Imperial College of Science, Technology & Medicine
University of London

STRUCTURAL AND FUNCTIONAL ANALYSIS OF THE BACTERIAL TRANSCRIPTION FACTOR SIGMA 54 (SIGMA N)

Siva R Wigneshweraraj

Submitted for the degree of Doctor of Philosophy, 2001

ABSTRACT

The DNA dependent RNA polymerase (RNAP) holoenzyme of *Escherichia coli* consists of the core enzyme with the subunit composition of $\alpha_2\beta\beta'\omega$ and a sigma (σ) subunit, which directs the core enzyme to initiate transcription at specific promoter sites on DNA. Many bacteria adapt to changing growth conditions by replacing the σ-factor subunit of the RNAP holoenzyme as a way of changing gene expression patterns. In *E. coli*, seven different molecular species of σ-factors are categorised into two classes. The σ^{70}-class includes the vegetative σ^{70} subunit as well as five different alternative σ-factors (σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{18}). The σ^{54}-class is uniquely represented by σ^{54}, that differs functionally and structurally from the σ^{70}-class members.

In marked contrast to the RNAP holoenzyme containing σ^{70}-class σ-factors, genes under the control of the σ^{54}-RNAP holoenzyme are subjected to a regulatory step after promoter binding. Transition of the σ^{54}-RNAP holoenzyme closed promoter complex to a transcription competent open complex is dependent upon γ-β bond nucleoside triphosphate hydrolysis by an activator protein that binds to enhancer-like DNA sequences 100-200bp upstream of the transcription start site. The enhancer binding activator proteins interact with the σ^{54}-RNAP holoenzyme to initiate transcription by looping out the intervening DNA.

In this work the structural interfaces of σ^{54}-core RNAP and σ^{54}-promoter DNA have been characterised using tethered footprinting methodologies. Using tethered iron chelate derivatives of the enhancer-dependent σ^{54} several sites of proximity to the β and β' subunits of the core RNAP were mapped. With limited, but interesting exceptions, most sites localised to those previously identified as close to the enhancer-independent σ^{70}-class of σ-factors. Small-angle X-ray scattering studies on σ^{54} have not only allowed us to develop the first low resolution model for this σ-factor, but confirmed that σ^{54} and σ^{70} share a significant structural similarity in the interface that interacts with the core RNAP.

Promoter DNA footprinting using iron chelate tethered σ^{54} proteins indicates the existence of a "regulatory centre". This centre is close to the active site of the RNAP, the start site proximal promoter region and includes σ^{54} regulatory sequences needed for DNA melting upon activation. Upon activation, the relationship between the centre and the promoter DNA changes, suggesting that this centre includes a target for the activator.

Together with these structural approaches, genetic approaches were used to determine and understand the role of σ^{54} in converting the core RNAP to an holoenzyme form that requires an enhancer binding activator protein coupled to nucleotide hydrolysis for productive transcription initiation.

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor, Professor Martin Buck for his continuous support, encouragement and willingness to provide invaluable help and advice throughout the project. I would also like to thank Professor Akira Ishihama for inviting me to work in his laboratory and introducing me to protein tethered footprinting methodologies; for his help and hospitality throughout my stay in Japan. I am also thankful to Dr Nobuyuki Fujita, Tasuka Nomura and members of the Bacterial and Viral RNA Polymerase Group at the National Institute of Genetics, Japan, for their advice and help at various stages of the experiments. I am also indebted to Dr Dimitri Svergun and Dr Marc Malfois for inviting me to the EMBL-Hamburg and assisting me with small angle x-ray scattering measurements.

I would like to thank all the past and present members of Martin Buck's laboratory for putting up with me during the last three years and the years to come. Especially, Matthew Chaney and David Studholme for teaching me good laboratory practice!!!! Mari-Trini Gallegos and Paul Casaz for technical help, helpful discussions and general advice. Wendy Cannon for help at various stages of this work and for proof-reading this thesis. Sarah Elderkin for providing me with a constant supply of sterile Eppendorf tubes and pipette tips (without her knowledge) in the final stages of the thesis. Robert Finn, Melinda Pitt, Elaine Benelli, Jose Oguiza and Ian Morris for general help. Finally, a very big thank you to ALL of the above and Cheesemaster for their friendship and making my decision very easy to continue staying in this laboratory.

A big thank you to you Sharon for your understanding and support throughout and at the end of everything for marrying me. Finally, a very big thank you to my parents – simply for everything. Especially, my father who has always taught me to do things quickly, efficiently and thoroughly as possible – an advice which came very useful during this project.

Table of Contents

CHAPTER THREE - TETHERED FOOTPRINTING AS A METHODOLOGY FOR STUDYING THE σ^{54}-CORE RNA POLYMERASE INTERFACE

CHAPTER FOUR - CONSERVATION OF SIGMA-CORE RNA POLYMERASE PROXIMITY RELATIONSHIPS BETWEEN THE ENHANCER INDEPENDENT AND ENHANCER DEPENDENT SIGMA CLASSES

CHAPTER FIVE - REGULATORY SEQUENCES IN SIGMA 54 LOCALISE NEAR THE START OF DNA MELTING

List of Figures

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

CHAPTER SEVEN

CHAPTER EIGHT

List of Tables

CHAPTER SEVEN

CHAPTER EIGHT

CHAPTER ONE

INTRODUCTION

1.1 RNA Polymerases

Gene expression requires the transfer of genetic information from DNA into RNA molecules, the direct templates for protein synthesis. The machinery which catalyses DNA-dependent RNA synthesis is the enzyme DNA-dependent RNA polymerase. Multisubunit DNA-dependent RNA polymerases are evolutionarily conserved among the three phylogenetic domains of *Eubacteria*, *Archaea*, and *Eucarya* and are found in plant chloroplasts and some viral species. Not all RNA polymerases are multisubunit enzymes. Some phages, like T7, T3, SP6 or K11 encode RNA polymerases which consist of only a single polypeptide. These single polypeptide RNA polymerases appear to be composed of several distinct domains, as revealed by the crystal structure of T7 RNA polymerase and mutation studies (Cheetham & Steitz, 2000; Sousa *et al.*, 1993; Delarue *et al.*, 1990). In addition to the RNA polymerases transcribing the nuclear genes, eukaryotic cells also require RNA polymerases to transcribe the genes of the mitochondrial genome and, in plants, of the chloroplast genome. Despite their eubacterial origin, mitochondria employ a bacteriophage T3/T7-like single subunit RNA polymerase (Tracey & Stern, 1995). In contrast, chloroplast genomes encode a eubacterial-like multisubunit RNA polymerase. Transcription in plant chloroplasts is mediated by at least two distinct transcription systems: one by the primarily plastid encoded RNA polymerase (PEP), and the other by the nuclear-encoded RNA polymerase (NEP). The NEP RNA polymerase is a single subunit bacteriophage-type enzyme. In contrast to the simple composition of the NEP RNA polymerases, the PEP RNA polymerase is a multisubunit enzyme complex (Gray & Lang, 1998). The polypeptides comprising the core PEP enzyme have been identified in chloroplast protein fractions exhibiting transcription activity from many plant species (Igloi *et al.*, 1992). These purification studies, in combination with sequence information

for the chloroplast genomes of several photosynthetic eukaryotes have indicated that the core PEP RNA polymerase is similar in subunit composition to the RNA polymerases from eubacteria (Figure 1A and see later). In Table 1.1, eubacterial, archaeal and eukaryotic multisubunit RNA polymerases from *Escherichia coli*, *Sulfolobus acidocaldarius* and *Saccharomyces cerevisiae* are compared. In eubacteria and archaea, a single multisubunit RNA polymerase is responsible for transcription of the major classes of genes including rRNA, mRNA and tRNA. In eukaryotic species, three distinct multisubunit RNA polymerases are found within the cell nucleus. RNA polymerase I synthesises rRNA, RNA polymerase II synthesises mRNA and some small nuclear RNAs, and RNA polymerase III synthesises 5S rRNA, tRNAs, and some small nuclear RNAs.

Bacterial RNA Polymerase (*E. coli*)	β (150.6)	β' (155.2)	αI (36.5)	αII (36.5)	ω (10)	-
***Archaeal* RNA Polymerase** (*S. acidocaldarius*)	A' / A'' (101 / 44)	B (122)	D (30)	L (10)	K (9.7)	6 other subunits
Eukaryotic RNA Polymerase I (*S. cerevisiae*)	2	1	5	9	6	9 other subunits
Eukaryotic RNA Polymerase II	RPB1 (190)	RPB2 (140)	RPB3 (35)	RPB11 (14)	RPB6 (18)	7 other subunits
Eukaryotic RNA Polymerase III	2	1	5	9	6	11 other subunits

Table 1.1. Compositions of eukaryotic and archaeal RNA polymerases compared to eubacterial RNA polymerases. Numbers in parentheses indicate molecular weights in kilodaltons. Subunits in the same colour have sequence homology (adapted from Ebright, 2000).

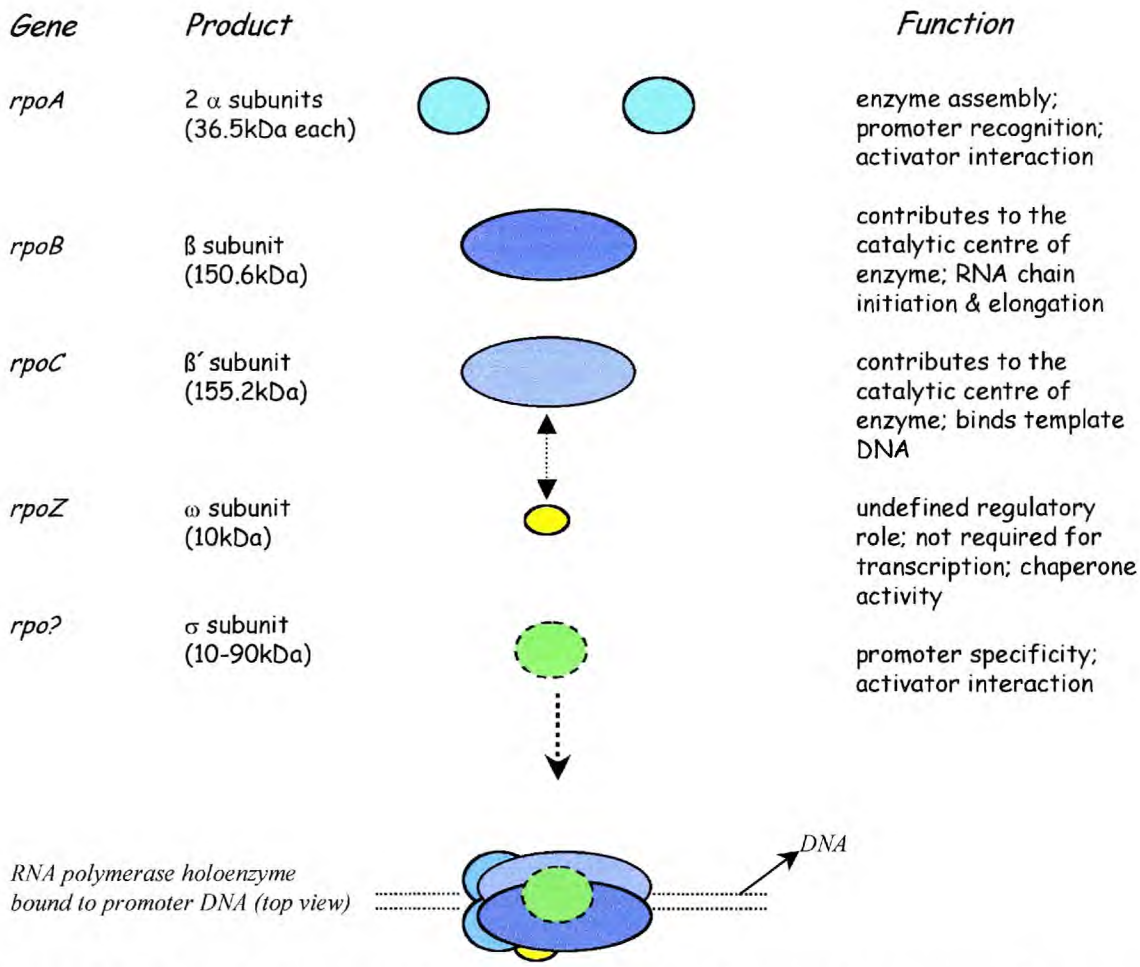


Figure 1A. The subunit composition of *E. coli* RNA polymerase holoenzyme. The cartoon illustrates the order in which the RNA polymerase enzyme is assembled (see text below for more details). rpo? indicates one of a number of sigma factors.

In bacteria, however, there is only one type of core RNA polymerase which has to transcribe all the different RNA species. This means that, in addition to the control of RNA polymerase activity, it is also necessary that the specificity of the enzyme has to be regulated such that it is able to carry out transcription of fundamentally different RNA molecules. The core RNA polymerase (RNAP) of *E. coli*, which is the best-studied enzyme of its class, contains two large subunits, named β and β', and two smaller α subunits. An ω subunit is also identified in the enzyme, but ω is not required for cell viability or transcription. The core $\alpha_2\beta\beta'$ RNAP supports RNA elongation but is incapable of accurate initiation. Promoter recognition is conferred on the core RNAP by binding of a σ subunit,

and the initiating $\alpha_2\beta\beta'\sigma$ RNA polymerase is referred to as holoenzyme (Figure 1A). Table 1.1 shows that regions of sequence similarity have been noted between α, Rpb3 and Rpb11, between β and Rpb2, and between β' and Rpb1. Consistent with the conservation of these structural elements, similar modes of interaction with nucleic acids in the vicinity of the active site have been proposed for the eukaryotic and bacterial enzymes (recently reviewed in Ebright, 2000; Cramer *et al.*, 2000; Zhang *et al.*, 1999).

1.2 Bacterial RNA Polymerase Structure-Function Relationship

The functional cycle of RNAP consists of transcription initiation, processive transcription elongation, and transcription termination (Figure 1B). During transcription the RNAP processively moves in the 3' to 5' direction along the template DNA strand, assembling a complementary, antiparallel strand of RNA that grows from its 5' terminus in a 3' direction. As indicated in Figure 1B, the RNAP catalyses the reaction

$$RNA_n + P_\gamma P_\beta P_\alpha N \rightarrow RNA_{n+1} + P_\beta P_\alpha i \ (\rightarrow 2Pi)$$

in which ribonucleoside triphosphate precursors ($P_\gamma P_\beta P_\alpha N$) are hydrolysed into nucleoside monophosphates as they are polymerised into a covalent chain. Hydrolysis of the γ-phosphate ester bond is a highly exergonic reaction with its equilibrium far towards the direction of the pyrophosphate ($P_\beta P_\alpha i$) formation. The reaction is driven further in this direction by hydrolysis of the pyrophosphate by a phosphatase to form organic phosphate ions.

Much valuable information on the function of the individual subunits of the *E. coli* RNAP has been obtained in the past from the study of mutant subunit genes or from direct biochemical analysis. Results of these studies have allowed the correlation of specific sites in the subunits with a particular function of the RNAP. Recently, two seminal papers by Zhang *et al.*, (2000) and Finn *et al.*, (2000) revealed the structures of the core RNAP from thermophilic *Thermus aquaticus* at 3.3Å, and that of the *E. coli* core RNAP and that of the

holoenzyme at 11.0Å and 9.5Å, respectively. Together, the structure-function and genetic studies of bacterial RNAPs have begun to provide insights into the basic mechanism of initiation of RNA synthesis (transcription initiation). In this section, emphasis is placed on the α, β and β' and σ subunits of the *E. coli* RNAP and their contribution and role in transcription initiation by the RNAP holoenzyme.

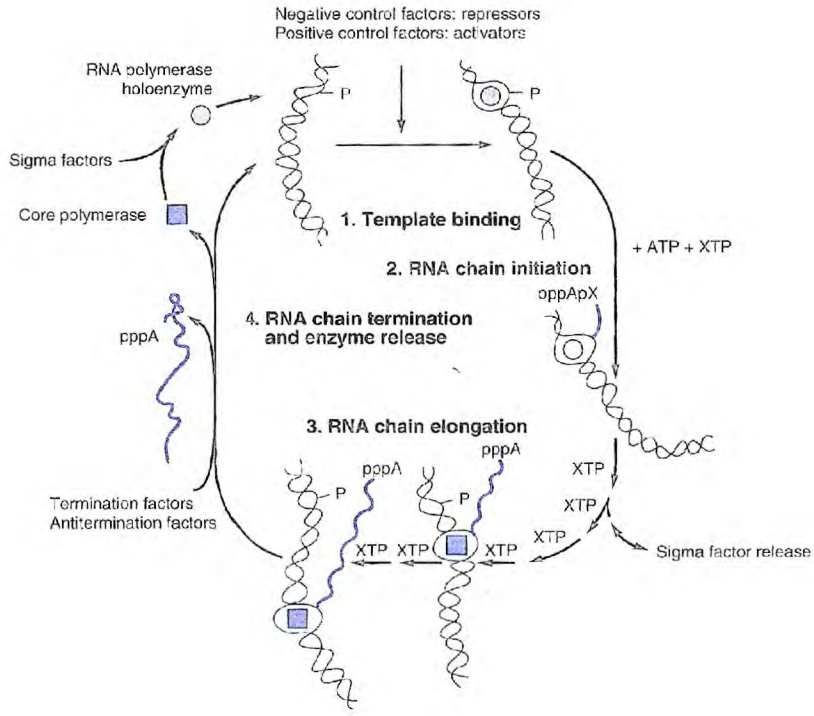


Figure 1B. Cartoon showing the various stages of bacterial transcription. (see text for details).

1.2.1 The alpha (α) subunit

The 329-amino-acid (36.5kDa) α subunit from *E. coli* contains two independently folded domains (Blatter *et al.*, 1994) whose structures have been determined at high resolution: an amino-terminal domain (αNTD) extending to residue 240 (Zhang & Darst, 1998) and a carboxyl-terminal domain (αCTD) extending from residue 249 to the end of the protein.

The linker connecting the two domains is unstructured and flexible (Blatter *et al.*, 1994) (Figure 1C).

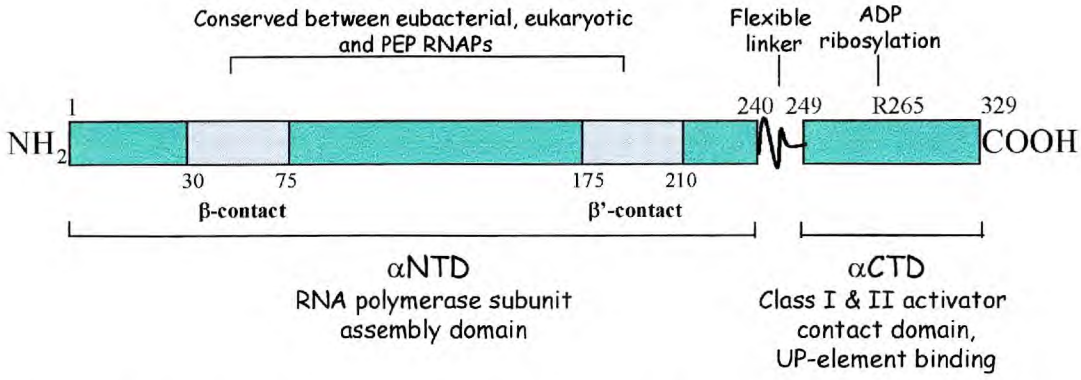


Figure 1C. Domain organisation of the *E. coli* RNAP α subunit (see text for details)

The assembly of the different subunits for the formation of a functional core RNA polymerase is not an arbitrary process but occurs sequentially in the following order scheme, in which two α subunits dimerise to form a molecular platform for β and β' binding:

$$2\alpha \rightarrow \alpha_2 \rightarrow \alpha_2\beta \rightarrow \alpha_2\beta\beta'$$

However, binding of β and β' subunits to α occurs in an asymmetrical way, with one α subunit in contact with β (αI) and the other with β' (αII). No contact regions have been established between the α and σ subunits, however cross-linking and genetic studies have indicated putative α-σ interactions (Estrem *et al.*, 1999; McMahan & Burgess, 1994). The αNTD domain (residues 1-240) is sufficient for subunit assembly and contains two conserved structural segments, which are also found in eukaryotic, archaeal and PEP RNAPs, that are believed to interact with β (residues 30-75) and β' subunits, respectively (Igarashi & Ishihama, 1991; Heyduk *et al.*, 1996). The αCTD domain (residues 249-329) is necessary for interaction with activator proteins for regulated transcription initiation

from class I and class II promoters (see later). At class II promoters, a second site of interaction between the activator (the cAMP receptor protein, CRP) and α subunit has been identified by genetic and cross-linking studies, which is located at the αNTD (see later) (Ishihama, 1993; Igarashi & Ishihama). The αCTD also contains determinants required for dimerisation and binding to several promoters which contain AT-rich sequence regions (UP elements) upstream of the core promoter (reviewed in Grouse *et al.*, 2000). Thus, although segments responsible for UP element binding and activator interaction at class I and II promoters are located in the αCTD domain, several different residues in the αNTD are also involved in both of these functions (reviewed in Ishihama, 1993). The flexible linker between the αNTD and αCTD (residues 240-249) is considered to facilitate the necessary motion between the two α-subunit domains to fulfil their function in regulating RNAP activity (Meng *et al.*, 2000). A further involvement of the αCTD in controlling RNAP activity is in T4 *early* transcription upon infection of *E. coli* cells by the T4 phage. The addition of an ADP-ribosyl group to a residue R265 alters the transcriptional properties of the RNAP in favour of the *early* transcription of the T4 phage genes.

1.2.2 The beta (β) subunit

The 1342-amino-acid (150kDa) β subunit is the second largest subunit of the *E. coli* RNAP holoenzyme and is encoded together with the β' subunit from one transcription unit. Bacterial β subunits have considerable homology to the Rpb2 subunits of all eukaryotic multisubunit RNAPs. Structural comparisons also show their relatedness (reviewed Ebright, 2000). The regions of sequence conservation are distributed in clusters over the whole β subunit and have been grouped into nine domains (Figure 1D).

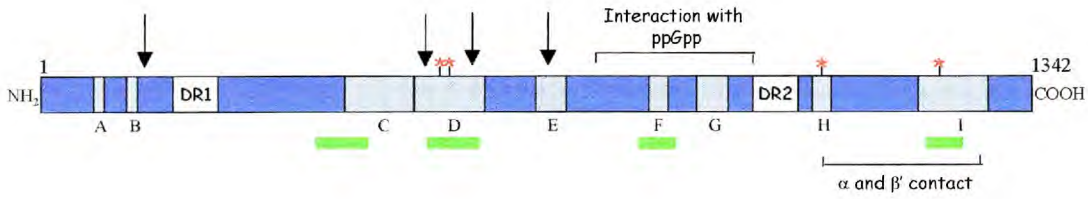


Figure 1D. Domain organisation of the *E. coli* RNAP β subunit. Black arrows indicate the Rif binding sites, red stars regions that cross-link to the 5' end of nascent RNA, and green bars regions involved in transcription termination (see text for details).

Together with the β' subunit, the β subunit forms the catalytic centre of the RNAP. In the *E. coli* β subunit, large deletions in the non-conserved regions centred around residues 300 and 1000 do not prevent the basic transcription function of RNAP and are defined as the dispensable regions (DR) 1 and 2, respectively (Severinov *et al.*, 1994, 1992). The carboxyl-terminal conserved elements H and I have been proposed to make subunit-subunit contact to both the α and the β' subunits (Heyduk *et al.*, 1996). The antibiotic rifampicin (Rif) inhibits the phosphodiester bond formation between the first nucleotides of a growing RNA chain during transcription initiation. Thus, it is believed that the Rif binding site is close to the site of nucleotide addition in the catalytic centre of the RNAP. Mutations in the β subunit that confer Rif resistance locate within three distinct conserved regions (B, D, and E) of the β subunit (Jin & Cross, 1988). Cross-linking studies have identified residues K1065 and H1237 in conserved regions H and I, and a region between residues D516 and D540 in domain D, which are close to the 5' end of the nascent transcript. Furthermore, the existence of numerous mutants with altered termination and pausing properties documents that the β subunit is also involved in pausing and termination. These mutations map in the conserved regions C, D, F and I. Several mutations in the β subunit have been described which affect the stringent response (reviewed in Cashel *et al.*, 1996). These mutations map to a region between residues 736-906, proximal to the Rif binding site (Reddy *et al.*, 1995). Such mutations are believed to affect the interaction of the effector nucleotide guanosine tetraphosphate (ppGpp) with the β subunit. Thus, bacterial cells harbouring such stringent response mutant β subunits display a "relaxed phenotype", which means that stable RNA synthesis is not as severely

reduced as in the wild-type under conditions that normally induce the stringent response, for example when cells are starved of amino acids.

1.2.3 The beta prime (β') subunit

The 1407-amino-acid (155kDa) β' subunit is the largest subunit of the *E. coli* RNAP holoenzyme. It shares considerable homology with the Rpb1 subunit of all eukaryotic multisubunit RNAPs and regions of conservation are distributed in clusters over the whole β' subunit and have been grouped into eight domains (Figure 1E). Similar to the β subunit, a non-conserved region of 250 amino acids centred around position 1050 in the β' subunit is dispensable for basic transcription initiation *in vitro* (but not elongation) (Zakharova *et al.*, 1998).

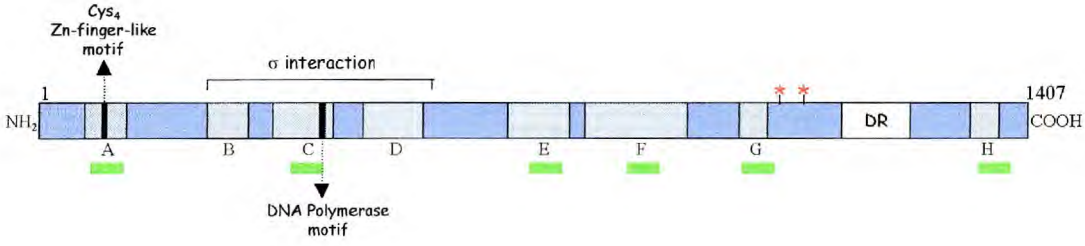


Figure 1E. Domain organisation of the *E. coli* RNAP β' subunit. Red stars indicate the region that cross-link to the 3' end of nascent RNA, and green bars regions involved in transcription termination (see text for details).

The amino-terminal proximal conserved segment A contains a Cys_4Zn finger-like motif, which is assumed to be involved in DNA binding. Equally, within the conserved segment C a region of homology to the DNA binding cleft of the *E. coli* DNA polymerase I suggests the presence of a DNA binding motif in this region. Deletion mutagenesis and Far-western protein-protein interaction assays localise the σ^{70}-type σ subunit (see later) interaction domain of β' to lie within amino acids 260-309 (Arthur *et al.*, 2000; Luo *et al.*, 1996). Studies using protein-protein hydroxyl-radical footprinting further confirm that this segment of β' subunit serves as the main attachment site to σ^{70}-type σ in the holoenzyme

(Travigilia *et al.*, 1999; Owens *et al.*, 1998). As mentioned above, the β' subunit participates in the constitution of the RNAP catalytic centre. Further evidence for this not only comes from the elucidation of the crystal structure of the *T. aquaticus* core RNAP, but also from cross-linking studies, in which the 3' end of a nascent transcript has been crosslinked to a region proximal to conserved segment G (residues M932-W1020) of the β' subunit (Borukhov *et al.*, 1991). Mutations in the β' subunit that affect RNAP pausing and termination are spread through several conserved segments (C, E, F, G and H) suggesting an involvement of the β' subunit in transcription termination (Weilbacher *et al.*, 1994). Studies with temperature sensitive mutants of the β' subunit and protein-protein interaction assays indicate that two regions, a central region between residues 515 and 842, and a second carboxyl terminal proximal region downstream from residue 1141, are involved in interaction with the $\alpha_2\beta$ sub-assembly (Katayama *et al.*, 2000; Nedea *et al.*, 1999).

1.2.4 The omega (ω) subunit

The 10kDa ω subunit is the smallest and the only subunit of the *E. coli* RNAP which is dispensable for cell viability. *In vitro* studies show that RNAP lacking ω is indistinguishable from the control enzyme in transcription experiments (Gentry *et al.*, 1989; Kashlev *et al.*, 1996). In the *E. coli* β' subunit, the last 52 carboxyl terminal amino acids are important for tight anchoring of the β' subunit on the $\alpha_2\beta$ sub-assembly (Nedea *et al.*, 1999). In the *T. aquaticus* core RNAP crystal structure, the ω subunit wraps around the carboxyl terminus of the β' subunit (Zhang *et al.*, 1999). Therefore, it is plausible that ω helps the β' subunit to anchor itself on the $\alpha_2\beta$ sub-assembly, consistent with biochemical data suggesting that the ω subunit functions as a chaperone and promotes RNAP assembly *in vitro* (Mukherjee & Chatterji, 1997).

1.2.5 The sigma (σ) subunit

The σ subunit is the promoter specificity subunit of bacterial RNAPs which confers on the core RNAP the ability of promoter specific transcription initiation. Functional differentiation of RNAP holoenzyme by replacement of the σ subunit was first found in *Bacillus subtilis*, and the sequential pathway of σ replacement during sporulation has been established (see Sullivan & Maddock, 2000; Kroos *et al.*, 1999 for recent reviews). In *E. coli* binding of one of the seven (σ^{70}, σ^{54}, σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{18}) σ factors converts the core RNAP into the holoenzyme, which can recognise different promoters (see later). Based on sequence similarity and mode of function *E. coli* σ subunits are grouped into two classes: all except σ^{54} belong to the σ^{70}-class, whereas σ^{54} belongs to a class of its own. Most genes expressed in exponentially growing cells are transcribed by the RNAP holoenzyme containing the σ^{70} subunit. Thus, in this section emphasis is placed on the *E. coli* σ^{70} subunit which has been the subject of several structural, biochemical and genetic studies.

Sequence comparison of the σ^{70}-family of proteins from different bacteria have led to the identification of four highly conserved amino acid regions (Helmann & Chamberlin, 1988), which are further subdivided into subregions (Lonetto *et al.*, 1992) (Figure 1F).

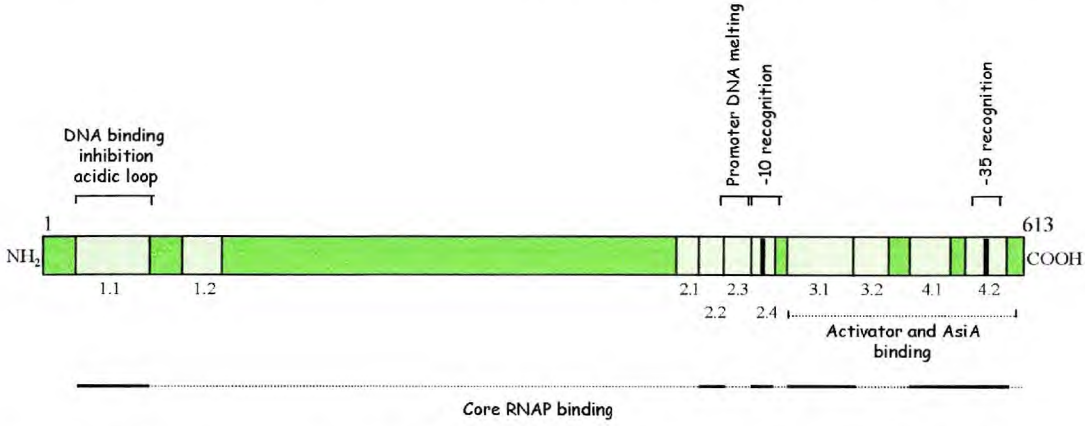


Figure 1F. Domain organisation of the *E. coli* RNAP σ^{70} subunit. The striped bars indicate the conserved regions (see text for details)

σ^{70} confers upon core RNAP the specificity of initiation at promoters with consensus sequences centred around position –35 (5'-TTGACA-3') and –10 (5'-TATAAT-3'), respectively from the transcription start site. The conserved subregion 2.4 is α-helical and recognises the –10 promoter element, whereas a helix-turn-helix motif located in subregion 4.2 recognises the upstream –35 promoter element. 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 amino terminal conserved region 1 and the carboxyl terminal region 4, and that this inhibition is relieved upon formation of the holoenzyme (Malhotra *et al.*, 1997; Dombroski *et al.*, 1992, 1993). The interface of σ^{70} with core RNAP is extensive, conserved and functionally specialised (Sharp *et al.*, 1999). Mutational data suggest that residues in subregions 1.1, 2.2, 2.4, 3.1, 4.1 and 4.2 reduce the interaction of σ^{70} with core RNAP. The locations of these residues are ideally positioned to communicate information concerning DNA binding and transcription initiation from σ^{70} to core RNAP and back. Region 3.1 is in close proximity to the initiating nucleotide (Severinov *et al.*, 1994) and subregions 2.2, 2.4, 4.1 and 4.2 are in close proximity to the two promoter binding determinants (see above). Deletion analysis showed that subregion 2.3 also contains core RNAP binding determinants. Deletion of 25 amino acids from subregion 2.3 causes reduced affinity of the mutant σ^{70} for core RNAP binding (Hernandez & Cashel, 1995).

Subregion 2.3 is highly conserved and contains a high proportion of solvent-exposed amino acids. Also, being adjacent in primary structure to subregion 2.4 that recognises the –10 promoter element in which strand separation is likely to originate, subregion 2.3 is strategically positioned for DNA melting function. Callaci & Heyduk (1998) showed that the *E. coli* RNAP holoenzyme selectively binds short single stranded oligonucleotides that contain the non-template –10 consensus sequence. This selectivity was mapped to region 2.3 (Huang *et al.*, 1997). The tertiary structure of subregion 2.3 shows the side-chains of several aromatic amino acid residues protruding from an α helix (Malhotra *et al.*, 1996). Thus this side of the helix has been proposed to form a "melting motif" that allows the juxtaposition of aromatic amino acid residues to the non-template

single stranded DNA so that they can stabilise the melted state (Fenton *et al.*, 2000; Juang & Helmann, 1995, 1994). Recently, it has been shown that these aromatic amino acids collectively participate in DNA melting and open promoter complex formation (Panaghie *et al.*, 2000).

