Core–sigma interaction: probing the interaction of the bacteriophage T4 gene 55 promoter recognition protein with *E.coli* RNA polymerase core

Jean-Paul Léonetti^{1,2}, Kevin Wong¹ and E.Peter Geiduschek

Department of Biology and Center for Molecular Genetics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0634, USA

²Present address: Faculté de Pharmacie, CNRS UMR 9921, 15 Avenue Charles Flahault, 34060 Montpellier Cedex 2, France

¹Corresponding authors

e-mail: jp@pharma.univ-montpl.fr; or kwong@biomail.ucsd.edu

The bacterial RNA polymerase sigma subunits are key participants in the early steps of RNA synthesis, conferring specificity of promoter recognition, facilitating promoter opening and promoter clearance, and responding to diverse transcriptional regulators. The T4 gene 55 protein (gp55), the σ protein of the bacteriophage T4 late genes, is one of the smallest and most divergent members of this family. Protein footprinting was used to identify segments of gp55 that become buried upon binding to RNA polymerase core, and are therefore likely to constitute its interface with the core enzyme. Site-directed mutagenesis in two parts of this contact surface generated gene 55 proteins that are defective in polymerase-binding to different degrees. Alignment with the sequences of the σ proteins and with a recently determined structure of a large segment of σ^{70} suggests that the gp55 counterpart of σ^{70} regions 2.1 and 2.2 is involved in RNA polymerase core binding, and that σ^{70} and gp55 may be structurally similar in this region. The diverse phenotypes of the mutants implicate this region of gp55 in multiple aspects of σ function.

Keywords: bacteriophage T4 late genes/RNA polymerase/sigma proteins/T4 gp55/transcriptional regulation

Introduction

Eubacterial RNA polymerase holoenzymes share four highly conserved subunits, α (present as a dimer), β , β' and σ . The σ subunits bind reversibly to the core (α_2 , β , β'), confer promoter specificity on the holoenzymes and, as a consequence, play a central role in the control of gene expression. Sigma factors are multi-functional: in addition to core enzyme binding, they recognize sequence-specific double-stranded DNA, interact with certain transcriptional activators, determine promoter melting by interacting with single-stranded DNA, inhibit non-specific transcription, play a key role in promoter clearance and are released from the polymerase core at an early transition step of RNA chain elongation (for reviews, see Gribskov and Burgess, 1986; Helmann and Chamberlin, 1988; Gross *et al.*, 1996). It is likely that the interface between a σ

factor and its RNA polymerase core plays a major role in these different processes. This interface is probably where most of the information is exchanged between the promoter and the polymerase, and where the resulting conformational changes of the σ factor, the polymerase and DNA are determined. In spite of the importance of this interface, little is known about it.

Bacterial σ factors can be divided into two families, the primary σ proteins, such as σ^S and σ^{70} , which share relatively extensive identity and homology, and the diverged σ proteins, which are not identified by standard similarity searching methods (σ^E and Fecl, for example). The most conserved segments of the primary bacterial σ factors are located in the four subdivisions of region 2 (Gribskov and Burgess, 1986; Helmann and Chamberlin, 1988; Gross *et al.*, 1996). Point mutations and deletions in region 2.1 implicate this domain in core binding (Lesley and Burgess, 1989; Shuler *et al.*, 1995), and a single point mutant in region 2.2 of σ^{32} is also partially defective in core binding (Joo *et al.*, 1997). The recently solved structure of a large part of σ^{70} reveals that regions 2.1 and 2.2 form two α -helices, linked by a loop and interacting mostly through hydrophobic contacts (Malhotra *et al.*, 1996). Since these two regions are close together in space, strong conformational changes due to deletions or point mutations in one of the α -helices can strongly affect the other one, and in this way influence polymerase binding indirectly. Moreover, it is conceivable that this structure changes as a consequence of the interaction with RNA polymerase core. As a consequence, it is difficult to define the contact surface and the strength of contacts between σ^{70} and RNA polymerase core without additional information.

Our interest in this subject focuses on the T4 gene 55 protein, gp55, one of the smallest and least conserved members of the σ protein family. T4 gp55 enables *Escherichia coli* RNA polymerase to recognize T4 late promoters, which consist of an 8 bp TATA box, TATAAATA, centered ~10 bp upstream of the transcriptional start site. T4 late promoters entirely lack a –35 sequence element (Brody *et al.*, 1995); for its part, T4 gp55 is missing σ region 4, which, in other proteins of the family, binds to –35 promoter elements through its region 4.2 helix–turn–helix domain (Helmann and Chamberlin, 1988).

T4 gp55 is a target of a unique mechanism of transcriptional activation that connects T4 late transcription to DNA replication through the action of the T4 gene 45 protein (gp45), which is the sliding clamp of the T4 DNA polymerase holoenzyme (Kuriyan and O'Donnell, 1993; Nossal, 1994; Young *et al.*, 1994). Gp45, which is topologically but not physically linked to DNA, binds directly to gp55 and to gp33, a second T4-encoded RNA polymerase core-binding protein, apparently through C-terminal segments of similar sequence in gp55 and

gp33 (Sanders *et al.*, 1997). [Gp33 is a co-activator of T4 sliding clamp-mediated activation of T4 late transcription (Herendeen *et al.*, 1990).]

In the work presented here, we analyze the interaction of T4 gp55 with *E.coli* RNA polymerase core by protein footprinting, which probes for changes of the gp55 surface that are generated by binding to RNA polymerase core and identifies those parts of the surface that are likely to be in the most intimate contact. An alignment of the amino acid sequence of gp55 with that of σ^{70} also identifies a candidate for regions 2.1 and 2.2 in gp55. Since this is the region of tightest gp55–core contact, we explore it further by creating gp55 point mutants that are analyzed for RNA polymerase core binding and for transcription. Our findings yield a self-consistent view suggesting, first, that gp55 and σ^{70} may be structurally related in their regions 2.1 and 2.2; second, that region 2.1 and, to a lesser extent, region 2.2 are involved in gp55–core binding; and third, that some point mutants in this region primarily affect this binding while others are more defective in transcription than in core binding. The diversity of phenotypes implicates this region of gp55 in multiple aspects of σ function.

