J. Mol. Biol. (1996) **263**, 637–647

JMB

Domain Organization of the *Escherichia coli* RNA Polymerase σ^{70} Subunit

Elena Severinova[1], Konstantin Severinov[1], David Fenyö[2] Michael Marr[3], Edward N. Brody[4], Jeffrey W. Roberts[3], Brian T. Chait[2] and Seth A. Darst[1]*

[1]*Laboratory of Molecular Biophysics and* [2]*Laboratory of Mass Spectrometry and Gaseous Ion Chemistry The Rockefeller University 1230 York Avenue New York, NY 10021, USA*

[3]*Section of Biochemistry Molecular, and Cell Biology Cornell University, Ithaca NY 14853, USA*

[4]*SUNY Buffalo, Department of Biological Sciences Cooke Hall, Buffalo NY 14260, USA*

**Corresponding author*

We used limited trypsin digestion to determine the domain organization of the *Escherichia coli* RNA polymerase σ^{70} subunit. Trypsin-resistant fragments containing σ^{70} conserved region 2 (σ^{70}_2), and carboxy-terminal fragments containing conserved regions 3 and 4 (σ^{70}_{3-4}) were identified by a combination of amino acid sequencing and mass spectrometry. The domains were studied for partial biochemical functions of $\sigma^{70}\cdot\sigma^{70}_2$ bound core RNA polymerase competitively with intact σ^{70}. In contrast to σ^{70}_2 alone, the RNA polymerase holoenzyme formed with σ^{70}_2 specifically bound a single-stranded DNA oligomer with a sequence corresponding to the non-template strand of the –10 promoter element (the Pribnow box). σ^{70}_2 also forms crystals that are suitable for X-ray analysis. σ^{70}_{3-4} bound the T4 AsiA protein with high affinity. The epitope for T4 AsiA on σ^{70} was further localized to within σ^{70}[551–608], comprising σ conserved region 4.2.

Keywords: sigma factor; *E. coli* RNA polymerase; AsiA; transcription; limited proteolysis

Introduction

Escherichia coli core RNA polymerase (RNAP), containing two α subunits, one β, and one β' subunit (with molecular masses of 36.5, 151, and 155 kDa respectively), is fully active in RNA polymerization but is incapable of promoter recognition and specific initiation (Burgess *et al*., 1987). Specific initiation of transcription is dependent upon binding of σ factors to the core RNAP to form the RNAP holoenzyme. Different σ factors promote transcription from different sets of promoters (Helmann & Chamberlin, 1988). It has been proposed that in the holoenzyme, σ factors are responsible for the specific interactions with promoter sequences (Losick & Pero, 1981).

The predominant σ factor in *E. coli*, σ^{70} (70.2 kDa), is also the best studied. The possibility that σ^{70} comprises independent domains is suggested by the following observations. (1) Sequence comparisons among a large number of σ factors from diverse bacteria show that the σ^{70} family of proteins comprises four highly homologous regions which are separated by regions of variable length and sequence (Helmann & Chamberlin, 1988; Lonetto *et al*., 1992). (2) Genetic studies indicate that part of σ^{70} conserved region 2 interacts with the –10 consensus region (the Pribnow box) of the promoter (Siegele *et al*., 1989; Waldburger *et al*., 1990), while part of the C-terminal conserved region 4 recognizes the –35 consensus (Gardella *et al*., 1989; Waldburger *et al*., 1990). However, isolated σ^{70} in solution does not bind DNA. It has been proposed that the DNA binding activity of σ^{70} is blocked by an inhibitory interaction between the N-terminal conserved region 1 and the C-terminal region 4, and that this inhibition is relieved upon formation of holoenzyme (Dombroski *et al*., 1992, 1993). This model suggests a substantial degree of conformational mobility in the σ^{70} molecule, which could be

Permanent address: E. Severinova, Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, Russia.

Abbreviations used: RNAP, RNA polymerase; MALDI-MS, matrix-assisted laser desorption/ionization mass spectrometry; TPCK, L-1-tosylamido-2-phenylethyl chloromethylketone; IPTG, isopropyl-β-D-thiogalactopyranoside; PMSF, phenylmethylsulfonyl fluoride; NTA, nickel-nitrilotriacetic acid.

0022–2836/96/450637–11 $25.00/0

realized by structurally independent domains connected by flexible linkers. (3) High resolution structural studies of σ factors have been hampered by the inability to obtain suitable crystals despite considerable effort. This is consistent with σ^{70} comprising independent, mobile domains connected by flexible linkers, which could lead to a structure recalcitrant to crystallization.

Limited proteolysis has often been used to define the domain organization of proteins (Wilson, 1991). Early studies of σ^{70} identified a fragment resistant to trypsin with an apparent mobility by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) corresponding to about 40 kDa (Lowe *et al*., 1979; Gribskov & Burgess, 1983). More recently, the structural and functional domain organization of other σ factors has been studied in more detail (Chang & Doi, 1990; Cannon *et al*., 1995; Chen & Helmann, 1995).

We have probed the domain architecture of *E. coli* σ^{70} using limited trypsin digestion. Trypsin-resistant fragments were subjected to N-terminal sequencing and/or matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS; Hillenkamp *et al*., 1991) to identify the N and C termini within the σ^{70} sequence. Domains containing conserved region 2 and conserved regions 3 and 4 together were identified and studied for partial biochemical functions of σ^{70}. The domain containing σ^{70} conserved region 2 binds core RNAP competitively with intact σ^{70} and the holoenzyme thus formed specifically binds a single-stranded DNA oligomer with a sequence corresponding to the non-template strand of the –10 promoter consensus region. This domain also forms crystals that are suitable for X-ray analysis. The domain containing conserved regions 3 and 4 binds the T4 AsiA protein with high affinity.

Results

Identification of σ^{70} structural domains by limited trypsin digestion

Trypsin degradation of σ^{70} proceeds in a highly ordered manner (Severinov *et al*., 1994). SDS/PAGE of purified σ^{70} treated with increasing amounts of trypsin revealed the appearance of discrete bands that were stable over a wide range of trypsin concentrations (Figure 1). We used a combination of N-terminal–sequencing and MALDI-MS, along with a consideration of the cleavage specificity of trypsin (C-terminal of Arg or Lys residues), to identify the products of σ^{70} trypsinolysis precisely. The data that led to the identification of the σ^{70} fragments are tabulated in Table 1. An overall summary is illustrated in Figure 2.

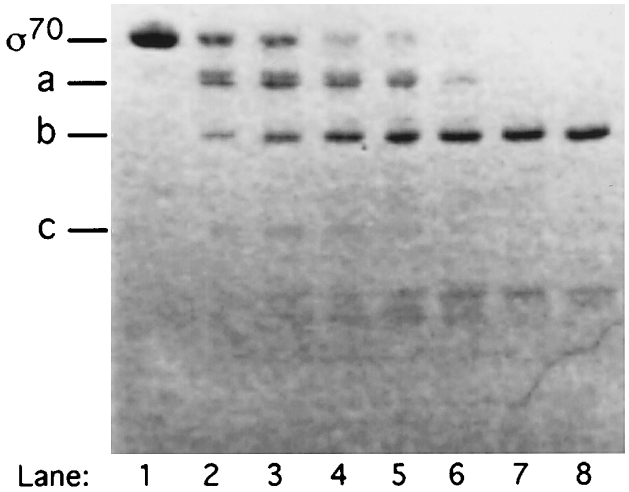


Figure 1. Limited trypsinolysis of σ^{70} analyzed by SDS–PAGE on an 8% to 25% gradient PhastGel (Pharmacia). Reactions contained 60 pmol of σ^{70} and 0, 0.15, 0.19, 0.3, 0.4, 0.6, 1.2, or 3 pmol of trypsin (lanes 1 to 8, respectively).

