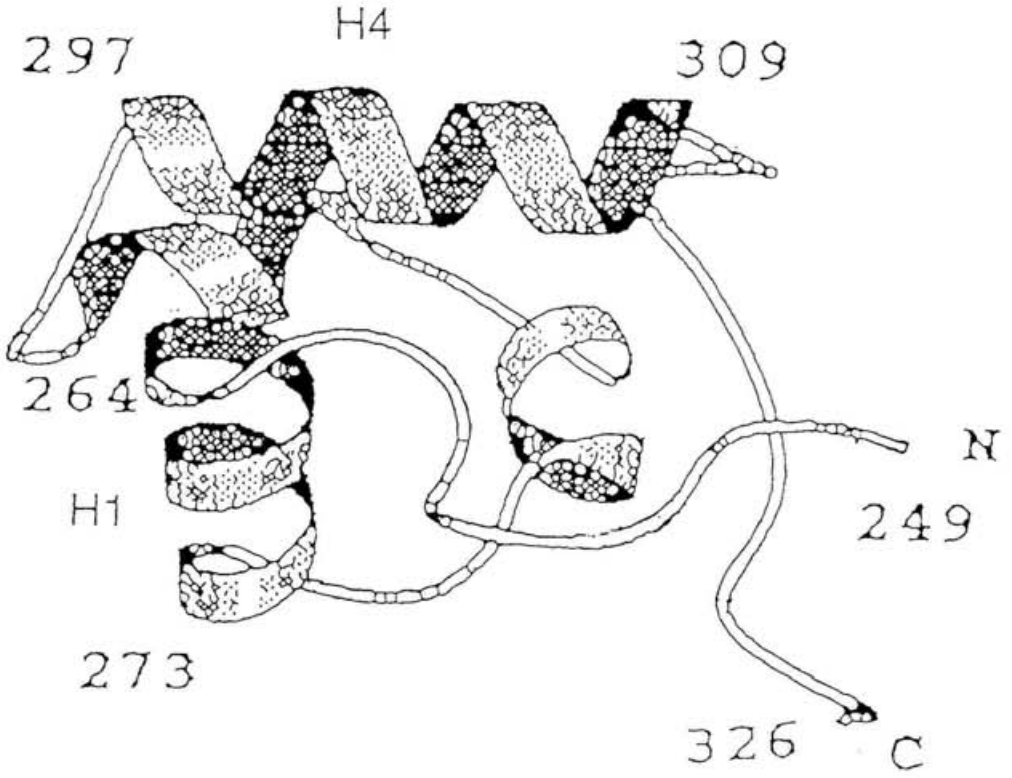


Figure 10. Ribbon diagram representing the folding αCTD, from Phe249 to Ile326. Numbers refer to the first and last residues in the helix1 and helix4. The structure of αCTD was determined by multidimensional heteronuclear magnetic resonance spectroscopy. The root-mean-square deviation (RMSD) for the backborn heavy atoms of 50 structures from Phe249 to Ile326 is 0.67Å.