At class II promoters, activator binding sites overlap the –35 region: in these cases, the activator can contact the σ^{70} subunit and αNTD and αCTD (see later). A small patch adjacent to subregion 4.2 of σ^{70}, located between residues 590 and 603, is involved in interaction with activators (Lonetto *et al.*, 1998). Activators that contact the patch in σ^{70} include AraC , MalT, FNR, CRP and PhoB. Interestingly, these activators affect promoters which do not have a strong –35 consensus element. Therefore, it is assumed that activator-σ^{70} interaction compensates for the absence of the normal interaction between σ^{70} and the consensus –35 promoter element. In addition, the AsiA protein of bacteriophage T4 functions as an anti-sigma factor (see later) that binds to σ^{70} and inhibits transcription initiation by the σ^{70} RNAP holoenzyme (Colland *et al.*, 1998; Orsini *et al.*, 1993). The AsiA protein has been shown to associate with regions 3 and 4 of σ^{70} – specifically with the carboxyl terminal 63 amino acids of subregion 4.2 (Colland *et al.*, 1998; Severinova *et al.*, 1996).

1.2.6 RNA Polymerase: Structure-Function Relationship

The high degree of sequence similarity between the *T. aquaticus* RNAP and the prototypical RNAP from *E. coli* allows the superimposition of biochemical and genetic data from the *E. coli* core enzyme and the structural data from the *T. aquaticus* core RNAP (Zhang *et al.*, 1999). The *T. aquaticus* RNAP has a crab claw like shape, with a deep channel separating the jaws of the claw (Figure 1G-A). The upper jaw is composed mostly of the β subunit and the lower jaw of the β' subunit. Portions of conserved segment I and segments C, G and H of β together with conserved segments H and D of β' interact to form the catalytic centre of the core RNAP. Three aspartic acid residues in the absolutely conserved –NADFDGD- in segment D of β' chelate a Mg^{+2} ion. Interactions between β

segments H and I and β' segment D are responsible for positioning this motif within the catalytic centre. The width of the channel separating the two jaws is ~27Å and appropriate to accommodate the double stranded DNA template. The upper jaw is bilobate: the first lobe (left of the Mg^{+2}) corresponds to β DR1 and portions of conserved segment C, the second lobe (right of the Mg^{+2}) corresponds to the remainder of segment C and segment D. In addition to the main DNA channel between the two jaws, the lower jaw contains a second channel which runs through the molecule. In *E. coli*, a fusion between the carboxyl terminus of β and amino terminus of β' shows no defects *in vitro* or *in vivo* (Severinov *et al.*, 1997). In the *T. aquaticus* structure the β carboxyl terminus and β' amino terminus are adjacent to one another and make extensive contact with each other at the base of the channel. The α_2 dimer interacts with the base of the channel, consistent with the main role of the α_2 dimer in RNAP assembly (see above). On the *T. aquaticus* structure β regions F, G, H and I contact the NTD of αI with the primary interface being region H. β' regions C, D, G and H contact exclusively the NTD of αII. The interactions that β and β' make with the α subunits closely match the hydroxyl radical footprinting data (Heyduk *et al.*, 1996). The ω subunit wraps around portions of the lower jaw, which corresponds to a region in the carboxyl terminus of the β' subunit.

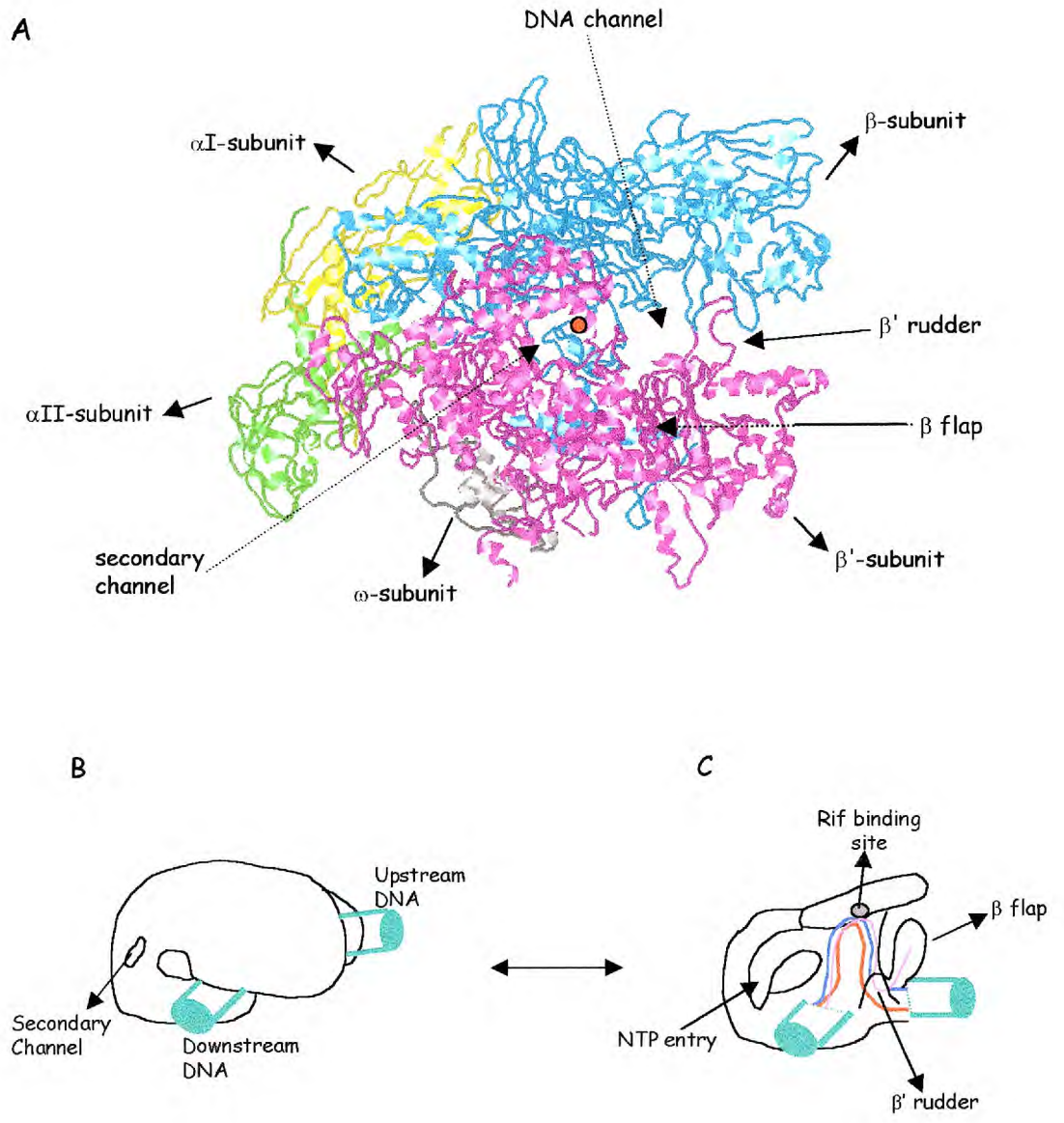


Figure 1G. The structure of *T. aquaticus* core RNAP. (**A**) Crystal structure at 3.3Å (Zhang *et al.*, 1999). Subunits are as indicated. Red dot represents the Mg^{+2} ion in the catalytic centre. (**B**) and (**C**) Schematic model of the structure of a transcription elongation complex (adapted from Zhang *et al.*, 1999). Shown in pink is the RNA transcript, in red the template strand and in blue the non-template strand (see text for details).

The shape and size of the *T. aquaticus* core RNAP correspond well with the molecular contour of the *E. coli* core RNAP elucidated by cryo-electron microscopy (Finn *et al.*, 2000). In a model recently described (Mooney & Landick, 2000; Zhang *et al.*, 1999), the area of the channel positioned left of the Mg^{+2} is said to contact the downstream

double stranded DNA, whereas the area on the right makes contacts with the DNA-RNA heteroduplex in the transcription complex. A structural element, called the "β flap", consisting of β subunit conserved segment G is loosely connected to the rest of the enzyme. In *E. coli*, the β flap is supposed to interact with the termination hairpin of the nascent RNA (Wang *et al.*, 1994). Another structural element called the "β' rudder", consisting of portions of β' conserved segment B and the area between segments B and C, protrudes from the lower jaw into the DNA channel. It is proposed that the β flap and the β' rudder determine the RNAP holoenzyme's processivity during transcription. For instance, the β flap may close the DNA binding channel and determine the RNAP holoenzyme's processivity during elongation, whereas the β' rudder is said to initiate localised melting of DNA during transcription initiation or also to separate the RNA transcript in the RNA-DNA heteroduplex during transcription elongation (Korzheva *et al.*, 2000). The σ subunit makes extensive contacts with the core RNAP polymerase β and β' subunits as determined by genetic, biochemical and structural studies (Finn *et al.*, 2000; Sharp *et al.*, 1999; Owens *et al.*, 1998 and see later). The floor of the upstream portion of the lower jaw (β' subunit) has a platform like coiled-coil structure (Zhang *et al.*, 1999). This platform corresponds to residues 260-309 of the β' subunit and has been shown to interact with the subregion 2.1 of the σ^{70} subunit (Owens *et al.*, 1998; Burgess *et al.*, 1998; Arthur *et al.*, 1998; Luo *et al.*, 1996). The crystal structure of the σ^{70} shows that subregion 2.1 is very close to the subregion 2.3 which is involved in promoter DNA melting (Panaghie *et al.*, 2000; Malhotra *et al.*, 1996). Given that β' residues 260-309 are adjacent to residues that form the β' rudder, it is plausible that subregion 2.3 and the β' rudder form a "melting motif" that is involved in the nucleation of promoter melting. The function of the secondary channel is thought to supply the NTP substrates to the catalytic centre (Zhang *et al.*, 1999).

In the transcription elongation complex (TEC), the nascent RNA-DNA hybrid is separated by the β' rudder and the RNA is routed through a pocket in β that binds rifampicin (an antibiotic that inhibits transcription initiation by bacterial RNAPs), thus explaining why TECs are insensitive to this antibiotic (Korzheva *et al.*, 2000). The

separated single stranded RNA is suggested to exit through a groove underneath the β flap (Zhang *et al.*, 1999) and this view is supported by cross-linking studies (Korzheva *et al.*, 2000). A schematic model of the TEC is shown in Figure 1G-B and 1G-C.

A major constraint in using atomic structures to model functional data is that conformational changes of the RNAP are not considered. It seems doubtless that major conformational changes will occur in the core RNAP during transcription initiation and in TECs. For instance, binding of σ^{70} to the core RNAP induces a ~100° rotation of the β' conserved segments G and H away from the catalytic centre (Finn *et al.*, 2000). Nevertheless, emerging structural information of the bacterial and eukaryotic RNAPs by x-ray crystallography, cryo-electron microscopy, cross-linking and footprinting studies is beginning to unveil the mechanism of transcription initiation by multisubunit RNAPs.

1.3 Transcription Initiation: A Two-Step Chemical Reaction

Transcription initiation is a complicated process involving the participation of the RNAP in a series of molecular and chemical reactions, involving protein-protein and protein-DNA interactions: promoter location by RNAP, formation of a competent initiation complex, synthesis of the initial phosphodiester bonds, and movement of RNAP from the promoter as it enters the elongation phase. A major part of the control of gene expression in prokaryotes and eukaryotes occurs at the level of transcription initiation – the first and generally the rate limiting step of transcription. Transcription "through put" for a given promoter is driven by a mixture of thermodynamics and kinetics: thermodynamics because to maximise transcript yield a promoter must initially bind a RNAP (in competition with other promoters and non-specific DNA sites), and kinetics because it must then move this promoter bound RNAP through the initiation phase of transcription into elongation as rapidly as possible to "clear" the promoter and make it available for re-use. Thus, in its simplest form, the process of transcription initiation can be regarded as a two-step chemical reaction that includes the formation of only one transition state (Figure 1H).

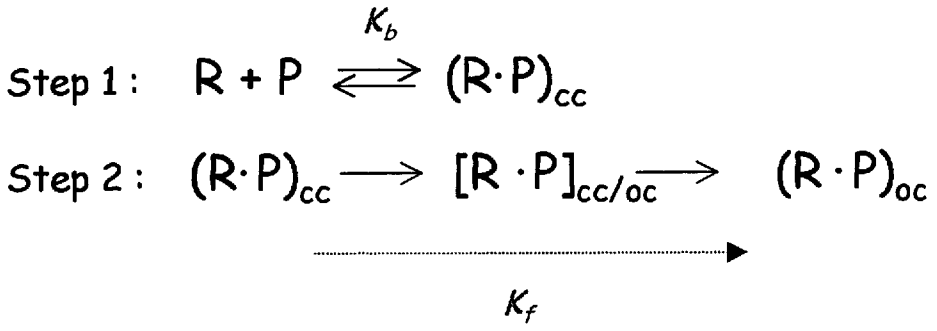


Figure 1H. Transcription initiation is a two step chemical reaction (see text for details).

The first step of the reaction involves the initial binding of the RNAP (R) to the promoter (P) to form a stable closed complex ($R{\cdot}P_{cc}$). K_b is an equilibrium binding constant which is characterised as a reversible, RNAP concentration dependent step. The second reaction is the isomerisation of the closed complex to an open complex, ($R{\cdot}P_{oc}$), via the transition state ($R{\cdot}P_{cc/oc}$), with an irreversible rate constant K_f.

The regulation of gene activity is often the result of regulating the frequency of transcription initiation by the RNAP holoenzyme. The total number of core RNAPs in a growing *E. coli* cell is ~2000 (Ishihama , 1981, Ishihama *et al.*, 1976), which is less than the total number of genes (~4000) on the *E. coli* genome (Blattner *et al.*, 1997). These observations accentuate how important it is for RNAP to choose which genes to transcribe and how often. The functional specificity and activity of the RNAP is modified at two levels: (1) through the action of regulatory proteins that stimulate (activate) or inhibit (repress) transcription initiation, and (2) by the modification of the core RNAP σ subunit (reviewed in Ishihama, 2000).

1.4 Regulation of Gene Expression: Modulation of RNAP by Transcription Regulatory Factors

Each RNAP holoenzyme recognises and transcribes a different set of genes but transcription of some genes additionally requires accessory proteins or nucleotide factors.

The functional specificity of RNAP holoenzyme is modulated through the activity of one or (sometimes) more of 100-150 different transcription factors, most which act from specific binding sites on DNA, and most of which interact with one or more of the RNA polymerase subunits (α, β, β' and σ) to influence the rate constant K_b or K_f (Figure 1H). Transcription factors are classified as activators, that increase K_b or K_f, or as repressors if they are known to inhibit transcription. Transcription factors are classified into four classes (Class I-IV) based on their contact subunit(s) on the RNAP (Figure 1I). In the following subsections the action of transcription activators and repressors in regulating gene expression by modulating K_b or K_f is very briefly touched upon.

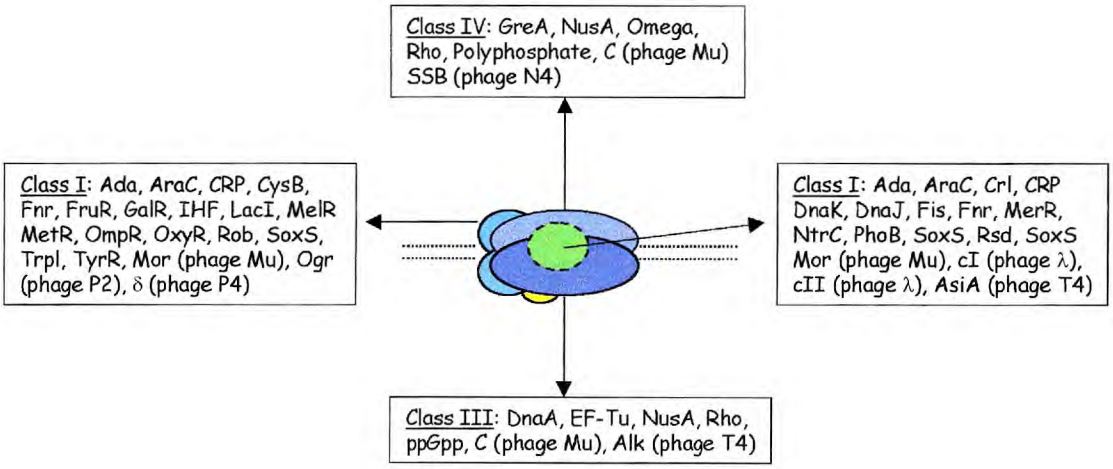


Figure 1I. Classification of *E.coli* transcription factors. Transcription factors have been classified based on their contact subunit of the RNAP. Shown in dark blue is the β subunit, light blue β' subunit, in cyan the α subunits and in green the σ subunit (adapted from Ishihama, 2000b).

1.4.1 Simple Transcription Regulation by Class I Transcription Activators

Transcription regulatory proteins which activate transcription initiation can be categorised into one of four classes, depending on the location of their target contact site in the RNAP holoenzyme (Ishihama, 2000a, 2000b; Busby and Ebright 1994). Class I activators contact target sites in the C-terminal domain of the RNA polymerase α subunit (αCTD), whilst Class II activators contact target sites in the RNAP σ subunit. Activator

proteins which contact the β or β' subunit are designated as Class III and Class IV activators, respectively. The most studied example of a Class I activator is the cyclic AMP receptor protein (CRP) of *E. coli* when it is activating transcription from positions upstream of the promoter as at *lac* P1. In response to elevated intracellular cAMP levels, CRP binds to a 22bp 2-fold symmetrical consensus sequence centred at –61.5 with respect to the transcription start site and enhances the *lac* P1 promoter strength (see reviews by Kolb *et al.*, 1993; Ebright, 1993). A CRP-induced DNA bend of ~90° is seen in the CRP-DNA crystal structure (Schutz *et al.*, 1991). Several mutants deficient in the production of functional CRP have been isolated and shown to be also defective in transcription activation at the *lac* promoter but not defective in DNA binding and DNA bending (Bell *et al.*, 1990; Zhou *et al.*, 1993a). In addition, several lines of evidence show that transcription activation at the *lac* promoter involves protein-protein interactions between a surface loop of the carboxyl-terminal domain of CRP and the α-subunit CTD of RNAP (Igarashi and Ishihama, 1991, Zhou *et al.*, 1993b; Heyduk *et al.*, 1993; Niu *et al.*, 1994). It is believed that the CRP-induced bending facilitates this physical contact by placing the activator in a favourable position for interacting with RNAP or by stabilising the RNAP-CRP interactions. Therefore, cooperativity between CRP and RNA polymerase at the *lac* promoter increases K_b (Figure 1J-A).

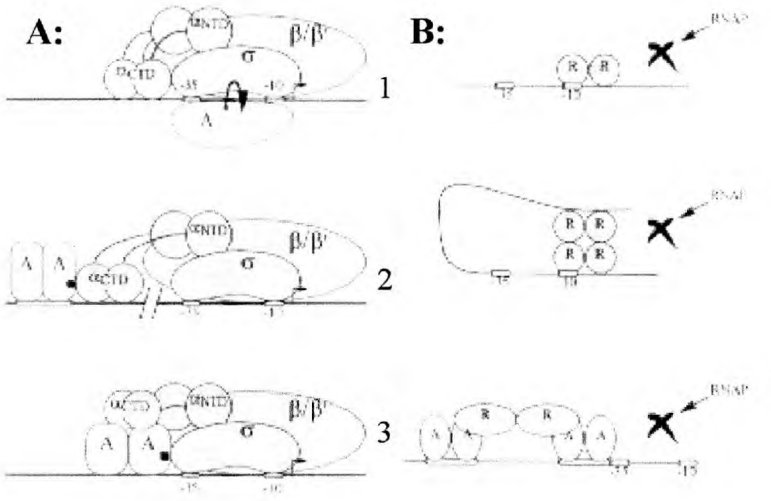


Figure 1J. Schematic diagram of three forms of transcriptional regulation of gene expression. (**A**) Simple transcription activation: (1) by binding to promoter DNA, (2) by interacting with RNA polymerase core subunits (shown here: α interaction) (class I), and (3) by interacting with the sigma subunit (class II). (**B**) Transcription repression: (1) by simple and complex (2) interaction with promoter DNA, and (3) by binding to transcription activator proteins. (Busby *et al.*, 1998).

1.4.2 Simple Transcription Regulation by Class II Transcription Activators

A well studied example of a Class II activator, the bacteriophage lambda cI protein, is essential for open complex formation at the lambda P*rm* promoter and accelerates the isomerisation of the closed complex to an open one, i.e. increases K_f, at the phage lambda P*rm* promoter. The cI protein interacts with the σ subunit and enables it to bind to the -35 promoter element, consequently tightening the binding of RNAP to the promoter (Hochschild *et al.*, 1983; Li *et al.*, 1994), thereby also increasing K_b. A mutant σ which contains a substitution near the carboxyl terminus has been reported to be able to suppress an activation defect caused by a mutation in the activating region of the cI protein (Li *et al.*, 1994; Kuldell & Hochschild, 1994).

CRP activates the *gal*P1 promoter from a binding site centred at –41.5 with respect to the transcription start site, which therefore overlaps with the (here extremely weak) –35 binding site for σ^{70}. The footprint of RNAP in the presence of CRP shows that RNAP is

bound to the DNA immediately upstream and downstream of the CRP site (Belyaeva *et al.*, 1996). This close proximity between CRP and RNAP allows the establishment of three direct contacts between the CRP and the RNAP: between the (1) α-subunit carboxyl terminus and CRP (Class I), (2) α-subunit amino terminus and the CRP (Class I), and (2) σ-subunit and CRP (Class II) (Busby and Ebright, 1997; Kumar *et al.*, 1994).

A distinctive group of activators function at σ^{54}-dependent promoters by modulating the isomerisation step (K_f), by contacting the σ^{54} subunit in the RNAP holoenzyme (Cannon *et al.*, 2000; Wedel & Kustu, 1994). These activators bind to sites located ~100bp upstream of the transcription start site and contact the σ^{54} subunit by means of looping out the intervening DNA between the activator binding site and the promoter. They cause promoter DNA isomerisation by hydrolysing nucleoside triphosphates. Transcription activation at σ^{54}-dependent promoters will be discussed in detail in the later sections of this chapter.

An additional type of transcription activation involves the binding of an activator protein to promoter DNA to aid initiation of transcription by altering the conformation of promoter DNA so that the -10 and -35 promoter elements become aligned correctly to facilitate RNAP binding (Ansari *et al.*, 1992; Summers, 1992). For example, at the *pmerT* promoter the activator protein MerR in the presence of Hg^{+2} twists the promoter DNA appropriately, so that the -10 and -35 elements become aligned correctly, to enable RNA polymerase binding.

1.4.3 Complex Transcription Regulation

Three forms of transcription activation by transcription activators are illustrated in Figure 1J-A. The fact that a variety of different activator-RNAP interactions can activate transcription suggests that no special protein-protein interaction is required for recruiting the RNAP to the promoter DNA. This is supported by two further observations. First, in the case of the ribosomal RNA promoters, the activator-RNAP interaction can be replaced by a protein-DNA interaction: these promoters work at a high level in the absence of activators, an effect attributable to a DNA UP element (AT-rich sequence of 20bp) that is

specifically recognised by the αCTD (Chen *et al.*, 1994). Direct physical contact between the UP element and the α-subunit has also been demonstrated using a fluorescent probe (Ozoline *et al.*, 1991) and by protein tethered chelate methods employing FeBABE (Murakami *et al.*, 1997). Secondly, activator-polymerase contacts have synergistic effects. That is, the level of transcription elicited by two contacts is significantly greater than the sum of the levels elicited by each single contact. Dove *et al.* (1997) showed this by replacing the α-carboxyl domain with the carboxyl domain of the lambda cI protein. cI is a two-domain protein that binds DNA as a dimer, and pairs of dimers bind co-operatively to adjacent operator sites. The amino-terminal domain contacts the DNA and can interact with the σ subunit (see above). The carboxyl-terminal domain mediates both dimer formation and the dimer-dimer interaction that results in cooperativity (Sauer *et al.*, 1990). When the RNAP consisting of the chimeric α-subunit is used for transcription at the lambda P*rm* promoter, the cI protein makes two contacts: (1) the natural contact between its amino terminal domain and the σ subunit, and (2) an artificial contact involving the lambda repressor tetramerisation interaction. Together these two interactions elicit a much greater level of activation than either single interaction (Dove *et al.*, 1997).

The regulation of most promoters is complex and expression is not simply dependent on one activator. In many cases two activators are involved: one activator is a "global" regulator, sensing a global metabolic signal and interacting at a large number of promoters, whilst the other is a "specific" regulator, triggering specific responses to a specific inducer at a very small number of promoters (Busby *et al.*, 1998). However, apparent co-dependence may sometimes be due not to both activators binding at the promoter, but due to the synthesis of one activator being dependent on the activity of the other. The expression of the *mal* and *mel* operons, encoding genes for the catabolism of maltose and melibose, respectively, is dependent on the global regulator CRP and operon-specific regulators MalR and MelR. CRP binding to DNA is triggered by elevations in cAMP levels (due to glucose starvation) and MalT and MelR are activated by maltose and melibiose, respectively. The expression of the *melAB* operon (encoding an α-galactosidase and a melibiose transport protein) is dependent on CRP and MelR, and yet the *melAB* promoter is activated by MelR alone (Webster *et al.*, 1988). However, the *melR* promoter

is totally dependent on CRP, thus explaining the dependence of *melAB* expression on two activators. In contrast, although the synthesis of MalT, the maltose-specific activator is similarly dependent on CRP, expression from several promoters of the maltose regulon requires simultaneous binding of both CRP and MalT for transcription initiation.

1.4.4 Transcription Regulation by Repression

Regulatory proteins which inhibit the initiation of transcription are known as repressors. Repressors work to prevent or interfere with the mechanism of transcription initiation. The manner in which repressors affect the mechanism of transcription initiation is similar to that used for the transcriptional activators, mainly affecting the K_b or K_f. The mechanism of their action can be categorised into two types: (1) simple repression, and (2) complex repression (Busby *et al.*, 1998). In the first case, transcription initiation is inhibited by the binding of the repressor directly to a promoter element which (the repressor protein) occludes the binding of the RNAP holoenzyme to that promoter (Muller-Hill, 1996). The *lac* repressor can be considered as a paradigm for this type of simple repression. In some instances interaction between a repressor bound to a promoter element and a distally-bound repressor enhances repression at the promoter site via a looping mechanism (Phillips and Stockley, 1996; Muller-Hill, 1996). A characteristic example of this type of repression includes the binding of the AraC protein (in the absence of arabinose) to distant sites and binding to the -10 promoter element via DNA looping (Aki *et al.*, 1996). In the complex repression events, the repressor acts as an anti-activator, in which case it binds to the activator protein and inhibits transcription activation by the activator protein. The best studied example of this type of repression is CytR, which represses the expression of genes, scattered at approximately 10 different locations on the chromosome, involved in nucleoside catabolism and transport (Sogaard-Andersen *et al.*, 1991). The three modes of repression of transcription are shown in Figure 1J-B.

1.5 Regulation of Gene Expression: Modulation of RNAP Specificity by the Sigma Subunit

1.5.1 Sigma Factors

The *E. coli* core RNAP is converted into different forms of holoenzyme by binding one of the seven different σ subunits (Table 1.2). This replacement of the σ subunit of RNAP is the most efficient way of altering its promoter recognition properties. The ability of RNAP to initiate selective transcription of genes from promoters is dependent on the σ subunit associated with it. Thus, gene expression regulated at the level of transcription initiation is dependent upon the type of σ factor polypeptide associated with the core RNAP hololenzyme. There are two classes of bacterial σ factors (reviewed in Helmann & Chamberlin 1988). The σ^{70}-class of σ-factors includes σ^{70}, the major vegetative σ-factor encoded by the *rpoD* gene in *E.coli*, and other structurally and mechanistically related members, which are known as alternative σ factors (in *E.coli* these include: σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{18}) within the σ^{70}-class. The *rpoN* (also *ntrA*, *glnA*) encoded σ^{54} factor bears little obvious resemblance to the members of the σ^{70}-class of σ-factors in its primary structure, function (other than core RNAP binding) or mode of action (it requires a specialised activator protein and nucleotide hydrolysis – see later).

Sigma Species	Size (aa)	Gene	[fmol/mg] Expon.	[fmol/mg] Stat.	Targets
σ^{70} (σ^{D})	613	*rpoD*	150	150	exponential growth related genes
σ^{54} (σ^{N})	477	*rpoN*	35	35	various metabolic functions
σ^{38} (σ^{S})	362	*rpoS*	-	55	stationary-phase genes
σ^{32} (σ^{H})	284	*rpoH*	-	-	heat-shock genes
σ^{28} (σ^{F})	239	*rpoF*	90	90	flagella-chemotaxis genes
σ^{24} (σ^{E})	202	*rpoE*	-	-	extracytoplasmic and periplasmic genes
σ^{18} (σ^{FecI})	173	*fecI*	-	-	extracytoplasmic and ferric-uptake genes

Table 1.2. *E. coli* σ-factors. Shown in green are those belonging to the σ^{70}-class, and in red those belonging to the σ^{54}-class. The intracellular concentrations of all 7 σ subunits were determined for both exponential and stationary phases of *E. coli* W3110 strain (Maeda *et al.*, 2000; Jishage *et al.*, 1996; Jishage & Ishihama, 1995).

Alternative σ-factors and σ^{54} are involved in both reversible (adaptive) changes and irreversible (developmental) pathways of gene expression. The former category includes changes in gene expression which adapt cells to alterations in the environment. The latter category includes programmed changes in gene expression which transform cells of one cell type into another cell type. A paradigm for programmed changes in gene expression is sporulation in *B. subtilis*, where a cascade of alternative σ-factors is involved (see Sullivan & Maddock, 2000; Kroos *et al.*, 1999 for recent reviews). In *E. coli*, most housekeeping genes expressed during exponential growth are transcribed by the RNAP holoenzyme containing the σ^{70} subunit ($E\sigma^{70}$). The holoenzyme $E\sigma^{54}$ is involved in transcribing genes of various metabolic functions, historically those associated with deficiency of nitrogen (reviewed in Buck *et al.*, 2000). The transcription of stationary phase and stress related genes is dependent upon the RNAP holoenzyme containing the σ^{38} subunit (reviewed in Ishihama, 2000). $E\sigma^{32}$ transcribes the genes for heat shock proteins (Yura *et al.*, 1993); $E\sigma^{28}$ is needed for the expression of some of the flagella genes and chemotaxis genes (Helmann 1991); and $E\sigma^{24}$ is responsible for the transcription of genes that are needed to cope with damage to extracytoplasmic proteins caused by heat or other factors (Raina *et al.*, 1995). σ^{FecI}, like σ^{24} is a member of the extracytoplasmic function subfamily of σ factors and is involved in transcriptional activation of the *fec* operon (Angerer *et al.*, 1995).

The total number of σ molecules of all types is less than the total number of core RNAP molecules (Jishage *et al.*, 1996, Jishage & Ishihama, 1995). This is consistent with the fact that the RNAP is not always associated with the σ subunit during processive RNA chain elongation. The concentration of each σ subunit is subject to variation depending on cell growth, although the concentration of core RNAP enzyme stays constant at a level characteristic of the rate of cell growth. The intracellular concentration is highest for the σ^{70} subunit in both exponential and stationary phases of cell growth (Table 1.2 and Maeda *et al.*, 2000; Jishage *et al.*, 1996, Jishage & Ishihama, 1995). In exponential phase, two of the alternative σ subunits, σ^{54} and σ^{24}, are present in significant concentrations. The level of σ^{38} becomes detectable only in the stationary phase (Maeda *et al.*, 2000a; Jishage *et al.*, 1996, Jishage & Ishihama, 1995). In growing *E. coli* cells, the total number of RNAP is

2000 molecules/cell, of which only 1300 molecules are involved in transcription (Maeda *et al.*, 2000b). Studies on the binding affinities of all seven *E. coli* σ factors to the core RNAP lead to the conclusion that of the 700 "free" RNAPs, 78% (550 molecules) must exist as $E\sigma^{70}$, 8% (55 molecules) as $E\sigma^{54}$ and 14% (95 molecules) as $E\sigma^{24}$ (Maeda *et al.*, 2000b). It seems doubtless that upon entering into stationary phase or encountering stress conditions this distribution will change; the complex regulation leading to a large transient increase in σ^{32} activity is a clear example (Yura *et al.*, 1993). Thus, σ-factors are widely used as a regulatory device of gene expression in bacteria. The association of the σ-factor with core RNAP is an economical way to regulate the expression of large numbers of co-ordinately controlled genes by generating effectively different RNAPs from a single core enzyme. Additionally, the σ-factor replacement not only serves to switch on new gene expression, but also to reduce the potentially "wasteful" expression of genes.

1.5.2 Anti-Sigma Factors

Anti-σ factors function to negatively regulate the activity of the cognate σ subunit by forming a binary complex with it and thereby inhibiting its function. A paradigm for anti-σ factor activity is exemplified in *B. subtilis* during sporulation control (Kroose *et al.*, 1999; Brown *et al.*, 1995). For the seven *E. coli* σ-factors, only three anti-σ factors (for σ^{70}, σ^{28} and σ^{24}) have been to date identified (Jishage & Ishihama, 1999). A protein called Rsd (regulator of sigma D) forms a binary complex with the σ^{70} and inhibits its function. Jishage & Ishihama (1998, 1999) found that Rsd is not expressed during exponential growth, that is, under conditions in which most of the free RNAP is in the $E\sigma^{70}$ form (see above & Maeda *et al.*, 2000b), ready to be used immediately for transcription of growth related genes. During stationary phase, the levels of Rsd increases and inhibits σ^{70} function by binding to region 4 of σ^{70}. Thus, it seems that in the stationary phase, some unused σ^{70} is stored in an inactive form ready to be used when growth conditions improve for the expression of growth related genes (Akira Ishihama, personal communication).

The *rseA* gene, located downstream of the *rpoE* gene (which encodes σ^{24}) within the same operon negatively regulates σ^{24} activity. It has been shown that the amino terminus of RseA interacts directly with σ^{24} (Missiakas, *et al.*, 1997). The FlgM protein forms a binary complex with σ^{28} resulting in the inactivation of the σ^{28} subunit. The contact site for FlgM is located in region 4 of the σ^{28} subunit (Iyoda & Kutsukake, 1995).