A primary site of trypsin attack is between Arg448 and Thr449, near the C-terminal end of conserved region 2.4, which comprises residues 435 to 456 (Lonetto *et al*., 1992). This results in the appearance of bands a and c (Figure 1). Band a was identified as σ^{70} residues 1 to 448, containing conserved region l and all but the eight C-terminal residues of conserved region 2. For convenience, we will refer to this fragment as $\sigma^{70}_{1\text{-}2}$, where the subscripts refer to the conserved regions contained within the fragment. Band c ($\sigma^{70}_{3\text{-}4}$) contains completely conserved regions 3 and 4.

$\sigma^{70}_{1\text{-}2}$ (band a) was degraded eventually to an exceptionally trypsin-resistant fragment with an apparent mobility of 40 kDa (band b, Figure 1) that corresponds to the trypsin-resistant fragment observed by Gribskov & Burgess (1983). This fragment was biochemically isolated and shown by

Table 1. Identification of σ^{70} tryptic fragments

Band[a]	Protein	N-terminal sequence	Observed mass (Da)	σ^{70} residues	Calculated mass (Da)
σ^{70}	σ^{70}	MEQNPQSQLK	70,340 (±120)	1–613	70,263
a	$\sigma^{70}_{1\text{-}2}$	MEQNPQS	51,270 (±60)	1–448	51,257
b	σ^{70}_{2a}	—	39,780 (±30)	104–448[b,c]	39,759
b	σ^{70}_{2b}	—	38,640 (±30)	114–448[b,c]	38,629
c	$\sigma^{70}_{3.4}$	XIRIPVHMIE	18,452 (±12)	449–608	18,446
d	σ^{70}_{4}	IAKEPISMET	12,355 (±7)	500–608	12,350

[a] As labeled in Figures 1 and 2.
[b] C terminus confirmed by sub-digestion with CNBr.
[c] N terminus confirmed by sub-digestion with endoproteinase Glu-C.

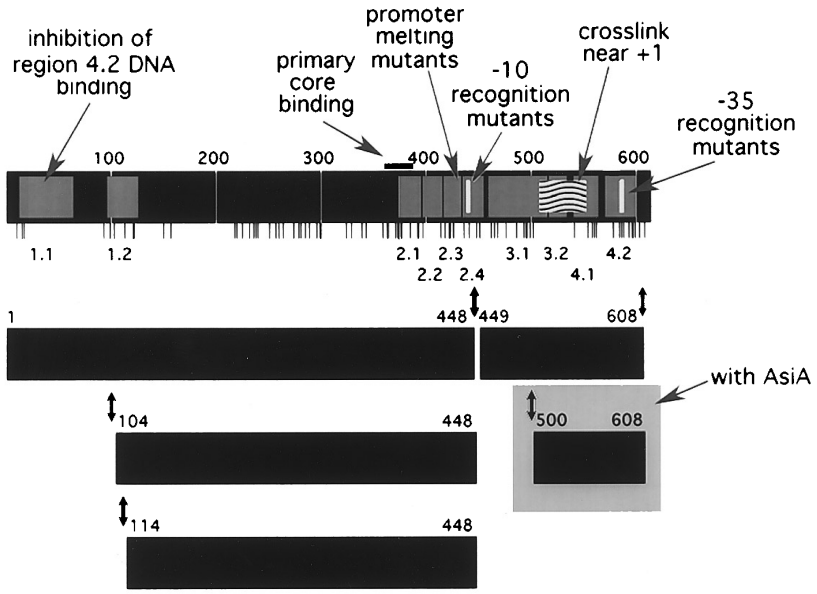


Figure 2. Schematic illustration of the σ^{70} primary structure (top) and the trypsinolysis results (beneath). The horizontal black bar at the top represents the σ^{70} primary sequence with amino acid numbering shown above the bar. Evolutionarily conserved regions are shaded grey and numbered (below the bar) according to Lonetto *et al*. (1992). The hash marks below the bar denote potential trypsin cleavage sites. Functional regions defined by genetic or biochemical studies are labeled (see text for details and references). The major trypsin resistant fragments are schematically illustrated underneath.

MALDI-MS to be a mixture of two polypeptides, which we will refer to as σ^{70}_{2a} and σ^{70}_{2b} (Table 1).

A complex between σ^{70} and the bacteriophage T4 anti-σ^{70} protein (AsiA; Stevens, 1977; Orsini *et al*., 1993) was also subjected to trypsin treatment (Figure 3). We reasoned that the interaction between σ^{70} and AsiA might alter the degradation pathway in an informative way. In the presence of AsiA, the overall dependence of σ^{70} degradation on trypsin concentration is only slightly altered (note the persistence of intact σ^{70} in the presence of AsiA). However, two σ^{70} fragments, $\sigma^{70}_{3\text{-}4}$ and its degradation product σ^{70}_{4}, were dramatically stabilized in the presence of AsiA (bands c and d, Figure 3).

Functional studies of σ^{70} domains

The σ^{70} domains (σ^{70}_{2a}, σ^{70}_{2b}, $\sigma^{70}_{3\text{-}4}$, and σ^{70}_{4}), identified above based on trypsin resistance despite the presence of numerous trypsin cleavage sites within them (Figure), were studied for partial biochemical functions of σ^{70} itself. Models of σ^{70} function in binding to core RNAP (Lesley & Burgess, 1989), promoter recognition (Gardella *et al*., 1989; Siegele *et al*., 1989; Waldburger *et al*., 1990), and promoter melting (Helmann & Chamberlin, 1988; Juang & Helmann, 1994; Waldburger & Susskind, 1994) have emerged from sequence analysis, genetic, and biochemical studies (summarized in Figure 2). Functions for regions 1.2 or 3 have not been proposed but crosslinking studies have established that a fragment of σ^{70} containing conserved region 3.2 must be within about 5 Å of the γ-phosphate of the priming nucleotide (Severinov *et al*., 1994). Because it was difficult to obtain sufficient amounts of the σ^{70} fragments in purified form from trypsinolysis of intact σ^{70}, we sub-cloned these fragments into the pET-15b (Novagen) overexpression system (Studier *et al*., 1990). The final products were easily purified by virtue of N-terminal His_6-tags, which were subsequently removed by thrombin digestion, resulting in the σ^{70} fragment and four N-terminal residues (GSHM) remaining from the vector (see Materials and Methods). In no case in our functional studies described below did we observe effects from the His_6-tags or from the four vector residues, although we sometimes observed steric effects due to immobilization of His_6-σ^{70} fragments on Ni^{2+}-(NTA) agarose beads, as described below. For both σ^{70}_{2a} and σ^{70}_{2b}, about 20% of the total overexpressed protein was found in, and purified from, the soluble fraction of the cell lysates. Both $\sigma^{70}_{3\text{-}4}$ and σ^{70}_{4} were renatured from inclusion bodies. With these purified, recombinant fragments we performed functional tests to investigate the role of the different σ^{70} domains in binding to core RNAP, to the T4 AsiA protein, and to promoter elements. Since in many of our investigations σ^{70}_{2a} and σ^{70}_{2b}

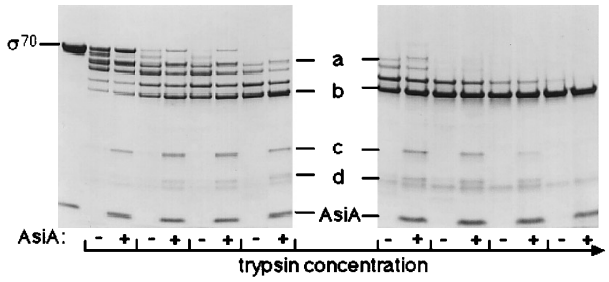


Figure 3. Limited trypsinolysis of σ^{70} in the presence and absence of T4 AsiA, analyzed by SDS-PAGE on a 10% Tris-tricine gel (Novex). Reactions contained 100 pmol σ^{70}. Trypsin concentrations were 0 (leftmost lane) and (for each pair of lanes from left to right) 0.01, 0.02, 0.025, 0.034, 0.04, 0.063, 0.1, 0.15 pmol. At each trypsin concentration, identical pairs of reactions were performed with 0 or 120 pmol AsiA as indicated at the bottom. The leftmost lane also contains 120 pmol AsiA.