Results

Protein footprint of the gp55–RNA polymerase interaction

The recently developed iron-chelate methods of protein footprinting (Rana and Meares, 1991; Ermacora *et al.*, 1992; Platis *et al.*, 1993; Heyduk and Heyduk, 1994) can be used to map the surface of a protein and to detect changes of surface accessibility upon formation of more complex multi-protein and multi-protein–nucleic acid assemblies (Greiner *et al.*, 1996; Heyduk *et al.*, 1996). The untethered iron technique of hydroxyl radical footprinting (Heyduk and Heyduk, 1994) has been used successfully to study the effects of σ^{70} binding on the surface of the *E.coli* RNA polymerase (Greiner *et al.*, 1996; Nagai and Shimamoto, 1997).

The sensitivity of peptide bonds in N-terminally kinase-tagged and ^{32}P -labeled gp55 to cleavage by hydroxyl radicals was compared between the protein free in solution and bound to the core polymerase. The presence of a kinase tag at the N-terminus of the protein did not affect binding to RNA polymerase core, or activity in basal transcription of supercoiled DNA (data not shown).

In order to increase the sensitivity and the reliability of the analysis, the complex between gp55 and histidine-tagged core polymerase was purified away from unbound gp55 on Ni(NTA)–agarose. (Unbound gp55 might include incorrectly folded protein.) The hydroxyl radical sensitivity of the protein bound to the polymerase was compared with the sensitivity of the same protein bound (and therefore active according to that criterion) and subsequently displaced into solution with an excess of σ^{70} (Figure 1A). In the conditions of the analysis, 80% or more of gp55 was released from the core polymerase by σ^{70} (data not shown). A limited cleavage of gp55 with cyanogen bromide was used to generate fragments serving as molecular weight standards (Figure 1B). We estimate that uncertainties in assigning location in the footprint (Figure 1C) are as high as ± 5 –10 amino acids, depending

Fig. 1. The footprint of gp55 in its complex with *E.coli* RNA polymerase core. (A) Scans of radioactivity profiles of $\cdot\text{OH}$ cleavage patterns generated in free (top line) and bound (middle line) N-terminally ^{32}P -labeled gp55. The profile of uncut gp55 is also shown (lowest line). Electrophoretic migration is from left to right. (B) Calibration: electrophoretic mobility of N-terminally ^{32}P -labeled gp55 subjected to partial cleavage at M152, M136, M71 and M48 with cyanogen bromide. (C) Protection of gp55 from $\cdot\text{OH}$ cleavage by binding to RNA polymerase core. The cleavage profiles of free and bound gp55 were corrected for background. This panel reports the background-corrected ratio (cleavage of free gp55 minus cleavage of bound gp55 divided by cleavage of bound gp55) as a function of electrophoretic mobility. The locations of amino acids 48, 71, 136 and 152 (from B) are indicated by the downward-pointing arrows. Regions of strongest, intermediate and weakest protection are indicated at the bottom of the panel by wide, intermediate and narrow bars, respectively.

upon the position in the gel. Assignments are highly accurate at locations that are very close to the size markers; the M48 and M71 markers are especially helpful in that regard.

The above way of generating the reference, free state of gp55 is particularly relevant for a protein that is purified

Fig. 2. Alignment of T4 gp55 with *E. coli* σ^{70} . The complete sequence of the 185 amino acid gp55 is shown (bottom rows of one letter symbols) and indexed. The alignment with σ^{70} residues 381–456 (top rows of letters) is based principally on the highly conserved amino acids among the primary σ proteins (Helmann and Chamberlin, 1988) indicated in bold type; numbers in brackets indicate location in the σ^{70} sequence. The extent of domains 2.1, 2.2, 2.3, 2.4 and 3 is also indicated, and the extent of α -helical segments in the recently determined structure (Malhotra *et al.*, 1996) is shown by underlining. Segments of gp55 predicted to be α -helical are indicated by underlining. The principal features of the gp55 footprint, indicated in Figure 1, are transposed to the sequence (the widest bar indicating the strongest protection from peptide bond cleavage in the RNA polymerase core–gp55 complex, as in Figure 1; the alignment of the protein footprint with the gp55 sequence is most accurate in the vicinity of the marker methionines 48, 71, 136 and 152).

from an inclusion body in denatured form and then refolded (see Materials and methods). Sigma proteins such as σ^{70} and σ^{32} are also purified in this way, and are prone to self-association, particularly during refolding (Burgess, 1996). On the other hand, straightforward interpretation of the footprint requires that the reference, free state of the footprinted protein be well defined. Further examination of the input gp55 showed it to be monomeric even in advance of selection by polymerase core binding (see Materials and methods). Interpretation of hydroxyl radical protein footprints can also be complicated by unintended site-specific cleavage due to chelation of Fe^{2+} at specific locations either in the end-labeled protein or in its interacting partner. Site-specifically tethered Fe^{2+} has been shown to yield sharply defined cleavage patterns (Rana and Meares, 1991). The relatively smooth and broad contours of the cleavage patterns in Figure 1A suggest that this is not a complication of the gp55 footprinting analysis.

The quantitative pattern of susceptibility to hydroxyl radical cleavage and the calculation of the protection ratio along the gp55 polypeptide chain revealed two major protected regions corresponding to T42–K56 and K68–K81 (Figure 1C). Two regions of minor protection were also reproducibly detected between G82 and M136, and between I148 and T158. We verified that the protection of gp55 bound to the polymerase was not due to the steric effects of immobilization on Ni(NTA)–agarose, since a similar pattern was obtained when the footprinting was done in the presence of 100 mM imidazole, which elutes the polymerase from the solid support (data not shown). The partial digestion of gp55 with cyanogen bromide allowed a relatively precise match of the protected regions to the amino acid sequence of gp55, especially in the region of particular interest surrounding M48 and M71. (Thus, the segments of greater and intermediate protection

are located more accurately than the segment of lower protection in Figure 2.)

The effect of adding gp33 to the gp55–core complex was also examined, but no change in the gp55 footprint was detected (data not shown). If additional close contacts with gp55 are made by gp33, they lie outside the span of gp55 examined (i.e. they are confined to the very N-terminus or to the 25 C-terminal amino acids).