Even though σ^{54} is not essential for growth, levels of σ^{54} in *E. coli* seem to remain constant throughout the growth transfer from exponential phase to stationary phase (Table 1.2). In comparison with the number of genes under the control of σ^{54} however, the molar ratio of σ^{54} to σ^{70} is high (Jishage *et al.*, 1996). Therefore, it is possible that σ^{54} may complex with another as yet unidentified protein as in the case of σ^{70}, σ^{28} and σ^{24}. It is postulated that one or more additional factor(s) may check the ability of σ^{54} to initiate transcription from its promoters by preventing σ^{54} interacting with the core RNAP, promoter DNA or activator protein (Mike Jishage, personal communication). Carmona & de Lonrenzo (1999) showed that *E. coli* cells lacking the *ftsH* gene, which encodes FtsH protease, express σ^{54} at normal levels, but behave *in vivo* as if the σ^{54} subunit were inactive, suggesting a link between the FtsH protease and σ^{54} such that the FtsH protease might control σ^{54} activity.

Overall, it appears that anti-σ factors represent a new chapter in the versatile mechanisms employed in regulating bacterial gene expression.

1.6 Regulation of Gene Expression: Other Regulatory Systems

Various other mechanisms exist which regulate transcription in bacteria. One of these, attenuation, allows transcription to begin but when active as a regulatory mechanism cuts transcription off almost immediately, before the first coding sequence of the operon is reached (Yanofsky *et al.*, 1996). Attenuation is particularly important as a regulatory mechanism of many operons encoding enzymes that synthesise amino acids (Antson *et al.*, 1995). Another system, antitermination, works in long, complex operons that contain

internal signals for termination of transcription in addition to their final terminators (Ruthberg *et al.*, 1997) When active the antitermination mechanism prevents termination at one or more of the internal terminators, causing the RNAP to read further downstream in the operon and include more of its coding sequences.

Bacterial genes are also regulated by mechanisms not directly involving positive or negative regulatory factors, sigma factors, attenuation or antitermination. The degree of supercoiling, for example, regulates some bacterial operons (Wang & Lynch 1993, Lilley *et al.*, 1996). Okamoto *et al.*, (1988) found that transcription of the *E. coli* gene encoding the cAMP receptor protein (CRP) is inhibited by an RNA molecule instead of regulatory proteins. This RNA molecule is complementary (antisense RNA) to sequences in the 5′ region of the mRNA copied from the gene. The CRP, when present in high quantity, stimulates transcription of the antisense RNA, which slows or stops transcription of the CRP gene. Presumably, the antisense RNA pairs with the 5′ end of the CRP mRNA, forming a secondary structure which stops transcription at this point.

The variety and multiplicity of mechanisms controlling bacterial gene expression at the level of transcription ties the regulatory mechanisms into networks that provide a balanced, precise, and sensitive adjustment of the entire cell to suit environmental conditions and allow maximum efficiency in the use of available nutrients. The entire mechanism is aided by the fact that bacterial mRNAs are very short-lived. This permits the cytoplasm to be cleared quickly of the mRNAs produced by one set of operons as they are replaced in activity by another set, and reduces the lag time for a cellular response to an environmental change.

The remainder of this chapter is dedicated to the σ^{54} RNAP subunit which confers on the bacterial RNAP holoenzyme a unique form of regulatory mechanism that is common in eukaryotic RNAP II transcription systems, and which is the main subject of my Ph.D. research.

1.7 The RNAP Sigma Subunit Sigma 54

1.7.1 Mechanism of Function

The σ^{54} subunit, in marked contrast to σ^{70} and σ^{70}-class members, binds promoter DNA in the absence of core RNAP (Buck & Cannon, 1992). The σ^{70} class of σ-factors confer promoter specificity to core RNAP by enabling it to recognise promoters with consensus –35 and –10 sequences, whereas σ^{54} allows the RNAP to recognise promoters with the consensus (underlined) TGGAC-N5-TTGCa/t at –24 and –12, respectively from the transcription start site (Barrios *et al.*, 1999). This property implies that the σ^{54} RNAP holoenzyme is involved in a transcription cycle which is different from that of the RNAP holoenzyme containing the σ^{70} subunit or one of its class members (reviewed in Buck *et al.*, 2000). In contrast to the σ^{70}-type RNAP holoenzyme, the holoenzyme containing σ^{54} binds to the promoter DNA in a transcription incompetent closed complex. Isomerisation of this closed complex to form a transcription competent open complex is only initiated in the presence of an activator protein (Figure 1K).

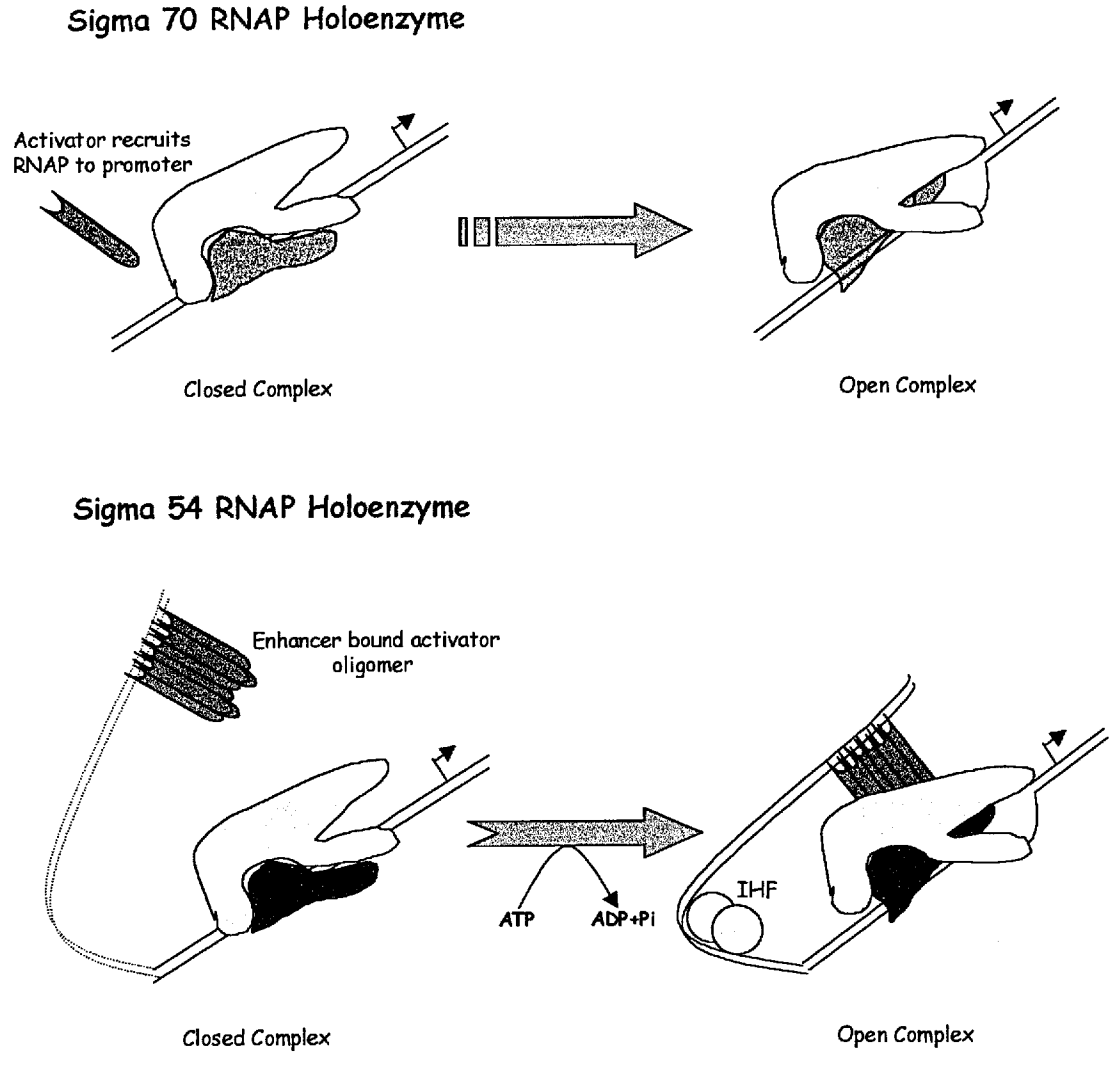


Figure 1K. Cartoon illustrating the functional differences in σ^{70} dependent and σ^{54} dependent transcription initiation (see text for details).

At σ^{70}-type promoters, activation of transcription results from recruitment of the RNAP and stabilised DNA binding (Ptashne *et al.*, 1997 and see above). In contrast, σ^{54} containing RNAP holoenzyme binds to its cognate promoters in the absence of activator protein and other factors to form stable and inactive closed complexes, both *in vivo* and *in vitro* (Buck & Cannon, 1992; Sasse-Dwight & Gralla, 1990). Thus, the function of σ^{54}-dependent activator proteins is not one of recruitment, but to catalyse post-binding steps. The rate limiting step in σ^{54} mediated transcription is the isomerisation of the closed

complex to an initiation competent open complex. Isomerisation requires nucleoside triphosphate hydrolysis by the activator protein and is mediated by direct activator-σ^{54} interactions within the closed complex (Guo *et al.*, 2000; Cannon *et al.*, 2000; Wyman *et al.*, 1997; Wedel & Kustu, 1995; Chaney *et al.*, in preparation). Open complex formation is initiated when oligomers of the activator protein, bound to an enhancer-like sequence located more than 100bp upstream of the transcription start site, interacts with the σ^{54} RNAP holoenzyme at the promoter site by DNA looping (Wyman *et al.*, 1997; Popham *et al.*, 1989). For this reason of similarity to the eukaryotic RNAP II activation system, σ^{54}-dependent activators are called prokaryotic enhancer binding proteins (EBPs). At some promoters, the DNA looping event is facilitated by the DNA bending protein called the integration host factor (IHF) (Hoover *et al.*, 1990). This effect can be mimicked by HU or even the mammalian non-histone chromatin protein HMG-1 and can be bypassed by intrinsically curved DNA (Perez-Martin & Lorenzo, 1997, 1995). Interestingly, IHF has been proposed to stimulate recruitment of σ^{54} RNAP holoenzyme at at least one promoter (Bertoni *et al.*, 1998), but this seems a rare example.

The progression from the closed to an open complex proceeds through numerous conformational adjustments in the RNAP holoenzyme and promoter DNA. In particular, σ-core RNAP and RNAP holoenzyme-DNA interfaces are likely to be significantly altered to allow correct positioning of the catalytic centre upon open complex formation. σ^{54} and σ^{70} are clearly distinct in their mechanisms of open complex formation. The association of σ^{54} causes the RNAP holoenzyme to change its regulatory properties and to assume an enhancer responsive mechanism. Both σ^{54} and σ^{70} promoters have two consensus elements, although the recognition sequences are entirely different (Barrios *et al.*, 1999; Gross *et al.*, 1992). These two elements allow recognition of the DNA in double-stranded form. In both cases the subsequent DNA melting originates within the promoter proximal of the two elements (Cannon *et al.*, 2000; Chen & Helmann, 1997; Morris *et al.*, 1994). Recent work from our laboratory (Cannon *et al.*, 2000; Cannon *et al.*, 1999) and others (Guo *et al.*, 2000; Fenton *et al.*, 2000; Guo *et al.*, 1999; Guo & Gralla, 1998) with stably melted heteroduplex promoter DNA fragments or fork junction promoter DNA probes, that mimic the various transition stages of the promoter DNA during open complex formation,

have revealed that both types of holoenzymes bind optimally to a fork junction DNA structure at a common –12/-11 location, relative to the transcription start site. Therefore, all σ-factors have a common ability to direct selective recognition of a fork junction structure that must be created when DNA melting originates. However, the mode of binding to this fork junction structure and regulating DNA melting differs significantly between σ^{70} (and its class members) and σ^{54}. At the initial stages of DNA melting, both RNAP holoenzymes are inhibited from proceeding with the required conformational changes to fully melt the DNA. This inhibition is very strong for σ^{54} RNAP holoenzyme, and weak for the RNAP holoenzyme harbouring the σ^{70} subunit. The strong inhibition of the σ^{54} RNAP holoenzyme may be due to a lack of interaction with non-template strand DNA adjacent to the –12/-11 fork junction structure. This interaction is a critical feature of open complex formation for the σ^{70} RNAP holoenzyme. The two types of holoenzyme differ with respect to which of the two single stranded DNAs that lead out of the fork junction structure are tightly bound. The σ^{70} RNAP holoenzyme binds tightly to the non-template strand. In contrast, the σ^{54} RNAP holoenzyme prefers binding to the template strand. Since extensive interactions along the non-template strand are important for the promoter DNA melting process leading to transcription initiation (Marr & Roberts, 1997), this single-strand DNA binding preference highlights an important functional difference between σ^{54} and σ^{70}. For σ^{70}, elevated temperature can promote the conformational changes to interact with non-template strand DNA downstream of the –12/-11 fork junction structure. σ^{70} promoters do not necessarily require activators, and when they do, the assistance is either in promoter binding or RNAP isomerisation (Hochschild & Dove, 1998; McClure, 1985). For σ^{54}, interaction with the activator coupled to nucleoside triphosphate hydrolysis serves to relieve the inhibition of melting, helping the σ^{54} RNAP holoenzyme to establish (via the σ^{54} subunit) the missing interaction with non-template strand sequences downstream of the –12/-11 fork junction structure by unmasking protein determinants required for non-template single-stranded DNA interaction.

1.7.2. Domain Structure, Organisation and Function

To date, there is no structural information for σ^{54}. From the 36 available rpoN genes sequenced (Appendix A), there is a clear indication of three distinct regions in σ^{54}. The purified σ^{54} from *Klebsiella pneumoniae* is acidic (Cannon *et al.*, 1995), which is largely due to the numerous aspartic and glutamic acid residues of Region II which links the conserved amino terminal Region I to the conserved carboxyl terminal Region III. Region II is not conserved. Each of these three regions (I-III) is required for various steps during the regulated transcription process. The schematic structure of σ^{54} is depicted in Figure 1L.

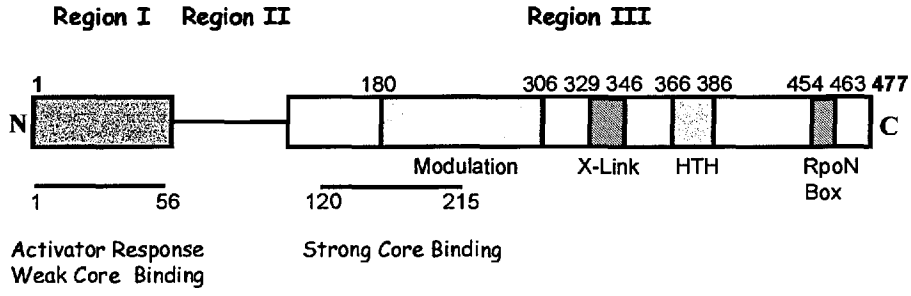


Figure 1L. *K. pneumoniae* σ^{54} domain organisation (see text for details)

Region I. The functional domain organisation of σ^{54} is complex and appears to differ significantly from that of σ^{70}. In particular, the overall DNA binding by σ^{70} is modulated by amino terminal sequences, whereas the amino terminal sequences in σ^{54} are required for activation. Extensive deletion and mutational analyses of σ^{54} have allowed functions to be assigned to different regions of the protein (Buck *et al.*, 2000; Gallegos & Buck, 1999; Cannon *et al.*, 1995; Wong *et al.*, 1994) The amino terminal Region I of σ^{54} is regulatory and closely implicated in polymerase isomerisation and DNA melting (Cannon *et al.*, 2000; Guo *et al.*, 2000; Cannon *et al.*, 1999; Gallegos *et al.*, 1999; Guo *et al.*, 1999). It plays a central role in mediating the response to activator proteins and in binding to early melted DNA structures (Cannon *et al.*, 2000), but is dispensable for overall core

RNAP and DNA binding (Cannon *et al.*, 1999; Wang *et al.*, 1995). Mutant σ^{54}s with certain point mutations in Region I or removal of Region I ($\Delta RI\sigma^{54}$) display an "activator-bypass" phenotype, that is, closed complexes formed with such mutants can initiate transcription in the absence of activator proteins and nucleotide hydrolysis (Casaz *et al.*, 1999; Wang *et al.*, 1995, 1997), suggesting a role for Region I in inhibition of polymerase isomerisation and transcription initiation under normal conditions. Polymerase isomerisation that occurs with the $\Delta RI\sigma^{54}$-RNAP holoenzyme is reversed when Region I is added *in trans*, suggesting that activator overcomes the inhibition caused by Region I *in cis* in a reaction involving nucleoside triphosphate hydrolysis (Gallegos *et al.*, 1999). Some properties of the RNAP holoenzyme that depend upon Region I are likely to rely upon the interaction Region I makes with core RNAP (Gallegos & Buck, 1999); others may depend upon more direct domain communication between Region I and other parts of σ^{54} or σ^{54} and promoter DNA. Hydroxyl-radical footprinting experiments on σ^{54} have shown that the conformation of the carboxyl terminal DNA binding domain is changed when Region I is deleted (Casaz & Buck, 1999). Ortho-copper phenanthroline footprinting of closed complexes with the wild-type σ^{54}-RNAP holoenzyme detects a hypersensitive site around position –12 (Morris *et al.*, 1994). This hypersensitive site is absent in footprints with $\Delta RI\sigma^{54}$-RNAP holoenzymes (Cannon *et al.*, 1995), but is present after the addition of Region I *in trans*, suggesting that Region I is responsible for altering the interactions between holoenzyme and the –12 promoter DNA element.

Region II. Region II which lies between two conserved regions in σ^{54} (Region I and Region III), is very variable both in length and in sequences, although it is characterised by a predominance of acidic residues (Appendix A). Early *in vivo* studies on Region II dealt with a conserved subregion of high amino acid homology and negative charge density in which every third amino acid is an acidic residue. This motif is termed the acidic trimer repeat (ATR) (Hsieh *et al.*, 1994; Wong & Gralla., 1992). The ATR has been implicated in promoter melting at physiological temperatures (Wong & Gralla, 1992). Mutant σ^{54} proteins with deletions of the ATR showed reduced promoter DNA melting. In contrast, mutant σ^{54} proteins with duplications of the ATR showed elevated DNA melting

and transcription initiation. Recent *in vitro* studies on a Region I+II deleted σ^{54} or with mutants containing partial deletions within Region II showed that Region II is dispensable for polymerase isomerisation, but is required for holoenzyme formation and the DNA binding properties of the holoenzyme (Southern & Merrick, 2000; Cannon *et al.*, 1999).

Core RNAP binding. The core RNAP binding interface of σ^{54} comprises at least two functionally distinct sequences: a 95 amino acid residue sequence (120-215) within Region III, which binds core RNAP strongly, and a second sequence within Region I that binds core more weakly (Gallegos & Buck, 1999). Comparison of hydroxyl-radical mediated cleavage profiles of wild-type and Region I deleted proteins indicate that the overall solvent exposed surface of σ^{54} is not greatly altered by deletion of Region I (Casaz & Buck, 1999). Residues protected from hydroxyl-radical mediated cleavage in the holoenzyme made with $\Delta RI\sigma^{54}$ largely coincide with residues protected specifically by DNA in the wild-type holoenzyme, implying a role for Region I in establishing an appropriate conformation of the DNA-binding domain in the wild-type holoenzyme (Casaz & Buck, 1999). Therefore, it is possible that the interaction of Region I with core RNAP is important for controlling DNA binding domain conformation. Residues in the DNA binding domain of σ^{54} exist (see later) that are required for activator-dependent isomerisation of the RNAP holoenzyme (Wang & Gralla, 2000; Chaney *et al.*, 2000; Chaney & Buck, 1999; data in chapters 4, 5 & 7), suggesting that they contribute to the interfaces of σ^{54} that interact with the core RNAP and probably also the DNA, and which may change upon activation. Thus, it appears that the interface between σ^{54} and core RNAP is probably very extensive, as indicated by protein footprinting experiments (Casaz & Buck, 1997; data in chapter 4).

Region III & promoter DNA binding. The primary DNA binding functions of σ^{54} are associated with the highly conserved Region III. Purified fragments yielded by proteolysis of the *K. pneumoniae* σ^{54} protein yielded a minimal 16kDa DNA binding fragment consisting of residues 329-477 which interacts with promoter DNA in a similar way to the wild-type protein (Cannon *et al.*, 1995). Two highly conserved motifs exist within this 16kDa peptide: (1) a putative helix-turn-helix (HTH) motif (residues 366-386)

and (2) a σ^{54} diagnostic motif known as the RpoN box (residues 454-463). A small sequence of amino acids from residues 329-346 spans a region, called the X-link, which UV-crosslinks to promoter DNA (Cannon *et al.*, 1993). A discrete domain between residues 180-306 has been identified which fails to interact with the core RNAP, but greatly enhances the DNA binding of the 16kDa fragment. The mechanism of –12 promoter element recognition by σ^{54} is believed to be complex, involving the putative HTH motif and Region I sequences (Merrick & Chambers, 1992; Gallegos *et al.*, 1999). The –24 recognition is presumed only to occur via determinants in the carboxyl terminus of σ^{54} (Taylor *et al.*, 1996; Guo & Gralla, 1997), although direct evidence is still lacking. Existing data, although indirect, suggest that the RpoN box most likely has a direct role in recognition of the –24 promoter element. The –24 region interactions are known to be dominant for DNA binding by σ^{54}. That is, there are no known σ^{54} mutants that destroy-24 binding, but still allow the holoenzyme to become bound to the DNA. This is in contrast to the existence of many mutations in the carboxyl terminal domain that interfere with –12 element binding but still allow the holoenzyme to become promoter bound (Wang & Gralla, 2000; Chaney *et al.*, 2000; Chaney & Buck, 1999; data in chapter 7). Many mutations that alter –12 recognition also have regulatory defects, which are not seen with any of the RpoN box mutants or other Region III mutants (Wang & Gralla, 2000). This is consistent with the critical fork junction binding by σ^{54} needed to silence transcription prior to activation. Carboxyl terminal mutant σ^{54} proteins with alanine substitutions at residues F318 (data in chapter 7), R336 (Chaney & Buck, 1999), and K388 (Wang & Gralla, 2000) form holoenzymes that are deregulated for polymerase isomerisation (activator bypass mutants). Since the activator bypass phenotype is characteristic of many Region I mutants, it is possible that subdomains in Region III function, perhaps in conjunction with Region I regulatory sequences, to control the regulated properties of the σ^{54} holoenzyme.

The –12 promoter element has multiple effects in σ^{54} dependent transcription. Changes in this promoter region cause changes in transcription levels (Buck & Cannon, 1989; Buck *et al.*, 1985) and *in vivo* work has shown that the sequence at –12 can determine how sensitive the promoter is to a particular activator (Buck *et al.*, 1985; Ray *et*

al., 1990). These and the observation that certain –12 sequences increase "activator bypass" transcription *in vitro* suggest a multiple role for the –12 promoter element in regulated transcription (Wang & Gralla, 1998). In a recent *in vivo* study, in which a consensus σ^{54} promoter sequence ($^{-28}$CT**GG**CACA^{-21} $^{-18}$ATTT**GC**(A/T)T^{-11}) was derived from the alignment of σ^{54} promoters from three related bacteria, it was found that the consensus promoter was not the strongest promoter (Wang & Gralla, 1998). The authors explained this by suggesting that the binding of σ^{54} to the promoter could be too tight for optimal function in the context of a fully consensus promoter. That is, as long as σ^{54} becomes bound to the promoter, other sequences (or sequence combinations which could have been missing from the consensus) may influence how many transcripts are produced by the bound holoenzyme. In the same study, mutational analysis of the –12 consensus promoter showed it to consist of two subregions that behave differently when mutated. Single changes in the upstream TTT consensus subregion led to increased transcription levels, whereas single changes in the downstream GC(A/T) led to decreased transcription, indicating a requirement for some minimal match to the consensus to fully bind the σ^{54} holoenzyme (Wang & Gralla, 1998). This and previous studies clearly suggest that recognition of the –12 promoter region by σ^{54} is complex and is influenced by determinants in σ^{54}, promoter DNA and to some extent environmental cues (Wang & Gralla, 1998).

1.7.3 σ^{54} biology

σ^{54} is implicated in expressing genes whose products have a wide range of different functions. σ^{54} was initially identified as a positive transcription factor for the expression of the *glnA* gene product (glutamine synthetase) in enteric bacteria (Garcia *et al.*, 1977). Later it was discovered that σ^{54} is required for the expression of genes for nitrogen assimilation, hence σ^{54} is also known as σ^{N}. Examination of all gene functions under the control of σ^{54} does not suggest any unifying area of cellular functions in which σ^{54} operates (reviewed in Studholme & Buck, 2000). The σ^{54} dependent promoters in *E.*

coli indicate that σ^{54} is involved in the expression of genes with diverse physiological functions (Table 1.3). Other examples include chemotaxis in *Rhodospirillum capsulatus*, flagellation and cell-cycle dependent stalk biosynthesis in *Caulobacter crescentus* and *Vibrio cholerae*, phenol utilisation in *Pseudomonas putida*, histidine transport in *Salmonella typhimurium*, urea utilisation in *Pseudomonas spp.* and many other cellular functions in many Gram-positive and Gram-negative bacterial species (reviewed in Studholme & Buck, 2000).

Promoter	Gene Function
argT	amino acid transport
fdhF	formate dehydrogenase H
glnA	glutamine synthetase
glnH	glutamine transport
hycA	formate hydrogenase regulatory protein
hypA	hydrogenase synthesis
nac	nitrogen assimilation control
pspA	phage shock
rtcBA	RNA metabolism

Table 1.3. Biological functions for which there is evidence of σ^{54} dependent transcriptional control in *E. coli* (Studholme & Buck, 2000).

The majority of bacterial pathogen genomes investigated so far encode σ^{54}. Examples include *Neisseria gonorrhoeae* (pilin synthesis), *Erwinia chrysanthemi* (protein secretion and pathogenicity), *Pseudomonas aeruginosa* (adhesion to mucin), *Borrelia burgdorferi* and *Enterobacter cloacae* (stress response), *Vibrio anguillarum* (flagellin synthesis), *Chlamydia trachomatis* (as yet unknown developmental process), *Treponema pallidum* (bacterioferritin) and *Haemophilus influenzae* (as yet unknown pathogenic function). However, virtually all genes expressed under a σ^{54} dependent promoter have in common that they are not essential for growth. An exception, has to be made for some genes in *Myxococcus xanthus* because the σ^{54} in this bacterium appears to be essential for survival (Keseler and Kaiser, 1997). Interestingly, σ^{54} is also associated with the expression of several σ^{70}-class σ-factors. Pallen (1999) proposed a role for σ^{54} in transcriptional

regulation of the heat shock σ-factor, σ^{32}. In *Pseudomonas syringae* and *B. burgdorferi* the stationary phase σ^{38} is potentially regulated by σ^{54} (Studholme & Buck, 1999).

The distribution of the *rpoN* gene throughout phylogenetically diverse bacteria underlines its biological importance (Studholme & Buck, 2000). It has been found in Gram-positive bacteria (*B. subtilis*, *Clostridium difficile*), the hyperthermophile *Aquifex aeolicus*, the planctomycete *Planctomyces limnophilus*, chlamydias (*Chlamydia trachomatis*), spirochaetes (*T. pallidum*) and the green sulphur bacterium *Chlorobium tepidum*. Most σ^{54} containing bacterial species contain one single *rpoN* gene, but *Bradyrhizobium japonicum* and *Rhizobium sphaeroides* contain two different copies of *rpoN* genes (Powell *et al.*, 1995; Kulik *et al.*, 1991). Inactivation of one of these genes in *B. japonicum* and *R. sphaeroides* has no influence on the phenotype of these bacteria. Expression of *rpoN* is usually constitutive but in some cases, such as in *R. sphaeroides*, it is regulated by oxygen and nitrogen levels. In *C. crescentus* it is regulated temporally during the cell cycle (Bruno & Sharpio, 1992). In *B. japonicum*, *rpoN1* is regulated in response to oxygen while *rpoN2* is negatively autoregulated (Powell *et al.*, 1995). Negatively autoregulated *rpoN* genes are also found in *K. pneumoniae* (Grande *et al.*, 1999), *P. putida* (Kohler *et al.*, 1994), *Azotobacter vinelandii* (Merrick *et al.*, 1987) and *Rhizobium etli* (Michiels *et al.*, 1998).

1.7.4 In a class of its own - σ^{54} Transcription

Bacterial σ-factors carry out functions that involve many different polypeptides in eukaryotic RNAP II dependent transcription. The σ^{54} factor forms a holoenzyme with some of its functional and regulatory properties shared by eukaryotic multisubunit RNAPs. These include (1) the requirement to be activated by activator proteins bound to enhancer-like sequences, (2) the ability to form stable closed complexes that require an activator for isomerisation, (3) the requirement for activators that probably couple their nucleotide hydrolysis to promote open complex formation, and (4) some suggested structural similarities with eukaryotic activator proteins (Sasse-Dwight *et al.*, 1990; Merrick, 1993). These properties can be summarised by proposing that the mechanism of transcription

initiation and regulation of gene expression by σ^{54} are, at least conceptually, an intermediate step between those mediated by the σ^{70} type of σ-factors, and those involved in eukaryotes. These unique functional features of σ^{54} and its existence in a structural class of its own, provide a model system to study the regulation of gene expression in bacteria and eukaryotes.

Multisubunit RNAPs are targets of sophisticated signal transduction pathways that link environmental and temporal cues to changes in gene expression. Cannon *et al.*, (2000) recently showed that promoter bound σ^{54} undergoes a nucleotide hydrolysis dependent isomerisation on DNA which involved conformational changes in protein and DNA, thus suggesting that the σ^{54} is the target of the activator protein. Current work in the laboratory (Chaney *et al.*, in preparation) confirm that this is indeed the case. Together, these studies provide clear evidence that activator driven changes in σ^{54} conformation trigger the conversion of a transcriptionally silent RNAP conformation to one able to interact productively with template DNA. As in the case for σ^{70} holoenzyme, activation of σ^{54} holoenzyme occurs by signal transduction pathways using numerous and diverse activators (reviewed in Shingler, 1996). Although these pathways are diverse, the most well-studied of σ^{54} activators, exemplify a commonly used gene regulatory mechanism – the two component signal transduction system.

1.8 Two Component Signal Transduction System

1.8.1 Definition and Genomic Distribution

Bacteria are remarkable for their ability to sense and respond rapidly to the chemical and physical composition of their environment, which often results in changes in gene expression. One widely distributed and core mechanism which contributes to this capacity is the two-component signal transduction system. The prototypical two-component signal transduction system consists of a histidine protein kinase (HK), containing a conserved kinase core, and a response regulator protein (RR), containing a

conserved regulatory domain. Extracellular stimuli are sensed by, and serve to modulate the activities of the HK. The HK transfers a phosphoryl group to the RR, in a reaction catalysed by the RR. Phosphotransfer to the RR results in activation of a downstream effector domain that elicits the specific response (Figure 1M).

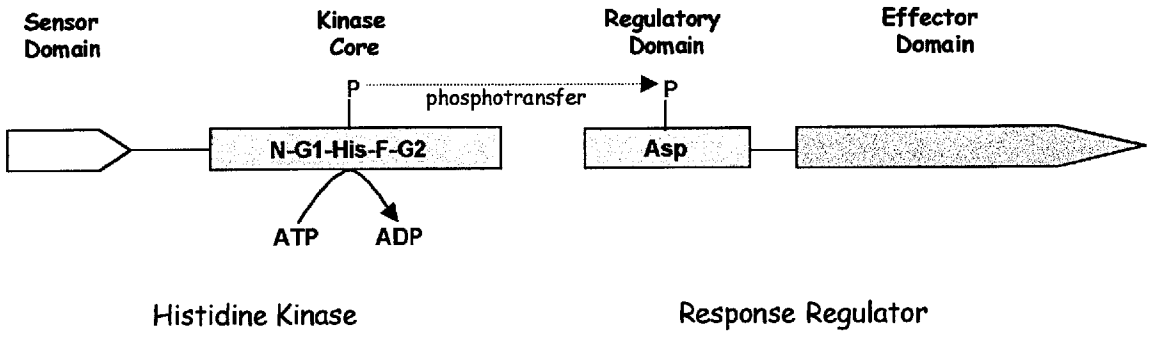


Figure 1M. The prototypical two component signal transduction system (see text for details).

Two-component systems are found in organisms of all domains: eubacteria, archaea and eukaryotes. However, their abundance in each domain differs substantially. In *E. coli* there are 30 HKs (5 of which are hybrid HKs, which contain a RR domain – see later) and 32 RRs (Mizuno *et al.*, 1996). However, the number of two-component system proteins differs greatly in different bacteria, ranging from 0 in *Mycoplasma genitalium* to 80 in *Synechocystis sp.*, in which these proteins account for ~2.5% of the genome (Mizuno *et al.*, 1996). Preliminary analyses of other completed bacterial genomes have estimated the numbers of two-component proteins to be as follows: 70 in *B. subtilis* (Fabret *et al.*, 1999); 9 in *Haemophilus influenzae* (Fabret *et al.*, 1999); 11 in *Helicobacter pylori* (Mizuno, 1998), and 19 in *Thermotoga maritima* (Nelson *et al.*, 1999). Similar analyses of two archaea have estimated 24 two component proteins in *Methanobacterium thermoautotrophium*, and 0 in *Methanococcus jannaschii* (Mizuno, 1998). Only a limited number of two-component proteins have been found in eukaryotes. In the genome of the budding yeast *S. cerevisiae*, there is only one phosphorelay system involved in osmoregulation (Maeda *et al.*, 1994). The fission yeast *Schizosaccharomyces pombe* contains an RR that regulates a stress-activated MAP kinase cascade (Shieh *et al.*, 1997).

Two-component proteins are not only limited to eukaryotic microorganisms; they have also been found in plants such as *Arabidopsis thaliana* (Chang *et al.*, 1993). However, several features distinguish eukaryotic two-component systems from those of prokaryotes. Hybrid HKs that contain RR domains are rare in prokaryotes (5 of 30 in *E. coli*), whereas eukaryotic HKs are almost exclusively hybrid HKs. Prokaryotic RRs are predominantly transcription factors (at least 25 of 32 in E. coli), whereas there is only one known eukaryotic RR with a DNA-binding domain (Brown *et al.*, 1993).