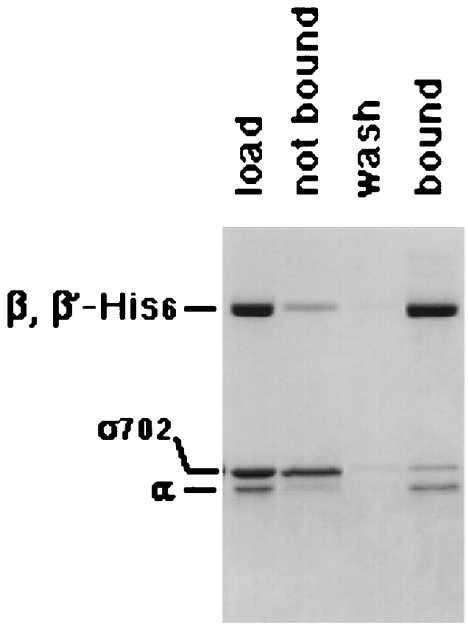


Figure 4. Binding of σ_2^{70} to immobilized His_6-core RNAP. His_6-core RNAP ($\alpha 2\beta\beta'$ with a His_6-tag at the C terminus of β') and a molar excess of σ_{2b}^{70} (load) were incubated with Ni^{2+}-NTA agarose beads in buffer containing no imidazole. The beads were then washed with buffer containing 0.5 mM imidazole (not bound), then with buffer containing 10 mM imidazole (wash), then eluted with buffer containing 100 mM imidazole (bound). The presence of σ_2^{70} in the bound fraction in amounts comparable to α (which is present at two copies/core molecule) indicates strong binding. In separate control experiments, σ_2^{70} was not bound to the beads in the absence of His_6-core RNAP.

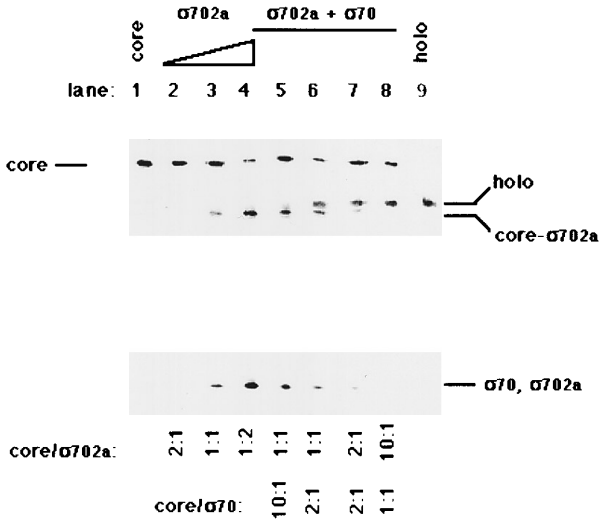


Figure 5. Competitive binding to core RNAP by σ_2^{70} and intact σ^{70}. Core RNAP (15 pmol in each reaction) was incubated with the indicated molar ratios of σ^{70} and σ_{2a}^{70} and then analyzed by native PAGE on a 4% to 15% gradient PhastGel (Pharmacia).

behaved essentially identically except when noted below, we will use the term σ_2^{70} to denote both of the fragments.

Functional studies of σ_2^{70}

σ_2^{70} binds to core RNA polymerase and competes with intact σ^{70}

The trypsin-resistant fragments σ_{2a}^{70} and σ_{2b}^{70} contain within them the primary determinant of core RNAP binding (Lesley & Burgess, 1989), σ^{70}[361-390] (Figure 2). We used three different assays, native gel-shift analysis, Ni^{2+}-NTA agarose co-immobilization, and gel filtration, to investigate the binding of these fragments to core RNAP.

σ_2^{70} was able to bind His_6-core RNAP (carrying a His_6-tag at the C terminus of the β' subunit; Kashlev *et al*., 1993) immobilized on Ni^{2+}-NTA agarose beads, resulting in co-immobilization of the σ^{70} fragment (Figure 4). However, immobilization of His_6-σ_2^{70} did not result in detectable co-immobilization of wild-type core RNAP. Based on native gel shift analysis, His_6-σ_2^{70} bound core RNAP in solution with similar affinity to σ_2^{70} (data not shown), suggesting that the failure of immobilized His_6-σ_2^{70} to bind core RNAP was due to steric hindrance from the close proximity and/or orientation of the N terminus of the σ^{70} fragment to the agarose beads.

The binding of σ_2^{70} to core RNAP was competitive with intact σ^{70}. This was most clearly demonstrated using native PAGE (Figure 5). As σ_2^{70} was added to core RNAP, the band corresponding to core RNAP decreased in intensity and a new band with mobility slightly higher than that of RNAP holoenzyme appeared (Figure 5, lanes 2 to 4). When increasing amounts of intact σ^{70} were added, the band corresponding to the core RNAP-σ_2^{70} complex disappeared and a complex with an apparent mobility of σ^{70}-holoenzyme emerged (Figure 5, lanes 5 to 8), indicating that σ_2^{70} and intact σ^{70} bound competitively to core RNAP, as expected. This interpretation of the native PAGE data was confirmed by an analysis of the protein content of individual bands cut from the native gel by SDS-PAGE (data not shown, see Materials and Methods).

Finally, mixtures of core RNAP and excess amounts of σ_2^{70} were analyzed by gel filtration on a Superose 6 FPLC gel-filtration column. A significant fraction of the σ_2^{70} eluted in the same fractions as the RNAP and was well separated from a later peak containing σ_2^{70} alone (data not shown), further indicating the formation of a core RNAP-σ_2^{70} complex.

σ_2^{70} is unable to promote transcription

σ_2^{70} was not able to promote transcription of DNA fragments containing a strong σ^{70}-holoenzyme promoter (T7A1), nor could it support transcription from a consensus ''extended –10'' promoter (P_{RE}, provided by R. Hayward; Kumar *et al*., 1993; Keilty & Rosenberg, 1987). DNase I footprints of σ_2^{70} alone

or in complex with core RNAP did not reveal any protection of promoter DNA (data not shown). The fragment also did not stimulate transcription of poly[dAdT] by core RNAP (data not shown).

Since σ_2^{70} competes with σ^{70} for core RNAP binding but does not support transcription itself, we performed competition experiments to test for inhibition of σ^{70}-dependent transcription initiation by σ_2^{70}. When σ_2^{70} was added at a 30-fold molar excess over σ^{70}, transcription from phage T7 DNA (Chamberlin *et al*., 1979) was inhibited about 30%. Abortive initiation both from a ''–10/–35'' promoter (T7A2; McClure *et al*., 1978) and from an ''extended –10'' promoter (galP1) was inhibited about 50% by a 100-fold molar excess of σ_2^{70} over σ^{70}.