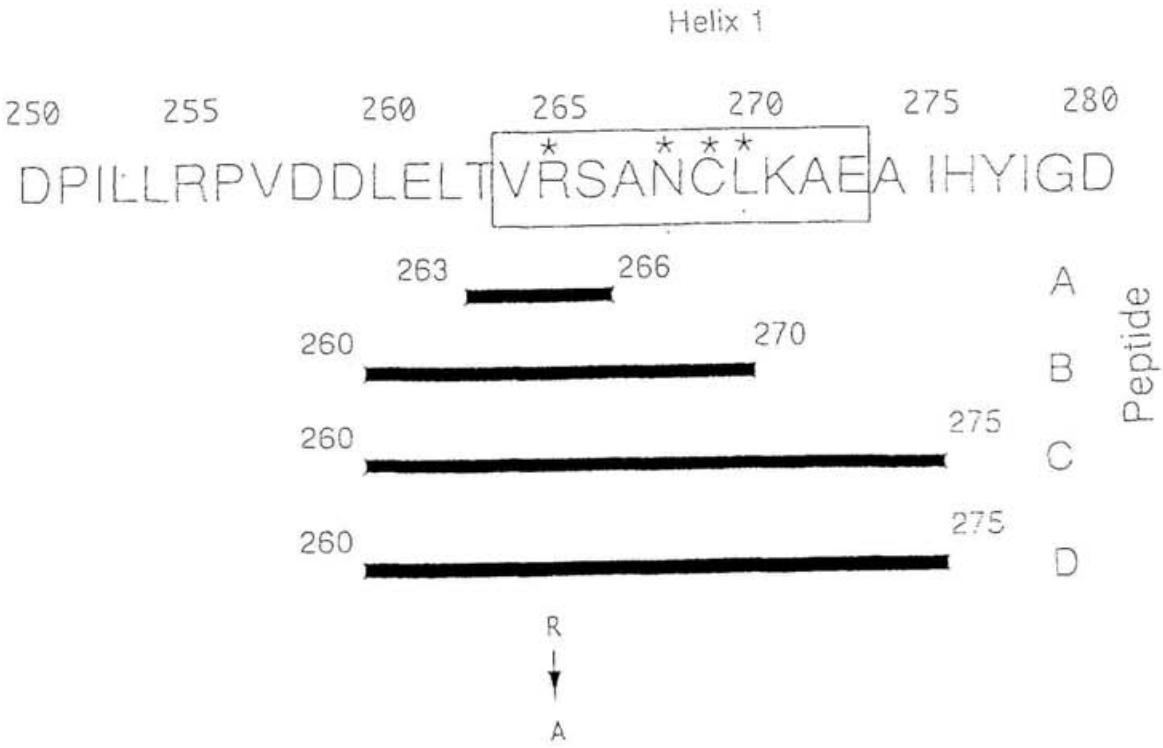


Figure 11. Preparation of alpha peptides. Four species of the alpha peptides (A to D) were used in this study, each covering the indicated sequence. Helix 1 consists of a sequence from Val264 to Glu273. Box means helix 1, and asterisks indicate important residues for contact with CRP. CRP contact site is between aa residue Arg265 to Leu270. Preparation methods are in the text.

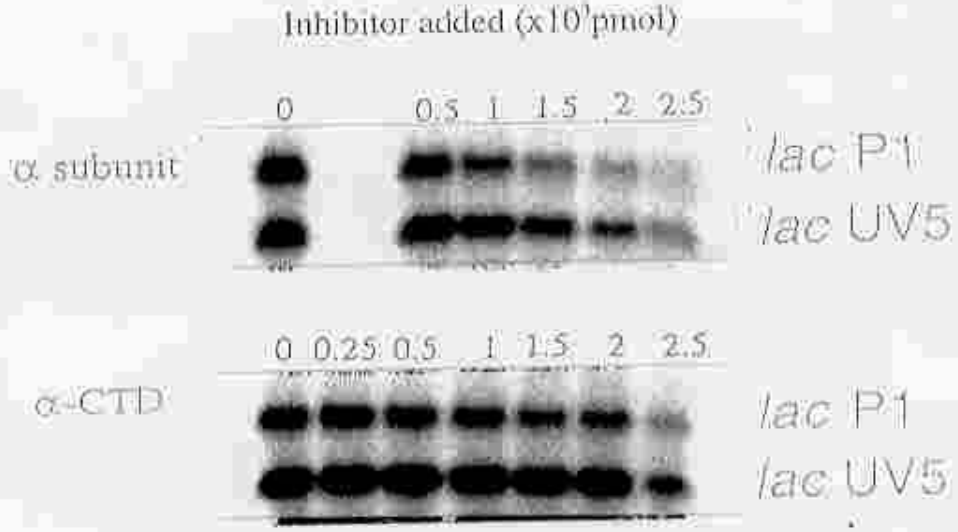


Figure 12. Inhibition of *lac* transcription by α subunit and αCTD. Template DNA (*lac*P1, 0.1 pmol; *lac*UV5, 0.02 pmol) and α subunit or αCTD were mixed for 3 min at 37°C, and after addition of CRP (2.5 pmol) and cAMP (10μM), incubated for another 3 min at 37°C. After addition of $E\sigma^{70}$ holoenzyme (1pmol), preincubation was carried out for 5 min at 37°C. Finally a substrate mixture containing heparin (final 200μ/ml) was added and RNA synthesis was allowed for 5 min at 37°C. Labelled transcripts were analyzed by electrophoresis on polyacrylamide gels containing 8 M urea. The gel was exposed to imaging plate and analyzed with BAS-2000 image analyzer (Fuji).