1.8.2 Structure and Function of HKs

Both prokaryotic and eukaryotic HKs contain the same basic signalling components, namely a diverse sensing domain and a highly conserved kinase core that has a unique fold. The overall activity of the kinase domain is modulated by input signals from the sensing domain. HKs undergo an ATP-dependent autophosphorylation at a conserved His residue in the kinase core. Autophosphorylation is a bimolecular reaction between homodimers, in which one HK monomer catalyses the phosphorylation of the conserved His residue in the second monomer. The RR stoichiometrically transfers the phosphoryl group from the phospho-HK to a conserved Asp residue in its own regulatory domain. Therefore, control in two-component pathways is accomplished through the ability of the HK to regulate the phosphorylation state of the downstream RR. Besides directing forward phosphorylation reaction, many HKs possess a phosphatase activity, enabling them to dephosphorylate their cognate RRs (Keener & Kustu, 1998). These bifunctional HKs are commonly present in phosphotransfer pathways that need to be shut down quickly.

The kinase core of HKs is composed of a dimerisation domain and an ATP/ADP-binding phosphotransfer or catalytic domain (Lau *et al.*, 1997). There are five conserved amino acid motifs present in both eukaryotic and prokaryotic HKs (Stock *et al.*, 1989). The conserved His substrate is the central feature in the H-box, whereas the N, G1, F and G2 boxes define the nucleotide binding cleft (Figure 1M). The H-box is part of the dimerisation domain. Sequence and structure analyses of HKs show that the nucleotide binding cleft is in a highly flexible region of the protein. This flexibility may reflect

conformational changes that accompany ATP binding. The amino terminally located sensing domains of HKs share little primary sequence similarity, thus supporting the idea that they have been designed for specific ligand/stimulus interactions.

Most HKs, as exemplified by the *E. coli* osmosensor EnvZ, are periplasmic membrane receptors; EnvZ has two transmembrane regions that separate the protein into a periplasmic amino-terminal sensing domain and cytoplasmic carboxyl terminal catalytic domain that contains the kinase core. Other HKs, such as the *S. meliloti* FixL, contain multiple transmembrane segments (Lois *et al.*, 1993). Not all HKs are membrane bound. The *E. coli* nitrogen regulatory protein NtrB (MacFarlane & Merrick, 1985) is an example of a soluble cytoplasmic HK. Soluble HKs, such as NtrB are regulated by intracelluar stimuli and/or interactions with cytoplasmic domains of other proteins. In transmembrane HKs, the sensing domain is connected to the cytoplasmic kinase core through a transmembrane helix and a cytoplasmic linker. Computational analyses of transmembrane HKs have suggested that this linker may adopt a coiled coil-like motif. It is proposed that the coiled-coils might be used as a structural relay, sensing conformational changes in the periplasmic region of the HK and communicating this signal to the kinase core (Singh *et al.*, 1998).

1.8.3 Structure and Function of RRs

In most bacterial two-component signal transduction systems, RRs are the terminal component of the pathway, functioning as phosphorylation-activated switches to effect the adaptive response. The RR catalyses phosphoryl transfer from the phospho-His of the HK to a conserved aspartic acid residue in its own regulatory domain. Small molecules such as acetyl phosphate, carbamoyl phosphate, imidazole phosphate, and phosphoramidate can serve as phosphodonors to RRs, demonstrating that RR can catalyse phosphoryl transfer independently of assistance from an HK (Lukat *et al.*, 1992). Most RRs also catalyse autodephosphorylation, limiting the lifetime of the activated state. Most RRs consist of two domains: a conserved amino terminal regulatory domain and a variable carboxyl terminal effector domain. The majority of bacterial RRs are transcription factors with DNA-binding

effector domains (25 out of 32 in *E. coli*). Some RRs lack the carboxyl terminal effector domain altogether.

The single domain *E. coli* chemotaxis protein CheY has served as representative model for RR regulatory domains (Stock *et al.*, 1990). The site of phosphorylation in CheY is D57. This residue lies adjacent to other acidic residues, D12 and D13 in CheY. Two other highly conserved residues, T87 and K109, complete the cluster of conserved residues surrounding the active site of the regulatory domain. The carboxylate side chains of the acidic cluster are involved in co-ordination of the Mg^{+2} that is required for phosphoryl transfer and dephosphorylation (Lukat *et al.*, 1990). A majority of effector domains have DNA-binding activity and function to activate and/or repress transcription of specific genes. Detailed analyses of individual RRs have revealed a great deal of complexity in the functioning of these transcription factors, and therefore RRs are divided into three major subfamilies. The *E. coli* OmpR protein belongs to the largest subfamily of RRs, and functions as both an activator and repressor to regulate differentially the expression of the *ompC* and *ompF* genes that encode outer membrane porin proteins. Transcriptional activation by OmpR involves interaction with the α subunit of the RNAP. A second RR subfamily is represented by NarL, a transcription factor that both activates and represses genes involved in nitrate and nitrite metabolism (Iuchi & Lin, 1987). The most structurally and perhaps functionally complex RR subfamily is that represented by the nitrogen regulatory protein NtrC, a transcriptional enhancer that activates the RNAP holoenzyme containing the σ^{54} subunit. The effector region of this subfamily consists of two domains: an ATPase domain and an HTH DNA binding domain (Morret & Segovia, 1993). NtrC dimers, capable of binding to DNA (Weiss *et al.*, 1992), oligomerise into octamers upon phosphorylation (Wyman *et al.*, 1997). Oligomerisation stimulates ATP hydrolysis (Austin & Dixon, 1992; Weiss *et al.*, 1991), which provides energy for open complex formation and activation of transcription (Wedel & Kustu, 1995). The structure of the NtrC amino terminal domain and how it changes on phosphorylation are known, but the way in which phosphorylation controls the oligomerisation of NtrC is not fully understood. Oligomerisation is required for NtrC to hydrolyse ATP and activate the σ^{54} holoenzyme, but not to bind ATP.

The regulatory domains of RRs are thought to exist in equilibrium between two conformation states, inactive and active. Phosphorylation of the regulatory domain shifts the equilibrium towards the active form. The different molecular surfaces displayed in the two forms can facilitate specific protein-protein (or possibly protein-DNA) interactions. In some cases, activation involves a relief of inhibition, as observed in RRs that can be activated by removing the amino terminal regulatory domain (Kahn & Ditta, 1991). In other cases, the phosphorylated regulatory domain plays an active role. Phosphorylation can promote dimerisation (Fiedler & Weiss, 1995), higher-order oligomerisation (Wyman *et al.*, 1997), or interaction with other proteins (Blat & Eisenbach, 1994) or DNA (Aiba *et al.*, 1989). Some proteins use a combination of these mechanisms (Anand *et al.*, 1998). Recently determined structures of the phosphorylated regulatory domains of NtrC (Klose *et al.*, 1993) and FixJ (Birck *et al.*, 1999) show that phosphorylation causes long-range structural perturbations, altering the molecular surface of the regulatory domain. However, phosphorylation does not alter the overall fold or result in any substantial changes in secondary structure. Rather, the secondary structure elements are slightly repositioned, causing backbone deviations of only a few angstroms. These changes, however, dramatically affect the molecular surface, altering both topological and electrostatic features. The phosphorylation induced conformational change affects a large face of the regulatory domain, providing ample molecular surface to be exploited by multiple protein-protein interactions. RRs that are members of the NtrC family also function as ATPases, thus providing energy for open complex formation and activation of transcription by the σ^{54} containing RNAP holoenzymes. Emerging data from our laboratory imply that the ATPase activity induces a conformation change in the effector domain of the RR that enables a σ^{54}-interacting domain within the latter to interact with the σ^{54} subunit and thereby stimulate open complex formation (Chaney *et al.*, in preparation).

1.8.4 The NtrC-subfamily of RRs in *E. coli* : Activators of the σ^{54} RNAP holoenzyme

The σ^{54}-dependent family of prokaryotic transcriptional activators is unique in that a promoter utilising the σ^{54} RNAP holoenzyme requires an activator capable of

hydrolysing ATP to promote RNAP isomerisation and open complex formation. Regulators of this family have a typical domain structure shown in Figure 1N.

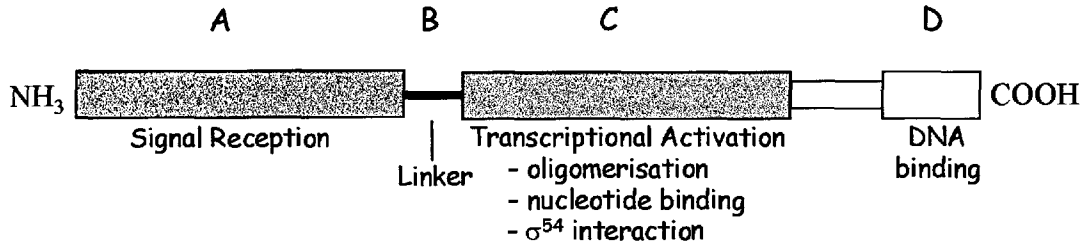


Figure 1N. Domain organisation of a σ^{54}– dependent activator

The central C-domain provides the ATP binding and hydrolysis activity and interacts with the σ^{54} RNAP holoenzyme (Figure 1N). This domain provides the essential transcriptional activation properties common to all the family and therefore exhibits the greatest conservation (Morett & Segovina, 1993). Sequence alignment of C-domains of σ^{54} activators shows that the conserved residues lie in seven regions (C1 to C7). Region C1 is glycine rich and contains a conserved GESGTGK sequence known as Walker A domain and is implicated in mononucleotide binding (Walker *et al.*, 1982). The lysine residue in the GESGTGK motif is suggested to be in close contact with the α-phosphoryl group of ATP (Fry *et al.*, 1986). Region C3 (also known as Switch I region) contains a GAFTGA sequence that is conserved in many σ^{54} activators and is suggested to be involved in σ^{54} RNAP holoenzyme interactions (Morett & Segovia, 1993). Region C4 (switch II) is glycine rich and is implicated in sensing the regulatory signal. In NtrC and DmpR (and ligand response regulator in *Pseudomonas* spp. – see later) mutations in this region lead to constitutively active proteins (Flashner *et al.*, 1995; Shingler & Pavel, 1995). Region C6 is positively charged and rich in aromatics and prolines and region C7 has a core region of eight conserved amino acid residues, has been predicted to be involved in nucleotide binding and is probably involved in binding the nucleotide base (Weiss *et al.*, 1991).

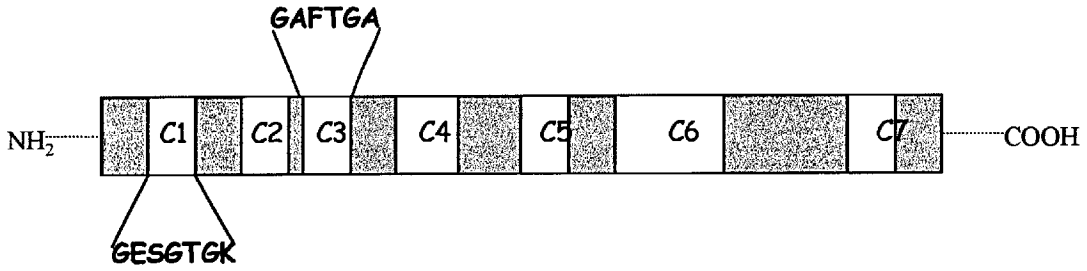


Figure 1O. The central domain of σ^{54} dependent activators (see text for details)

The D-domain contains the DNA binding determinants and shows less homology among family members. It is involved in binding enhancer like sequences known as upstream activator sequences (UAS) located 100-200 base pairs upstream of the transcription start site. The UAS are composed of inverted repeats, which require positioning on the same face of the DNA helix to facilitate the oligomerisation of the activators (Wyman *et al.*, 1997). The amino terminal A-domain has the greatest divergence in protein sequence and this is reflected in distinct activation mechanisms of the family, namely activation by phosphorylation, removal of inhibitory protein-protein contacts or direct binding of small effector molecules (reviewed in Shingler, 1996). The B-domain linker separates the A- and C- domains, and is assumed to be a flexible loop in NtrC and NifA (Wootton & Drummond, 1989). However, the function of a subset of the family is sensitive to mutations in the B-domain, suggesting its significance in their protein function (Perez-Martin & De Lorenzo, 1995). Recent mutational analysis of the B-domain in DmpR of *Pseudomonas* spp. showed that (i) the B-linker itself actively represses ATP hydrolysis, and (ii) that allosteric changes upon effector binding are transduced to alleviate both B-linker repression of ATP hydrolysis and A-domain repression of transcriptional activation (Eric O' Neill, Ph.D. thesis, Ronan O' Toole, personal communication). The three mechanisms by which σ^{54}-dependent activators can sense signals and activate transcription in *E. coli* are briefly mentioned below, i.e. via phosphorylation, protein-protein contacts or direct effector activation.

In *E. coli*, based on the genome sequence and prior genetic and physiological studies, there appear to be 11 transcription activator proteins which interact with the σ^{54}

RNAP holoenzyme to initiate transcription of genes of diverse physiological function (Table 1.4). At least three of these activator proteins (AtoC, HydG and NtrC) are RRs and belong to two-component signal transduction systems (Shingler, 1996 and references within). AtoC and HydG are both strongly similar to NtrC, in that they possess an amino terminal regulatory domain that can be phosphorylated by a HK partner and a HTH DNA binding domain in their carboxyl terminus. In the case of NtrC, the membrane bound HK NtrB controls the level of phosphorylation at D54 in the A-domain of NtrC, in response to nitrogen limitation conditions. The gene sets which are activated by the hypothetical activators YfhA, YgaA and YgeV are, to date, not identified. However, comparison of their amino acid alignments with the other σ^{54} activators show that these proteins contain an amino terminal regulatory domain (with a phosphorylation site) and a central domain which contains a nucleotide binding and σ^{54} interacting domain (Appendix B).

FhlA, PspF, PrpR, RtcR and HyfR lack the amino terminal regulatory domain and thus do not have a cognate HK partner. An example of a protein-protein regulated σ^{54}-dependent activator is PspF, the regulator of the phage shock proteins. This system is positively responsive to such signals as phage infection, heat shock and osmolarity. PspF promotes transcription of the *pspABCDE* operon in a constitutive manner as it completely lacks a regulatory A-domain (Jovanovic *et al.*, 1996). However, negative control of the activator function is imparted by the product of the *pspA* gene, through binding and inhibition of PspF activity (Dworkin *et al.*, 2000; Elderkin & Buck, unpublished data). The third group of σ^{54} dependent activators have the ability to respond directly to their activating signal, that is, in the presence of specific ligands in the medium or cytoplasm, without the requirement for auxiliary sensory proteins. In *E. coli*, the formate-responsive FhlA protein represents this small ligand responsive subclass of σ^{54} dependent activators (Hopper & Bock, 1995). In the absence of its effector molecule, the ATPase and transcriptional promoting activity of FhlA is repressed by A/C-domain interactions. Upon binding to its effector by the A-domain, the A/C-domain interactions are broken and FhlA is free to hydrolyse ATP and activate transcription. Studies on the DmpR protein (which is another member of the ligand responsive subclass of σ^{54} dependent activators) in *Pseudomanas* spp. showed that the action of binding of effectors can be mimicked by

deletion of the repressive A-domain, leading to a constitutively active effector-independent derivative (O' Neill *et al.*, 1998).

Response Regulator	Gene Set(s) Activated	Histidine Kinase
AtoC	acetoacetate metabolism	AtoS
HydG	hydrogenase activity	HydS
NtrC	nitrogen assimilation	NtrB
FhlA	carbon fermentation	not applicable
PspF	phage shock response	not applicable
PrpR	propionate catabolism	not applicable
RtcR	putative *rtcAB* genes	not applicable
HyfR	hydrogenase 4 structural genes	not applicable
YfhA	not known	putative
YgaA	not known	putative
YgeV	not known	putative

Table 1.4. σ^{54} dependent *E. coli* transcription activator proteins. Shown in red are response regulators that belong to two component signal transduction systems, in black are putative two component members, and in green are activator proteins that do not belong too the two component signal transduction systems (see text for details).

1.8 Objectives

The domain organisation and specific DNA binding of σ^{54} and its role in transcription initiation upon activation in response to environmental cues are clearly complex and do not obviously closely resemble the conventional bacterial gene expression and regulatory mechanisms. The general aim of my research was to provide functional and structural information on σ^{54} in a structural framework that will enhance our understanding of how different σ^{54} parts orchestrate the regulation of gene expression.

Initially, two core RNAP binding fragments of σ^{54} (residues 70-324 and 1-306) were used to assess the suitability of a cysteine tethered iron chelate methodology (FeBABE methodology) (reviewed in Ishihama, 2000c) to study the protein-protein and protein-DNA interactions that occur in σ^{54} dependent transcription.. The promising results of this initial study prompted us to make single cysteine mutants of σ^{54} at strategic

positions in Regions I and III of σ^{54}. The FeBABE derived σ^{54} proteins were used as chemical proteases and nucleases to elucidate the core RNAP and promoter DNA binding interface of σ^{54}, respectively.

A σ^{54} protein FeBABE modified at residue R383 in the HTH motif failed to show any proximity to promoter DNA and core RNAP, even though previous genetic studies implicated R383 as having a role in recognising the G of the consensus GC element at the –12 promoter element (Merrick & Chambers, 1992). We made the R383A and R383K mutants to biochemically characterise the role of residue R383 in σ^{54} function *in vitro*.

Casaz & Buck (1997) showed that σ^{54} undergoes specific alterations in protease sensitivity at different steps on the pathway leading from free σ^{54} to one engaged in a transcription complex. We mutated surface exposed residues in Region III of σ^{54} that were protected from protease attack by the core RNAP in the holoenzyme or by promoter DNA in the closed complex to alanine and glycine in order to seek a correlation between the protein footprinting and the functional importance of the amino acids. The *in vivo* and *in vitro* activities of these mutant σ^{54} proteins were studied in an attempt to identify potentially significant residues for FeBABE footprinting work.

Finally, small-angle X-ray Scattering (SAXS) was used to evaluate the envelope shape and tertiary domain organisation of σ^{54}. We used SAXS measurements on σ^{54} fragments comprising residues 70-324, 1-306, 117-477 and the full-length protein 1-477 to yield low resolution structural information on σ^{54}. The availability of structural information on the tertiary domain organisation of σ^{54} provided us with a valuable tool for interpreting FeBABE footprinting data.

The data presented in the remainder of this thesis involving the FeBABE methodology and SAXS were results of collaborative efforts with Professor Akira Ishihama at the National Institute of Genetics (NIG) at Mishima, Japan and Dr Dimitri Svergun from the Non-Crystalline Structure (NCS) outstation of the European Molecular Biology Laboratory (EMBL) in Hamburg, Germany, respectively, over the past three years. All FeBABE work on the core RNAP was done at the NIG in Japan and the SAXS

studies on σ^{54} were conducted at the EMBL Outstation in Germany using the synchrotron radiation facilities of the Deutsches Elektronen Synchrotron (DESY).

CHAPTER TWO

Materials and Methods

2.1 Working Conditions and Materials

A facemask was worn when weighing out hazardous solids or dispensing toxic or volatile liquids. Exposure to ultraviolet radiation was minimised by wearing a full-face visor in conjunction with a laboratory coat and gloves. All waste materials and media were autoclaved prior to disposal in accordance with Imperial College Safety Regulations. Strict sterile techniques were practised when handling bacterial cultures and during all experimental work.

Radioactive isotopes were handled behind a perspex screen wearing disposable gloves in conjunction with a laboratory coat and safety goggles; working areas were monitored with a hand-held Geiger-Müller counter. Personal exposure to radioactivity was monitored by a film-badge, which was developed and checked every three months. All radioactive waste was disposed and handled in accordance with Imperial College Safety Regulations.

2.2 General Materials

2.2.1 Composition of Growth Media and Growth Conditions

All bacterial growth media were autoclaved for 20 minutes at 120°C for sterilisation. The types of media used and their composition per litre are shown in Table 2.1.

Medium	Composition	Application
Luria-Bertani (LB) Medium (Miller 1972)	10g tryptone, 5g yeast extract, 5g NaCl, adjust pH to 7.3 with 1M NaOH	General
Luria-Bertani Agar (LA) (Miller 1972)	10g tryptone, 5g yeast extract, 5g NaCl, 15g agar, adjust pH to 7.3 with 1M NaOH	General
2xYT Medium	16g tryptone, 10g yeast extract, 10g NaCl, adjust pH to 7.3 with 1M NaOH	Protein Overexpression

Table 2.1. Composition of bacterial growth media per litre

The sterilised media were supplemented with appropriate antibiotics and nutritional supplements to select for the bacterial strain/plasmid of interest. Detailed information on the antibiotics is summarised in Table 2.2. All inoculated agar plates and liquid media, unless otherwise stated, were incubated in a 37°C incubator or a shaking incubator rotating at 200rpm, respectively, unless otherwise stated.

Antibiotic	Preparation of Stock	Final Concentration	Storage
Ampicillin	200mg in sterile water to make a stock at 200mg/ml	200µg/ml	filter sterilised and at –20°C in 1ml aliquots
Kanamycin	50mg in sterile water to make a stock at 50mg/ml	50µg/ml	filter sterilised and at –20°C in 1ml aliquots
Chloramphenicol	50mg in methanol to make a stock at 50mg/ml	50µg/ml	room temperature
Nalidixic acid	5mg in sterile water to make a stock at 5mg/ml; adjust pH to 11 with 1M NaOH	15µg/ml	filter sterilised and at –20°C in 1ml aliquots

Table 2.2. Preparation of antibiotics

2.2.2 General Laboratory Reagents and Solutions

- ***24:1*** : 24 parts of chloroform to 1 part of isoamyl alcohol.
- ***Laemmli Buffer x10*** : for 2L, 60g of Tris base, 280g of glycine and 20g of SDS.
- ***PAGE Solution 2*** : 1.5M Tris-HCl, pH 8.8, 0.3% (w/v) SDS.
- ***PAGE Solution 3*** : 0.5M Tris-HCl, pH 6.8, 0.3% (w/v) SDS.
- ***Phenol*** : 100g phenol, 125ml distilled water, 8-hydroxyquinoline to mop up free radicals.
- ***Phosphate buffered saline (PBS) x10*** : 80g NaCl, 2g KCl, 11.5g $Na_2HPO_4 \bullet 7H_2O$, 2g KH_2PO_4. Add water to 1 litre.
- ***STA Buffer x1*** : 25mM Tris-acetate, 8mM magnesium acetate, 10mM KCl, 1mM dithiothreitol (DTT), 3.5% (w/v) polyethyleneglycol (PEG), pH 8.0.
- ***TBE Buffer x10*** : 108g Tris base, 55g boric acid, 40ml 0.5M EDTA, pH 8.0. Add water to 1 litre.
- ***TE Buffer, pH 7.4, 7.5 or 8.0*** : 10mM Tris-Cl at pH 7.4, 7.5 or 8.0, 1mM EDTA at pH 8.0.
- ***Tris-Cl (trishydroxymethylaminomethane) 1M*** : dissolve 121g Tris base in 800ml of water. Adjust pH to desired value with concentrated HCl. Mix and make up to 1 litre with water.
- ***Tris-Glycine Buffer x10*** : 30.3g of Tris and 150g of glycine. Add water to 1 litre. pH 8.6.

2.2.3 Bacterial Strains and Plasmid Constructs

All bacterial strains used for cloning, protein expression and *in vivo* activity assays are listed in detail in Table 2.3. The plasmid constructs used in this work are summarised in Table 2.4.

Strain	Genotype	Application	Reference
E. coli B834(DE3)	F⁻ *ompT hsdS$_B$(r^-_Bm^-_B) gal dcm met* (DE3)	General Expression[1] Host	Novagen
E. coli BL21(DE3) pLysS	F⁻ *ompT hsdS$_B$(r^-_Bm^-_B) gal dcm* (DE3) pLysS	High Stringency[2] Expression Host	Novagen
E. coli JM109	*endA*1 *recA*1 *gyrA*96 *thi hsdR*17 (r^-_K m^+_K) *relA*1 *supE*44 Δ(*lac-proAB*) [F´*traD*36 *proAB lacI^{q}ZΔM15*]	Cloning	Promega
E. coli XL2B	*recA*1 *endA*1 *gyrA*96 *thi*-1 *hsdR*17 *supE*44 *relA*1 *lac* [F´*proAB lacI^{q}ZΔM15*Tn *10*(Tet^r) Amy Cam^r]	Mutagenesis	Stratagene
E. coli TH1	*ΔrpoN2518 endA1 thi-1 hsdR17 supE44 ΔlacU169*	*in vivo* activity assay	Oguiza & Buck, 1997
K. pneumoniae UNF2792	*hisD2 ΔrpoN71::kan, recA56, sbl300::Tn10*	*in vivo* activity assay	Coppard & Merrick, 1991

[1] Expression means that the strain is a λDE3 lysogen, i.e. it carries the gene for T7 RNA Polymerase under lacUV5 control. It is therefore suited to expression from T7 promoters.

[2] High-stringency means that the strain carries pLysS, a pET-compatible (see later) plasmid that produces T7 lysozyme, thereby reducing basal expression of target genes. Even greater stringency is provided by pLysE hosts.

Table 2.3. Host strains used in this work

Plasmid	Background	Application	Antibiotic Resistance	Reference
pMTHσ	pET28b+	- wild-type σ^{54} overexpression[1] - template for mutagenesis	Km	Gallegos & Buck, 1999
pMJ15	pQE-32	PspFΔHTH overexpression	Amp	Jovanovic *et al.*, 1996
pMTH1-306	pET28b+	overexpression of the 1-306aa σ^{54} fragment	Km	Gallegos & Buck, 1999
pMB7.7:760	pET28b+	overexpression of the 70-324aa σ^{54} fragment	Km	Missailidis *et al.*, 1997
pSRW117-477	pET28b+	overexpression of the 117-477aa σ^{54} fragment	Km	This Work
pSRW36	pET28b+	overexpression of σ^{54} mutant Cys36	Km	This Work
pSRW346	pET28b+	overexpression of σ^{54} mutant Cys198	Km	This Work
pSRW336	pET28b+	overexpression of σ^{54} mutant Cys336	Km	This Work
pSRW198	pET28b+	overexpression of σ^{54} mutant Cys346	Km	This Work
pSRW383	pET28b+	overexpression of σ^{54} mutant Cys383	Km	This Work
pSRWCys(-)	pET28b+	- Cys(-) mutant overexpression - template for mutagenesis	Km	This Work
pSRWR383A	pET28b+	overexpression of σ^{54} mutant R383A	Km	This Work
pSRWR383K	pET28b+	overexpression of σ^{54} mutant R383K	Km	This Work

Plasmid	Background	Application	Antibiotic Resistance	Reference
pSRWE36G(83)	pMM83[2]	*rpoN* with E36G mutation	Cm	This Work
pSRWF318A(83)	pMM83	*rpoN* with F318A mutation	Cm	This Work
pSRWE325G(83)	pMM83	*rpoN* with E325G mutation	Cm	This Work
pSRWE410A(83)	pMM83	*rpoN* with E410A mutation	Cm	This Work
pSRWE414A(83)	pMM83	*rpoN* with E414A mutation	Cm	This Work
pSRWE431A(83)	pMM83	*rpoN* with E431A mutation	Cm	This Work
pSRWDM(83)	pMM83	*rpoN* with EE410/414AA mutation	Cm	This Work
pMM83	pHSG575	*rpoN* template for mutagenesis	Cm	Merrick & Gibbins, 1985
pSRWE36G	pET28b+	overexpression of σ^{54} mutant E36G	Km	This Work
pSRWF318A	pET28b+	overexpression of σ^{54} mutant F318A	Km	This Work
pSRWE325G	pET28b+	overexpression of σ^{54} mutant E325G	Km	This Work
pSRWE410A	pET28b+	overexpression of σ^{54} mutant E410A	Km	This Work
pSRWE414A	pET28b+	overexpression of σ^{54} mutant E414A	Km	This Work
pSRWE431A	pET28b+	overexpression of σ^{54} mutant E431A	Km	This Work
pSRWDM	pET28b+	overexpression of σ^{54} mutant EE410/414AA	Km	This Work

[1] Overexpression means that the proteins were expressed as amino-terminal 6His-tagged proteins (see later, section 2.2). [2] pMM83 is a low copy number plasmid harbouring the *K. pneumoniae rpoN* gene

Table 2.4. Plasmid constructs used in this work

2.3 General DNA Methods

2.3.1 Preparation and Purification of Plasmid DNA

Plasmid DNA for manipulation and analysis was obtained using the alkaline lysis method. Typically, a 5-50ml (depending on plasmid copy number) culture is grown to saturation, cells are pelleted and the supernatant discarded. The pellet from a 50ml culture is then resuspended in 2ml of cold GTE buffer (50mM glucose, 25mM Tris-HCl, 10mM EDTA, pH 8.0) and left at room temperature for 10mins. 4ml of cold lysis buffer (0.2M NaOH, 1%(w/v) SDS) is added, mixed gently and incubated on ice for 5mins to lyse the cells. Then, 3ml of 5M potassium acetate at pH 5.5 is added, mixed by vortexing for 5-10secs and stored on ice for at least 10mins before being centrifuged at 15000*g* for 10mins. The supernatant is transferred to a fresh tube and 10µg of DNase-free RNase A is added to the supernatant and incubated at room temperature for at least 5 minutes. The supernatant is then phenol/24:1 extracted and precipitated by adding 0.25 volumes of cold 10M ammonium acetate and 2.5 volumes of cold 100% ethanol and stored at –80°C for 1 hour. The supernatant is centrifuged at 15000*g* for 10mins. Following centrifugation, the supernatant is discarded and the pellet washed with 80% ethanol, re-centrifuged at 15000*g* for 10mins and air-dried. The dried pellet is resuspended in 0.25-1ml of TE buffer to obtain the desired concentration of plasmid DNA. The concentration is determined by running 1-5µl of plasmid DNA preparation on an agarose gel alongside a known concentration of marker DNA and by spectrophotometrically at absorbance (*A*) 260nm. The concentration of double-stranded DNA is calculated as follows:

$$[\text{dsDNA}] = A_{260} \times \text{dilution factor} \times 50 = \mu\text{g/ml}.$$

High quality plasmid DNA for site-directed mutagenesis and sequencing was prepared using QIAGEN anion exchange columns. The plasmid purification protocol is based on an alkaline lysis procedure, followed by binding of plasmid DNA to the anion exchange resin under appropriate low salt and pH conditions. RNA, proteins and other low molecular weight impurities are removed by a medium salt wash. Plasmid DNA is eluted from the columns using TE buffer (Qiagen Plasmid Purification Handbook).

2.3.2 Restriction Analysis of Plasmid DNA

Restriction enzyme digestions were carried out in a total reaction volumes of 10-50µl. Typically, a reaction consisted of the plasmid DNA to be digested, the restriction enzyme(s) and the appropriate reaction buffer. All enzymes and buffers were purchased from New England Biolabs. The reaction was incubated at 37°C for 1-1.5 hours and was then analysed by agarose gel electrophoresis. Typically, 10µl of digest were mixed with 2µl of x10 agarose loading dye (0.2M EDTA pH 8.3, 50% (v/v) glycerol, 0.05%(w/v) bromophenol blue).

2.3.3 Agarose Gel Electrophoresis

Solid agarose (BDH, Molecular Biology Grade) was added to x1 TBE buffer at the desired concentration (w/v) as indicated in Table 2.5 and microwaved at medium power for 5-10mins for the agarose to dissolve. The molten agarose was left on the bench for a few minutes to cool and poured into a gel-casting tray, where it was allowed to set prior to use.

% (w/v) Agarose	Optimum Resolution of DNA (kbp)
0.8	12 to 0.8
1	10 to 0.5
1.5	3 to 0.7
2	1-0.1

Table 2.5. Concentration of agarose gel to use for optimum resolution of DNA

2.3.4 Isolation of DNA Fragments from Agarose Gel

The DNA fragment(s) of interest was excised out of the agarose using a fine, clean, sharp blade, and excess agarose was removed carefully without disrupting the fragment. The fragment was placed in a Petri dish and stored at –80°C for at least 60mins. Using a clean and sharp blade, the solidified agarose fragment was chopped into very fine pieces, which are then placed in a 1.5ml Eppendorf tube. 100µl of sterile distilled water and 100µl

of buffer saturated phenol were added to the tube containing the chopped pieces, vortexed briefly for 10secs, and frozen in liquid nitrogen for 1min. The tube is placed on the bench to thaw. Once thawed, the mixture is vortexed for 10secs and microcentrifuged for 10mins at full-speed. The supernatant with the recovered DNA is transferred into a fresh Eppendorf tube, extracted with phenol/24:1, precipitated with 0.25 volumes of cold 10M ammonium acetate and 2.5 volumes of 100% ethanol and stored at –80°C for 1hr. The supernatant is microcentrifuged at 15000*g* for 10mins. Following centrifugation, the supernatant is discarded and the pellet washed with 80% ethanol, re-microcentrifuged and the pellet air-dried. The pellet was resuspended in the appropriate volume of sterile distilled water to obtain the desired concentration of plasmid DNA. The concentration was determined by running 1-5µl of the plasmid DNA preparation on an agarose gel alongside a known concentration of marker DNA or by spectroscopy.

2.3.5 DNA Ligation Reactions

Ligation reactions were carried out in a 20µl reaction volume using 1 unit of T4 DNA ligase (New England Biolabs), 1mM ATP, and T4 DNA ligase buffer (50mM Tris-Cl, pH 7.8, 10mM $MgCl_2$, 1mM DTT). The reaction containing the insert fragment in 10 fold molar excess over the vector fragment was heated at 65°C for 2-3 minutes and cooled on ice prior to adding the T4 DNA ligase. The reaction was then incubated at 4°C for 16-24 hours. For transformation, half (10µl) of the ligation reaction was used with 150µl of competent cells. Positive transformants were screened by colony screening PCR (see below).