σ_2^{70}-Holoenzyme specifically recognizes single-stranded DNA containing the non-template –10 consensus

Conserved region 2.4 contains residues implicated by genetic analysis in recognition of the –10 element of promoters (Waldburger *et al*., 1990). The adjacent region 2.3 contains a high proportion of conserved aromatic and basic amino acid residues, leading to the proposal that this region may be involved in promoter opening and single-stranded DNA binding (Helmann & Chamberlin, 1988; Juang & Helmann, 1994; Rong & Helmann, 1994). σ^{70}-Holoenzyme has been shown to specifically recognize the non-template strand bases of the –10 consensus sequence (Ring & Roberts, 1994; Ring *et al*., 1996; Roberts & Roberts, 1996). Correspondingly, *E. coli* holoenzyme shows a specific interaction with single-stranded DNA oligonucleotides containing non-template consensus –10 elements by gel mobility shift assay (M. Marr & J. W. Roberts, unpublished results). The σ_2^{70} fragments reconstituted specific binding when complexed with coreRNAP (Figure 6). Lanes 1 and 2 show σ^{70}-holoenzyme binding to oligo 1 (Figure 6), which contains a consensus –10 element (TATAAT). Core RNAP, either alone or as a contaminant in the holoenzyme preparations, also bound the oligonucleotides at the highest concentration of RNAP, although this binding was non-specific (the band corresponding to core RNAP is of equal intensity in lanes 5 to 7). Lanes 5 to 7 show σ_{2a}^{70}-holoenzyme binding to the consensus element oligonucleotide (oligo 1), an anti-consensus element (oligo 2) containing the least likely nucleotide at each position of the –10 hexamer (Hawley & McClure, 1983), and a point mutation in the –10 hexamer (C-T at –12; oligo 3). σ_{2a}^{70}-Holoenzyme bound both oligos containing consensus elements, although it bound the mutant (oligo 3) less efficiently and did not bind the anti-consensus oligo (oligo 2). These results mimic the binding specificity of σ^{70}-holoenzyme (M. Marr & J. W. Roberts, unpublished results). σ_{2b}^{70}-Holoenzyme bound the oligos with the same specificity as σ_{2a}^{70}-holoenzyme but the level of binding was much weaker (not shown). It is noteworthy that the fragments alone (σ_{2a}^{70} and σ_{2b}^{70}) bound all three oligos weakly with no apparent specificity (band labeled ''σ fragments''), as revealed by silver-staining of the gels. The relevance of this non-specific binding is unclear. We do not know the nature of the highest bands in lanes 5 to 7, which appear to be specific and may be multimers of the core-σ_{2a}^{70} complex.

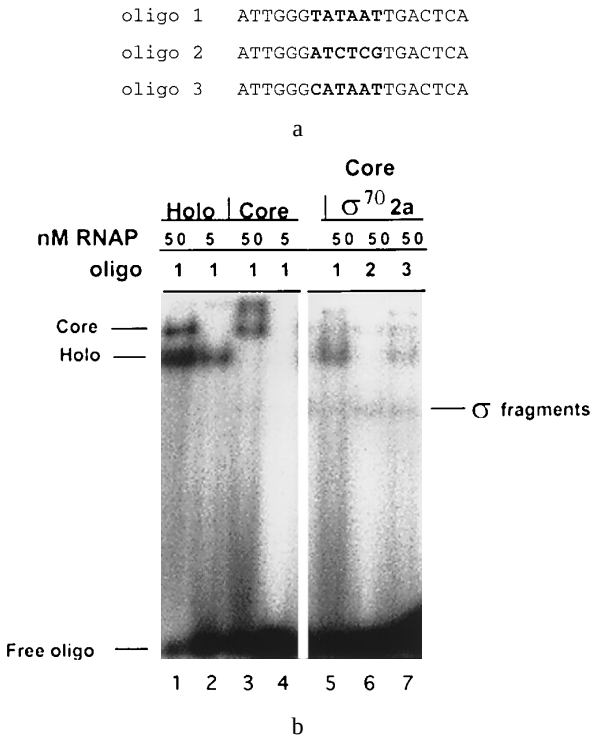


Figure 6. Binding of σ^{70}-holoenzyme, core, and σ_{2a}^{70}-holoenzyme to single-stranded oligonucleotides. a, Single-stranded oligonucleotides used; oligo 1 contains the –10 consensus element; oligo 2 contains the least likely nucleotide at each position of the –10 hexamer (Hawley & McClure, 1983); oligo 3 contains a point mutation (C–T at –12) within the –10 hexamer. b, Gel mobilty shift analysis by non-denaturing 5% PAGE. Lanes 1 and 2, σ^{70}-holoenzyme; lanes 3 and 4, core RNAP; lanes 5 7, σ_{2a}^{70}-holoenzyme.

Crystallization of σ_2^{70}

Possibly one of the most sensitive indicators of the structural integrity and homogeneity of a protein is its crystallization. σ Factors have proven to be stubbornly refractory to crystallization, possibly due to an architecture of relatively independent domains connected by flexible linkers, which is consistent with our proteolysis results. σ_2^{70} is resistant to very high concentrations of trypsin, suggesting a high degree of structural integrity and stability, thus making it an attractive candidate for crystallization.

Several crystal forms of σ_2^{70} prepared from trypsin treated σ^{70} (probably a mixture of σ_{2a}^{70} and σ_{2b}^{70}), as well as pure, recombinant σ_{2a}^{70} and σ_{2b}^{70}, were grown using the method of hanging-drop vapor diffusion.

Table 2. Crystal data

Unit cell (Å)		Resolution	Completeness		R_{merge}
a, *b*	*c*	(Å)	(%)	I/σ_I	(%)
79.093	133.879	2.9	92.6	21.3	5.2
		(2.97–2.9)	86.9	5.0	17.4

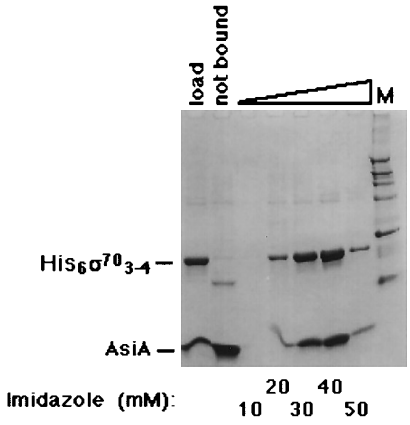
Only one crystal form suitable for X-ray structure determination was obtained. These crystals, of recombinant σ^{70}_{2b}, were grown as described in Materials and Methods. At room temperature, diffraction from the crystals, as well as the physical appearance of the crystals, degraded severely within a few hours, making it impossible to collect a data set. A native data set (Table 2) was collected from a cryo-protected, frozen crystal at –180°C. The crystals are tetragonal and systematic absences in the diffraction data indicate they belong to the space group $P4_12_12$ or $P4_32_12$. Using a protein density of 1.3 g/cm^3 and the unit-cell volume of 840,000 Å^3, it is clear that there is one protein molecule per asymmetric unit. The solvent content then is 53%. The calculated cell volume per unit mass, V_M, is then 2.7 Å^3, which is within the range found for other protein crystals (Matthews, 1968).

Functional studies of σ^{70}_{3-4} and σ^{70}_4

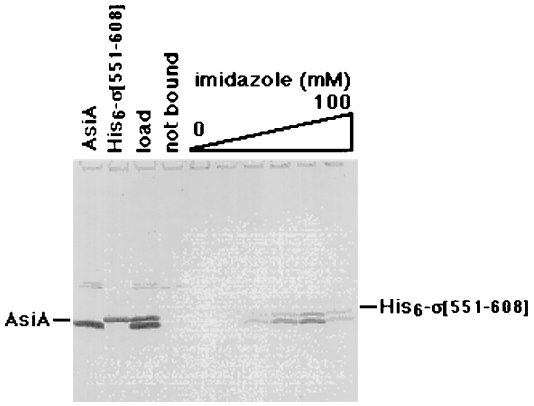
σ^{70}_{3-4} and σ^{70}_4 bind the T4 AsiA protein

The proteolysis experiments (Figure 3) demonstrated that the carboxy-terminal fragments of σ^{70}, σ^{70}_{3-4} and σ^{70}_4, were markedly stabilized in the presence of the T4 AsiA protein. This stabilization is likely due to an interaction of AsiA with the carboxy-terminal region of σ^{70}. To investigate this more directly, the interaction of AsiA with recombinant His$_6$-σ^{70}_{3-4} (Figure 7a) and His$_6$-σ^{70}_4 was monitored by affinity chromatography on Ni^{2+}-NTA agarose. Separate control experiments showed that AsiA did not bind the beads by itself. Based on this assay, both fragments interacted strongly with AsiA. σ^{70}_2 did not interact with AsiA in either Ni^{2+}-NTA agarose affinity chromatography assays or by native PAGE analysis (data not shown). These results are further supported by the following experiment. Complete reaction mixtures from trypsin digestion experiments were mixed with T4 AsiA and then loaded onto an HQ FPLC column (PerSeptive Biosystems) at low NaCl. The bound proteins were then eluted with a NaCl gradient. The various σ^{70} fragments were well resolved over the gradient, but only σ^{70}_{3-4} and σ^{70}_4 co-eluted with AsiA.