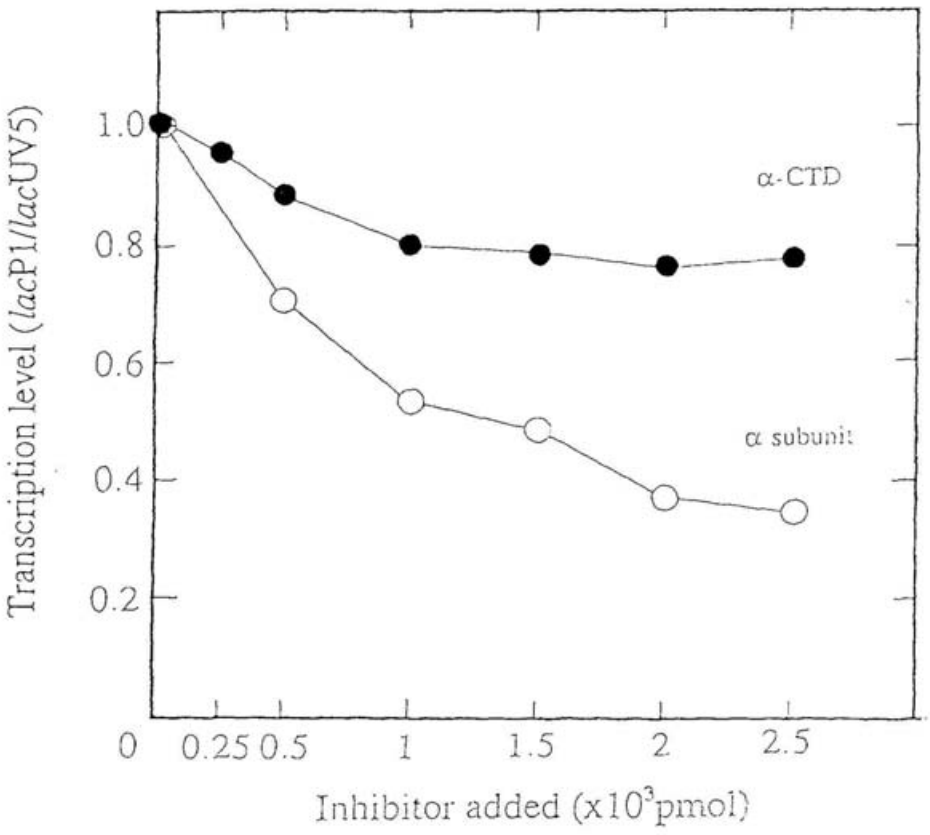


Figure 13. Inhibition of *lac* transcription by α subunit and αCTD. Gel patterns are shown in Fig. 12. Details of the reaction conditions are shown in the text and legend for Fig. 12.

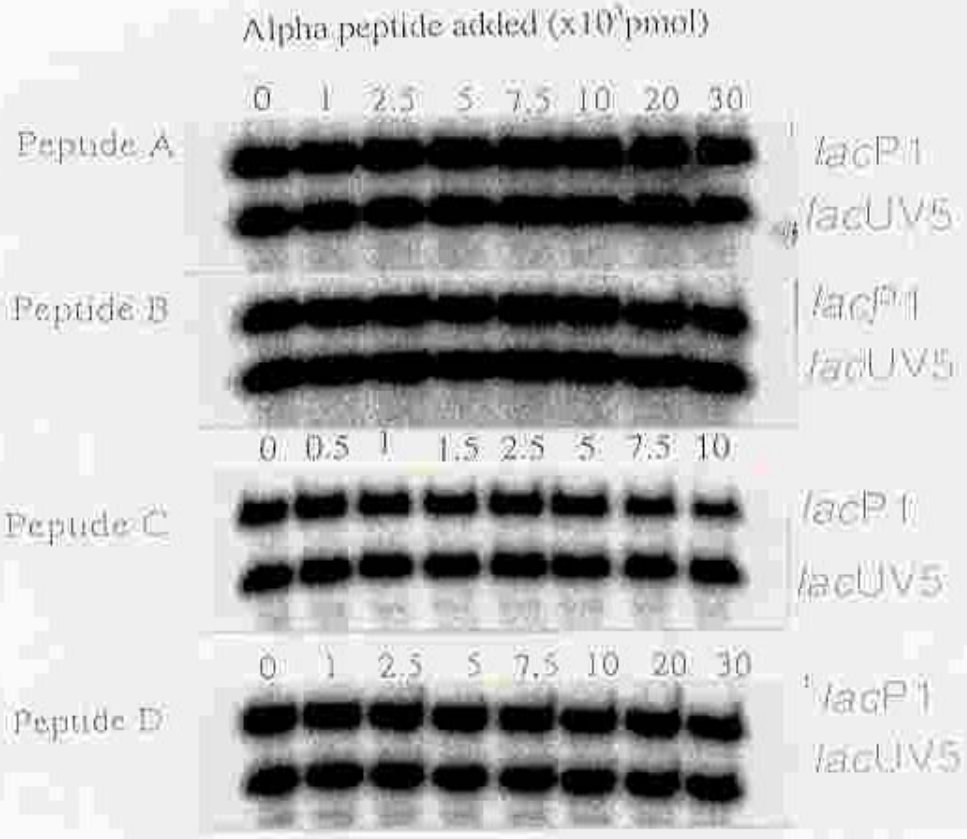


Figure 14. Inhibition of *lac* transcription by alpha peptides. Template DNA (*lac*P1, 0.1 pmol; *lac*UV5, 0.02 pmol) and alpha peptide (0.5 to 30 x 10^3 pmol) were mixed for 3 min at 37°C, and after addition of CRP (2.5 pmol) and cAMP (10μM), incubated for another 3 min at 37°C. After addition of $E\sigma^{70}$ holoenzyme (1pmol), preincubation was carried out for 5 min at 37°C. Finally a substrate mixture containing heparin (final 200μ/ml) was added and RNA synthesis was allowed for 5 min at 37°C. Labelled transcripts were analyzed by electrophoresis on polyacrylamide gels containing 8M urea. The gel was exposed to imaging plate and analyzed with BAS-2000 image analyzer (Fuji).