Alignment of the amino acid sequences of gp55 and σ^{70}

Significant direct homologies have been noted previously between some members of the σ^{70} family (i.e. SPO1 gp34, σ^{29} , σ^{32}) and gp55 in regions 2.1, 2.2 and 2.3. Multiple sequence alignments have been proposed (Gribskov and Burgess, 1986; Helmann and Chamberlin, 1988): according to these indirect alignments, gp55 and σ^{70} appear to be diverged members of the same family. The two alignments differ in region 2.4, and no significant homology was detected in regions 1, 3 and 4. [Gp55 has no –35 DNA-binding site, but can instead form an additional topological linkage to DNA through its interaction with the T4 gp45 sliding clamp transcriptional activator (Tinker *et al.*, 1994; Tinker-Kulberg *et al.*, 1996).] Regions 2.1 and 2.2, which form two interacting α -helices linked by a loop in σ^{70} (Malhotra *et al.*, 1996), are at the center of the most conserved region of the σ family. As already pointed out (Malhotra *et al.*, 1996), if the ternary structure of σ^{70} is not modified during binding to polymerase core, an exposed hydrophobic face of this helix–loop–helix domain is likely to face the polymerase, while the other side would face region 2.4 and DNA.

Secondary structure prediction programs identify several regions of gp55 with very high probability of forming an α -helix. The alignment shown in Figure 2 locates this

Fig. 3. σ^{70} and gp55. The left-hand picture shows the structure of amino acids 370–446 of σ^{70} taken from Malhotra *et al.* (1996), with regions 2.1 (red), 2.2 (yellow), 2.3 (blue) and 2.4 (white) indicated. Amino acids N383 and L386 (region 2.1), and L404 and I405 (region 2.2) are indicated. At the right, the footprint of gp55 is mapped onto the structure of this segment of σ^{70} , following the Figure 2 alignment, with the most protected region of gp55, aligned to amino acids ~379–392 of σ^{70} , indicated in red; the somewhat less protected segment, aligned to amino acids ~402–414 of σ^{70} , is indicated in dark blue, and the segment that is less protected still, and aligns to amino acids ~415–446 of σ^{70} , is in yellow. It is proposed that amino acids N383, L386, L404 and I405 of σ^{70} (identified at the left) correspond with I47, I50, M71 and I72 of gp55, respectively. This postulated correspondence has been indicated by the arrows in the right-hand panel.

predicted secondary structure of gp55 quite close to the corresponding secondary structure elements of σ^{70} in regions 2.1, 2.2, 2.3 and 2.4. Highly conserved residues of the σ^{70} family also align adequately with assigned correspondents in gp55 (shown in bold type in Figure 2). We also note a very hydrophilic patch of amino acids located on one face of σ^{70} and a similar patch for gp55 in region 2.2; in addition, strong similarities can be seen in the distributions of the hydrophobic amino acids in region 2.1 of σ^{70} and the aligned segment of gp55. Two aromatic amino acids in region 2.3 of σ^{70} separated by 10 other residues (F419 and Y430) are characteristic of the sigma family (Gribskov and Burgess, 1986), and are identical in gp55 (F86 and Y97). Mutations of these two amino acids affect, respectively, the protein folding and promoter melting functions of σ^{70} (Juang and Helmann, 1994); alanine substitutions at F86 and Y97 were found to abolish the T4 late transcription activity of gp55 (J.-P.Léonetti and K.Wong, unpublished observations). Therefore, it is likely that the diverged sequences of these two primary σ factors are nevertheless compatible with similar secondary structure. Local conservation of secondary structure implies probability of similar tertiary structure (cf. Fischer and Eisenberg, 1997). We postulate that this structure, more than sequence, has been conserved (Figure 3). Experiments in which the gp55 structure was probed by limited proteolysis supported this conjecture by exposing a proteolysis-resistant core comprising regions 2.1–2.4, comparable with what has been shown previously for σ^{70} (Severinova *et al.*, 1996), and a proteolysis-exposed C-terminus (J.-P.Léonetti, unpublished observations).

Mutagenesis of the contact region

The identification of a region of gp55–core polymerase interaction by footprinting (Figure 1) and the implications of the sequence and potential structure alignment were pursued by site-directed mutagenesis. Four targets—the hydrophobic amino acids 47, 50, 71 and 72—were chosen primarily on the basis of the sequence alignment (Figure 2) and with an eye to the σ^{70} structure (Figure 3), and only secondarily on the basis of the footprint. (As noted in Materials and methods, indexing of the footprint is not reliable to the single amino acid, although highly accurate in the immediate vicinity of the markers at amino acids 48 and 71.) Genes encoding C-terminally His-tagged mutants of gp55 were expressed in *E.coli*, and the proteins were purified to homogeneity as described in Materials and methods. These mutant proteins were tested for polymerase binding by affinity chromatography and for transcription activity.

The affinity chromatography assay used here is sensitive enough to detect mutants that are partially defective in polymerase binding. We were interested to observe that the assay is only compatible with a C-terminal placement of the His tag; N-terminally His-tagged gp55 bound RNA polymerase core less efficiently, and the resulting core–gp55 complex bound extremely poorly to Ni(NTA)–agarose (data not shown).

In the presence of 0.1% Tween-20 and 100 mM NaCl, binding of gp55 to the core polymerase was efficient and very specific; no core enzyme binding to Ni(NTA)–agarose was detected in the absence of gp55. Of these single amino acid substitution proteins, only gp55-I47K was

completely defective in polymerase binding at low salt concentration (data not shown). However, when the NaCl concentration was increased to 250 and 400 mM, gp55-I50K, -M71A and -I72K were also found to be defective for binding to the core enzyme (Figure 4 and Table I). Gene 55 proteins deleted for amino acids 46–52 and 66–72, respectively, failed to bind RNA polymerase at the lowest salt concentration tested (100 mM NaCl; data not shown).

All these mutants were tested for their ability to tran-

Fig. 4. Affinity chromatography of *E.coli* RNA polymerase core on immobilized wild-type and mutant gp55. C-terminally His-tagged gp55 wild-type, -I47K, -I50K, -M71A and -I72K (A–E, respectively) was bound to Ni(NTA)-agarose. RNA polymerase core was applied (load in lanes 1, 11 and 21; flow-through in lanes 2, 6, 12, 16 and 22, respectively) and washed with buffer AC with 400 mM NaCl (lanes 3 and 4, 7 and 8, 13 and 14, 17 and 18, 23 and 24, respectively) and finally eluted with the same buffer containing 200 mM imidazole (lanes 5, 9, 15, 19 and 25, respectively). Proteins are identified on the left. (The asterisks mark a band that is probably due to gp55 dimer and is not consistently observed.)

scribe linear DNA containing a late T4 promoter in conjunction with core polymerase in the presence of 240 mM K acetate (Figure 5). Gp55-I47K, which was defective for polymerase binding at low salt concentration, was completely defective for transcription in these conditions. Gp55-I50K and -M71A were also strongly, but not absolutely, defective for transcription (~10% of the wild-type). Surprisingly, transcription activity was less strongly affected by the I72K mutation, and this was confirmed by titration (Figure 5B).