In order to further localize the AsiA epitope within σ^{70}, we also investigated the binding of AsiA to a smaller recombinant σ^{70} fragment, His$_6$-σ^{70}[551-608], comprising essentially conserved region 4.2, also by affinity chromatography on Ni^{2+}-NTA agarose (Figure 7b). The elution profile of both proteins from the Ni^{2+}-NTA agarose beads was identical. These results localize the binding epitope for AsiA to within σ^{70} residues 551 to 608.



a



b

Figure 7. a, Binding of T4 AsiA to immobilized His$_6$-σ^{70}_{3-4}. His$_6$-σ^{70}_{3-4} and a molar excess of AsiA (load) were incubated with Ni^{2+}-NTA agarose beads in buffer containing no imidazole. The beads were then washed with buffer containing 0.5 mM imidazole (not bound). Subsequent elutions were then performed with buffer containing the indicated imidazole concentrations. The presence of AsiA in the elution fractions, along with the similarity of the elution profiles of the two proteins, indicates strong binding. In separate control experiments, AsiA alone was not bound to the beads. b, Binding of T4 AsiA to immobilized His$_6$-σ^{70}[551-608]. The experiment was performed as in a.

Discussion

Despite the central importance of σ factors in the control of bacterial gene expression, a fundamental understanding of their mechanism of action, their regulation, and their role in such processes as promoter melting and promoter clearance, is lacking. This is due, in large part, to an almost total lack of structural information. A number of observations, including our inability to crystallize σ^{70}, have led us to assume that σ^{70} is composed of independent domains connected by linkers, leading to a structure recalcitrant to crystallization. We have used limited proteolysis with trypsin, com-

bined with MALDI-MS and N-terminal sequencing, to identify these domains within the σ^{70} sequence for further structural and functional studies. To assess whether the trypsin-resistant fragments identified were, in fact, independently folded domains, the fragments were isolated and investigated for partial functions expected of σ^{70} itself. Furthermore, one of the trypsin-resistant fragments crystallized in a form suitable for X-ray analysis, indicating that it adopts a compact, folded structure.

The assumption that σ^{70} is composed of independent, compactly folded domains (which are relatively resistant to protease cleavage) connected by flexible linkers (which are generally sensitive to protease cleavage) is supported by the observation that σ^{70} degrades into smaller fragments that are relatively stable to further digestion despite containing numerous potential cleavage sites (Figure 2). This is most apparent with σ^{70}_{2b}, which appears to be almost completely resistant to further trypsin cleavage. Moreover, the trypsin-resistant σ^{70} fragments exhibit partial functions expected of σ^{70} itself.

The primary site of trypsin attack of σ^{70} (between residues 448 and 449) is near the C-terminal end of conserved region 2. Other sensitive sites are located near the N-terminal end of conserved region 1.2 (between 103 and 104) and, in the presence of AsiA, near the boundary between conserved regions 3.1 and 3.2 (between 499 and 500). The domain structure of *Bacillus subtilis* σ^{D} was recently investigated by limited proteolysis. Two regions of preferential protease attack were identified, one near the C-terminal end of conserved region 2 and the other in the boundary between conserved regions 3.1 and 3.2 (Chen & Helmann, 1995). Since *B. subtilis* σ^{D} does not contain the N-terminal conserved region l.l, the trypsin-sensitive site found in σ^{70} at the N-terminal end of conserved region 1.2 would not be present. Thus, the primary sites of protease sensitivity observed in the two proteins are essentially the same, suggesting similar structures, which is expected from the high degree of sequence conservation within the σ^{70} family of proteins (Lonetto *et al*., 1992).

Our results establish that σ^{70} comprises at least two structurally independent and functionally distinct domains, σ^{70}_2 and $\sigma^{70}_{3\text{-}4}$. The σ^{70}_2 domain is remarkably resistant to trypsin degradation and is able to bind core RNAP and compete with σ^{70}. Deletion analysis has identified a region of σ^{70}, σ^{70}[361-390], contained within σ^{70}_2, that is necessary and sufficient for core RNAP binding (Lesley & Burgess, 1989). Other regions of σ^{70} must contribute directly or indirectly to core RNAP binding, however. For instance, a C-terminal truncation mutant of σ^{70}, σ^{70}[1-574], which lacks essentially region 4.2, binds core RNAP with reduced affinity (Kumar *et al*., 1993, 1994). A mutant *E. coli* σ^{32} with a deletion in region 3 (the deletion corresponding to σ^{70} residues 503 to 520 in the aligned sequences) appears to function normally except that it exhibits a reduced affinity for core RNAP (Zhou *et al*., 1992). Also, the model of Dombroski *et al*. implies that region 1.1 binds to a site on core RNAP (Dombroski *et al*., 1992, 1993). Thus, all of the σ^{70} conserved regions may contribute to core RNAP binding affinity. These results are consistent with our finding that σ^{70}_2 binds core RNAP with an affinity about 30-fold less than σ^{70}. However, the C-terminal σ^{70} fragments ($\sigma^{70}_{3\text{-}4}$ and σ^{70}_4) do not bind core RNAP with an affinity sufficient to be observed by any of the methods employed in this study.

The necessity of a σ factor to impart specific promoter recognition to core RNAP, along with the observation that different σ factors confer specificity for different sets of promoters, led to the proposal that σ factors directly contact promoter DNA, providing specific recognition of conserved promoter elements (Losick & Pero, 1981). Support for this hypothesis came from the finding that, in promoter complexes with *E. coli* RNAP holoenzyme, only the σ^{70} subunit could be crosslinked to both the –10 and –35 promoter elements (Chenchick *et al*., 1981, 1982). Stronger support came from genetic studies, which have demonstrated allele-specific suppression of promoter mutations by specific mutations in the corresponding σ factor (Gardella *et al*., 1989; Siegele *et al*., 1989; Zuber *et al*., 1989; Daniels *et al*., 1990; Waldburger *et al*., 1990). These studies identified σ conserved regions 2.4 and 4.2 as specifying recognition of the –10 and –35 promoter elements, respectively. Only recently, however, have specific interactions between conserved regions of σ^{70} and promoter DNA been indicated by competitive filter retention assays on N-terminally truncated σ^{70} derivatives (Dombroski *et al*., 1992, 1993). This led to the hypothesis that the latent DNA binding activity of σ^{70} is inhibited by interaction with the N-terminal conserved region 1.1.

All of the sites of promoter recognition mutations found in region 2.4 are contained within σ^{70}_2. While we have been unable to demonstrate a specific interaction of σ^{70}_2 by itself with single or double-stranded DNA fragments containing a –10 promoter element, σ^{70}_2-holoenzyme interacts specifically with a single-stranded DNA fragment containing the non-template sequence of the –10 promoter element. This finding leads to the speculation that, in the open complex between RNAP holoenzyme and promoter DNA, σ region 2 functions in part as a sequence-specific single-strand DNA binding protein, which has been suggested earlier based on sequence similarity between σ conserved region 2.3 and single-strand DNA binding proteins (Helmann & Chamberlin, 1988; Juang & Helmann, 1994). Sequence-specific binding of the non-template strand would stabilize the transcription bubble in the open promoter complex and leave the template strand available for the RNAP catalytic machinery. The specific interaction of DNA with σ^{70}_2 alone is either of insufficient affinity to observe by the methods used, or residues N-terminal of conserved region 2 (σ^{70}[114-374] in σ^{70}_{2b}) are inhibitory to DNA binding

and this inhibition is relieved by the interaction with core RNAP, as suggested by other studies (Dombroski *et al*., 1992, 1993).