2.3.6 Preparation of Competent Cells & Transformation

The $CaCl_2$ method was used to prepare competent cells for transformation by heat-shock. A primary culture was prepared by inoculating 10ml of LB medium containing the appropriate antibiotics with a single colony of the appropriate strain, and grown at 37°C (shaking) to saturation. The primary culture was used to inoculate 100ml of LB medium

containing with the appropriate antibiotics, which was then grown for 2-3 hours until it reached mid-log phase (OD_{600nm} of 0.5-0.6). At this point, the cells were harvested by centrifugation at 15000*g* for 10mins at 4°C. The supernatant was discarded and the cell pellet was resuspended in 10ml of ice cold 50mM $CaCl_2$/15%(v/v) glycerol and left on ice for 30mins. The cells were then harvested by centrifugation at 15000*g* at 4°C. The cell pellet was resuspended in a 5ml solution of 50mM $CaCl_2$/15%(v/v) glycerol. The cell suspension was stored as 100µl aliquots at –80°C.

For transformation, typically 100µl of competent cells were mixed with 5-10µl of DNA (depending on concentration) and stored on ice for 30mins. The cells were then heat shocked by placing them in a 42°C water bath for 45secs and cooled on ice for a further 2mins. Then 1ml of LB liquid medium was added to the cells and incubated at 37°C (shaking) for 1-2 hours before plating all the cells onto LB agar plates containing the appropriate antibiotics and/or growth supplement. The plates were incubated at 37°C for 12-16 hours.

2.3.7 Polymerase Chain Reaction (PCR)

PCR reactions were performed using *Taq* polymerase (Promega). The standard reaction volume used was 50µl. Typically, a reaction mix consisted of 5µl of 10x *Taq* polymerase reaction buffer (20mM Tris-HCl, 100mM NaCl, 0.1mM EDTA, 1mM DTT, 50% (v/v) glycerol, 1% (v/v) Triton, pH 8.0), 0.2mM of each dNTP, 500pmoles of forward and reverse primers, 2pmoles of template DNA, 2.5U of *Taq* polymerase and sterile distilled water to a final volume of 50µl. Each amplification reaction was set up with a series of final $MgCl_2$ concentrations at 2mM, 4mM, 6mM and 8mM, to determine the optimal $MgCl_2$ concentration for that particular reaction. Each reaction was overlayed with 50µl of mineral oil to prevent evaporation during temperature cycling. The PCR program had the following parameters: 30 cycles, denaturation at 94°C for 45secs, annealing at 50-65°C (depending on primer melting temperature) for 30secs and extension at 72°C for 1min /kb of DNA to be amplified. The PCR product was checked by agarose gel electrophoresis. For cloning experiments, the PCR reaction was precipitated with 0.25

volumes of 10M ammonium acetate and 2.5 volumes of 100% ethanol at –80°C for 1hr, then microcentrifuged at full speed for 20mins. Following microcentrifugation, the supernatant was discarded and the pellet was washed with 80% ethanol, microcentrifuged at 15000*g* for 10mins and the pellet air dried. The pellet was resuspended in the appropriate volume of TE buffer.

Colony screening PCR was essentially done as described above. A single colony was picked from LB agar plates and inoculated into the PCR reaction buffer without the *Taq* polymerase. The reaction mixture was heated at 100°C for 5mins. Prior to temperature cycling, 2.5U of *Taq* polymerase were added and each reaction was overlayed with 50µl of mineral oil. The PCR reaction was checked by agarose gel electrophoresis.

2.3.8 Site-Directed Mutagenesis

All mutant σ^{54} proteins used in this work were made using a PCR based site-directed mutagenesis method (developed by Stratagene GmbH) which employs two enzymes: (1) *Pfu* DNA polymerase – a DNA polymerase from the hyperthermophilic Archae *Pyrococcus furiosus* with a blend of a thermostable factor that has x12 lower error rate than conventional thermostable polymerases, e.g. the *Taq* DNA polymerase; and (2) *DpnI* restriction endonuclease, which specifically digests methylated and hemimethylated DNA with the target sequence 5'Gm6ATC-3'. The basic procedure of this site-directed mutagenesis method is outlined in Figure 2.1. The protocol uses supercoiled, double stranded DNA as a template and requires two synthetic oligonucleotide primers containing the desired mutation. The mutagenic primers are complementary to each strand of the plasmid DNA and are extended during temperature cycling by the *Pfu* DNA polymerase. Following temperature cycling, the reaction mix (which will contain the *in vivo* methylated parental DNA and the *in vitro* synthesised mutated DNA) is treated with *DpnI*, which will digest away the parental strand.

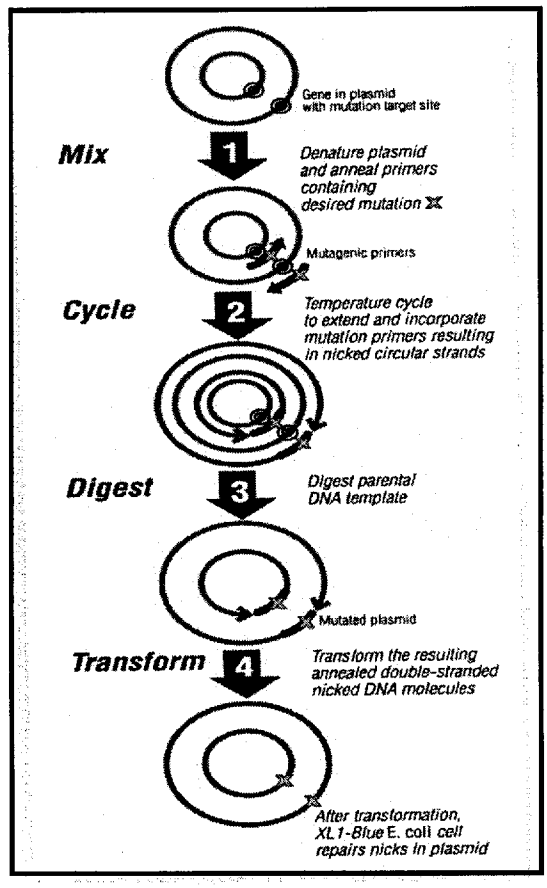


Figure 2A. Site-directed mutagenesis methodology. Adapted from Stratagene's catalogue 1999. Note that *E. coli* XL2B cells were used in this work for transformation, instead of the indicated XL1B cells.

The mutagenic oligonucleotide primers were synthesised and HPLC purified by Sigma Genosys. The parental plasmid DNA was prepared using Qiagen plasmid DNA preparation columns (see above) after growth in a *dam*$^+$ *E. coli* background (plasmid DNA isolated from *dam*$^+$ strains is methylated at the adenine base in the sequence 5'GATC-3', making it susceptible to digestion by *DpnI*). Sample reactions for the mutagenesis PCR contained 50ng of parental plasmid DNA, 125ng of each mutagenic oligonucleotide primer, 2.5mM of each dNTP, 5µl of x10 reaction buffer, 2.5U of Pfu DNA polymerase and sterile distilled water to a final volume of 50µl. Each reaction mix was overlayed with 30µl of mineral oil and subjected to PCR with cycling parameters as summarised below in Table 2.6.

Segment	Cycles	Temperature	Time
1	1	95°C	45 seconds
2	25	95°C	30 seconds
		55°C	60 seconds
		68°C	2 mins / kbp plasmid size

Table 2.6. Site-directed mutagenesis PCR parameters

Following the PCR reaction, the mixtures were placed on ice for 2mins to cool. 10U of *DpnI* restriction enzyme was directly added to each reaction below the mineral oil layer using a small, pointed pipette tip and mixed gently by pipetting the solution up and down several times. Each reaction was then incubated at 37°C for 2hrs to digest the parental (non-mutated) template DNA. Following incubation, 10µl of the reaction mix were used to transform competent *E. coli* XL2B cells. Positive transformants were selected for by plating the cells onto LB agar containing the appropriate antibiotics. The LB agar plates were incubated at 37°C for >16hrs.

Single colonies were picked from the transformation plates and inoculated into 10ml of LB medium containing the appropriate antibiotics. The mutant plasmid DNA was isolated using the Qiagen plasmid preparation columns (see above) and the *rpoN* gene sequenced prior to further application.

2.3.9 DNA Sequencing

All DNA sequencing reactions were performed using the "ABI Prism Big Dye Terminator Cycle Sequencing Reaction Kit" (Perkin Elmer). This system uses a variation of the chain termination method developed by Sanger *et al.*, 1997. The DNA to be sequenced acts as a template for the enzymatic synthesis of new DNA starting at a defined primer binding site. A mixture of both deoxynucleotides and fluorescent labelled dideoxynucleotides is used in the reaction, at concentrations which create a finite probability that a dideoxynucleotide will be incorporated at each position in the growing chain. Incorporation of a dideoxynucleotide blocks further chain elongation, resulting in a population of truncated fragments of varying lengths. The identity of the chain terminating

dideoxynucleotide is determined by separating the fragments by size on a denaturing high resolution polyacrylamide gel and using a laser to detect the fluorescent label at the end of the fragment. The sequence is subsequently specified by correlating the order of the bands on the gel with the dideoxynucleotide used to create each band.

The sequencing reactions were all performed in the laboratory strictly following the manufacturer's instructions. The electrophoresis and laser detection were done at the Biomedical Sciences (BMS) Sequencing Service, Department of Molecular Pathology, Imperial College.

2.4 General Protein Methods

2.4.1 Protein Overexpression

The Expression System

The pET system. The pET28b+ (**p**lasmid for **e**xpression by **T7 DNA-dependent RNA polymerase**) (Novagen) contains a T7 promoter region which is selectively (and with high specificity) recognised by the bacteriophage T7 DNA dependent RNA polymerase. The multiple cloning into which the target DNA can be inserted, is directly downstream of the T7 promoter. The pET28b+ plasmid has a Km^R gene which confers kanamycin resistance to the bacterial cell transformed with this plasmid, a feature which can be used to select transformed colonies of *E. coli* expression strains on agar plates containing kanamycin. The pET28b+ plasmids also contain a consecutive stretch of six histidine residues which are expressed at the amino terminal end of the target protein. This histidine sequence allows the target protein to be purified by metal chelating chromatography. Induction and overexpression of target protein from the pET28b+ plasmids occurs as follows: transcription from the *lac* promoter is controlled by the *lac* repressor, the product of the *lac*I gene. In the absence of the inducer, lactose or IPTG, the *lac* repressor inhibits transcription from the *lac* promoter by binding to the operator site

and thereby preventing the proper binding of the RNA polymerase to the promoter region so that transcription of downstream genes does not occur. However, there is a low basal level of transcription of these genes, due to the equilibrium which exists between the bound and unbound repressor molecules, which causes the operator site not to be continuously occupied by the *lac* repressor. When the cell encounters lactose or IPTG, these will bind to and inhibit the binding of the *lac* repressor to the operator region. As a consequence, transcription of the downstream genes will increase many fold. In DE3 (see blow), the T7 DNA dependent RNA polymerase is under the control of the *lacUV5* promoter. The *lacUV5* promoter contains a mutation in the consensus region of the *lac* promoter that increases promoter strength, and therefore has been used in place of the wild-type *lac* promoter in DE3. All *E. coli* strains used for protein overexpression are lysogenised with the bacteriophage DE3, which is a λ derivative (see Table 2.3). In the presence of lactose or IPTG, high level expression of the T7 RNA polymerase consequently results in the expression of the target proteins from pET28b+ plasmids containing the target protein under the control of the T7 promoter.

The pQE system. The pQE overexpression vector series (Qiagen) provides high level expression of target proteins containing a 6xHis affinity tag. It contains a promoter/operator element consisting of the *E. coli* T5 promoter and two lac operator sequences. When the *lac* repressor is present no expression of the target protein occurs. Target protein expression is rapidly induced by the addition of IPTG. A synthetic ribosome binding site encoded by this plasmid, RBSII, allows optimal mRNA recognition and binding during protein expression. Since the expression from the T5 promoter is extremely efficient, it must be controlled by the presence of high levels of *lac* repressor, which are provided by the pREP4 plasmid. The multiple copies of pREP4 present in this expression system ensure the presence of high levels of *lac* repressor and tight regulation of protein expression. The double operator system in the pQE vectors, in combination with the high levels of *lac* repressor provided by pREP4, permit strict control over the level of expression of the target protein. In principle, any *E. coli* strain containing both the expression (pQE) and the repressor (pREP4) plasmids can be used for the overproduction

of target proteins. pQE and pREP4 are selected for by kanamycin and ampicillin, respectively.

The Expression Protocol

The following protocol has been optimised to maximise recovery in the soluble fraction (>95%) of the overexpressed target protein, usually a σ^{54} derivative. Colonies from freshly (overnight) transformed plates were used to inoculate 2xYT medium with the appropriate antibiotics to select for the transformants. Typically, 50 colonies were used to inoculate 1L of medium. The cells were pre-grown in a 37°C shaking incubator for 2-3 hours until the culture reached mid-log phase (OD_{600nm}= 0.5). The cells were shifted to a 25°C shaking incubator and the culture was cooled to 25°C for 30mins. The cells were induced with 1mM of IPTG. Overexpression of the target protein was stopped after 2-3 hours by adding frozen blocks of PBS buffer. The cells were harvested at 4°C by centrifugation at 10000*g* for 10mins. Following harvesting, the cells were resuspended in a lysis buffer solution (composition depending on purification method used – see below) containing a cocktail-tablet (1 tablet / 50ml of lysis buffer solution) of protease inhibitors (Boehringer Mannheim) against proteases in pancreas extract (0.015mg/ml), pronase (0.0015mg/ml), thermolysin (0.0008mg/ml), chymotrypsin (0.0015mg/ml), trypsin (0.0002mg/ml) and papain (1mg/ml). The cells were lysed using a French press, which achieves cell lysis by placing the cell sample under high pressure (1400 kg/cm^3) followed by a sudden release into atmospheric pressure. The rapid change in pressure causes the cells to burst. The lysate (L) was sampled before analysis by SDS polyacrylamide gel electrophoresis (PAGE) and centrifuged at 4°C for 45mins at 18000*g*. The supernatant (S1) was collected in a fresh tube and stored at 4°C. The pellet (P1) was resuspended in lysis buffer, sampled for SDS-PAGE analysis and centrifuged at 4°C for 45mins at 18000*g*. The resulting supernatant (S2) was collected in a fresh tube and stored at 4°C. The pellet (P2) was resuspended in lysis buffer and sampled for PAGE analysis. All samples were subjected to PAGE to determine whether the overexpressed protein was predominantly in

soluble form, i.e. in the soluble fractions S1 and S2, or in inclusion bodies, i.e. in the insoluble fractions P1 and P2.

2.4.2 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

All protein samples were checked and analysed by SDS-PAGE using the BioRad Mini-Protean II gel system. The percentage of gel chosen for SDS-PAGE analysis depends on its molecular weight of the protein of interest. Table 2.7 summarises the composition of the SDS gel:

Component	5%	7.5%	10%	12.5%	15%	20%
Acrylamide	1.7ml	2.5ml	3.25ml	4.1ml	5.0ml	6.7ml
Solution 2	2.7ml	2.5ml	2.5ml	2.5ml	2.5ml	2.5ml
Water	5.7ml	5.0ml	4.2ml	3.4ml	2.5ml	0.8ml
10% (w/v) APS	100µl	100µl	100µl	100µl	100µl	100µl
TEMED	10µl	10µl	10µl	10µl	10µl	10µl

Component	Volume to add
Acrylamide	0.75ml
Solution 3	1.25ml
Water	3.0ml
10% (w/v) APS	50µl
TEMED	5µl

Table 2.7. Compositions of the 2x SDS-PAGE separating gel (top table) and stacking gel (bottom table).

In preparation for loading (5µl) the protein samples onto the SDS-gel, the samples (10µl) were mixed with 1 volume of x2 SDS sample buffer (Sigma), heated at 95°C for 5mins and microcentrifuged at full speed for 10mins. The gels were run at 200V in x1 Laemmli running buffer and stained with Coomassie blue stain (minimum sensitivity: 30ng protein).

2.4.3 Protein Purification

All proteins purified for this work and the purification methodologies applied are listed in Table 2.8. A detailed description of each purification method is given below.

Protein	Plasmid	Purification Step 1	Purification Step 2	Purification Step 3
General				
$H\sigma^{54}$	pMTHσ	Ni	Hep	-
PspFΔHTH	pMJ19	Ni	-	-
Chapter 3				
70-324	pMB7.7:760	Ni	-	-
1-306	pMTH1-306	Ni	-	-
Chapter 4&5				
Cys36	pSRW36	Ni	-	-
Cys198	pSRW346	Ni	-	-
Cys336	pSRW336	Ni	-	-
Cys346	pSRW198	Ni	-	-
Cys383	pSRW383	Ni	-	-
Cys(-)	pSRWCys(-)	Ni	-	-
Chapter 6				
R383A	pSRWR383A	Ni	Hep	MQ
R383K	pSRWR383K	Ni	Hep	MQ
NtrC	pDW78	Ni	-	-
Chapter 7				
E36G	pSRWE36G	Ni	Hep	-
F318A	pSRWF318A	Ni	Hep	-
E325G	pSRWE325G	Ni	Hep	-
E410A	pSRWE410A	Ni	Hep	-
E414A	pSRWE414A	Ni	Hep	-
E431A	pSRWE431G	Ni	Hep	-
EE410/414AA	pSRWDM	Ni	Hep	-
Chapter 8				
70-324	pMB7.7:760	DEAE	Ni	MQ
1-306	pMTH1-306	DEAE	Ni	MQ
117-477	pSRW117-477	DEAE	Ni	MQ

Table 2.8 (previous page). List of proteins used in this work. DEAE means (diethylaminoethyl)-Cellulose anion exchange chromatography; Hep means Heparin column affinity chromatography; Ni means Ni-column affinity chromatography; and MQ means Mono Q ion exchange chromatography (all column material were from Pharmacia).

(Diethylaminoethyl)-Cellulose Anion Exchange (DEAE) Chromatography

Principle. DEAE chromatography is an ion-exchange (IEX) chromatography method which separates biomolecules on the basis of charge characteristics. Charged groups on the surface of a protein interact with oppositely charged groups immobilised on the IEX medium (solid phase). The solid phase of an IEX packing consists of functional groups that have either a positive charge (anion exchange), used to separate negatively charged target molecules, or a negative charge (cation exchange), used to separate positively charged target molecules. IEX binding occurs when the salt concentration (ionic strength) of the mobile phase is reduced to the point that the ionic groups on the target molecule begin to serve as the counter ions for the charged groups on the solid phase. This causes the target molecules to bind to the surface. Elution of the bound target molecule takes place when the ionic strength of the mobile phase is increased. As this happens, the salt molecules (counter-ions) displace the bound target molecule back into the mobile phase. At the same time, salt molecules (co-ions) with the same charge as the IEX medium (solid phase) bind to the charge groups of the target molecules, hence elution of the target molecule occurs. Thus, DEAE is an anionic IEX medium and elution of the target molecule occurs when the Cl^- ions (counter-ions) bind to DEAE, and Na^+ ions (co-ions) bind to the target molecule.

Protocol. For DEAE IEX chromatography the protein solution was buffer exchanged (by dialysis) into buffer A_{DEAE} (20mM imidazole, 100mM NaCl, 1mM DTT, 0.1mM EDTA, 5% (v/v) glycerol, pH 7, filter sterilised). DEAE IEX chromatography was usually the first purification step. Thus, the overexpressed cells were resuspended and lysed in buffer A_{DEAE}. The lysate was ammonium sulphate (in order to "salt out" (concentrate) the protein) and streptomycin sulphate precipitated (in order to remove nucleic acids by precipitation) to concentrate and remove contaminating proteins and nucleic acids, respectively. Following the precipitation steps the lysate was centrifuged at

18000*g* for 45mins at 4°C. The resulting pellet was resuspended in buffer A_{DEAE} and dialysed against buffer A_{DEAE} to remove the ammonium/streptomycin sulphate, overnight at 4°C. The dialysate was centrifuged at 18000*g* for 45mins at 4°C and was loaded (at a flow rate of 1ml/min) onto a 150ml DEAE column previously equilibrated with 5 column volumes (CV) of buffer A_{DEAE}, 5CV of buffer B_{DEAE} (same as buffer A_{DEAE}, but with final concentration of 1M NaCl) and 10CV of buffer A_{DEAE}. The flow-through (FT) was collected and the column was washed with 5CV of buffer A_{DEAE}. The washes were collected as 20ml fraction. The protein was eluted with 800ml linear gradient, 0.1-1M NaCl gradient in buffer A_{DEAE} at a rate of 1ml/min. Elution fractions of 5ml were collected and analysed by SDS-PAGE for the presence of the target protein.

Heparin Column Chromatography

Principle. Immobilised heparin has two main modes of interaction with proteins. It can operate as an affinity ligand; or, as used here it can function as a high capacity cation exchanger, due to its anionic sulphate groups (solid phase). For the elution of bound target molecules a salt gradient is used, as described above for DEAE IEX chromatography.

Protocol. For heparin chromatography the protein solution was buffer exchanged (dialysed) into dialysis buffer (10mM Tris-Cl, 50mM NaCl, 0.1mM EDTA, 5% (v/v) glycerol, 1mM DTT, pH 8). The dialysate was centrifuged at 18000*g* for 45mins at 4°C and loaded (at a flow rate of 1ml/min) onto a 5ml heparin column (Pharmacia) previously equilibrated with 5CV of buffer A_{HEP} (10mM Tris-Cl, 50mM NaCl, 0.1mM EDTA, 5% (v/v) glycerol, 1mM DTT, 10mM $MgCl_2$, pH 8), 5CV of buffer B_{HEP} (same as buffer A_{HEP} but with 1M NaCl), and 10CV of buffer A_{HEP}. The FT was collected and the column was washed with 5-10CV of buffer A_{HEP}. The washes were collected as 5ml fractions. The protein was eluted with an 80ml linear gradient, 0.05-1M NaCl in buffer A_{HEP} at a rate of 1ml/min. Elution fractions of 1ml were collected and analysed by SDS-PAGE.

Mono Q IEX Chromatography

Principle. Mono Q, like DEAE is an anionic IEX chromatography material. Mono Q is a strong anionic exchanger based on a beaded hydrophilic resin with a very narrow particle size distribution, designed for high resolution. A quaternary amine group, $-CH_2N^+(CH_3)_3$, is attached to the resin which binds positively charged target molecules. The positively charged quaternary amine groups remain equally charged over the entire useful pH range of 3-11. As described above for DEAE chromatography, elution of the target molecule occurs when the Cl^- ions (counter-ions) bind to Mono Q matrix, and Na^+ ions (co-ions) bind to the target molecule, thereby displacing the target molecule from the Mono Q matrix.

Protocol. For Mono Q chromatography the protein sample was buffer exchanged (by dialysis) into buffer A_{MQ} (10mM Tris-Cl, 50mM NaCl, 0.1mM EDTA, 5% (v/v) glycerol, 1mM DTT, pH 7, filter sterilised). The dialysate was centrifuged at 18000*g* for 45mins at 4°C and loaded (at a flow rate of 1ml/min) onto a 5ml Mono Q column (Pharmacia) previously equilibrated with 5CV of buffer A_{MQ}, 5CV buffer B_{MQ} (same as buffer A_{MQ}, but with a final concentration of 1M NaCl), and 10CV of buffer A_{MQ}. The FT was collected and the column was washed with 5CV of buffer A_{MQ}. The washes were collected as 5ml fractions. The protein was eluted with an 80ml linear gradient, 0.05-1M NaCl gradient in buffer A_{MQ} at a rate of 1ml/min. Elution fractions of 1-5ml were collected (depending on protein preparation) and analysed by SDS-PAGE.

Metal Affinity Chromatography

Principle. Proteins containing 6His affinity tags, located at either the amino or carboxyl terminus of the target protein, can bind to Ni-NTA (Ni-Nitrilotriacetic acid, QIAGEN) groups immobilised on a matrix. Binding of the 6His tag does not depend on the three-dimensional structure of the protein. Even when the 6His tag is not completely accessible it will bind as long as more than two histidine residues are available to interact with the nickel ion; in general, the smaller the number of accessible histidine residues, the

weaker the binding will be. Untagged proteins (other host proteins) that have histidine residues in close proximity on their surfaces will also bind to the Ni-NTA resin, but in most cases this interaction will be much weaker than the binding of the 6His tag. Any host protein that binds non-specifically to the Ni-NTA resin is washed away under relatively mild conditions that do not affect the binding of the 6His tagged target proteins. Elution of the target protein is achieved with an imidazole gradient chromatography. The imidazole ring is part of the structure of histidine. The imidazole rings in the histidine residues of the 6His tagged proteins bind to the nickel ions immobilised by the NTA groups on the matrix. Imidazole itself can bind to the nickel ions and disrupt the binding of histidine residues to the Ni-NTA resin. At sufficiently high imidazole concentrations the 6His tagged target protein can no longer bind to the nickel ions and will dissociate from the Ni-NTA resin. Thus, elution of the 6His tagged target protein occurs. The Ni-NTA resin was packed into 5-20ml of chromatography column (Pharmacia) and the chromatography was done on a Fast Protein Liquid Chromatography (FPLC) machine (Pharmacia).

Protocol. For the metal affinity chromatography the protein solution was buffer exchanged (by dialysis) into buffer A_{NI} (20mM sodium phosphate, 0.5M NaCl 5% (v/v) glycerol, pH 7, filter sterilised). If the metal affinity chromatography was the first purification step, the overexpressed cells were resuspended and lysed in buffer A_{NI}. The protein sample (dialysate / lysate) was loaded (at a flow rate of 1ml/min) onto a 5-20ml Ni column previously equilibrated with 5CV of buffer A_{NI}, 5CV of buffer B_{NI} (same as buffer A_{NI}, but with final concentration of 1M imidazole) and 10CV of buffer A_{NI}. The FT was collected and the column was washed with 5CV of 5% buffer B_{NI}. The washes were collected as 5ml fractions. The protein was eluted with a 40ml imidazole gradient (5%→80% buffer B_{NI} in A_{NI}) at a flow rate of 1ml/min. Elution fractions of 5ml were collected and analysed by SDS-PAGE for the presence of the target protein. For the purification of NtrC, buffer A_{NI} contained 50mM Tris-Cl, 300mM NaCl, 20mM imidazole, 10% (v/v) glycerol, pH 8), buffer B_{NI} contained the same components as buffer A_{NI} but with 1M imidazole.

2.4.4 Protein Handling & Storage

All purified σ^{54} derivatives and PspFΔHTH were buffer exchanged into storage buffer (10mM Tris-HCl, 50mM NaCl, 0.1mM EDTA, 1mM DTT, 50%(v/v) glycerol, pH 7) at 4°C. Stock solutions were stored at –80°C; working solutions, at –20°C. PspFΔHTH was stored at –80°C in 10µl aliquots. NtrC was stored in NtrC storage buffer (50mM Tris-Cl, 300mM NaCl, 20mM imidazole, 10%(v/v) glycerol, pH 8) at –80°C in 10µl aliquots.

2.4.5 Measuring Protein Concentration by the Bradford Assay

The Bradford assay for protein concentration determination is a rapid and reliable dye-based assay for determining protein content in solution. Prior to measuring the concentration of the target protein, a standard curve with samples of known protein concentration is prepared in parallel with the unknown samples. The standard curve is essential for the quantitative assessment of protein concentration. Bovine serum albumin (BSA) is the most commonly used standard protein. The absorbance values for a range of BSA concentrations (1.0µg/ml to 10µg/µl) was used to generate the standard curve.

2.5 *In vivo* Activity Assays

2.5.1 Promoter Activation Assay (β-galactosidase Assay)

Activator constitutively expressed. In vivo promoter activation assays were done in *E. coli* TH1 containing the plasmid pMB221 (Table 2.9), which carries the β-galactosidase gene under the control of the *K. pneumoniae nifH* promoter and constitutively expresses *K. pneumoniae* activator NifA. Plasmids harbouring mutant *rpoN* genes were transformed into *E. coli* TH1 cells containing the pMB221 plasmid. Three or four independent colonies were grown at 30°C in LB with antibiotics to an OD_{600nm} of 0.6-1.0. The cultures were then placed on ice for 10mins before 1ml was taken to measure cell density at 600nm. 50µl of

culture was added to 950µl of Z-buffer (for 1L: 16.1g $Na_2HPO_4 \cdot 2H_2O$, 5.5g $NaH_2PO_4 \cdot H_2O$, 0.75g KCl, 0.246g $MgSO_4 \cdot 7H_2O$, 2.7ml β-mercaptoethanol) to make a final volume of 1ml for each reaction. The cells were lysed by adding 2 drops of chloroform and 1 drop of 0.1%(w/v) SDS solution and vortexing for 10sec. The reaction tubes were then placed in a 28°C water bath for 5mins before the reaction was started by adding 200µl of freshly prepared ONPG (*o*-Nitrophenyl β-D-galactopyranoside) from a stock solution of 4mg/ml in Z-buffer to each tube, with vortexing for 1sec. Once the solution had turned pale yellow, the time was noted and the reaction stopped by adding 500µl of 1M sodium carbonate solution. The absorbance at 420nm was measured. ß-galactosidase activity (in Miller units) was calculated using the following equation:

$$\text{Units of } \beta\text{ - galactosidase Activity} = \frac{1000 \times OD_{420}}{T \times V \times OD_{600}}$$

where, T = reaction time in minutes, V = volume of culture (ml) used in the reaction, and OD_{600} is that of the culture when sampled.

Activator chromosomally expressed. Assays were done in *K. pneumoniae* UNF2792 cells harbouring either pMB1 or pMB210 (Table 2.9), which carry the β-galactosidase gene under the control of the *K. pneumoniae nifH* and *S. meliloti nifH* promoters, respectively. For activation, these promoters are dependent on the chromosomally expressed NifA protein which is induced under nitrogen limiting and anaerobic conditions. Compatible plasmids harbouring mutant *rpoN* genes were transformed into the *K. pneumoniae* UNF2792 cells containing the pMB1 or the pMB201 plasmid. Three or four independent colonies were grown at 37°C in nutrient broth (NB) (15g in 1L H_2O) overnight. 100µl of these cultures were used to inoculate 3mls of nitrogen free Nitrogen-Free Davis & Mingoli Medium (NFDM) (For 1L: 3.4g $K_2HPO_4 \cdot 3H_2O$, 12.05g KH_2PO_4 , 0.025g Na_2MoO_4 , 0.025g $FeSO_4 \cdot 7H_2O$, 7g glucose) and grown under anaerobic conditions for 18hrs. Chromosomal *ntrC* and *nifA* expression was activated by supplementing NFDM with 100µg/ml of aspartic acid (and in control cultures repressed with 200µg/ml of ammonium sulphate). Following this incubation period, the assay was conducted essentially as described above.

Plasmid	Promoter	Activation System	Resistance	Reference
pMB1	*K.p nifH*	Cx; NifA	Ampicillin	Cannon & Buck, 1989
pMB210	*S.m nifH*	Cx; NifA	Tetracyclin	Better *et al.*, 1985
pMB221	*K.p nifH*	Px; NifA	Chloroamphenicol	Cannon & Buck, 1992

Table 2.9. Reporter plasmid system used for the *in vivo* β-galactosidase assays. Keys: K.p = *K. pneumoniae*; S.m = S.meliloti; Cx = chromosomally expressed; Px = expressed constitutively from plasmid.

2.5.2 Immunoblotting

The *in vivo* stabilities of mutant σ^{54} proteins were measured in *E. coli* TH1 or *K. pneumoniae* UNF2792 strains, which have a deletion of the chromosomal *rpoN* gene. Mutant plasmids were transformed into TH1 or UNF2792 cells and grown to late exponential phase (OD_{600nm} ~ 1). Cells (1ml) were collected by centrifugation and resuspended in 60µl of 10mM Tris-Cl, 0.1mM EDTA, pH 7.9. 10µl of concentrated cells were lysed with 10µl of 2 x SDS sample buffer, heated at 95°C and 6µl of each loaded. Proteins were separated on denaturing SDS 7.5% PAGE minigels and blotted onto PVDF membranes (BioRad) as follows: The gel, a similar sized piece of methanol soaked PVDF membrane and 6 gel sized pieces of 3MM Whatman filter paper were equilibrated in transfer buffer (25mM Tris-Cl, 190mM glycine, 20% (v/v) methanol, pH 8) for a few minutes. Three pieces of filter paper, the PVDF membrane, the gel and then three more layers of filter paper were assembled on the immunoblotter (BioRad). The immunoblot was run for 30mins at 10V. The membrane was blocked with 5%(w/v) dry milk powder solution (in water) for 1hr. Following this, the membrane was incubated with anti-σ^{54} antibodies (Jishage *et al.*, 1996) (diluted 1:3000 in 5% (w/v) dry milk powder solution) for 1hr, then the membrane was washed in Tween buffer (for 1L: 3g Tris, 8g NaCl, 0.2g KCl, 1% (v/v) Tween-20) for 2 x 15mins. The washed membrane was then incubated with anti-rabbit alkaline phosphate conjugated antibodies (Promega) (diluted 1:7500 in 5% dry milk powder solution). The membrane was then washed in Tween buffer for 2 x 15mins and developed with 10ml of developing buffer (100mM Tris-HCl, 100mM NaCl, pH 9.5)

containing 100µl of NBT (nitro blue tetrazolium) and 50µl of BCIP (5-bromo-4-chloro-3-indolyl-phosphate) (both from Promega). The phosphatase reaction was stopped by washing the membrane with sterile distilled water. The membrane was then air dried.

2.6 *In Vitro* Activity Assays

2.6.1 Core RNAP Binding Assays

Native gel holoenzyme assembly assays were used to detect complexes forming between core RNAP and σ^{54} based on the different mobility of free core RNAP versus holoenzyme. Reactions were done in Tris-NaCl buffer (40mM Tris-HCl, 0.1mM EDTA, 1mM DTT, 100mM NaCl, 10% (v/v), pH 8). *E. coli* core RNAP (Epicentre Technologies) (250nM) and different amounts of σ^{54} proteins / fragments were mixed together and incubated at 30°C for 10mins in a total reaction volume of 10µl. The reactions were stopped by the addition of 2µl x5 native loading dye (STA buffer with 50% glycerol and 0.05%(w/v) bromophenol blue). Samples were loaded onto native 4.5% polyacrylamide Bio-Rad Mini-Protean II gels and run at 50V for 2hrs at room temperature in x1 Tris-glycine buffer. Complexes were visualised by Coomassie blue staining.