The buffer of the preceding transcription analysis contained 240 mM K acetate, which is less destabilizing for gp55 binding to core polymerase than is the same molar concentration of NaCl. Higher NaCl concentrations cannot be used for basal unenhanced late transcription, as the polymerase dissociates from DNA. Accordingly, the analysis was extended in two ways. First, transcription was examined with 100 mM NaCl and yielded a very similar outcome to the analysis with 240 mM K acetate (Table I), in that gp55-I47K was inactive, the -I50K and -M71A proteins were poorly active and gp55-I72K was partially active. Second, the core affinity chromatography was repeated in the presence of 240 mM K acetate, and showed gp55-I50K, -M71A and -I72K binding detectably but comparatively weakly relative to the wild-type, while gp55-I47K did not bind (data not shown). These findings indicate that the I50K, M71A and I72K mutations affect two aspects of gp55 function differently. We discuss the implications of this below.

The complete inactivity of the I47K protein prompted concern for its state of aggregation and solubility when renatured out of guanidine (as described in Materials and methods). As an indication of aggregation on the Ni(NTA) column, we compared the imidazole elution patterns of wild-type gp55 and gp55-I47K. When His₆ tags are placed at both ends of a protein, the affinity for Ni(NTA) increases so that higher concentrations of imidazole are required for elution (G.A.Kassavetis, personal communication). We reasoned that extensive self-association would generate effectively multi-dentate ligands for the Ni(NTA) column eluting at higher imidazole concentrations and, in the extreme case, polymerization products that would be intertwined with the column matrix and incapable of elution with imidazole. The imidazole elution patterns of wild-type gp55 and gp55-I47K deposited on the Ni(NTA) column under the conditions of the RNA polymerase core-binding experiments (Figure 4 and Materials and methods) showed most of the mutant protein and almost all of the wild-type protein eluting with 100 mM imidazole; a

Table I. RNA polymerase core-binding and transcriptional activity of gp55 mutants

Protein	Binding at			Relative transcription efficiency ^a	
	100 mM NaCl	250 mM NaCl	400 mM NaCl	240 mM Na acetate	100 mM NaCl
K-gp55	+	+	+	1.0	1.0
-I47K	– ^b	–	–	0.02	0.01
-I50K	+	–	–	0.16	0.14
-M71A	+	+	–	0.11	0.10
-I72K	+	±	–	0.72	0.41

^aYield of transcripts relative to the wild-type protein (data for 240 mM K acetate from Figure 5A).

^bRNA polymerase core binding relative to His-tagged wild-type K-gp55-H6 at the same NaCl concentration.

Fig. 5. Single-cycle analysis of T4 late transcription with wild-type and mutant gp55. **(A)** The 420 nucleotide T4 late transcript of plasmid pDH310 is indicated. The prominent upper band (marked read-through) is the run-off transcript initiating at the same late promoter. The recovery marker for quantifying relative yields of transcripts is also indicated at the right side. Transcription was with 5 nM linearized pDH310, 50 nM RNA polymerase core and 300 nM gp55, as specified in Materials and methods. **(B)** Titrations with gp55. Molar ratios of gp55 to RNA polymerase core (50 nM) are indicated on the abscissa. The yield of 420 nucleotide transcript with gp55 wild-type (○), gp55-I72K (□), -I50K (△) and -M71A (●) is reported on the ordinate.

comparable, small proportion of both proteins eluted with 200 mM imidazole. A fraction of gp55-I47K (less than one-half) was, however, retained on the column in 200 mM imidazole and only released with SDS-EDTA (data not shown). We conclude that, although gp55-I47K is prone to self-aggregation under the conditions of the Figure 4 experiment, most of it is not occluded to polymerase binding through precipitation on the column.

The aggregation of gp55-I47K was also examined in solution by gel exclusion chromatography. Gp55 eluting from the Ni(NTA) column with 100 mM imidazole was mixed with marker proteins and analyzed on a Superose 12 sizing column, as described in Materials and methods (protein footprinting section). While the wild-type protein eluted from the sizing column as a monomer, gp55-I47K

eluted heterogeneously, indicating the presence of only a small proportion of monomer, a substantial fraction of dimers to tetramers, and a larger fraction of higher aggregates. We suggest that the especially strong polymerase-binding defect of gp55-I47K may be due to local misfolding of the projecting surface of this part of gp55 (Figures 1 and 3). Such a local structural change could also account for the tendency of gp55-I47K to aggregate. However, we cannot exclude the possibility that this mutation simply increases the natural propensity of gp55 to aggregate at higher concentrations.

Discussion

Sigma factors are multi-functional; they bind both the core polymerase and the promoter, are closely involved in all steps of promoter-specific transcriptional initiation and even have a role in promoter clearance. The interface with the polymerase core must be the place where σ proteins exert most of their diverse effects.

Our interest in the σ proteins focuses on gp55 because of its role in gene regulation and replication-coupled transcription (Brody *et al.*, 1995; Sanders *et al.*, 1997). Sequence alignment of regions 2.1, 2.2 and 2.3 clearly identifies gp55 as a member of the σ family. Gp55 and σ^{70} are known to compete for a common core-binding site (Williams *et al.*, 1987). Thus, although the sequences of σ^{70} and gp55 are highly divergent elsewhere, the secondary and tertiary structures are likely to be conserved in these three regions of these two proteins. That view is supported by secondary structure prediction of gp55 (Figure 2), and by a comparison of the putative α -helices of gp55 (Figure 2) with those identified in σ^{70} (Figure 3).