The AsiA protein of bacteriophage T4 is a 10 kDa protein that binds tightly to σ^{70} and inhibits σ^{70}-holoenzyme transcription (Stevens, 1977; Orsini *et al*., 1993). The T4 AsiA protein binds tightly to both $\sigma^{70}_{3\text{-}4}$, σ^{70}_{4}, and to σ^{70}[551-608], but not to other σ^{70} fragments. The epitope for T4 AsiA on σ^{70} must thus be contained within σ^{70}[551-608], comprising essentially σ conserved region 4.2. This leads to two conclusions. First, the binding of these σ^{70} fragments containing region 4.2 to T4 AsiA is a biochemical function expected of σ^{70} itself, suggesting that $\sigma^{70}_{3\text{-}4}$, σ^{70}_{4}, and σ^{70}[551-608] are properly folded in solution and can, in that sense, be called domains. Second, the binding of T4 AsiA to σ^{70} conserved region 4.2 suggests that the inhibition of σ^{70}-holoenzyme transcription by T4 AsiA occurs through the inhibition of the interaction between σ^{70} region 4.2 and the –35 promoter consensus element, and not by inhibition of the σ^{70}-core RNAP interaction, as has been suggested (Orsini *et al*., 1993). It will be interesting to investigate whether T4 AsiA inhibits σ^{70}-holoenzyme transcription from extended –10 promoters (Keilty & Rosenberg, 1987), which do not require σ^{70} conserved region 4 for activity (Kumar *et al*., 1993).

It has been shown that the N-terminal 529 residues of σ^{70} (σ^{70}[1-529]) was able to support transcription from an extended –10 promoter (Kumar *et al*., 1993), indicating that σ conserved region 4 is not essential for specific transcription initiation. Moreover, σ conserved region 1.1 is present only in primary σs and in *E. coli* σ^{S} (Mulvey & Loewen, 1989; Lonetto *et al*., 1992) and so region 1.1 is not essential for specific transcription initiation by a σ factor. Thus, the inability of σ^{70}_{2} to support transcription from –10/–35 promoters, extended –10 promoters, or even poly[dAdT], indirectly suggests that σ^{70} conserved region 3 is required for this function of σ factors.

The σ^{70}_{2} domain contains a part of conserved region 1.2, and all but the C-terminal eight residues of conserved region 2, which is the most highly conserved region in the σ^{70} family (Lonetto *et al*., 1992). Conserved region 2 has been implicated in interactions with core RNAP (Lesley & Burgess, 1989), binding to the –10 promoter element, and promoter melting (Helmann & Chamberlin, 1988; Juang & Helmann, 1994; Rong & Helmann, 1994). The structural analysis of this fragment, made possible by the crystals described herein, will be an important step towards understanding the detailed role of σ factors in directing bacterial transcription initiation.

Materials and Methods

Proteins

Recombinant σ^{70} was purified from an overexpression strain as described (Gribskov & Burgess, 1983). RNAP core enzyme was purified from *E. coli* MRE600 as described (Polyakov *et al*., 1995). His_6-tagged core RNAP was purified using the same procedure from *E. coli* RL324, which contains an altered chromosomal *rpoC* gene coding for His_6 at the C terminus of the RNA polymerase β' subunit (provided by R. L. Landick).

T4 AsiA protein was purified as described (Orsini *et al*., 1993) except a final purification step was introduced to prepare the protein to greater than 95% homogeneity. AsiA (6 mg) was loaded onto a phenyl-Toyopearl column equilibrated with 50 mM Tris-HCl (pH 8.0), 50 mM NaCl, 1.65 M $(NH_4)_2SO_4$, 1 mM β-mercaptoethanol, and 4% ethylene glycol. The column was washed, then eluted with an $(NH_4)_2SO_4$ gradient from 1.65 to 0 M. Homogeneous AsiA protein eluted from the column at approximately 50 mM $(NH_4)_2SO_4$.

Limited digestion with trypsin

Trypsin digestion reactions contained 20 mM Tris-HCl (pH 7.9), 50 mM NaCl, 5% glycerol, 0.1 mM EDTA, 1 mM DTT, 100 pmol σ^{70}, and 0.01 to 0.14 pmol L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK) treated trypsin (Sigma) in 14 µl. Reactions were allowed to proceed for 30 minutes at 25°C and stopped by addition of Laemmli loading buffer and immediate boiling. Reaction products were analyzed on 10% Tris-tricine gels (Novex). The σ^{70}-AsiA complex was formed by incubating 100 pmol σ^{70} with 120 pmol AsiA for ten minutes at 37°C prior to the addition of trypsin.

Matrix-assisted laser desorption mass spectrometry

Matrix-assisted laser desorption mass spectra (Hillenkamp *et al*., 1991) of the tryptic fragments of σ^{70} were collected using a time-of-flight mass spectrometer constructed at the Rockefeller University (Beavis & Chait, 1989). The tryptic fragments were mixed with α-cyano-4-hydroxycinnamic acid (10 g/l in formic acid/water/isopropanol (1:3:2, by vol.)) to obtain a final protein concentration of 2 pmol/µl. An aliquot (0.5 µl) was placed on the mass spectrometer probe tip and air-dried. The sample was irradiated with 10 ns duration laser pulses (355 nm wavelength) from a Nd(YAG) laser. The resulting ions were accelerated in an electrostatic field and their time-of-flight was measured with a LeCroy 8828D transient digitizer. The observed masses are listed in Table 1. To obtain additional confirmation of the identification of σ^{70}_{2a} and σ^{70}_{2b}, the purified fragments were digested with an excess of CNBr in 0.1 M HCl for one hour or with endoproteinase Glu-C in 50 mM NH_4HCO_3 for two hours (1:50, enzyme:protein ratio). The digestion reactions were subsequently lyophilized, redissolved in α-cyano-4-hydroxycinnamic acid, and mass analyzed as described.

Cloning and purification of σ^{70} fragments

The σ^{70} fragments were sub-cloned from pMRG8 (Gribskov & Burgess, 1983) by PCR into the pET15b expression plasmid (Novagen) and the final structures confirmed by DNA sequencing. The constructs were then transformed into *E. coli* BL21(DE3) cells. Transformants were grown in LB with ampicillin (100 µg/ml) to an A_{600} of 0.7 to 0.9 and expression was induced by the addition of isopropyl-β,D-thiogalactopyranoside (IPTG) to 1 mM.

After three hours, cells were harvested by centrifugation (5000 ***g***) and resuspended in Buffer A (20 mM sodium phosphate (pH 8), 0.5 M NaCl, 0.5 mM β-mercaptoethanol, 5% (v/v) glycerol) along with a mixture of protease inhibitors (174 mg/l phenylmethylsulfonyl fluoride (PMSF), 312 mg/l benzamidine, 10 mg/l aprotinin, 5 mg/l chymostatin, 5 mg/l leupeptin, and 1 mg/l pepstatin A). Lysates were prepared in a French press and clarified by centrifugation.

For σ^{70}_{2a} and σ^{70}_{2b} purification, the soluble fraction of the lysates was loaded onto a Poros MC20 column (PerSeptive Biosystems) loaded with Ni^{2+}. The column was washed with Buffer A + 20 mM imidazole. Bound proteins were eluted with Buffer A + 100 mM imidazole. For removal of the N-terminal His_6-tag, the proteins were digested with thrombin (Boehringer Mannheim) at a thrombin:protein molar ratio of about 1:1000. Digestion was performed in Buffer B (50 mM Tris-HCl (pH 8.4), 0.5 mM β-mercaptoethanol, and 5% glycerol) with 150 mM NaCl and 2.5 mM $CaCl_2$, at room temperature for five hours. The proteins were then diluted twofold with Buffer B and loaded onto a Poros HQ20 column (PerSeptive Biosystems). The σ^{70} fragments eluted at about 300 mM NaCl over a gradient from 0 to 500 mM. In cases where the N-terminal His_6-tag was removed by thrombin digestion, the fragments were loaded onto the Poros MC20-Ni^{2+} column a second time and collected in the flow-through. This effected the removal of trace contaminants that bound and eluted with the σ^{70} fragments through the first pass. The fragments were concentrated to about 25 mg/ml using a Centricon filter (Amicon) and dialyzed into storage buffer (20 mM Tris-HCl (pH 7.9), 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 5% glycerol).