2.6.2 DNA Binding Assays

DNA homoduplex & heteroduplex formation: ^{32}P end-labelled double-stranded homoduplex or mismatched heteroduplex DNA fragments consisting of the –60 to +28 *S. meliloti nifH* promoter sequence (unless otherwise stated) were used as templates for the gel mobility shift assays (see below). Complementary (for homoduplex) or non-template strand mismatched double stranded *S. meliloti nifH* promoter sequences (unless otherwise stated) were prepared by annealing of 5' ^{32}P end-labelled and unlabelled oligonucleotides. Size purified oligonucleotides (supplied from Sigma Genosys) were labelled with [γ-^{32}P]ATP and T4 polynucleotide kinase by the standard protocol. For preparing labelled

duplex, pairs of DNA strands with the unlabeled strand present at a twofold molar excess (10pmoles in 20µl) were heated at 95°C for 3mins in TM buffer (10mM Tris-HCl, 10mM $MgCl_2$, pH 8) and then chilled rapidly in iced water for 5mins to allow duplex formation. The DNA duplex was then stored at –20°C.

Gel mobility shift assays: The 10µl binding reactions contained 16nM DNA, and either 100nM holoenzyme (core RNAP:σ^{54} 1:2) (unless otherwise stated) or 1µM σ^{54} (unless otherwise stated) in STA buffer, and were incubated at 30°C for 10mins. Free DNA was separated from holoenzyme/σ^{54} bound DNA on 4.5% polyacrylamide gels run in x1 Tris-glycine buffer. For heparin stability assays, heparin (100µg/ml) was added for 5mins prior to gel loading. The gel was run at 60V for 80mins and complexes were visualised and quantified by phosphoimager analysis or autoradiography.

2.6.3 Holoenzyme Stability Assays on Heteroduplex DNA

A gel mobility shift assay (as described above) was used for holoenzyme stability assays on heteroduplex DNA. Typical reactions consisted of 100nM holoenzyme (core RNAP:σ^{54} 1:2) and 16nM of ^{32}P end-labelled heteroduplex DNA in STA buffer. Holoenzyme and DNA were pre-incubated at 30°C for 10mins and, then, if necessary, nucleotide and activator were added for 10mins followed by a glycerol bromophenol blue loading dye (final concentration 10% (v/v) glycerol) and, if required, heparin (at 100µg/ml), for 5mins prior loading. Then, samples were loaded on 4.5% native gels and run at 60V for 80mins at room temperature in x1 Tris-glycine buffer. Free DNA and proteins from protein-DNA complexes were detected by autoradiography and quantified by phosphoimager analysis.

2.6.4 Sigma – Isomerisation Assays

Typically, ^{32}P end-labelled *S. meliloti nifH* heteroduplex promoter DNA fragment containing a non template strand mismatch at positions –12 and –11 (respective to the transcription start site) was used as a template for the sigma isomerisation assays (unless

otherwise stated). σ^{54} (at 1μM) and DNA were pre-incubated at 30°C for 10mins. Hydrolysable nucleotide (at 4mM) and PspFΔHTH (at 4μM) were added for 10mins followed by a glycerol bromophenol blue loading dye (final concentration 10% glycerol) and, if required, heparin (at 100μg/ml) for a further 5mins prior loading. The samples were loaded on 4.5% native gels and run at 60V for 80mins at room temperature in x1 Tris-glycine buffer. σ^{54} bound to the –12 to –11 mismatched heteroduplex DNA responds to activator in a nucleotide hydrolysis-dependent reaction to form a new supershifted DNA complex with lower mobility than the initial σ^{54}-DNA complex (Cannon *et al.*, 2000) Free DNA and DNA-σ^{54} and the supershifted σ^{54}-DNA complex were detected by autoradiography and quantified by phosphoimager analysis.

2.6.5 Transcription Assays

Activator dependent assays: The template for the transcription assays was the supercoiled plasmid pMKC28, which contains the *S. meliloti nifH* promoter, or one of the pFC50 series of plasmids containing the *E. coli* wild-type and mutant *glnHp2* promoters (Chaney & Buck, 1999; Claverie-Martin & Magasanik, 1991). Transcription reactions (10μl final volume) were performed in STA buffer containing 100nM (unless otherwise stated) holoenzyme (core RNAP:σ^{54} 1:6), 10nM supercoiled plasmid template, 50μg/ml bovine serum albumin, 10 units of RNasin (Promega), 4mM ATP and 4μM PspFΔHTH. The reactions were incubated for 15mins at 30°C to allow initiated complexes to form. The remaining NTPs (100nM), 3μCi of [α-^{32}P]UTP and heparin (100μg/ml) were added and incubation continued for 15mins at 30°C. 4μl of x3 formamide loading buffer (3mg xylene cyanol, 3mg bromophenol blue, 0.8ml of 250mM EDTA in 10ml of deionised formamide) were added and 7μl of the samples loaded onto denaturing 6% sequencing gels. The dried gel was analysed and transcripts were quantified by phosphoimager analysis.

Activator independent assays: These were performed essentially as described above for activator dependent transcription, but in the absence of PspFΔHTH activator protein. The *S. meliloti nifH* start sequence is GGG (+1 to +3, non-template strand) and the *E. coli*

glnHp2 start sequence is TGT (+1 to +3, non-template). To allow activator independent initiated complex formation the holoenzyme was incubated with 4mM GTP (when using *S. meliloti nifH* as the template) or with 4mM GTP, 0.1mM UTP and 0.1mM CTP (when using *E. coli glnHp2* as the template) for 15mins at 30°C. The remaining nucleotides (100nM), 3μCi of [α-^{32}P]UTP and heparin (100μg/ml) were added and incubation continued for 15mins at 30°C. 4μl of formamide loading buffer were added and 7μl of the samples loaded onto denaturing 6% sequencing gels (Chaney & Buck, 1999). The dried gel was analysed and transcripts were quantified by phosphoimager analysis.

2.7 Chapters Three & Four Specific Methods

2.7.1 Construction of Single Cysteine σ^{54} Mutants

Construction of Cys198 and Cys346: pMTHσ (pET28b+) (Gallegos & Buck, 1999) containing the *K. pneumoniae rpoN* gene as a 1.68kbp NdeI-HindIII fragment was used to make the Cys198 and Cys346 protein donor plasmids by site-directed mutagenesis, using the HPLC purified primers purchased from Sigma Genosys (see Appendix C for sequences) to remove the other naturally occurring Cys-codons by C198A and C346A mutagenesis, respectively. The presence of the C198A (pSRW346) or C346A (pSRW198) mutations was confirmed by sequencing the whole *rpoN* gene in the pET28b+ vector using T7 and T7T primers (see Appendix C for sequences).

Construction of the Cys(-) protein: The Cys(-) mutant σ^{54} protein, donor plasmid, pSRWCys(-) was made by a double ligation procedure. The NdeI-BamH1 (770bp) and BamH1-HindIII (914bp) fragments containing the C198A and C346A mutations, respectively, were ligated. This ligation reaction was then digested with NdeI and HindIII for 1hr and the restriction enzymes were heat inactivated. The 1.68kbp ligated product (NdeI-HindIII) was then ligated into a pET28b+ vector previously digested with NdeI and HindIII. The ligation products were transformed into *E. coli* JM109 cells. Positive transformants were identified by colony PCR and the presence of the 1.68kpb NdeI-

HindIII fragment containing the C198A and C346A mutations was confirmed by DNA sequencing.

Construction of Cys36, Cys336 and Cys383 protein donor plasmid mutations: The mutations were made by site directed mutagenesis using pSRWCys(-) as the template DNA and primers purchased from Sigma Genosys (see Appendix C for sequences) The presence of the E36C, R336C and R383C mutations, respectively, were confirmed by sequencing the whole *rpoN* gene in the pSRWCys(-) derivatives.

2.7.2 Assay for Free Sulfhydryl Groups (CPM Test)

A modified version of the method by Parvari *et al.*, (1993) was used to fluorimetrically determine the presence of free sulfhydryl groups on σ^{54} single cysteine mutants. Briefly, β-Mercaptoethanol (BME) standard solutions were prepared in MOPS buffer (10mM MOPS (pH 8.0), 0.1mM EDTA, 50mM NaCl) at concentrations of 100, 50, 20, 10, 5, 1, 0.5 and 0.1μM. The single cysteine mutant σ^{54} was exchanged into MOPS buffer by dialysis at 4°C and protein concentration was determined by Bradford assay. A 15μl aliquot of 0.4mM CPM (7-diethylamino-3-[4'-maleimidylphenyl-4-methylcoumarin) (Molecular Probes) in DMF (dimethylformamide) (BDH) was added to 15μl of each standard and each protein sample. After 1 hour incubation at 37°C, the reaction was stopped by adding 3ml of 1%(v/v) Triton X-100. The intensity of fluorescence emission was measured on 1ml samples, using a Perkin-Elmer 2000 fluorescence spectrophotometer. The excitation wavelength was 390nm, and the emission wavelength was 473nm. A negative control reaction was also conducted using the cysteine-free σ^{54} mutant.

2.7.3 Conjugation of σ^{54} Single Cysteine Mutants with FeBABE

For conjugation with FeBABE [(p-bromoacetamidobenzyl)-EDTA Fe] the σ^{54} mutants/fragments were buffer exchanged into conjugation buffer (10mM MOPS, 0.2M

NaCl, 5% (v/v) glycerol, 2mM EDTA) by dialysis overnight at 4°C. Conjugation reactions were initiated by adding 10-fold molar excess of FeBABE (Dojindo Chemicals, Japan) to each σ^{54} single cysteine mutant/fragment. After 1hr incubation at 37°C, excess FeBABE was removed by overnight dialysis against storage buffer (10mM Tris-Cl, 0.1mM EDTA, 50mM NaCl, 50% (v/v) glycerol, pH 8). For conjugation with FeBABE under denaturing conditions, the protein was dialysed overnight at 4°C against conjugation buffer containing 6M deionized urea and treated essentially as described above. Control reactions consisted of using the cysteine-free σ^{54} for conjugation and carrying out the conjugation reaction with single cysteine σ^{54} mutants in the absence of FeBABE (unconjugated samples).

2.7.4 Estimation of FeBABE Conjugation Yield

The conjugation yield is estimated fluorimetrically by the CPM test by comparing the fluorescence intensity of conjugated versus unconjugated samples relative to the BME standard curve. To calculate the yield, the fluorescence value is plugged into the linear equation (y=mx+c) (see below) obtained for the BME standard curve and the x-value is calculated. This x-value represents the reactive concentration of cysteine on the protein, i.e. [cys]. Dividing the [cys] by protein concentration, i.e. [cys]/[protein] will give the fraction ratio of free cysteine on the protein. For unconjugated samples, this number will be close to one. Clearly, the [cys] should be lower after FeBABE conjugation (presuming some conjugation was achieved) than for the unconjugated protein. But note that the reactive cysteine concentration of a protein is usually less than the actual concentration of the protein (due to thiol oxidation, protein structure etc.). The % conjugation yield is $[cys]/[protein]_{unconjugated}-[cys]/[protein]_{conjugated}x100$.

2.7.5 Core RNAP cleavage by FeBABE modified σ^{54}

The *E. coli* RNAP holoenzymes were prepared by incubating core RNAP with FeBABE modified σ^{54} mutants/fragments (core RNAP:σ^{54} 1:2) at 30°C for 10mins. in cleavage buffer (10mM MOPS, 10mM $MgCl_2$, 10% (v/v) glycerol, 200mM NaCl, 2mM

EDTA). Holoenzymes formed with the cysteine-free σ^{54} (subjected to the FeBABE reagent) constituted the negative control reaction for the cleavage assays. Cleavage reactions were initiated by the rapid sequential addition of freshly prepared ascorbate (5mM final) and hydrogen peroxide (5mM final) and were allowed to proceed for 2mins at 30°C. Reactions were stopped with x5 SDS sample buffer (62.5mM Tris-Cl, 2% (w/v) SDS, 5% (v/v) BME, 10% (v/v) glycerol, 25mM EDTA, 0.02% (w/v) bromophenol blue) and immediately frozen in liquid nitrogen and stored at –80°C. The cleaved fragments were separated on 7.5%, 12.5% and 15% SDS polyacrylamide gels, blotted and visualised by immunostaining with affinity purified subunit β and β' N-terminus-specific antibodies (provided by Akira Ishihama)

2.7.6 Assignment of Cleavage Sites on β or β' Subunits

The β and β' subunits were cleaved at either cysteine or methionine residues using the sequence-specific chemical reagents NTCB (2-nitro-5-thiocyanobenzoate) and CNBr (cyanogen bromide), respectively. NTCB cleavage was performed as described previously (Jacobson *et al.*, 1973). For the NTCB cleavage, core RNAP was buffer exchanged into unfolding buffer (0.1M MOPS, 8M urea, pH 8.5) and incubated for 10mins at 37°C. Cleavage was initiated by adding a 5-fold molar excess of NTCB (in 0.1M MOPS buffer) over total sulfhydryl groups in core RNAP. After overnight incubation at 37°C, the cleavage reactions were stopped with 2 volumes of x5 SDS sample buffer (125mM Tris-Cl, 4% (w/w) SDS, 20% (v/v) glycerol, 10% (v/v) BME, 0.004% (w/v) bromophenol blue) and analysed by SDS-PAGE and immunostaining with β and β' N-terminus-specific antibodies. CNBr cleavage was performed essentially as in Grachev *et al.* (1989). Initially, core RNAP was brought to 10% SDS and incubated at 37°C for 30mins. For cleavage, 1M HCl and 1M CNBr in acetonitrile were then added and incubated at room temperature for 6hrs. The reaction was stopped with x5 SDS sample buffer and analysed as described above. The cleavage sites for each NTCB- and CNBr- generated fragment of β and β' were assigned by using third-order polynomial fit of log molecular weight versus migration

distance on SDS-PAGE of a known set of marker fragments, and the known β and β' amino acid sequences.

2.8 Chapter Five Specific Methods

2.8.1 DNA Cleavage of the *S. meliloti nifH* Promoter DNA

10µl σ^{54}-DNA (either homoduplex, early or late melted DNA) binding reactions were conducted in cleavage buffer (40mM HEPES (pH 8.0), 10mM $MgCl_2$, 5% (v/v) glycerol, 0.1M KCl, and 0.1mM EDTA) with 20nM DNA and 2µM σ^{54}, and incubated for 10 minutes at 30°C. For σ^{54}-isomerisation assays, 4mM dGTP (or where indicated, GTPγS) and 4µM PspFΔHTH were added for a further 10 minutes. Closed complexes were formed with 100nM holoenzyme (ratio 1:2, core RNAP to Fe-BABE modified σ^{54}). For open complex formation, 4mM dGTP (where indicated GTPγS) and 4µM PspFΔHTH were added and incubated for a further 10 minutes. Following DNA-binding or activation, cleavage of promoter DNA carrying FeBABE modified σ^{54} proteins or holoenzymes containing FeBABE modified σ^{54} proteins was initiated by the rapid sequential addition of 2mM sodium ascorbate (pH 7.0) and 1mM hydrogen peroxide. Reactions were allowed to proceed at 30°C for 10 minutes before quenching with 30µl stop buffer (0.1M thiourea and 100µg/ml sonicated salmon sperm DNA) and 80µl TE buffer. The stopped reactions were phenol:chloroform extracted, precipitated with ethanol and electrophoresed on 10% denaturing urea-polyacrylamide gels. The cleavage sites were determined by using ^{32}P end-labelled fragments of the *S. meliloti nifH* promoter DNA.

2.9 Chapter Six Specific Assays

2.9.1 Construction of R383A and R383K σ^{54} Mutants

These mutations were made by site directed mutagenesis using pMTHσ (Gallegos & Buck, 1999) as the template DNA and primers purchased from Sigma Genosys (see Appendix C for sequences) The presence of the R383A and R383K mutations, respectively, were confirmed by sequencing the whole *rpoN* gene in the pMTHσ derivatives using T7 and T7T primers (see Appendix C for sequences).

2.9.2 Making *E. coli glnHp2* Promoter Fragments by PCR

The *E. coli glnHp2* promoter fragments were obtained by PCR (denaturing: 96°C for 45s; annealing: 52°C for 45s; extension: 72°C for 30s; 30 cycles)) using pFC50 and pFC50-m12 (Claverie-Martin & Magasanik, 1991) as templates and the primers glnHp2-53 and glnHp2-35 (see Appendix C for sequences). The promoter fragments were gel purified and end-labelled with ^{32}P.

2.9.3 *In vitro* Transcription Assay with NtrC

These were conducted essentially as described in section 2.6.5, but using 100nM NtrC and 10mM carbamyl phosphate (for NtrC phosphorylation) for activation.

2.10 Chapter Seven Specific Methods

2.10.1 Construction of σ^{54} Mutants for *in vivo* Work

For *in vivo* promoter activation assays, the σ^{54} mutantations E36G, F318A, E325G, E410A, E414A, E431A and EE410/414AA were made by site-directed mutagenesis using

pMM83 as a template. The sequences of primers used for making these σ^{54} mutantions are given in Appendix C. pMM83 is a low copy number plasmid that contains a 1.98kbp fragment containing the *K. pneumoniae rpoN* gene under the control of its own promoter. pMM83 also contains the *cat* gene which confers resistance to chloramphenicol. Thus, pMM83 is a suitable vector for *in vivo* study of σ^{54} mutants in the *K. pneumoniae* UNF2792 background, which contains a kanamycin cassette inserted into the chromosomal rpoN gene, and thus is unsuitable for use with plasmid vectors which encode for kanamycin resistance, such as the pET family of vectors. The presence of the E36G, F318A, E325G, E410A, E414A, E431A and EE410/414AA mutations, respectively, were confirmed by sequencing the whole of the *rpoN* gene in the pMM83 derivatives using the primers MP64, N1780 and RpoN120 (see Appendix C for sequences).

2.10.2 Construction of σ^{54} Mutants for *in vitro* Work

For overexpression plasmids producing the σ^{54} mutants E36G, F318A, E325G, E410A, E414A, E431A and EE410/414AA were made by site-directed mutagenesis using pMTHσ (Gallegos & Buck, 1999) as the template and primers as indicated above. The presence of the E36G, F318A, E325G, E410A, E414A, E431A and EE410/414AA mutations, respectively, were confirmed by sequencing the whole of the *rpoN* gene in pMM83 using the primers T7 and T7T (see Appendix C for sequences).

2.11 Chapter Eight Specific Methods

2.11.1 Construction of 117-477aa σ^{54} Fragment-encoding Plasmid.

To express the 117-477aa fragment as an N-terminal 6His-tagged fusion protein a NdeI site was introduced in the DNA encoding amino acid residue 116 by PCR mutagenesis, using pDJS14.22 (provided by D. Studholme, which contains the 117-477aa fragment has a BamH1-HindIII fragment in a non-expression vector) and primers

RpoN117 and T7T (see Appendix C for sequence). The PCR product was gel purified and digested with NdeI and HindIII. The PCR fragment was then phenol:chloroform extracted, precipitated with ethanol and ligated into a pET28b+ vector previously digested with NdeI and HindIII to make pSRW117-477. The ligation products were then transformed into *E. coli* JM109 cells. Positive transformants were identified by colony PCR and the sequence of the His-tagged 117-477aa protein encoding region was confirmed by DNA sequencing prior to expression and purification.

2.11.2 Small Angle X-ray Scattering (SAXS) Measurements

For the SAXS measurements the protein samples were dialysed against a buffer solution containing 10mM Tris-Cl (pH 8), 5% glycerol, 75mM sodium thiocyanate at 4°C. For the measurements, protein concentrations ranging from 3-20mg/ml were used (protein concentrations were determined by the Bradford assay). All measurements (at λ=0.15nm) were done using the X33 camera of the EMBL on the storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY), Hamburg, Germany, essentially as described in Svergun *et al.*, (2000). The sample solution was placed in a quartz cell of 100µl volume, and the temperature was kept at 15°C by circulating water through the sample holder. The Q value (see Appendix D) was calibrated for forward scattering, I(0) using BSA as a standard. Scattering data for the Guinier plots and P(r) function were collected at separate, individual sample concentrations to increase the statistical value of the data. Analyses of the scattering data for shape determination by computer analysis were done by Dimitri Svergun and Michel Koch, essentially as described in Svergun *et al.*, (2000).

CHAPTER THREE

Tethered Footprinting as a Methodology for Studying the σ^{54}-core RNA Polymerase Interface

3.1 Introduction

The identification of protein-DNA, protein-RNA and protein-protein contact networks within the transcription apparatus is crucial for understanding the molecular events that occur in transcription and its regulation. Transcription initiation by σ^{54} involves multiple protein-protein (σ^{54}-core RNAP and σ^{54}-activator) and protein-DNA ($E\sigma^{54}$-DNA) interactions. A description of the interface σ^{54} adopts in these macromolecular interactions and target surfaces proximal to σ^{54} within the macromolecular complexes (σ^{54}-core RNAP and $E\sigma^{54}$-DNA) will greatly enhance our understanding of the role σ^{54} has in converting the core RNAP into a holoenzyme form that is dependent on enhancer binding activator proteins coupled to nucleoside triphosphate hydrolysis for transcription initiation.

In the past, the results of genetic studies, chemical cross-linking, DNA and protein footprinting have provided partial information on which parts of σ^{54} are involved in core RNAP binding, promoter DNA interaction and activator response. However, when interpreting the results of genetic studies (random or site-directed mutagenesis and deletion analyses) it is often difficult to discriminate between indirect effects of conformational changes introduced by the mutation or deletion and direct functional effects. In the case of footprinting data (both protein and DNA) specificity limitations of the footprinting reagent and effects of steric hindrance must be considered.

A way forward to overcome the limitations of genetic and footprinting approaches in elucidating the parts of σ^{54} that constitute an interface with core RNAP and/or promoter DNA is provided by a methodology known as protein tethered footprinting. This approach employs σ^{54} conjugated at a chosen location to a small metal chelate that cleaves peptide bonds and DNA, independent of amino acid or nucleic acid sequence, at sites determined

by their proximity to the chelate. The results of protein tethered footprinting provide direct information on the proximity of the tethered probes to particular peptide bonds or DNA sequence, easily exceeding the scope and resolution of chemical cross-linking reagents and conventional footprinting methods.

Furthermore, cleavage of protein or DNA by metal chelates is an important new approach to characterising important structural features of proteins and their complexes (for example with nucleic acids) in solution. Artificial proteolytic reagents such as FeBABE [iron (S)-1(p-bromoacetamidobenzyl)ethylenediaminetetraacetate] or 2IT-FeBABE [2-iminothiolane (S)-1(p-bromoacetamidobenzyl) ethylenediaminetetraacetate] are useful for dissecting protein structure and function and represent a new approach which could be employed for structural analysis and/or determining spatial arrangement of subunits within a macromolecular structure. Cleavage by such reagents is directed by proximity rather than by residue or sequence type.

3.2 Objectives

The core RNAP binding interface of σ^{54} consists of at least two functionally distinct sequences: a 95 amino acid residue sequence (120-215) within Region III which binds core strongly and a second sequence within Region I that binds core more weakly (Gallegos & Buck, 1999). Comparison of hydroxyl radical mediated cleavage profiles of wild-type and Region I-deleted ($\Delta R1\sigma^{54}$) proteins indicates that the overall solvent exposed surface of Regions II and III of free σ^{54} is not greatly altered by deletion of Region I (Casaz & Buck, 1999). Residues protected from hydroxyl radical mediated cleavage in the holoenzyme made with $\Delta R1\sigma^{54}$ largely coincide with residues protected specifically by DNA in the wild-type holoenzyme, implying a role for Region I in establishing an appropriate conformation of the DNA-binding domain in the wild-type holoenzyme (Casaz & Buck, 1999). Thus, it is possible that the interaction of Region I with core RNAP is important for controlling DNA binding domain conformation.

To date, little is known about the location of σ^{54} binding sites on the core RNAP. Recently, we showed by small-angle x-ray scattering studies that a core RNA polymerase binding fragment of σ^{54} (70-324aa) and the crystal structure of a σ^{70} fragment (114-448aa), which includes the main core-binding site have a similar envelope shape, leading to the suggestion that both σ classes might be located in similar places on the core RNAP (see Chapter 8; Svergun *et al.*, 2000). However the sequence differences that exist between σ^{54} and σ^{70} must also account for the markedly different properties of the two holoenzymes. The affinities of σ^{54} and σ^{70} for the core RNAP, measured *in vitro*, are similar (Gallegos & Buck, 1999). Both σ factors could use the same binding site on core RNAP, have partially overlapping sites or even utilise independent binding sites. Recent protein-protein interaction and footprinting studies have revealed that the β' subunit of core RNAP provides the major binding interaction for σ^{70} in the holoenzyme while the β subunit contributes a further binding interaction (Katayama *et al.*, 2000; Arthur & Burgess, 1998; Owens *et al.*, 1998).

The overall objective of the present work is to extend our understanding of the σ^{54}-core RNAP interface using FeBABE derivatives of σ^{54} to map core RNAP surfaces (of β and β' subunits) proximal to σ^{54}. Our approach involved the following experimental stages:

I. To test the suitability of FeBABE tethered footprinting as a method to describe the σ^{54}-core RNAP interface initially by modifing core RNAP binding σ^{54} fragments (1-306aa and 70-324aa), each of which contains a single cysteine, with the FeBABE reagent and mapping the surfaces on core RNAP subunits β and β' proximal to the FeBABE probe.

II. Construct mutant σ^{54} proteins with single cysteine residues located at strategic positions and characterise their phenotypes using *in vivo* and *in vitro* activity assays.

III. Derivatise the σ^{54} single cysteine mutants with the FeBABE reagent and use the products to map sites on the β and β' subunits of the core RNAP proximal to each FeBABE modified residue of σ^{54}.

The experimental stages I and II are presented in the this chapter

3.3 Chemistry of FeBABE

3.3.1 FeBABE modified Proteins as Chemical Proteases

FeBABE is derived from BABE, which is a bifunctional chelating agent that has a strong Fe binding moiety. The iron bound BABE (FeBABE) can be tethered covalently to sulphydryl groups of cysteine residues in proteins to make a FeBABE conjugated protein (Figure 3A). The FeBABE modified protein can be used as a molecular protease or nuclease that cleaves amino acid or DNA sequences that are within 12Å from the FeBABE attachment site, irrespective of the amino acid residues or DNA base pairs involved.

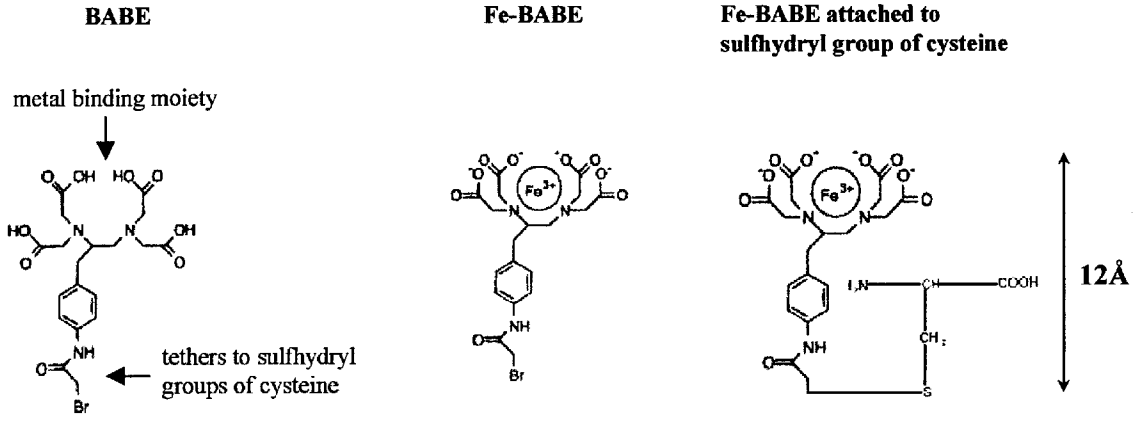


Figure 3A. Biochemical properties of the FeBABE reagent

Target protein cleavage by FeBABE modified proteins is not believed to result from the generation of a powerful nucleophile that selectively attacks carbonyl carbon, rather than from generation of hydroxyl radicals. In the presence of H_2O_2 (or O_2) and ascorbate, Fe-BABE forms an intermediate oxygen-activated complex that leads to

nucleophilic attack by oxygen on the carbonyl carbon of the peptide bonds within 12Å proximity from the FeBABE, and the chemistry is outlined in Figure 3B.

The peptide bond cleavage appears to be dominated by a powerful nucleophilic iron-peroxo complex. The first step involves the formation of intermediate A by the direct binding of H_2O_2 to the Fe chelate. Formation of intermediate A can also occur when O_2 is reduced to H_2O_2 by ascorbate in the presence of the Fe chelate. The coordinated peroxide acts as a nucleophile to attack the peptide bond, resulting in the formation of intermediate B. The final step (via intermediates C_1 and C_2) causes heterolytic cleavage of the C-N bond and of the peroxide bond to yield a new carboxyl terminnus. Using $[O^{18}]H_2O_2$, it has been shown that this chemistry results in the *transfer* of an oxygen atom from H_2O_2 to the newly formed C-terminal carboxyl group (Rana & Meares, 1991). Following cleavage, the metal complex is presumed to be converted back to its original form, since peptide chains were cleaved twice during the cleavage of bovine serum albumin (Rana & Meares, 1991). Evidently, the chelate generates reactive species at least twice, and possibly more times (Rana & Meares, 1991, 1990).

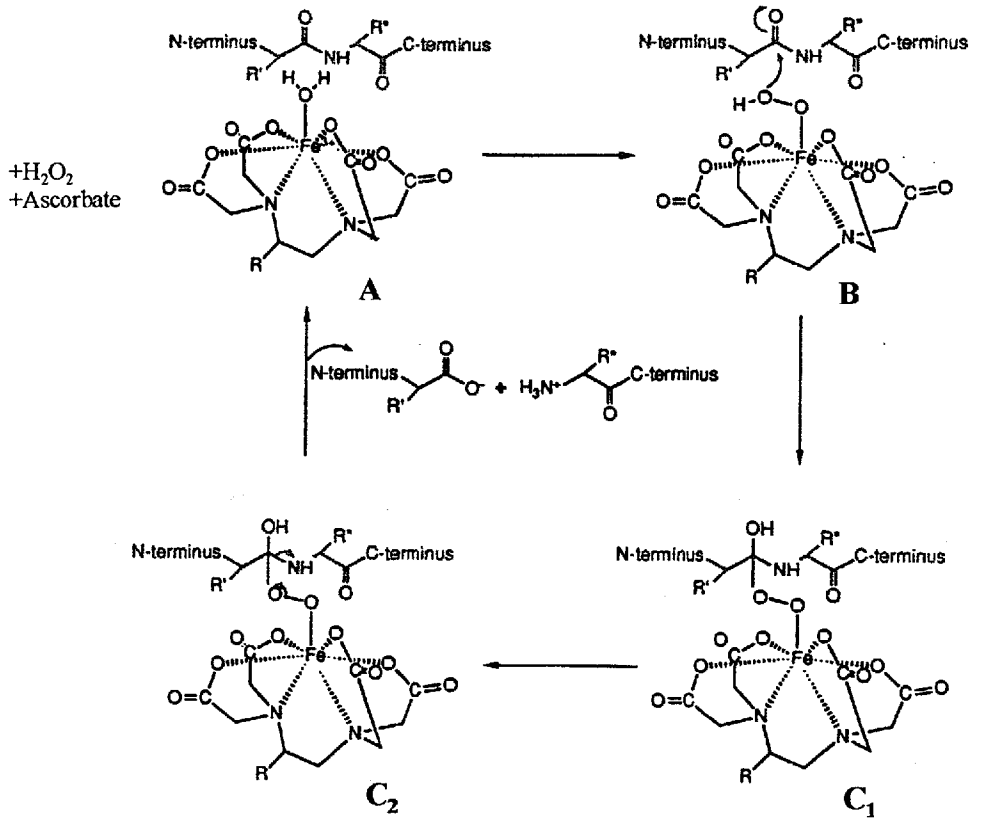


Figure 3B. Proposed mechanism of peptide bond cleavage by the FeBABE reagent (adapted from Rana & Meares, 1991). Shown in green is the target protein being cleaved by the FeBABE reagent (see text for details).

The oxidation state of the Fe in intermediate B is unknown. If it is in ferrous form (Fe^{+2}), intermediate B will decompose, resulting in hydroxyl radicals ($OH^{\bullet}$) formation. However, it is thought that the efficiency of the nucleophilic attack is such that it occurs before hydroxyl radical formation can take place, in which case the proximity and orientation of the reacting groups is crucial to the observed selectivity of the reaction. Furthermore, the proposed mechanism of nucleophilic cleavage also satisfies the experimental observations of 1:1:1 ascorbate : H_2O_2 : cleavage stoichiometry, O atom transfer to the new carboxyl group, extremely narrow geometric constraints, insensitivity to radical scavengers and the regeneration of the starting metal species (Rana & Meares, 1991 & 1990).

3.3.2 Derivatisation of Single Cysteines on Proteins with the FeBABE reagent

FeBABE is a sulphydryl specific protein modifying reagent. The general protocol for the derivatisation of proteins with FeBABE is illustrated in Figure 3C.