Its small size also recommends gp55 for structural and mutational analysis. It is less than one-third the size of σ^{70} ; indeed, 140 N-proximal amino acids suffice for directing promoter binding and accurate transcriptional initiation (Sanders *et al.*, 1997). Intramolecular, inter-domain interactions (involving regions 1.1, 2, 4 and the non-essential 245 amino acid segment between regions 1 and 2, for example) (Dombroski *et al.*, 1996; Wilson and Dombroski, 1997) that may complicate the interpretation of footprints of the σ^{70} -polymerase core interaction are likely to be lacking in the smaller and simpler gp55.

Protein footprinting was done on T4 gp55 in order to define the surface of its contact with the polymerase core. The small size of gp55 allows a relatively precise localization of polymerase contacts, and an indication of their intensity (Figure 1C): regions 2.1 and 2.2 of gp55 appear to be protected by binding to the core, indicating close proximity to the core polymerase. In σ^{70} , region 2.1 was the first candidate for core binding, identified by deletion (Lesley and Burgess, 1989) and by point mutation analysis (Shuler *et al.*, 1995). More recently, two mutations at a single site in region 2.2 of σ^{32} were shown to generate proteins that are partially defective in core binding (Joo *et al.*, 1997). Our data are consistent with these results, and provide evidence that the close proximity of these two regions to the core polymerase extends to a member of the σ family with extensively divergent primary amino acid sequence. The intensity of contacts can also be evaluated from the footprint: the vicinity of region 2.1 of

gp55 appears to be significantly more strongly protected than is region 2.2 (Figures 1 and 3).

Two other regions of gp55 are protected to a lesser extent by binding to polymerase. A low but reproducible protection was obtained on both sides of the M152 marker, around I148–T158. This negatively charged domain of gp55, located between region 2.4 and the C-terminal gp45-binding site (Sanders *et al.*, 1997; K.Wong, unpublished observations), is reminiscent of σ^{70} region 3, which is known to be located close to the polymerase catalytic center (Severinov *et al.*, 1994). A deletion in region 3 of σ^{32} is also known to partially destabilize its interactions with the core polymerase (Zhou *et al.*, 1992). Another weak but reproducibly observed protection, in regions 2.3 and 2.4, may reflect intra-subunit stabilization of these domains on the α -helices of regions 2.1 and 2.2 upon binding to core rather than a direct contact with the polymerase, since a seven amino acid deletion in this region (amino acids 107–114) does not drastically destabilize the interaction of gp55 with polymerase core (data not shown). A conformational change or stabilization in region 2.4 upon core binding may account for our failure, thus far, to detect a direct gp55–DNA footprint (K.Wong unpublished observations). It will be recalled that direct and specific binding of σ^{70} region 2 to promoter DNA is also relatively weak and has been detected, with technical difficulty, by filter binding rather than by footprinting (Dombroski *et al.*, 1992; Dombroski, 1996; C.A.Gross, personal communication). We cannot exclude the existence of additional contacts close to the N- or C-termini of gp55, which are not examined under our experimental conditions.

Core binding and transcription experiments conducted with point mutants in the most protected regions of gp55 support the information obtained from the protein–protein footprint. Mutations were introduced deliberately into hydrophobic amino acids, as the solvent-exposed hydrophobic face of σ^{70} is a good candidate for core polymerase binding (Malhotra *et al.*, 1996); all of them turned out to be core binding-defective. Despite the severity of the introduced changes, three of these binding-defective mutants retained some ability to transcribe, suggesting that their loss of polymerase binding is due to a specific loss of protein–protein contact and not simply due to complete misfolding. Deletions of multiple amino acids (46–52 and 66–72) generated binding-defective and transcriptionally inactive proteins (data not shown).

The phenotypes of the four point mutants are diverse. (i) Gp55-I47K does not bind the polymerase even at low salt stringency and, consequently, is completely inactive in transcription. Region 2.1 of σ^{70} has a long α -helix with a kink centered around N383, corresponding to I47 in gp55. It has been proposed that this kink could be important in core binding (Malhotra *et al.*, 1996). The severity of the core-binding defect of gp55-I47K and indications of aggregation may signify the loss of this structure and the associated loss of a specific core–gp55 contact. [It is necessary to point out that the direct interpretation of gp55 mutant phenotype in terms of σ^{70} structure becomes tenuous for I47, which is located at the edge of the σ^{70} alignment, and adjacent to the highly diverged N-terminus of gp55. Nevertheless, I47 is also adjacent to a stretch of amino acids that were shown to

be essential for core binding in σ^E (Shuler *et al.*, 1995).] (ii) Gp55-I72K binds the polymerase at low stringency, but binding is strongly destabilized at higher NaCl concentrations. Transcription activity is substantially restored at higher concentration of gp55-I72K (Figure 5B), suggesting that this mutation lowers the affinity for core polymerase. We note this primarily core binding-defective phenotype of gp55-I72K with interest because, in our alignment, I72 of gp55 corresponds to the conserved I405 in σ^{70} , which is part of a solvent-exposed hydrophobic patch (Malhotra *et al.*, 1996) that might be suited to protein–protein interaction. (iii) Two other mutants (gp55-I50K and -M71A) are essentially defective in transcription. These two proteins bind to the polymerase core at low stringency, but are unable to transcribe; core binding of both is destabilized at higher stringency (Table I).

The diverse properties of these mutants point to a diversity of roles of the interface between gp55 and the polymerase core. Some gp55 residues at this interface are more important for polymerase recognition, while other residues are more important for steps that follow the initial core binding. It is likely that the residues involved in gp55–core binding change during successive steps of the transcription cycle; mutants essentially deficient in function could be defective in binding to core during the later stages of forming the elongating transcription complex. The diversity of phenotypes observed in this work contrasts with a recent report of mutants in region 2.2 of σ^{32} (Joo *et al.*, 1997) that are more exclusively deficient in core binding; transcription *in vitro* with all these mutants is rescued by adding more mutant protein, implying a lower affinity for the σ^{32} –core interaction. The more restricted phenotype of these σ^{32} mutants may reflect their mode of isolation: they were selected as suppressors of a σ^{70} mutation conferring Lon protease-mediated instability at high temperature (Grossman *et al.*, 1985), precluding the generation of mutations entirely abrogating σ^{32} -dependent transcription. Concomitantly, the σ^{32} mutations that were found in this way are in polar amino acids (P→R, E→G and Q→R or N), while we largely opted for more drastic changes and for mutating hydrophobic amino acid codons of gene 55. A detailed mutagenesis of region 2.1 of σ^{70} has also led to the isolation of mutants whose functional defects are not solely due to defective core binding (L.Rao and R.R.Burgess, personal communication).