For σ^{70}_{3-4} and σ^{70}_{4} purification, pelleted inclusion bodies were suspended in Buffer A + 6 M guanidine-HCl and loaded onto the Poros MC20-Ni^{2+} column. The column was washed with Buffer A + 6 M guanidine-HCl + 25 mM imidazole, then eluted with Buffer A + 6 M guanidine-HCl + 125 mM imidazole. The eluted protein was diluted to less than 0.3 mg/ml with Buffer A + 6 M guanidine-HCl and dialyzed against renaturation buffer (25 mM Tris-HCl (pH 8.0), 200 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, 5 mM β-mercaptoethanol, and 20% glycerol). The remaining steps of purification (thrombin digestion to remove N-terminal His_6-tags, ion exchange chromatography, and second MC20-Ni^{2+} chromatography) were performed as described above. Finally, the proteins were dialyzed into storage buffer and stored at less than 0.4 mg/ml concentration to prevent precipitation.

σ^{70}_{2b} Crystallization

Crystals of σ^{70}_{2b} were grown by the hanging drop vapor diffusion method. 1.5 to 2 µl of protein (25 to 30 mg/ml) was mixed with an equal volume of crystallization solution (50 mM sodium acetate (pH 5.1), 0.5 to 0.7 M Li_2SO_4, 5 to 7% (w/v) polyethylene glycol 8000, 5 mM DTT) and incubated over the same solution at 15°C. Rectangular, brick-shaped crystals, up to 0.8 cm × 0.4 cm × 0.4 mm in dimensions, grew in one to two weeks. The crystals were very sensitive to changes in their environment (temperature, solution conditions) and were easily cracked. To prepare the crystals for flash-freezing, they were equilibrated at room temperature for several hours, then sequentially transferred in eight steps from crystallization solution to a final solution of 50 mM sodium acetate (pH 5.1), 0.5 M Li_2SO_4, 15% (w/v) polyethylene glycol 8000, 17.5% glycerol, and 10 mM DTT with incubation for at least one hour at room temperature between each step. After the final equilibration, the crystals were frozen in a loop of ophthalmological suture material (10 to 0 Ethilon, Ethicon) by plunging into liquid ethane at liquid nitrogen temperature. After a few minutes, the ethane froze in a solid block around the crystal and the frozen crystals could be stored indefinitely under liquid nitrogen. Diffraction data were collected on a Rigaku RAXIS-II imaging plate area detector equipped with mirror optics.

Core binding and AsiA binding experiments

Ni^{2+}-NTA agarose binding assays

Standard RNAP core binding reactions contained binding buffer (20 mM Tris-HCl (pH 7.9), 125 mM NaCl, 0.5 mM imidazole, 5% glycerol, 0.5 mM β-mercaptoethanol), 20 pmol core RNAP, and 50 to 200 pmol σ^{70} or σ^{70} fragments in 100 µl. Reactions were preincubated for 30 minutes at 30°C. The binding reactions were added to 40 µl of Ni^{2+}-NTA agarose beads (Qiagen; previously equilibrated with 1 ml of binding buffer) and incubated for 30 minutes at 4°C with gentle mixing of the beads. The beads were pelleted by brief centrifugation, and the unbound material in the supernatant was withdrawn. The beads were then washed twice with 500 µl of binding buffer, then with 500 µl of binding buffer + 10 mM imidazole, then eluted with 100 µl binding buffer + 100 mM imidazole. The protein samples were precipitated with 7% trichloroacetic acid and analyzed by SDS-PAGE.

Native polyacrylamide gel electrophoresis binding assay

Core RNAP (15 pmol) was mixed with σ^{70} or σ^{70} fragments in 10 µl of loading buffer (40 mM Tris-HCl (pH 8.0), 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, and 10% glycerol) and incubated for 20 minutes at 30°C. An amount of each reaction (0.5 to 1 µl) was loaded on a 4% to 15% (w/v) gradient PhastGel (Pharmacia) and run for 150 Vh at 15°C using buffer strips for native electrophoresis. To identify the protein content of bands observed in the native gel, 90 µg of core RNAP was radioactively labeled by phosphorylation using [γ-^{32}P]ATP and calf heart protein kinase (Sigma) at 30°C for 30 minutes. *E. coli* RNAP is efficiently phosphorylated at an intrinsic site and the phosphorylation does not appear to affect RNAP activity (V. Markovtzev & A. Goldfarb, personal communication). The labeled core RNAP was washed and concentrated using a Centricon 100 centrifugal filter (Amicon) into loading buffer. Complexes with σ^{70} and its fragments were formed as described above and loaded into the wells of an 80 mm × 80 mm × 1 mm polyacrylamide gel (29.2:0.8, acrylamide:methylene-bisacrylamide) comprising a 4% stacking gel (in 125 mM Tris-HCl (pH 6.8), 5% glycerol) and a 5% running gel (in 380 mM Tris-HCl (pH 8.9), 5% glycerol). The protein bands containing core RNAP were autoradiographed and bands of interest were cut out and soaked in 2 × SDS-loading buffer for 30 minutes at 37°C. The samples were then analyzed by SDS-PAGE and visualized by silver staining.

DNA binding experiments

σ_2^{70}

The σ^{70} fragments containing conserved region 2 were incubated at 172-fold molar excess over core RNAP for ten minutes at 4°C prior to band shift experiments. σ^{70}-Holoenzyme, σ_2^{70}-holoenzyme, or core RNAP were incubated with ^{32}P-end-labeled oligodeoxynucleotides (see Figure 6a) at 25°C for 30 minutes in binding buffer (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 5% glycerol, 100 μg / ml BSA). The samples were then analyzed by 5% PAGE under non-denaturing conditions (0.5 × TBE at 4°C) and exposed on either a phosphoimaging plate or X-ray film for visualization of the complexes. Following autoradiography the gels were silver-stained to identify protein bands.

Acknowledgements

We thank B. Berkowitz, who initiated the proteolysis studies as a summer student in the laboratory. We are extremely grateful to J. Kuriyan and S. K. Burley for the use of their X-ray equipment. We thank R. Landick for providing the *E. coli* RL324 from which His_6-RNAP was obtained, R. Hayward for providing the P_{RE} extended –10 promoter, E. Nudler for plasmid pT7tR2, and S. Busby for the galP1 promoter. K.S. is a postdoctoral fellow of the Jane Coffin Childs Memorial Fund for Medical Research. S.A.D. is a Lucille P. Markey Scholar and a Pew Scholar in the Biomedical Sciences. This work was supported in part by grants from the Lucille P. Markey Charitable Trust, the Irma T. Hirschl Trust, the Human Frontier Science Project, the Pew Foundation, GM53759 from the National Institutes of Health (to S.A.D.), and GM21941 from the National Institutes of Health (to J.W.R.).

References

Beavis, R. C. & Chait, B. T. (1989). Factors affecting the ultraviolet laser desorption of proteins. *Rapid Commun. Mass Spectrom.* **3**, 233–237.

Burgess, R. R., Erickson, B., Gentry, D., Gribskov, M., Hager, D., Lesley, S., Strickland, M. & Thompson, N. (1987). *RNA Polymerase and the Regulation of Transcription* (Reznikoff, W. S., Burgess, R. R., Dahlberg, J. E., Gross, C. A., Record, M. T. Jr & Wickens, M. P. eds), pp. 3–15, Elsevier Science Publishing Corp, New York.

Cannon, W., Missailidis, S., Smith, C., Cottier, A., Austin, S., Moore, M. & Buck, M. (1995). Core RNA polymerase and promoter DNA interactions of purified domains of σ^N: bipartite functions. *J. Mol. Biol.* **248**, 781–803.