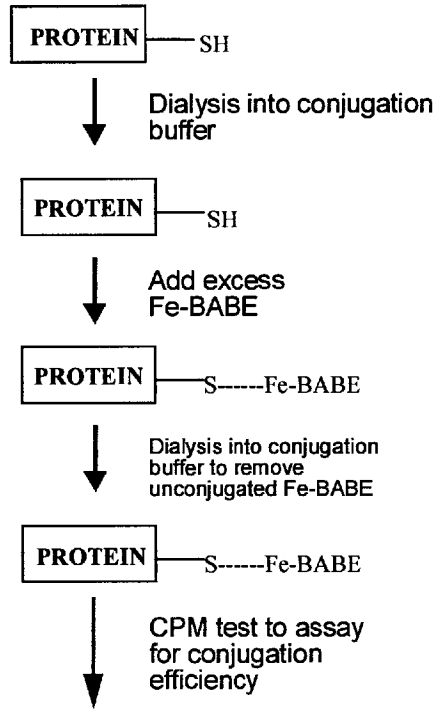


Figure 3C. General procedure for FeBABE conjugation of proteins (see text for details)

The procedure begins with dialysis of the protein against a Tris or MOPS based buffer containing 0.1mM EDTA. The presence of EDTA in the conjugation buffer is important as it is necessary to scavenge any unbound metal. The conjugation buffer also contains simple salts (NaCl or KCl) and glycerol for macromolecular stability. The buffer must not contain any free thiols. The amount of FeBABE added should be in 10-fold excess relative to the free thiol present in the protein sample. Covalent modification of sulphydryl groups with FeBABE is brought about by incubating the sample with the FeBABE reagent at 37°C for 1 hour or at room temperature overnight. The conjugation reaction is then stopped by dialysing the reaction against fresh conjugation buffer to remove free FeBABE from the

solution. The conjugation yield is fluorimetrically determined using the sulphydryl specific dye CPM (7-diethylamino-3-[4'-maleimidylphenyl]-4-methylcoumarin).

3.3.3 Using FeBABE Methodology to Characterise the σ^{54}-core RNAP interface

K. pneumoniae σ^{54} protein contains two native cysteine residues in the C-terminal Region III at positions 198 and 346. Core RNAP binding σ^{54} fragments (70-324aa and 1-306aa), each containing single cysteines at position 198, and the full-length wild-type σ^{54} containing an additional cysteine residue at position 346 were used for FeBABE modification (Figure 3D).

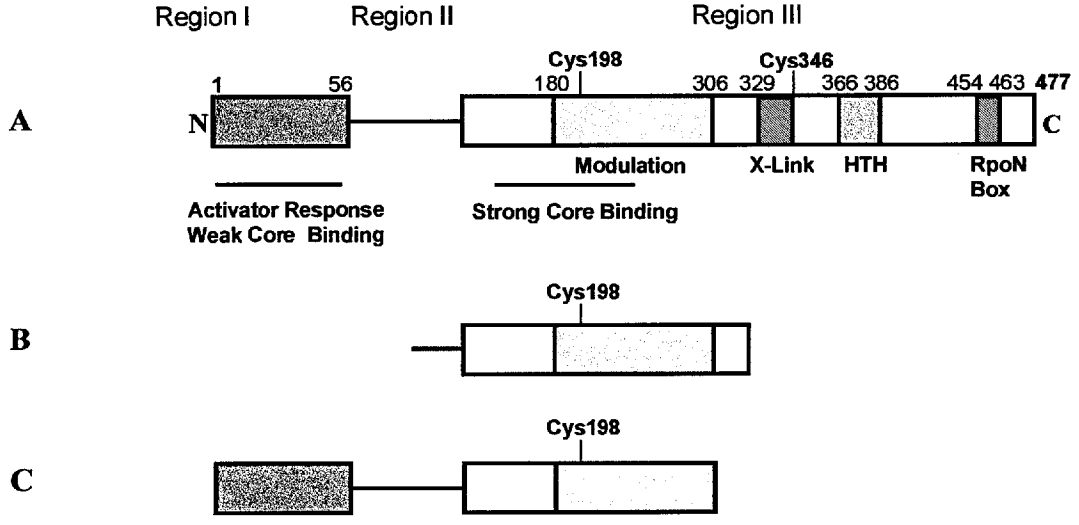


Figure 3D. Schematic diagram showing σ^{54} fragments used for FeBABE modification. (A) full-length, (B) 70-324aa fragment and (C) 1-306aa fragment. The native cysteines at positions 198 and 346 are shown.

The general approach taken to characterise the σ^{54}-core RNAP interface, initially using FeBABE modified σ^{54} fragments, is schematically shown in Figure 3E. This cartoon also serves to illustrate the limitations of the technology, which are mainly associated with the role and nature of the cysteine residue to which the FeBABE is tethered.

Stage 1: Initially, all native cysteine residues must be removed from the test protein (e.g. σ^{54} in Figure 3E). The cysteine-free mutant test protein can then be used as a control for FeBABE modification. Single cysteines can now be introduced where required.

Stage 2: Some proteins are sensitive to the removal or introduction of cysteine residues from or at, functional sites, respectively. It is therefore important to determine the functionality of the single-cysteine mutant test proteins (i.e. here does the single cysteine mutant σ^{54} complex with core RNAP to form a holoenzyme ?)

Stage 3: Once a functionally active single-cysteine mutant of the test protein is made, the cysteine residue could be unreactive towards alkylating reagents. For example, the cysteine residue could be structurally buried and therefore inaccessible for the FeBABE reagent. One commonly used approach to overcome this problem is to conjugate the test protein under denaturing conditions (see later).

Stage 4: A crucial experimental step when using the FeBABE methodology is to assess the functionality of the FeBABE modified test protein. The test protein could be functionally intolerant to the presence of the 490 Da FeBABE molecule.

Stage 5: Some single-cysteine mutant test proteins may become conjugated with the FeBABE reagent but produce no cleavage upon treatment with ascorbate and hydrogen peroxide.

Nevertheless, if the conjugation and cleavage reactions are successful, the results will provide direct information on the proximity of the tethered probe to particular peptide bonds within the target protein (e.g. core RNAP in Figure 3E), which can be mapped either by using terminal specific antibodies or Edman sequencing of the cleavage products. In this work, we aimed to modify two core RNAP binding σ^{54} fragments with FeBABE and use them as chemical proteases to preliminarily characterise the proximity of σ^{54} to the core RNAP subunits β and β' within the holoenzyme.

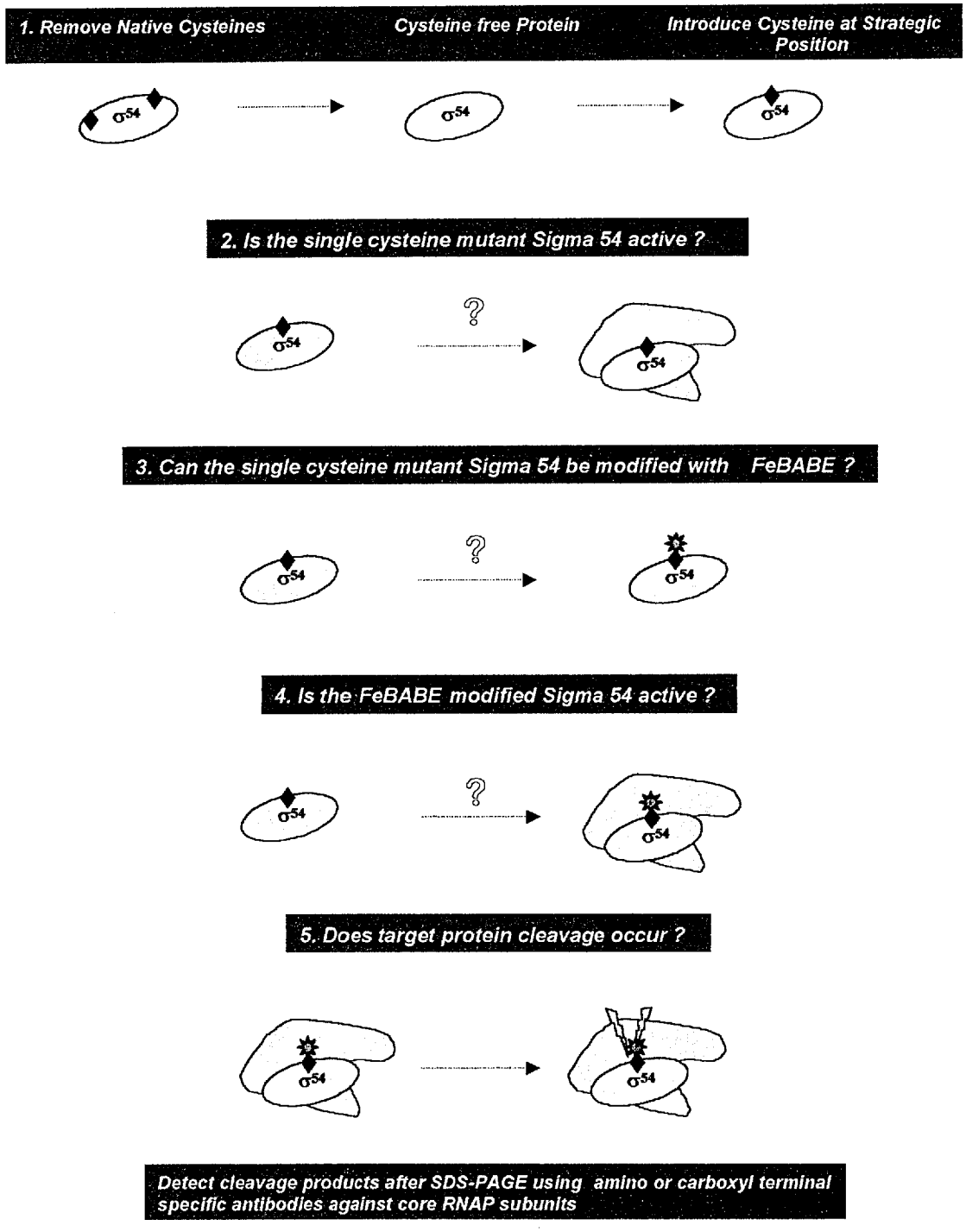


Figure 3E. Cartoon showing the experimental stages involved in using the FeBABE footprinting methodology for mapping surfaces on core RNAP proximal to σ^{54}. Black diamonds (♦) symbolise cysteine residues and a red star (✲) indicates the FeBABE reagent.

3.4 Modifying σ^{54} Fragments with FeBABE

3.4.1 FeBABE Modification

The σ^{54} fragments 70-324aa and 1-306aa and the full-length wild-type protein failed to become conjugated by FeBABE under native conditions, suggesting some burying of the native cysteine residues at positions 198 and 346, although not sufficient to stop access by the CPM reagent (see later). However, under denaturing conditions, that is, in the presence of 8M urea in the conjugation buffer, FeBABE modification of cysteines 198 and 346 occurred with high efficiency (Figure 3F)

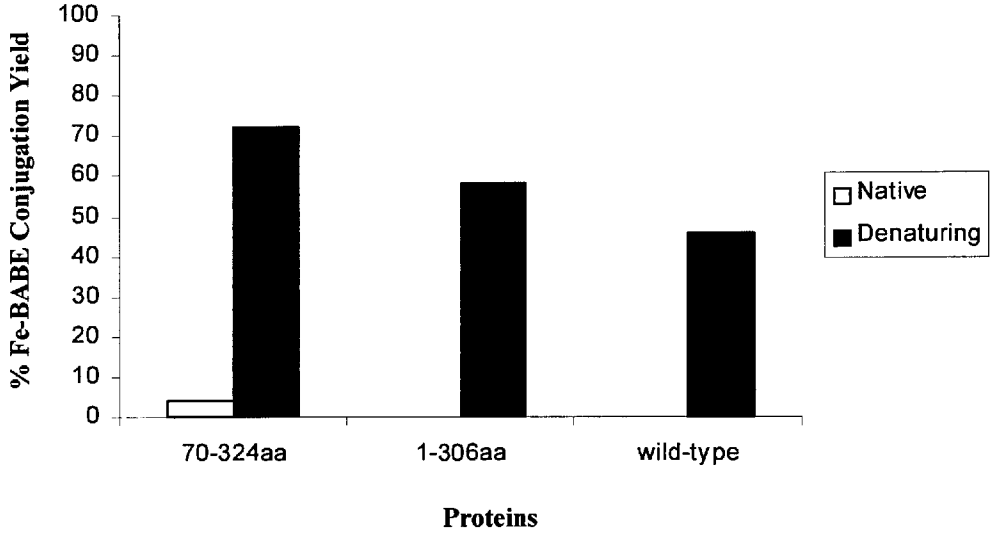


Figure 3F. FeBABE conjugation efficiency of the σ^{54} fragments 70-324aa and 1-306aa and the wild-type (1-477aa) protein. Shown are conjugation efficiencies when conjugated under native (white bars) and denaturing (black bars) conditions.

The great improvement of the FeBABE conjugation yield under denaturing (unfolded) conditions suggests that the cysteine residues at positions 198 and 346 are located in hydrophobic regions of the σ^{54} protein, because the hydrophobic reagent CPM reacted with the sulphydryl groups with higher yield under native conditions, whereas the

hydrophilic FeBABE did not (data not shown). A similar interpretation arose from a study by Miyake *et al.*, (1998) in which FeBABE conjugation of native cysteines in the *E. coli* core RNA polymerase α-subunit significantly improved under denaturing conditions.

3.4.2 Core RNAP binding of FeBABE modified σ^{54} fragments

Next, we assayed the ability of the renatured (by dialysing into urea-free storage buffer) FeBABE modified σ^{54} fragments to bind the core RNAP. The results showed that the two core RNAP binding fragments 70-324aa and 1-306aa with FeBABE at aa 198 and the wild-type σ^{54} harbouring FeBABE at positions 198 and/or 346, complexed with core RNAP to form holoenzyme at a 1:1 molar ratio (Figure 3G). The results imply that the FeBABE modified proteins still retain the ability to bind core RNAP. The presence of the 490 Da FeBABE reagent has not caused conformational changes in the σ^{54} core RNAP binding fragments (or the wild-type protein) that disrupt proper binding to core RNAP.

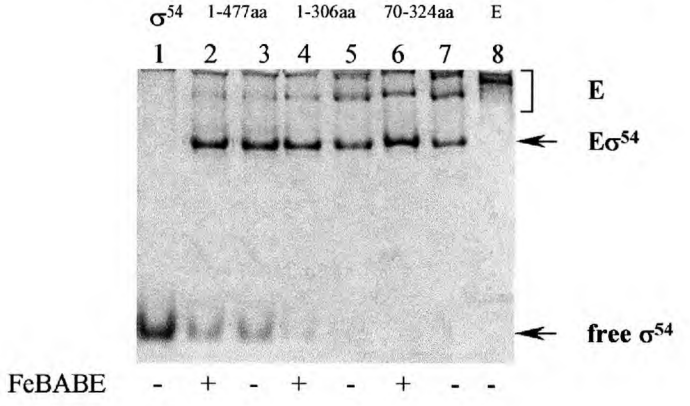


Figure 3G. Holoenzyme gel assembly assay with FeBABE modified (+, *lanes* 2, 4 and 6) and unmodified (-, *lanes* 3, 5 and 7) σ^{54}-fragments (70-324aa and 1-306aa) and the wild-type (1-477aa) protein. The migratory positions of the free core RNAP (*lane* 8) , holoenzyme ($E\sigma^{54}$) and free σ^{54} (*lane* 1) are as indicated.

3.4.3 Holoenzyme Cutting with FeBABE modified σ^{54}

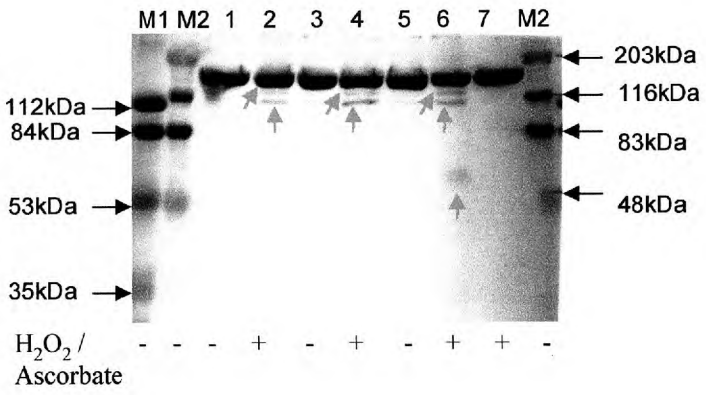
Holoenzyme cleavage by the FeBABE modified σ^{54} fragments and the full-length wild type σ^{54} was initiated by the sequential addition of ascorbate and H_2O_2. Conditions were developed wherein only a small percentage of the holoenzyme complexes was cleaved, that is, the cleavage reaction was only allowed to proceed for two minutes. Longer cleavage periods resulted in artifacts produced by secondary cleavage and the instability of the β and β' core RNAP subunits (personal communication, Akira Ishihama and data not shown). The cleavage products were separated by denaturing PAGE and blots were immunostained with affinity-purified β and β' amino terminal specific antibodies (Figure 3H). The summary of the cleavage data relating to the proximities of Cys198 and Cys346 to sites on β and β' is shown in Figure 3I, in which the σ^{54} cleavage sites are also compared to those obtained by using FeBABE modified σ^{70} proteins (Owens *et al.*, 1998).

Comparison of the relative molecular weights of the cleavage products with the amino acid sequence of the β and β' subunits reveals that Cys346 is proximal to the RifI cluster of the β subunit. Cleavage of the holoenzyme formed with the FeBABE modified (at Cys198 and Cys346) wild-type σ^{54} protein produces a ~50kDa β cleavage product which is absent in cleavage reactions in which holoenzyme was cleaved either with the 70-324aa or 1-306aa fragment, having only Cys198 conjugated with FeBABE (Figure 3H-A, compare lanes 2 and 4 with lane 6). Thus, we conclude that the 50kDa cleavage product is due to cleavage of β occurring from the FeBABE conjugated to the Cys346. However, we do not exclude the possibility that wild-type σ^{54} and the 70-324aa or 1-306aa fragments adopt different conformations when complexing with core RNAP to form the holoenzyme, hence explaining the different cleavage patterns. Cys198 appears to be closer to the carboxyl terminal end of the β subunit. It strongly cleaves the β' subunit at sites near its amino (most clearly for β' subunit) and carboxyl terminal ends. However, it is noteworthy that some inherent unstable sites (most clearly for β' subunit) become intensified by the addition of H_2O_2 and ascorbate. Similarly, we note a loss of cleavage at the most amino proximal site of β' when the intact σ^{54} is used, possibly indicating that the oritientation of

Cys198 w.r.t. β' is changed when present in the intact σ^{54} as opposed to in the 70-324aa and 1-306aa fragments.

The strong cleavage patterns observed with the σ^{54} fragments 70-324aa and 1-306aa harbouring a FeBABE modifed cysteine residue at position 198 are consistent with previous observations that Cys198 is located within a 95aa sequence (120-215aa) that binds core RNAP with high affinity (Gallegos & Buck, 1999). Even though Cys346 is located outside the minimal core RNAP binding domain in σ^{54}, it is adjacent to a region in σ^{54} that is protected by core RNAP when Region I is removed in hydroxyl radical footprinting experiments (Casaz & Buck, 1999). We also observe cleavage at sites proximal to the carboxyl terminal end of the β' subunit. In contrast, FeBABE modified σ^{70} proteins did not produce any cutting near the carboxyl terminal end of the β' subunit (Owens *et al.*, 1998). This is the main difference seen between the β and β' contact sites for σ^{54} and σ^{70}, which are otherwise strikingly similar.

A.



B.

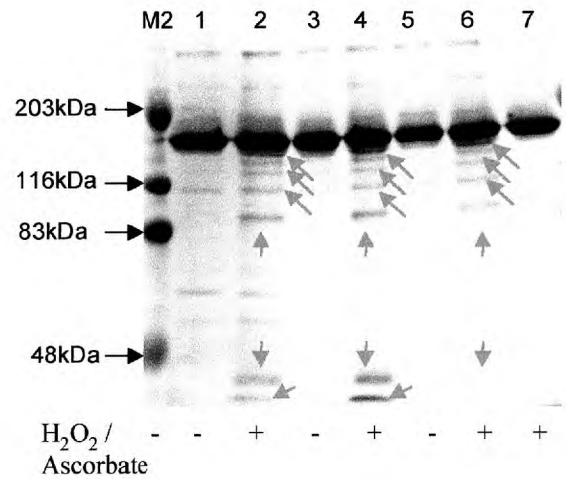
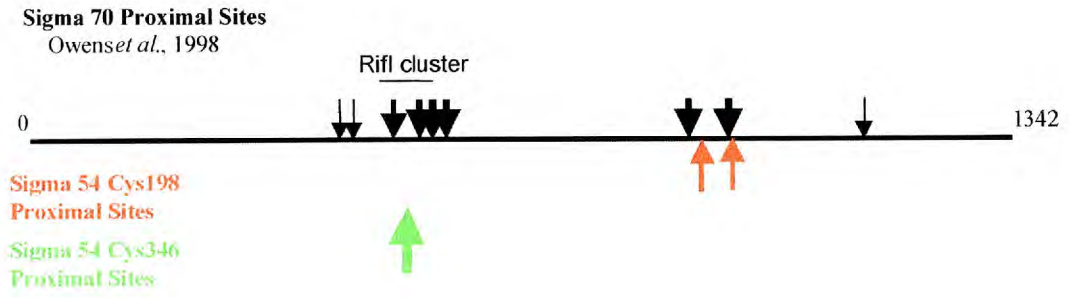


Figure 3H. Holoenzyme cutting by FeBABE modified σ^{54} fragments. Immunostained blots of SDS-PAGE detecting (**A**) β and (**B**) β'-subunit cleavage products using affinity purified N-terminal antibodies. *Lanes* 1&2: 70-324aa fragment conjugated and unconjugated, respectively; *lanes* 3&4: 1-306aa fragment conjugated and unconjugated, respectively; *lanes* 5&6: full-length σ^{54} conjugated and unconjugated, respectively; *lane* 7: core RNAP; M1 and M2 are prestained markers. Red arrows indicate cleavage products. All samples were treated (+) or untreated (-) with ascorbate and hydrogen peroxide.

A.



B.

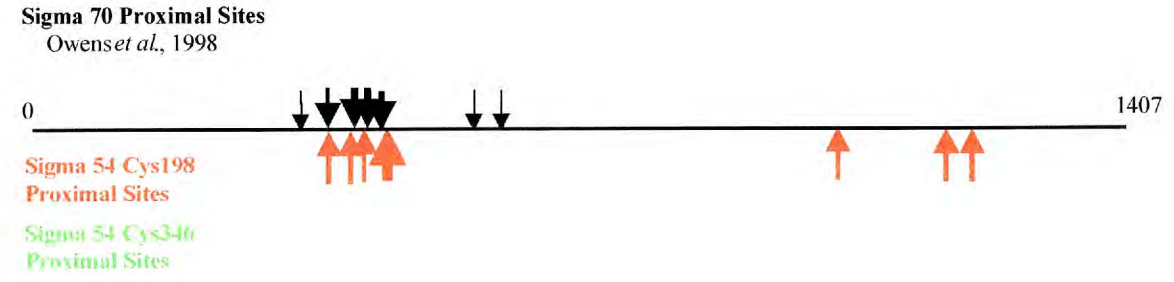


Figure 3I. Summary of cleavage data in Figure 3H, showing the proximities of Cys198 (red) and Cys346 (green) in σ^{54} to residues on (**A**) β and (**B**) β' subunits. The thickness of the arrows indicates the cleavage efficiencies. The σ^{70} cleavage sites, based on the results from Owens *et al.*, 1998, are shown for comparison.

3.5 Single Cysteine Mutants of σ^{54}

In order to expand the preliminary data obtained with FeBABE modified σ^{54} fragments and to further characterise the σ^{54}-core RNAP interface we constructed five single cysteine mutants of σ^{54}. Using information from extensive mutational analysis of σ^{54}, we selected five sites (E36, C198, R336, C346, and R383, described below) of

potential interest for tethering FeBABE. Figure 3J shows a schematic representation of the primary structure of σ^{54}, in which the positions of introduced cysteine residues are indicated.

Cys198 and Cys346. The σ^{54} from *K. pneumoniae* contains two highly conserved cysteine residues at amino acid positions 198 and 346. Cys198 is located in a 95 amino acid sequence (120-215) which binds core RNAP strongly and which is likely to include most of the fold critical for forming stable holoenzyme (Gallegos & Buck, 1999). Cys346 borders a DNA interacting patch in σ^{54} (Cannon *et al.*, 1994) and is strongly protected by core RNAP in hydroxyl radical cleavage experiments when Region I is absent (Casaz & Buck, 1999). By changing Cys198 and Cys346 to alanine, we produced two full-length mutant σ^{54} proteins with single cysteines at positions 198 (C346A) and 346 (C198A). The fully Cys(-)*rpoN* construct was then made, allowing further mutant construction, and also production of the cysteine-free double mutant (C198A and C346A) protein which was used as a negative control protein in the FeBABE conjugation and cleavage assays. The Cys(-)σ^{54} protein was found to maintain 90% of its transcriptional activity (see later), implying that substitution of the conserved Cys198 and Cys346 with alanine does not significantly affect the conformation of the mutant protein.

Cys36. In the next step, single cysteine residues were introduced into Cys(-)σ^{54}. Region I of σ^{54} has multiple roles in regulating σ^{54} function: (1) inhibiting polymerase isomerisation in the absence of activation and (2) directing fork junction DNA binding (Cannon *et al.*, 1999; Guo *et al.*, 1999; Syed & Gralla, 1998; Wang *et al.*, 1995), (3) stimulating initiation of open complex formation in response to activation (Casaz *et al.*, 1999; Sasse-Dwight & Gralla, 1990; Syed & Gralla, 1998; Gallegos & Buck, 2000). The low affinity of Region I sequences for core RNA polymerase and its protection from proteolysis by core RNAP (Casaz & Buck, 1997; 1999; Gallegos & Buck, 1999) prompted us to introduce a single cysteine residue in Region I in the Cys(-)σ^{54} protein. The increased protease sensitivity of residue E36 detected upon open complex formation suggested that a cysteine here might lie close to the core RNAP subunits in the closed promoter complex (Casaz & Buck, 1997).

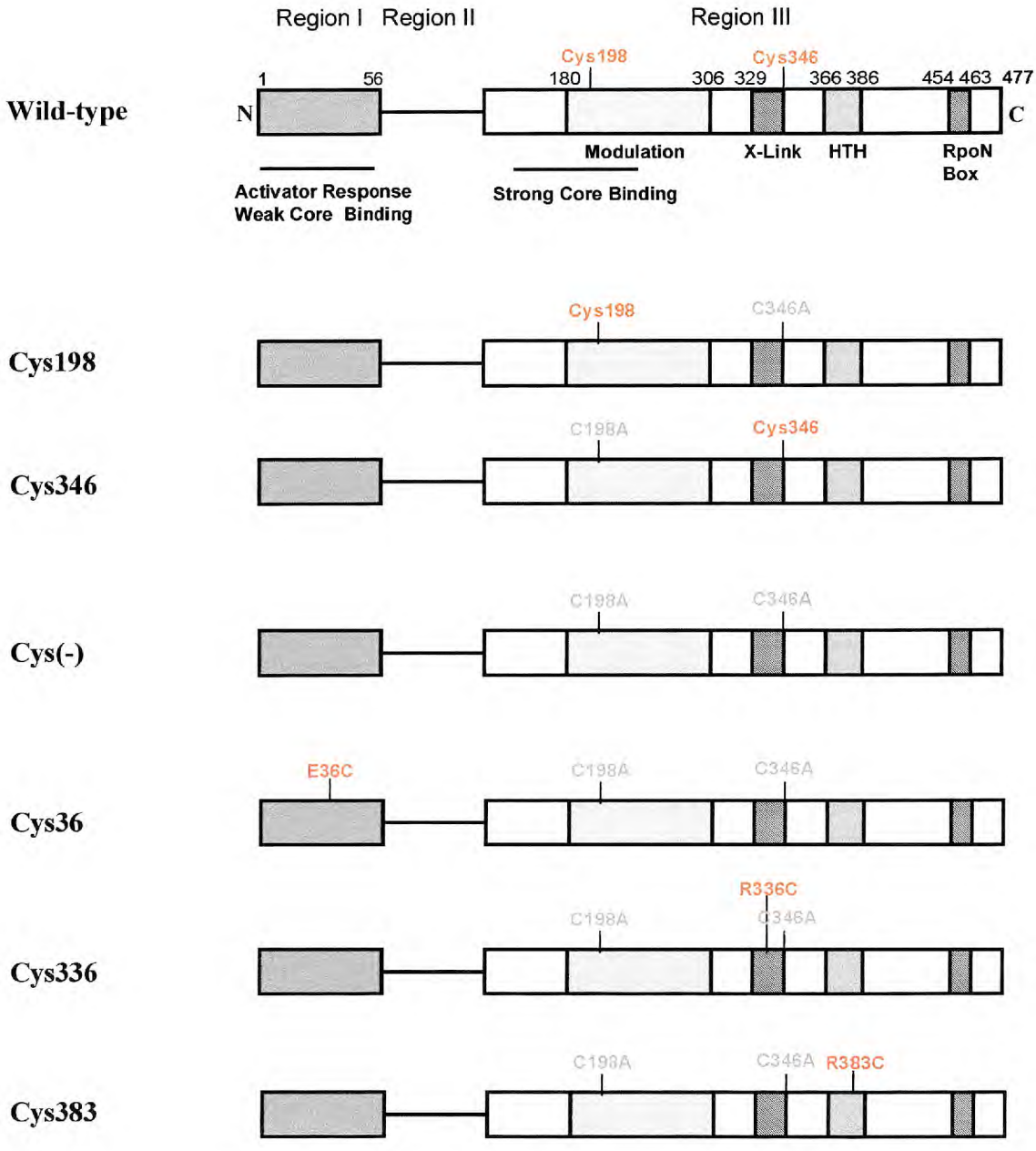


Figure 3J. Single cysteine mutants of σ^{54}. The sites chosen for introducing a cysteine residue are marked in red (see text for details)

Cys336. Recently it has been shown that a single amino acid substitution R336A in the Region III DNA binding domain of σ^{54} allows holoenzyme to isomerise and transcribe without activator (Chaney & Buck, 1999), a phenotype previously only associated with

Region I mutations (Casaz *et al.*, 1999; Syed & Gralla, 1998; Hsieh & Gralla, 1994; Hsieh *et al.*, 1994). We justify our choice of R336 for introducing a single cysteine residue by suggesting that R336 is involved in a network of interactions necessary for maintaining the transcriptionally silent state of the holoenzyme and thus might make close contact with core RNAP subunits.

Cys383. The highly conserved arginine at position 383 is located within the second helix of the proposed helix-turn-helix motif in Region III of σ^{54}, suggested to make direct contact with the conserved –12 promoter sequences (Merrick & Chambers, 1992; Coppard & Merrick, 1991). In light of recent evidence that conserved arginine residue in the Region III DNA binding domain of σ^{54} prevents polymerase isomerisation (Chaney & Buck, 1999) and the view that the –12 promoter sequence is part of a molecular switch that has to be thrown by the action of the activator to allow transcription initiation to proceed (Guo *et al.*, 1999) through a network of communication involving R383, other parts of σ^{54} and/or the core RNAP subunits, we targeted R383 to introduce a single cysteine residue.

3.6 σ^{54} Single Cysteine Mutants are Active for Transcription *in vivo* and *in vitro*

3.6.1 *In vivo* promoter activation assays

β-Galactosidase assays were used to assess the *in vivo* activity of the single cysteine mutants. *E. coli* TH1 cells, which lack functional σ^{54}, were transformed with the mutant *rpoN*⁻ containing vectors, and with pMB221, a combination reporter and activator plasmid (Cannon & Buck, 1992). The β-galactosidase gene in the reporter is under control of the σ^{54}-dependent *K. pneumoniae nifH* promoter, while the activator NifA is constitutively expressed from the *bla* promoter (Cannon & Buck, 1992). Figure 3K shows the *in vivo* β-galactosidase activity measured for each single cysteine mutant. Two different classes can be distinguished. The single mutants C198A and C346A, the double mutant

C198A/C346A (Cys(-)σ^{54}) and the triple mutant R383C (C198A/C346A) (hereafter called R383C) have near wild-type σ^{54} activity, suggesting substitution of native cysteines (C198 and C346A) with alanine and substitutions at the conserved arginine residue R383 have not caused conformational or other changes in the mutant proteins that affect *in vivo* function. The triple σ^{54} mutants E36C (C198A/C346A) and R336C (C198A/C346A) (hereafter called E36C and R336C, respectively) retain less than 20% of the wild-type level of activity. One explanation for the apparent low activity of the E36C and R336C mutants could be protein instability *in vivo*. The levels of σ^{54} expressed in *E. coli* TH1 cells under the same conditions as used in the activation assays were therefore measured by immunoblotting. R336C was found to be expressed at near wild-type levels in *E. coli* TH1 cells (data not shown). Thus, decreased protein stability *in vivo* does not adequately explain the inability of the R336C mutant to support transcription at the *K. pneumoniae nifH* promoter *in vivo*. However, in comparison to wild-type expression levels, the E36C mutant was less well expressed (~25% of wild-type level) (data not shown), and this observation supports the moderate levels of transcription by the E36C mutant *in vitro* (see later). Overall it appears that the R336C mutants is defective at one or more levels for activator dependent transcription initiation *in vivo*, and that E36C mutation causes (at the least) protein instability *in vivo*. These results are also consistent with previous observations that mutations in Region I between residues 33 and 38 reduce initial binding of σ^{54} to DNA and reduce *in vivo* expression and stability (Casaz *et al.*, 1999). These overall result, together with the location of residue R336 within a DNA crosslinking patch in Region III of σ^{54}, argue that in the mutants E36C and R336C some contacts between σ^{54} and DNA needed for transcription initiation may be lost, causing the apparent defects *in vivo*.

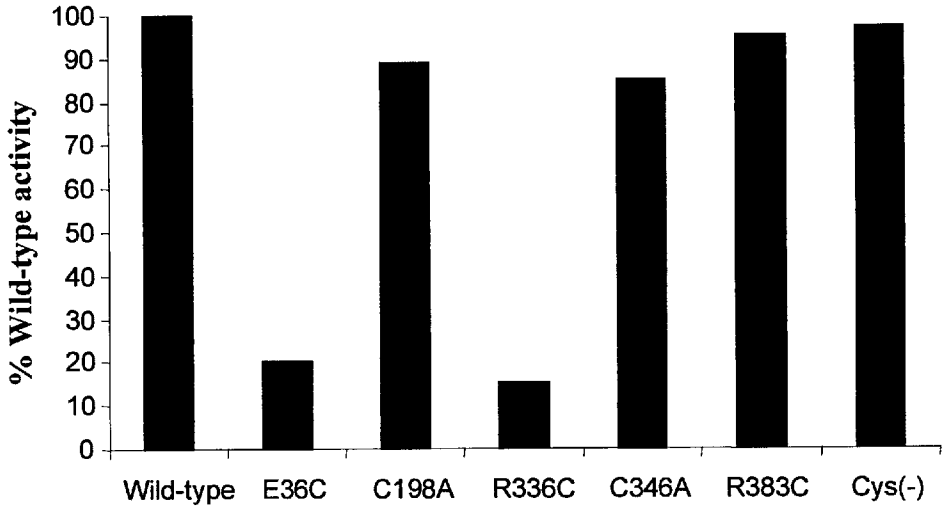


Figure 3K. *In vivo* activity of the single cysteine mutants. Shown are activity levels relative to the wild-type σ^{54} measured in a β-galactosidase assay in *E. coli* TH1 cells.