The Figure 1 footprint points to gp55–core contact in the region of gp55 that the Figure 2 sequence alignment associates with a part of σ^{70} that is already implicated in core binding. It appears very likely, therefore, that the diverged σ factors also interact with core through their regions 2.1 and probably to a lesser extent regions 2.2; for gp55, this conclusion is supported by site-directed mutagenesis (Figures 4 and 5; Table I). We suggest (Figure 3) that, in spite of sequence divergence, the general structure of the domain in contact with the polymerase core may be conserved between σ^{70} and gp55 and that, accordingly, it could be generally conserved among primary sigma proteins. The outcome of the transcription analysis of the gp55 mutants is also in keeping with the supposition that the function of regions 2.1 and 2.2 is more complex than just core polymerase binding. It is anticipated that the interactions and points of contact of

a σ factor with its core polymerase probably are not static, but progressively modified in succeeding steps of transcriptional initiation: first in generating polymerase isomerization, perhaps separately in facilitating promoter opening, and concluding with σ displacement from the elongating transcription complex.

Materials and methods

Cloning and protein purification

Gene K-gp55, encoding a derivative of the wild-type protein that is phosphorylated by bovine heart myokinase, was obtained by inserting codons for the amino acid sequence RRASV between the first and second codons of the wild-type gene 55, cloned in pET21.b. Gene K-gp55-H6 has six His codons inserted at the C-terminus of the K-gp55 open reading frame. The four mutations, I47K, I50K, M71A and I72K, and two mutants deleting amino acids 46–52 and 66–72, respectively, were obtained by oligonucleotide-directed mutagenesis of the K-gp55-H6 gene.

These genes were expressed in *E.coli* BL21 λ DE3 cells. Cells were harvested 4 h after induction with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG), lysed, and the proteins were extracted from inclusion bodies with 40 mM Tris-HCl, pH 8, 6 M guanidinium chloride, 10 mM β -mercaptoethanol and 10% (v/v) glycerol. Proteins K-gp55-H6, K-gp55-H6-I47K, -I50K, -M71A and -I72K were then submitted to Ni(NTA)-agarose chromatography in the same buffer, eluted with this buffer supplemented with 200 mM imidazole and stored in this buffer at -20°C . The purification of core RNA polymerase followed standard methods; purification of histidine-tagged core RNA polymerase from uninfected *E.coli* by T.-J.Fu followed an adaptation of a procedure provided by R.Landick (personal communication; see also Kashlev *et al.*, 1996).

Protein footprinting and structure prediction

Two hundred pmol of K-gp55 protein was diluted 50-fold directly out of its guanidine-containing storage buffer into a 50 μl volume of buffer containing 20 mM Tris-HCl pH 7.6, 100 mM NaCl and 10 mM MgCl_2 for phosphorylation with 5 U of bovine heart myokinase in the presence of 450 μCi of [γ - ^{32}P]ATP; free label was removed by gel filtration on Biogel P6 in buffer F [40 mM NaHEPES pH 7.8, 100 mM NaCl, 10 mM MgCl_2 , 10 mM β -mercaptoethanol, 10% (v/v) glycerol and 0.1% (w/v) Tween-20]. For protein footprinting, 0.8 pmol of [^{32}P]K-gp55 protein ($\sim 5 \times 10^5$ c.p.m.) were pre-incubated with 8 pmol of histidine-tagged RNA polymerase core for 5 min at room temperature in buffer F. The complex was then loaded onto a 30 μl Ni(NTA)-agarose column in order to remove unbound gp55, and the now immobilized E.gp55 was resuspended in 250 μl of buffer F without glycerol or β -mercaptoethanol. 50 μl aliquots of the slurry were used in the next steps. 20 pmol of σ^{70} , which binds to core polymerase more tightly than does gp55, were added in one sample to displace the bound K-gp55 from the polymerase. After 5 min incubation at 37°C , samples were stored on ice until cleavage, which followed Greiner *et al.* (1996). Separate 5 μl drops of 200 mM Fe(EDTA) complex, 40 mM Na ascorbate and 40 mM H_2O_2 , respectively, were placed on three sides of the tube containing the reaction sample, equilibrated at 37°C , and then shaken down to initiate the footprinting reaction, which proceeded for 90 s and was terminated by adding 30 μl of gel sample buffer. Samples were kept frozen until loaded on a 16% (w/v) acrylamide:0.5% (w/v) bisacrylamide, 6 M urea Tris-tricine gel (Shagger and Von Jagow, 1987). Gels were fixed with 20% trichloroacetic acid (TCA) before analysis with a phosphorimaging scanner. An average of 20–30% cleavage (i.e. much less than one cut per molecule of gp55 on average) was obtained. The different lanes of the footprint were aligned at the top (i.e. at the uncut gp55 signal). Additional alignments and normalizations were avoided, as gp55 is a small protein, and objective assignment of any part of its footprint as unchanged was not judged to be feasible. Reliance was placed instead on obtaining a reproducible cleavage pattern in four independent experiments. Locations along the gp55 polypeptide chain were determined against a partial digest of phosphorylated K-gp55 with cyanogen bromide.

The state of aggregation of (wild-type) gp55 was examined as follows: 150 pmol of protein was diluted out of storage buffer into 150 μl of buffer AC [40 mM Tris-HCl pH 7.8, 10 mM MgCl_2 , 10% (v/v) glycerol, 10 mM β -mercaptoethanol and 0.1% (w/v) Tween-20] with 100 mM

NaCl and loaded onto a 30 μl Ni(NTA) column pre-equilibrated with the same buffer. The column was washed with the same buffer and protein was eluted in two 150 μl steps with the same buffer also containing 100 and 200 mM imidazole, respectively. The great majority of the protein eluted with 100 mM imidazole. This material (125 μl) was mixed with an equal volume of marker proteins [cytochrome c, ovalbumin, bovine serum albumin (BSA), aldolase, catalase and ferritin] in buffer D+500 [20 mM NaHEPES pH 7.8, 7 mM MgCl_2 , 500 mM NaCl, 10% (v/v) glycerol, 0.006% (w/v) Tween-20, 0.004 mM EDTA, 10 mM β -mercaptoethanol, 5 μM phenylmethylsulfonyl fluoride (PMSF) and 0.5 $\mu\text{g/ml}$ each of leupeptin and pepstatin] and 200 μl was applied to Superose 12 pre-equilibrated with buffer D+500. Aliquots of fractions (350 μl) were prepared for analysis by SDS-PAGE, transferred to nitrocellulose, probed with rabbit antibody to gp55, and developed with alkaline phosphatase-conjugated second antibody, as described (Harlow and Lane, 1988). The wild-type gp55 eluted almost entirely as monomer (between cytochrome c and ovalbumin, mol. wts 12 500 and 45 000, respectively).