Chamberlin, M. J., Nierman, W. C., Wiggs, J. & Neff, N. (1979). A quantitative assay for bacterial RNA polymerases. *J. Biol. Chem.* **254**, 10061–10069.

Chang, B.-Y. & Doi, R. H. (1990). Overproduction, purification, and characterization of *Bacillus subtilis* RNA polymerase σ^A factor. *J. Bacteriol.* **172**, 3257–3263.

Chen, Y. F. & Helmann, J. D. (1995). The *Bacillus subtilis* flagellar regulatory protein σ^D: overproduction, domain analysis and DNA-binding properties. *J. Mol. Biol.* **249**, 743–753.

Chenchick, A., Beabealashvilli, R. & Mirzabekov, A. (1981). Topography of interaction of *Escherichia coli* RNA polymerase subunits with the lacUV5 promoter. *FEBS Letters*, **128**, 46–50.

Chenchick, A., Beabealashvilli, R., Mirzabekov, A. D. & Shik, V. V. (1982). Contact between subunits of *Escherichia coli* RNA polymerase and the nucleotides of the lacUV5 promoter. *Molekulyarnaya Biologiya*, **16**, 34–46.

Daniels, D., Zuber, P. & Losick, R. (1990). Two amino acids in an RNA polymerase σ factor involved in the recognition of adjacent base-pairs in the –10 region of a cognate promoter. *Proc. Natl Acid. Sci. USA*, **87**, 8075–8079.

Dombroski, A. J., Walter, W. A., Record, M. T., Siegele, D. A. & Gross, C. A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor sigma70 exhibit specificity of binding to promoter DNA. *Cell*, **70**, 501–512.

Dombroski, A. J., Walter, W. A. & Gross, C. A. (1993). Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes Dev.* **7**, 2446–2455.

Gardella, T., Moyle, T. & Susskind, M. M. (1989). A mutant *Escherichia coli* sigma 70 subunit of RNA polymerase with altered promoter specificity. *J. Mol. Biol.* **206**, 579–590.

Gribskov, M. & Burgess, R. R. (1983). Overexpression and purification of the sigma subunit of *Escherichia coli* RNA polymerase. *Gene*, **26**, 109–118.

Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucl. Acids Res.* **11**, 2237–2255.

Helmann, J. D. & Chamberlin, M. J. (1988). Structure and function of bacterial sigma factors. *Annu Rev. Biochem.* **57**, 839–872.

Hillenkamp, F., Karas, M., Beavis, R. C. & Chait, B. T. (1991). Matrix-assisted laser desorption / ionization mass spectrometry of biopolymers. *Anal. Chem.* **63**, 1193A–1203A.

Juang, Y.-L. & Helmann, J. D. (1994). A promoter melting region in the primary sigma factor of *Bacillus subtilis*: identification of functionally important aromatic amino acids. *J. Mol. Biol.* **235**, 1470–1488.

Kashlev, M., Martin, E., Polyakov, A., Severinov, K., Nikiforov, V. & Goldfarb, A. (1993) Histidine-tagged RNA polymerase: dissection of the transcription cycle using immobilized enzyme. *Gene*, **130**, 9–14.

Keilty, S. & Rosenberg, M. (1987). Constituitive function of a positively regulated promoter reveals new sequences essential for activity. *J. Biol. Chem.* **262**, 6389–6395.

Kumar, A., Malloch, R. A., Fujita, N., Smillie, D. A., Ishihama, A. & Hayward, R. S. (1993). The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an ''extended minus 10'' promoter. *J. Mol. Biol.* **232**, 406–418.

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.

Lesley, S. A. & Burgess, R. R. (1989). Characterization of the *Escherichia coli* transcription factor sigma 70: localization of a region involved in the interaction with core RNA polymerase. *Biochemistry*, **28**, 7728–7734.

Lonetto, M., Gribskov, M. & Gross, C. A. (1992). The σ^{70}

family: sequence conservation and evolutionary relationships. *J. Bacteriol.* **174**, 3843–3849.
Losick, R. & Pero, J. (1981). Cascades of sigma factors. *Cell*, **25**, 582–584.
Lowe, P. A., Hager, D. A. & Burgess, R. R. (1979). Purification and properties of the σ subunit of *Escherichia coli* DNA-dependent RNA polymerase. *Biochemistry*, **18**, 1344–1352.
Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
McClure, W. R., Cech, C. L. & Johnston, D. E. (1978). A steady state assay for the RNA polymerase initiation reaction. *J. Biol. Chem.* **253**, 8941–8948.
Mulvey, M. R. & Loewen, P. C. (1989). Nucleotide sequence of *katF* of *Escherichia coli* suggests that *katF* protein is a novel σ transcription factor. *Nucl. Acids Res.* **17**, 9979–9991.
Orsini, G., Ouhammouch, M., Le Caer, J.-P. & Brody, E. N. (1993). The *asiA* gene of bacteriophage T4 codes for the anti-σ^{70} protein. *J. Bacteriol.* **175**, 85–93.
Polyakov, A., Severinova, E. & Darst, S. A. (1995). Three-dimensional structure of *Escherichia coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell*, **83**, 365–373.
Ring, B. Z. & Roberts, J. W. (1994). Function of a nontranscribed DNA strand site in transcription elongation. *Cell*, **78**, 317–324.
Ring, B. Z., Yarnell, W. S. & Roberts, J. W. (1996). Function of *E. coli* RNA polymerase σ factor σ^{70} in promoter-proximal pausing. *Cell*, **86**, 485–493.
Roberts, C. W. & Roberts, J. W. (1996). Base-specific recognition of the non-template strand of promoter DNA by *E. coli* RNA polymerase. *Cell*, **86**, 495–501.
Rong, J. C. & Helmann, J. D. (1994). Genetic and physiological studies of *Bacillus subtilis* σ^{A} mutants defective in promoter melting. *J. Bacteriol.* **176**, 5218–5224.
Severinov, K., Fenyö, D., Severinova, E., Mustaev, A., Chait, B. T., Goldfarb, A. & Darst, S. A. (1994). The sigma subunit conserved region 3 is part of ''5'-face'' of active center of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **269**, 20,826–20,828.
Siegele, D. A., Hu, J. C., Walter, W. A. & Gross, C. A. (1989). Altered promoter recognition by mutant forms of the sigma 70 subunit of *Escherichia coli* RNA polymerase. *J. Mol. Biol.* **206**, 591–603.
Stevens, A. (1977). Inhibition of DNA binding by an RNA polymerase inhibitor from T4 phage-infected *Escherichia coli*. *Biochim. Biophys. Acta*, **475**, 193–196.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.
Waldburger, C. & Susskind, M. M. (1994). Probing the informational content of *Escherichia coli* sigma70 region 2.3 by combinatorial cassette mutagenesis. *J. Mol. Biol.* **235**, 1489–1500.
Waldburger, C., Gardella, T., Wong, R. & Susskind, M. M. (1990). Changes in conserved region 2 of *Escherichia coli* sigma 70 affecting promoter recognition. *J. Mol. Biol.* **215**, 267–276.
Wilson, J. (1991). The use of monoclonal antibodies and limited proteolysis in elucidation of structure-function relationships in proteins. *Methods Biochem. Anal.* **35**, 207–250.
Zhou, Y. N., Walter, W. A. & Gross, C. A. (1992). A mutant sigma 32 with a small deletion in conserved region 3 of sigma has reduced affinity for core RNA polymerase. *J. Bacteriol.* **174**, 5005–5012.
Zuber, P., Healy, J., Carter, H. L., III, Cutting, S., Moran, C. P., Jr & Losick, R. (1989). Mutation changing the specificity of an RNA polymerase sigma factor. *J. Mol. Biol.* **206**, 605–614.

Edited by M. Gottesman

(Received 8 April 1996; received in revised form 15 August 1996; accepted 20 August 1996)