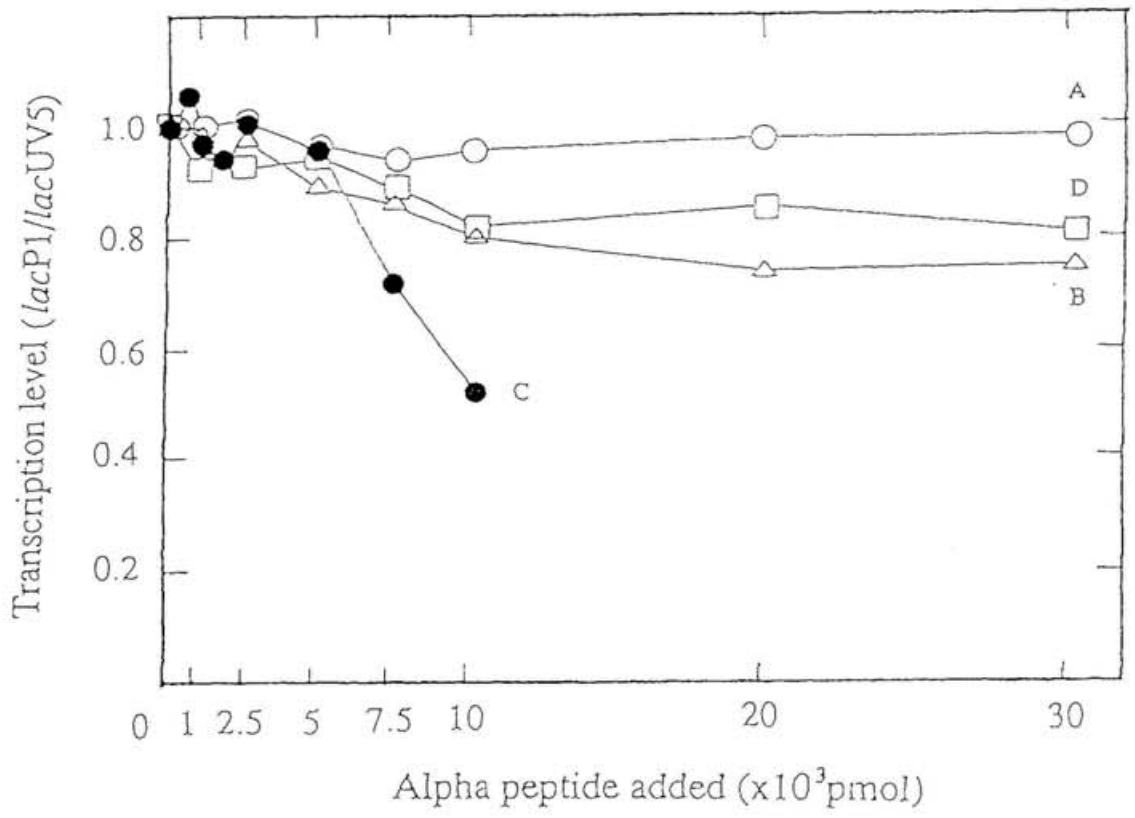


Figure 15. Inhibition of *lac* transcription by alpha peptides.
Gel patterns are shown in Fig. 14. Details of the reaction conditions are shown in the text and legend for Fig. 14.

8. Acknowledgements

I wish to thank Prof. Akira Ishihama, for his supports, guidances and discussions. I thank Dr. Nobuyuki Fujita, Dr. Masahiro Yamagishi, Dr. Tetsuya Toyoda, Dr. Hiroshi Mitsuzawa and Dr. Makoto Kimura for advices and discussions. I also thank Dr. Hideaki Tagami and Prof.Hiroji Aiba (Nagoya University) for preparation of CRP. I thank Dr. Young Ho Jeon, Dr. Masahiro Shirakawa, Dr. Toshio Yamazaki and Prof. Masayoshi Kyogoku (Osaka University) for NMR analysis of αCTD. Finally, I thank all members in the Department of Molecular Genetics, National Institute of Genetics.