For gp55 structure prediction, we used the Predict Protein server PHD at EMBL, Heidelberg (Rost *et al.*, 1994).

Assay for binding of gp55 to RNA polymerase core

Gp55 wild-type or mutant proteins with an N-terminal kinase tag and a C-terminal histidine tag were used in these experiments; all steps were performed at room temperature. 100 pmol of gp55 were diluted directly out of its storage buffer (400- to 700-fold according to initial concentration) into 100 μl of buffer AC with 100, 250 or 400 mM NaCl, and loaded on a 30 μl Ni(NTA)-agarose spin column pre-equilibrated in the same buffer. After 5 min incubation, the column was centrifuged and washed with the same buffer. Core polymerase (10 pmol in 50 μl of buffer AC) was added to the column, incubated for 5 min, the column was centrifuged and the flow-through liquid saved. After two washes (100 μl of buffer AC), bound proteins were eluted with 50 μl of buffer AC also containing 200 mM imidazole. 30 μl portions of each fraction were analyzed on a 12% SDS-PAGE gel. Gels were stained with Coomassie brilliant blue and recorded on a digital gel imaging system.

Single round transcription

The conditions for single round transcription essentially followed Sanders *et al.* (1995). For the experiment shown in Figure 5A, 0.75 pmol of RNA polymerase core was incubated at 0°C with 4.5 pmol of gp55 or gp55 mutant delivered to the reaction mixture by rapid serial (10- to 100-fold) dilution out of its storage buffer for 5 min, in a 10 μl volume of transcription buffer [240 mM K acetate, 33 mM HEPES pH 7.8, 10 mM Mg acetate, 1 mM dithiothreitol (DTT) and 150 $\mu\text{g/ml}$ BSA], transferred to 25°C , and 0.075 pmol of DNA in 5 μl of transcription buffer (pre-equilibrated at 25°C) was added. After 20 min of further incubation, 5 μl of NTP mix {5 mM GTP, 5 mM ATP, 0.5 mM [α - ^{32}P]UTP (4000 c.p.m./pmol), 0.5 mM CTP and 125 $\mu\text{g/ml}$ rifampicin in transcription buffer}, pre-equilibrated at 25°C was added. Transcription was stopped after 7 min with 150 μl of stop buffer [20 mM Na_3EDTA , 40 mM Tris-HCl pH 8, 250 mM NaCl, 0.4% (w/v) SDS and 250 $\mu\text{g/ml}$ yeast RNA] containing labeled DNA serving as a sample recovery marker. For the experiment shown in Figure 5B, the reaction volume was 20 μl , with 1.0 pmol of RNA polymerase core and the specified concentrations of gp55 proteins, and 0.1 pmol of DNA was added. Transcripts were purified by phenol extraction and ethanol precipitation and resolved on 5% polyacrylamide gels containing 7 M urea; transcription products were quantified using a phosphorimaging scanner.

Experiments otherwise identical with those reported in Figure 5A were also performed in transcription buffer containing 100 mM NaCl in place of 240 mM K acetate.

Acknowledgements

We are grateful to A.Kumar and G.A.Kassavets for their interest in this work and their advice, to T.-J.Fu for purified RNA polymerase, to H.Vu for help in generating one of the figures, to S.Darst for providing coordinates of σ^{70} , to R.R.Burgess and C.A.Gross for communicating unpublished findings and to S.Kolesky for comments on the manuscript. J.-P.L. was on leave from the CNRS Institut de Génétique Moléculaire de Montpellier, and acknowledges a NATO travel grant. K.W. acknowledges support from a USPHS training grant in Cell Biology, Molecular Biology and Genetics. Our research was supported by a grant from the NIGMS.

References

- Brody,E.N., Kassavetis,G.A., Ouhammouch,M., Sanders,G.M., Tinker,R.L. and Geiduschek,E.P. (1995) Old phage, new insights: two recently recognized mechanisms of transcriptional regulation in bacteriophage T4 development. *FEMS Microbiol. Lett.*, **128**, 1–8.
- Burgess,R.R. (1996) Purification of overproduced *Escherichia coli* RNA polymerase sigma factors by solubilizing inclusion bodies and refolding from Sarkosyl. *Methods Enzymol.*, **273**, 145–149.
- Dombroski,A.J. (1996) Sigma factors: purification and DNA binding. *Methods Enzymol.*, **273**, 134–144.
- Dombroski,A.J., Walter,W.A., Record,M.T., Jr, Siegele,D.A. and Gross,C.A. (1992) Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA. *Cell*, **70**, 501–512.
- Dombroski,A.J., Johnson,B.D., Lonetto,M. and Gross,C.A. (1996) The sigma subunit of *Escherichia coli* RNA polymerase senses promoter spacing. *Proc. Natl Acad. Sci. USA*, **93**, 8858–8862.
- Ermacora,M.R., Delfino,J.M., Cuenoud,B., Schepartz,A. and Fox,R.O. (1992) Conformation-dependent cleavage of staphylococcal nuclease with a disulfide-linked iron chelate. *Proc. Natl Acad. Sci. USA*, **89**, 6383–6387.
- Fischer,D. and Eisenberg,D. (1997) Assigning folds to the proteins encoded by the genome of *Mycoplasma genitalium*. *Proc. Natl Acad. Sci. USA*, **94**, 11929–11934.
- Greiner,D.P., Huges,K., Gunasekera,A. and Meares,C. (1996) Binding of σ^{70} protein to the core subunits of *Escherichia coli* RNA polymerase, studied by iron-EDTA protein footprint. *Proc. Natl Acad. Sci. USA*, **93**, 71–75.
- Gribskov,M. and Burgess,R. (1986) Sigma factors from *E.coli*, *B.subtilis*, phage SPO1, and phage T4 are homologous proteins. *Nucleic Acids Res.*, **14**, 6745–6763.
- Gross,C.A., Chan,C.L. and Lonetto,M.A. (1996) A structure/function analysis of *Escherichia coli* RNA polymerase. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **351**, 475–482.
- Grossman,A.D., Zhou,Y.N., Gross,C., Heilig,J., Christie,G.E. and Calendar,R. (1985) Mutations in the rpoH (htpR) gene of *Escherichia coli* K-12 phenotypically suppress a temperature-sensitive mutant defective in the sigma 70 subunit of RNA polymerase. *J. Bacteriol.*, **161**, 939–943.
- Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Helmann,J.D. and Chamberlin,M.J. (1988) Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.*, **57**, 839–872.
- Herendeen,D.R., Williams,K.P., Kassavetis,G.A. and Geiduschek,E.P. (1990) An RNA polymerase-binding protein that is required for communication between an enhancer and a promoter. *Science*, **248**, 573–578.
- Heyduk,E. and Heyduk,T. (1994) Mapping protein domains involved in macromolecular interactions: a novel protein footprinting approach. *Biochemistry*, **33**, 9643–9650.
- Heyduk,T., Heyduk,E., Severinov,K., Tang,H. and Ebright,R.H. (1996) Determinants of RNA polymerase alpha subunit for interaction with beta, beta', and sigma subunits: hydroxyl-radical protein footprinting. *Proc. Natl Acad. Sci. USA*, **93**, 10162–10166.
- Joo,D.M., Ng,N. and Calendar,R. (1997) A sigma32 mutant with a single amino acid change in the highly conserved region 2.2 exhibits reduced core RNA polymerase affinity. *Proc. Natl Acad. Sci. USA*, **94**, 4907–4912.
- Juang,Y.L. and 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., Nudler,E., Severinov,K., Borukhov,S., Komissarova,N. and Goldfarb,A. (1996) Histidine-tagged RNA polymerase of *Escherichia coli* and transcription in solid phase. *Methods Enzymol.*, **274**, 326–334.
- Kuriyan,J. and O'Donnell,M. (1993) Sliding clamps of DNA polymerases. *J. Mol. Biol.*, **234**, 915–925.
- Lesley,S.A. and 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.
- Malhotra,A., Severinov,E. and Darst,S.A. (1996) Crystal structure of a sigma 70 subunit fragment from *E.coli* RNA polymerase. *Cell*, **87**, 127–136.
- Nagai,H. and Shimamoto,N. (1997) Most conserved regions of the *E.coli* primary sigma factor are involved in interaction with RNA polymerase core enzyme. *Genes to Cells*, in press.
- Nossal,N. (1994) The DNA replication fork. In Karam,J.D. (ed.), *Molecular Biology of Bacteriophage T4*. American Society for Microbiology, Washington, DC, pp. 43–55.
- Platis,I.E., Ermacora,M.R. and Fox,R.O. (1993) Oxidative polypeptide cleavage mediated by EDTA-Fe covalently linked to cysteine residues. *Biochemistry*, **32**, 12761–12767.
- Rana,T.M. and Meares,C.F. (1991) Transfer of oxygen from an artificial protease to peptide carbon during proteolysis. *Proc. Natl Acad. Sci. USA*, **88**, 10578–10582.
- Rost,B., Sander,C. and Schneider,R. (1994) PHD—an automatic mail server for protein secondary structure prediction. *CABIOS*, **10**, 53–60.
- Sanders,G.M., Kassavetis,G.A. and Geiduschek,E.P. (1995) Rules governing the efficiency and polarity of loading a tracking clamp protein onto DNA: determinants of enhancement in bacteriophage T4 late transcription. *EMBO J.*, **14**, 3966–3976.
- Sanders,G.M., Kassavetis,G.A. and Geiduschek,E.P. (1997) Dual targets of a transcriptional activator that tracks on DNA. *EMBO J.*, **16**, 3124–3132.
- Severinov,K., Fenyo,D., Severinova,E., Mustaev,A., Chait,B.T., Goldfarb,A. and 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**, 20826–20828.
- Severinova,E., Severinov,K., Fenyo,D., Marr,M., Brody,E.N., Roberts,J.W., Chait,B.T. and Darst,S.A. (1996) Domain organization of the *Escherichia coli* RNA polymerase sigma 70 subunit. *J. Mol. Biol.*, **263**, 637–647.
- Shagger,H. and Von Jagow,G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368–379.
- Shuler,M.F., Tatti,K.M., Wade,K.H. and Moran,C.P., Jr (1995) A single amino acid substitution in sigma E affects its ability to bind core RNA polymerase. *J. Bacteriol.*, **153**, 597–603.
- Tinker,R.L., Williams,K.P., Kassavetis,G.A. and Geiduschek,E.P. (1994) Transcriptional activation by a DNA-tracking protein: structural consequences of enhancement at the T4 late promoter. *Cell*, **77**, 225–237.
- Tinker-Kulberg,R.L., Fu,T.-J., Geiduschek,E.P. and Kassavetis,G.A. (1996) A direct interaction between a DNA-tracking protein and a promoter-recognition protein: implications for searching DNA sequence. *EMBO J.*, **15**, 5032–5039.
- Williams,K.P., Kassavetis,G.A., Esch,F.S. and Geiduschek,E.P. (1987) Identification of the gene encoding an RNA polymerase-binding protein of bacteriophage T4. *J. Virol.*, **61**, 597–599.
- Wilson,C. and Dombroski,A.J. (1997) Region 1 of sigma70 is required for efficient isomerization and initiation of transcription by *Escherichia coli* RNA polymerase. *J. Mol. Biol.*, **267**, 60–74.
- Young,M.C., Reddy,M.K., Jarvis,T.C., Gogol,E.P., Dolejsi,M.K. and von Hippel,P.H. (1994) Protein-protein and protein-DNA interactions in the T4 DNA polymerase accessory protein complex. In Karam,J.D.(ed.), *Molecular Biology of Bacteriophage T4*. American Society for Microbiology, Washington, DC, pp. 313–321.
- Zhou,Y.N., Walter,W.A. and 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.

Received October 3, 1997; revised December 5, 1997;
accepted December 17, 1997