3.6.2 *In vitro* single-round transcription assays

Next, the activity of each single cysteine mutant and the Cys(-)σ^{54} protein was compared with the activity of wild-type σ^{54} in an *in vitro* activator-dependent single round transcription assay using the *E. coli glnHp2*-m12 promoter in the template (Claverie-Martin & Magasanik, 1992). Initially we tested whether the single-cysteine mutants and the Cys(-)σ^{54} were active for core RNAP binding. Native gel holoenzyme assembly assays showed that all mutant proteins formed holoenzyme at a 1:1 molar ratio of core RNAP to σ^{54} (data not shown; see Chapter 4). Transcription assays were conducted with sub-saturating amounts of holoenzyme to allow quantitative assessment of holoenzyme activities. The mutant proteins were all found to exhibit more than 50% the activity of the wild-type σ^{54} (Figure 3L).

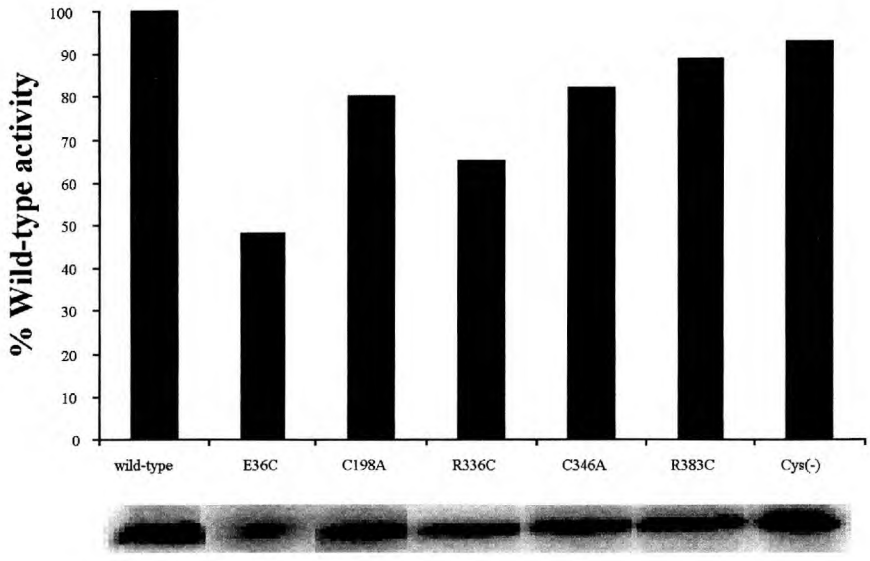


Figure 3L. Activator-dependent *in vitro* transcription activity of σ^{54} single cysteine mutants on supercoiled *E. coli glnHp2*-m12 promoter DNA. Transcript gels were scanned and % activity was determined by removing the background activity in each lane (not shown).

We therefore conclude that a replacement of both endogenous cysteines with alanine residues and the introduction of cysteine residues at alternative positions (see above) had no gross negative effect on the function of the σ^{54} protein. Also, because transcriptional activity of σ^{54} involves numerous unique stereospecific interactions, the preservation of transcriptional activity, both *in vivo* and *in vitro*, in the single cysteine mutants and the Cys(-)σ^{54} protein suggests that the conformation of the mutant proteins important for function is not significantly altered. The *in vitro* and *in vivo* activity of the single cysteine mutants are in good agreement with each other, except that the functional defects of E36C and R336C are more pronounced *in vivo*.

Region I sequences and R336 are involved in preventing unregulated RNAP isomerisation (Chaney & Buck, 1999; Casaz *et al.*, 1999; Syed *et al.*, 1998; Syed *et al.*, 1997). We therefore tested whether the single cysteine σ^{54} proteins were active for unregulated transcription (also called "activator bypass" transcription), that is, whether they allowed RNAP isomerisation in the absence of activator proteins. The protocol for these assays begins by incubating the holoenzyme with GTP, CTP and UTP. The first six

bases transcribed from the *E. coli glnHp2*-m12 template are UGUCAC (+1 to +6), so transcripts can be initiated under these conditions if an activator-independent open complex forms. ATP and [α-^{32}P]UTP are then added, together with heparin, to destroy residual closed complexes and unstable open complexes, and allow elongation of any initiated transcripts. E36C and R336C holoenzymes showed activator independent transcription activity (Figure 3M).

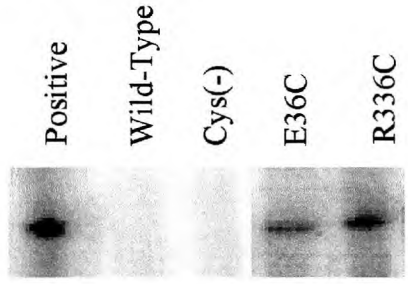


Figure 3M. Activator-independent transcription activity of σ^{54} single cysteine mutants and Cys(-)σ^{54} relative to the wild-type holoenzyme. Δ(21-27)σ^{54} (Gallegos & Buck, 2000) was used as the positive control (Positive) in the assays.

The lack of unregulated transcription activity from the Cys(-)σ^{54} protein suggests that the bypass phenotype of E36C and R336C is attributable to the substitutions at positions E36 and R336, and not to the replacement of the two endogenous cysteines at positions C198 and C346. Consistent with previous observations by us (Chaney & Buck, 1999; Casaz *et al.*, 1999) and others (Syed & Gralla, 1998; Hsieh & Gralla, 1994; Hsieh *et al.*, 1994) E36 and R336 in σ^{54} have a function in inhibiting RNAP isomerisation in the absence of activation.

3.7 Overview

The bifunctional chelating agent FeBABE covalently conjugates to the free sulphydryl groups of cysteine residues. After the addition of ascorbate and peroxide, the Fe-BABE generates nucleophiles which cleave polypeptide chains in proximity (~12Å) to the chelate, independent of the amino acid sequence involved (Rana & Meares, 1991;

Ishihama, in press). We have successfully applied this footprinting methodology to study σ^{54}-core RNAP interaction. Using two core RNAP binding fragments of σ^{54} harbouring a native cysteine at position 198 and the wild-type σ^{54} containing an additional native cysteine at position 346, the surfaces on core RNAP subunits β and β' proximal to Cys198 and Cys346 were investigated. The preliminary data indicate the use of the same or overlapping sequences in the β and β' subunits for interaction with the two σs (σ^{54} and σ^{70}). One site of σ^{54} proximity, towards the carboxyl end of to the β' subunit, is not evident with σ^{70}, and may represent a specialised interaction of σ^{54} with core RNAP.

In order to further evaluate the preliminary data and characterise the σ^{54}-core RNAP interface we constructed plasmids for production of five mutant σ^{54} proteins harbouring single cysteines at strategic locations. *In vitro* and in most cases *in vivo* activity assays revealed that the mutant proteins retain more than 50% the wild-type activity. Moreover, the *in vitro* activator-dependent promoter activation assays, revelaled that two of the single-cysteine mutant proteins (E36C and R336C) when complexed with core RNAP, yielded as predicted, holoenzymes that were deregulated for transcription initiation, that is the mutant holoenzymes were able to form stable open complexes in the absence of activator and nucleotide hydrolysis.

CHAPTER FOUR

Conservation of Sigma-Core RNA Polymerase Proximity Relationships between the Enhancer Independent and Enhancer Dependent Sigma Classes

4.1 Introduction

Two distinct classes of RNAP σ-factors exist in bacteria and are largely unrelated in primary amino acid sequence and their modes of transcription activation. Using tethered iron chelate (FeBABE) derivatives of two core RNAP binding fragments (70-324aa and 1-306aa) of the enhancer dependent σ^{54} we mapped sites of proximity to the β and β' subunits of the core RNAP. Remarkably most sites localised to those previously identified as close to the enhancer independent σ^{70}. Primarily, our initial data indicated a common use of sets of sequences in core RNAP for interacting with the two σ classes. In the present work, we aim to use five FeBABE modified single cysteine mutants (E36C, C198A, C346A, R336C, R383C) to map β and β' surfaces proximal to the FeBABE attachment site on σ^{54}. Some sites chosen in σ^{54} for modification with FeBABE (Cys36 and Cys336) were positions which when mutated deregulate the σ^{54}-holoenzyme and allow activator-independent initiation and holoenzyme isomerisation. We infer that these sites in σ^{54} may be involved in interactions with core RNAP that contribute to maintenance of alternate states of the holoenzyme needed for either the stable closed promoter complex conformation or the isomerised holoenzyme conformation associated with the open promoter complex.

4.2 FeBABE Modification of σ^{54} Single Cysteine Mutants

Previously we showed that the single cysteine mutants were active for transcription *in vivo* and *in vitro* directly implying that introduction of cysteine residues at functionally strategic positions has not resulted in conformational changes that inhibit σ^{54} function. Next, each single cysteine σ^{54} mutant and the Cys(-) protein was derivatised with chemical cleavage reagent FeBABE under native conditions and the conjugation yield was determined fluorimetrically by the CPM test (N-[4-[7-(diethylamino)-4-methylcoumarin-3-yl]phenyl]maleimide) (Greiner *et al.*, 1997) from the fluorescence difference between the conjugated and unconjugated proteins. The conjugation yield for the single cysteine mutants E36C, R336C, C198A and R383C was estimated to be >47% (Table 4.1). The mutant σ^{54} harbouring the native single cysteine at position C198 (the C346A mutation) proved difficult to be conjugated under native conditions (conjugation yield of 20%) implying that C198 is probably slightly buried or located in a hydrophobic region in the folded state because the hydrophobic reagent CPM reacted with high yield (data not shown) while the hydrophilic molecule FeBABE did not. When conjugation with C198 was performed under denaturing conditions, the yield increased to 50%. This is consistent with our previous observation with the 70-324aa and 1-306aa fragments in which FeBABE conjugation only occurred under denaturing conditions. The Cys(-) protein failed to detectably react with CPM and FeBABE under native and denaturing conditions, directly implying that under the conditions used FeBABE conjugation has only occurred at the free sulfhydryl side groups of cysteine residues, and not at other nucleophilic sites such as histidine or lysine side chains (Greiner *et al.*, 1997).

Table 4.1 Fe-BABE Conjugation Efficiency

Residue	Mutation	Conjugation Efficiency (%)
36	E36C	53
198	C346A	46*
336	R336C	51
346	C198A	47
383	R383C	76
Cys(-)	C198A/C346A	3 / 5*

* denotes conjugation under denaturing conditions

4.3 The FeBABE Derivatised Single Cysteine Mutants of σ^{54} are Active for Holoenzyme Formation and Transcription

We investigated the properties of the derivatised single cysteine mutants of σ^{54} proteins *in vitro* to determine if core RNAP binding and transcription were affected. Native gel holoenzyme assembly assays were used to detect complex formation between core RNAP and σ^{54}. The holoenzyme migrates slower than the core enzyme. With the exception of the conjugated R336C mutant, all derivatised σ^{54} mutant proteins bound core RNAP with the same apparent affinity as the non-conjugated wild-type σ^{54} (Figure 4A-A). Residue 336 lies outside the minimal core binding region defined by deletion mutagenesis (Gallegos & Buck, 1999), but adjacent to a region in σ^{54} (residues 290-310) that is protected by core RNAP from hydroxyl radical mediated cleavage (Casaz & Buck, 1999) and within a sequence (residues 325-440) strongly protected by core RNAP when Region I is removed (Casaz & Buck, 1999). Because the unconjugated R336C binds core RNAP with a similar affinity as the wild-type σ^{54}, we conclude that the presence of the 490Da FeBABE molecule at position R336 leads to steric hinderance of core binding, or causes the R336C protein to undergo conformational changes that diminish binding to the core RNAP. We also note that conjugated Cys198 binding to core RNAP produces two different

complexes. One has the mobility of normal holoenzyme; the nature of the other is unknown (Figure 4A-A, *lane 10*, marked with arrow).

To determine the effect of FeBABE conjugation on the transcriptional activity of each single cysteine σ^{54} mutant, single-round transcription assays were also performed after conjugation with the FeBABE probe. Assays were conducted with sub-saturating (50nM holoenzyme : 10nM promoter template) amounts of holoenzyme to allow detection of activity differences. Holoenzymes were formed at a 1:2 molar ration of core RNAP to σ^{54}. Only the wild-type, E36C and Cys(-) proteins show negligible effects of conjugation on activity. After correction for the presence of unconjugated protein, the remaining four conjugates show on average about half the activity of their unconjugated forms, with R336C (at 15% of the wild-type σ^{54} activity) the most severely affected, and the R383C (65% of wild-type) least (Figure 4A-B). The R336C result is consistent with our previous observation that the conjugated R336C mutant has reduced affinity to core RNAP. Nevertheless the conjugated R336C σ^{54} forms an holoenzyme that is active for a significant level (10%) of activated transcription.

We used the activator bypass transcription assay on the conjugated E36C and R336C mutants as a further indicator of conformational preservation of these proteins following conjugation. We reasoned that if the E36C and R336C mutants retained bypass activity following FeBABE conjugation then the presence of the FeBABE molecule at these positions did not drastically affect the conformation of the conjugated proteins. Having taken the presence of unconjugated σ^{54} into consideration, both mutants, E36C and R336C retained considerable activator bypass activity following conjugation with FeBABE (Figure 4A-C) In conclusion, the core RNAP binding assays, the activator dependent and the bypass transcription assays altogether suggest that the FeBABE conjugation did not grossly affect overall conformation and structural integrity of the mutant σ^{54} proteins.

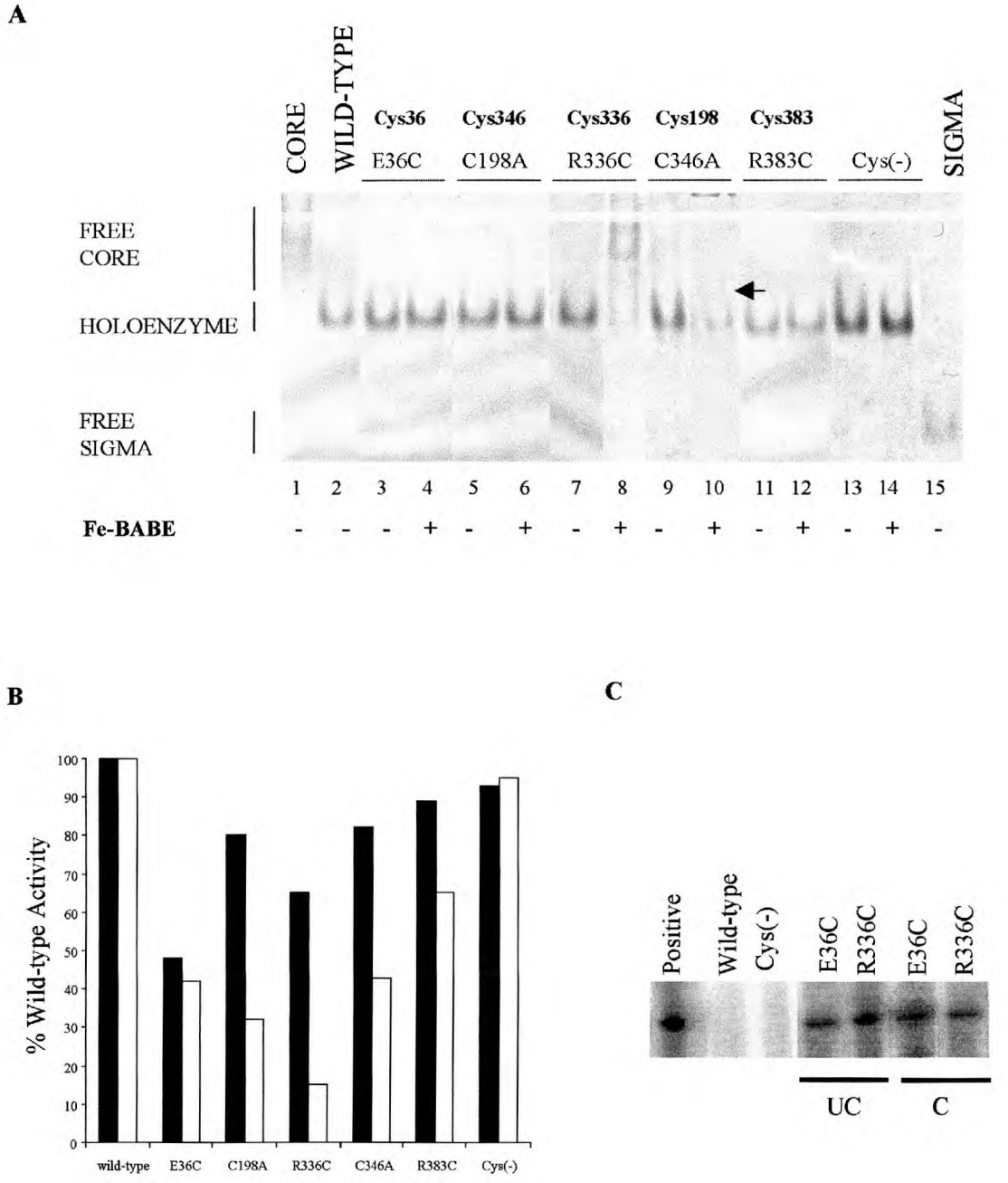


Figure 4A. (**A**) Formation of holoenzyme by the unconjugated (-) and FeBABE conjugated (+) σ^{54} single cysteine mutants at a molar ratio of 1:4 of core RNAP to σ^{54}. Migration positions of free core RNAP, holoenzyme and free σ^{54} proteins are indicated (see text for details). (**B**) *In vitro* activator dependent transcription assay. The activities of the σ^{54} single cysteine mutants before (black bars) and after (white bars) FeBABE conjugation are compared. The values for conjugated proteins are corrected for the presence of unconjugated proteins. (**C**) Activator bypass activity of the E36C and R336C σ^{54} proteins after FeBABE conjugation (C); the activator bypass activity of unconjugated proteins (UC) is shown for comparison. Positive control as in Figure 3M.

4.4 Enhancer Dependent σ^{54} and Enhancer Independent σ^{70} Interacting Sites on β and β' Core Subunits are Conserved

Holoenzymes were formed with conjugated and unconjugated single cysteine mutants of σ^{54} and cleavage reactions were carried out essentially as previously described (Owens *et al.*, 1998). In our approach, we assigned the cleavage sites on the β and β' subunits by comparing the relative migration of the FeBABE cleavage products with homologous markers generated by cleaving β and β' subunits at either cysteine or methionine residues (using 2-nitro-5-thiocyanobenzoate and cyanogen bromide, respectively). The cleavage sites and molecular sizes for each β and β' marker fragments were assigned, where possible, by calculating the molecular weight of each expected fragment from the known locations of cysteine or methionine residues in the amino acid sequence of β and β' subunits (Figure 4B) and using a plot of log molecular weight versus relative migration distance of a known set of marker proteins (data not shown)

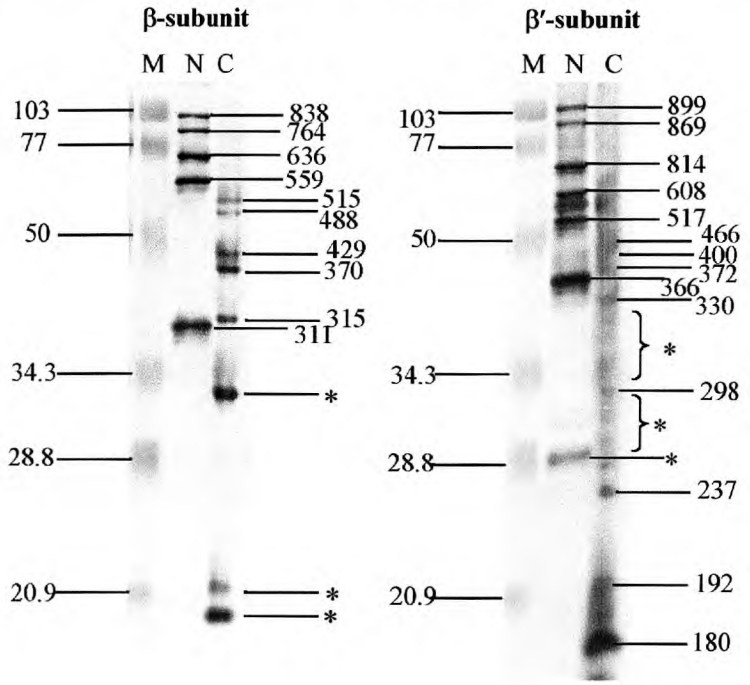


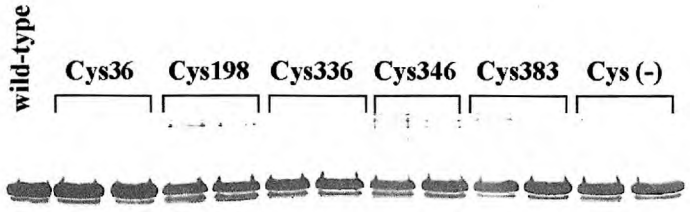
Figure 4B (from previous page). Immunoblot of SDS-PAGE fractionated markers generated by treating core RNAP with NTCB (N) and CNBr (C), respectively, using anti-β or anti-β' N-terminus-specific antibodies. Where possible, numbers indicate the positions where cleavage has occurred. * indicates cleavage products which could not be assigned a position/site of cleavage on the β and β' subunit, respectively; and have probably resulted from secondary cleavage events and/or inherent instability of the β and β' subunits. Because these lie outside the "cleavage area" obtained by FeBABE mediate cleavage using σ^{54}-derivative (see later), little attention is paid to their identities (Figure 4C-B and Figure 4C-C). Pre-stained molecular weight markers (M; BioRad, Low Range SDS-PAGE Standards) were used as reference molecular weight markers; their molecular sizes in kDa are shown on left of each pannel.

Initially we showed that self cleavage of σ^{54} did not occur (Figure 4C-A), indicating that cutting of β and β' was attributed to binding of intact σ^{54} proteins. Ascorbate and hydrogen peroxide treated and untreated reactions of each single cysteine mutant σ^{54}-holoenzymes were separated, blotted and visualised by immunostaining with affinity-purified β and β' amino terminal specific antibodies (Figure 4C-B and 4C-C, respectively). The lack of cleavage of free core RNAP and the holoenzyme formed with the Cys(-)σ^{54} (Figure 4C-B and 4C-C, *controls*) protein serves as an essential control experiment, demonstrating that under the conditions used there is scarely detectable artifactual cutting of the core RNAP subunits. Hence we believe that each strong cut seen in other tracks (Figure 4C-B and 4C-C) is attributable to a specific FeBABE form of σ^{54} created by derivatising a cysteine residues.

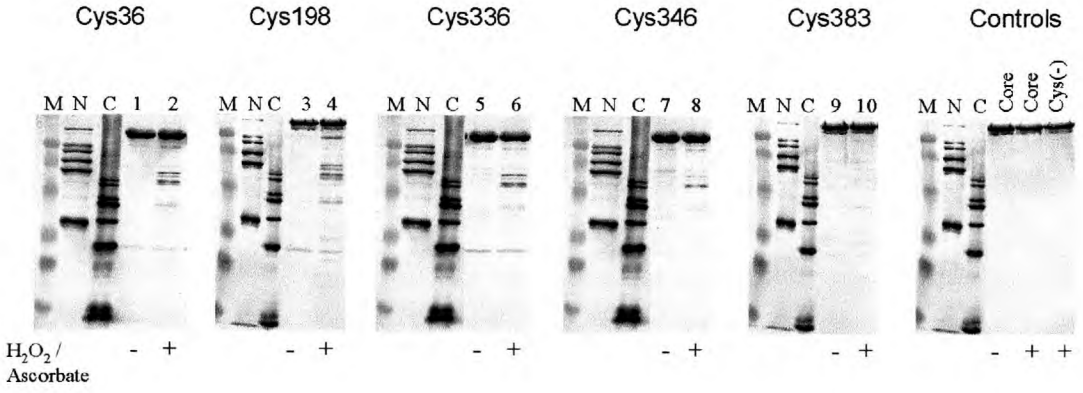
Cys198 and Cys346. Initially, we looked for the β and β' interactions made by the σ^{54} single cysteine mutants with FeBABE conjugated to the endogenous cysteines at positions Cys198 (C346A) and Cys346 (C198A), respectively. Cys198 cleaves β and β' at multiple sites. On the β subunit strong cleavage by Cys198 is seen within the Rif1 cluster (between conserved β regions C and D) and in conserved region G (Figure 4C-B, Cys198 *lane 4*). A few weak cleavages are also seen within the DR1 region. Moderate cleavage by Cys198 of the β' subunit occurs in conserved regions B and C and within region G (Figure 4C-C, Cys198 *lane 2*). In contrast, Cys346 only moderately cleaves β and β' at single sites, these map proximal to the conserved region D on the β subunit and within conserved regions C and D on the β' subunit (Cys346 *lane 6*, Figure 4C-B and 4C-C, respectively). These results are consistent with previous observations: Cys198 is located within a

sequence (residues 120-215) in Region III of σ^{54} which has a strong affinity to the core RNAP (Gallegos & Buck, 1999), while Cys346 is located within a region (residues 325-340) which is weakly protected by core RNAP during hydroxyl radical cleavage (Casaz & Buck, 1999), consistent with the differences in the β and β' cleavage patterns generated by the Cys198 and Cys346 conjugated σ^{54} proteins.

A



B



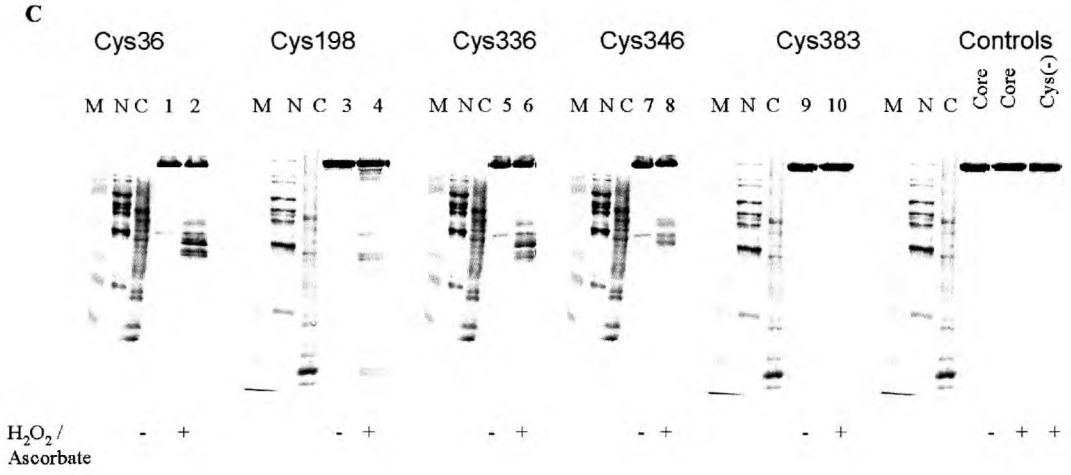


Figure 4C. (**A**) Self-cleavage immunoblot of σ^{54}-holoenzyme (formed with FeBABE conjugated His-tagged σ^{54} proteins) cleavage using anti-6His antibodies (Promega). The migration position of FeBABE unconjugated wild-type σ^{54} is shown for comparison. (**B**) and (**C**) Holoenzyme cutting by σ^{54} single cysteine FeBABE conjugates. Immunostained blots of 12.5% SDS-PAGE detecting either (**B**) β or (**C**) β' subunit fragments by using affinity-purified amino terminal antibodies. The σ^{54}-holoenzymes were either treated (+) or untreated (-) with ascorbate and hydrogen peroxide. The markers were as in Figure 4B.

Cys36 and Cys336. Next, we used the activator-independent single cysteine mutants of σ^{54} with FeBABE conjugated to Cys36 (E36C) and Cys336 (R336C) in the Cys(-) protein background to map their interaction sites on the β and β' subunits of core RNAP. The Cys36 is in Region I of σ^{54} which functions to keep the closed complexes in a transcriptionally silent state in the absence of activation (Gallegos *et al.*, 1999) and has been shown to have weak affinity to core RNAP (Gallegos & Buck, 1999). The core RNAP cleavage by Cys36 shows that Region I strongly interacts with β and β' subunits. Cleavage sites map to within the Rif1 cluster and the conserved region D on β and within conserved region C on β' (Cys36 *lane 2*, Figure 4C-B and 4C-C). Cys336 locates within a sequence in σ^{54} that lies outside the minimal DNA binding domain but nevertheless crosslinks to DNA (Chaney et al., 2000; Chaney & Buck, 1999). Remarkably, the cleavage data indicates that Cys336 interaction sites on both β and β' subunit unambiguously

overlap with those of Cys36. Given the shared activator independent transcription phenotypes of the E36C and R336C σ^{54} mutants, we rationalise this observation by suggesting that E36 and R336 belong to an interface between σ^{54} and core RNAP that acts to prevent polymerase isomerisation in the absence of activators, and propose that activator probably functions in changing the interface.

Cys383. Attempts to cleave core RNAP by Cys383 (R383C) conjugated σ^{54} revealed no discernible cleavage of either β or β' subunit (Cys383 *lane 10*, Figure 4C-B and 4C-C, respectively). In σ^{54}, Cys383 is proposed to fall within the recognition helix (helix 2) of the putative HTH and to contribute to the recognition of the -12 promoter element by the σ^{54}-holoenzyme (Merrick & Chambers, 1992; Coppard & Merrick, 1991). We can conclude that Cys383 is not located proximal (at least not within a radius of 12Å) to the core RNAP subunits β and β' and may be facing promoter DNA rather than core RNAP

A comparison to σ^{70}-core RNAP proximity. Because binding of σ^{54} to core RNAP results in the formation of a holoenzyme which has very distinct functional properties compared to the σ^{70}-holoenzyme, we compared our cleavage data with those obtained for σ^{70}-holoenzyme cleavage by FeBABE modified σ^{70} (Owens *et al.*, 1998) (Figure 4D).

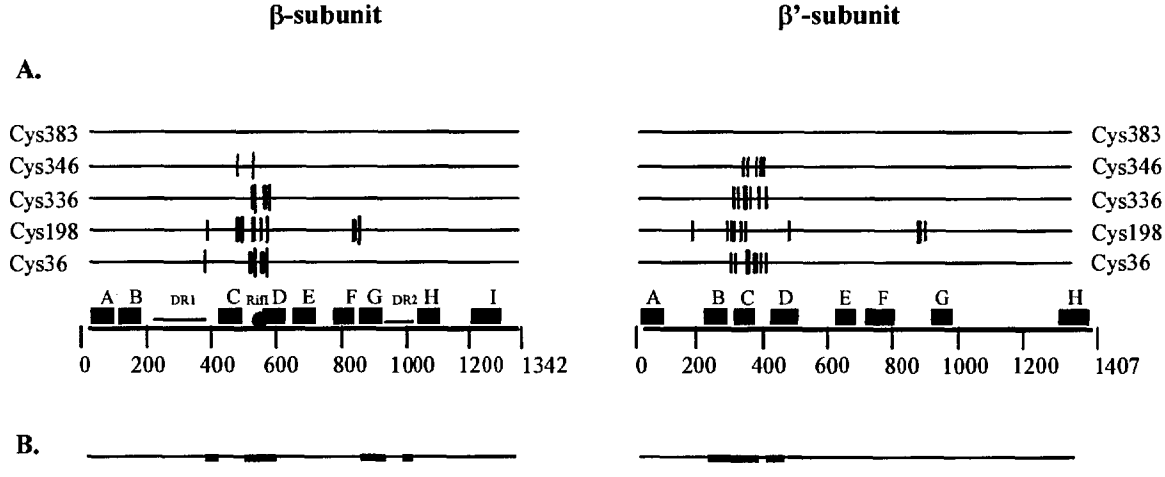


Figure 4D. (**A**) Summary of the cleavage data in Figure 4C-B & C-C showing the proximities of σ^{54} residues to surfaces on the β and β' subunits. The thickness of the lines indicates the cleavage

efficiency observed on immunoblots. (**B**) Surfaces (shaded boxes) proximal to σ^{70} as determined by FeBABE cleavage studies (Owens *et al.*, 1998). Conserved regions in β and β' are indicated along the horizontal axes.

Notably, cleavage by Cys198 is within β' subunit conserved region G whereas an interaction of σ^{70} with this region has not been observed. Several lines of evidence show that the amino terminal portion of the β' subunit is involved in the specific binding of the σ^{70} subunit (Katayama *et al.*, 2000; Arthur *et al.*, 1998; Owens *et al.*, 1998; Luo *et al.*, 1996). The interaction made by Cys198 within conserved β' region G may well be σ^{54} specific. Interestingly, a part of the sequence in β' involved in chelating Mg^{+2} to the active (catalytic) centre of RNAP, is in conserved region G (Zaychikov *et al.* 1996). We note that Cys198 is also proximal to the β' rudder, a region distant from the active centre of the RNAP (Zhang *et al.*, 1999; Figure 4D-A), suggesting significant conformational changes occurring in β' following σ binding. Remarkably, other surfaces of the core RNAP proximal to the two classes of σ factors appear to largely overlap (residues 480-600 and around residue 900±10 in the β subunit and between residues 200-350±20 in the β' subunit). Some are proximal to the parts of the RNAP catalytic centre, contributed by β conserved regions D and H and β' conserved regions D and H (Mustaev *et al.*, 1997). Locating the proximity sites of σ^{54} and σ^{70} within the recently resolved crystal structure of *T. aquaticus* core RNAP (Zhang *et al.*, 1999) serves to further illustrate the degree of overlap between the core surfaces that are proximal to the two classes of σ factors (Figure 4E).

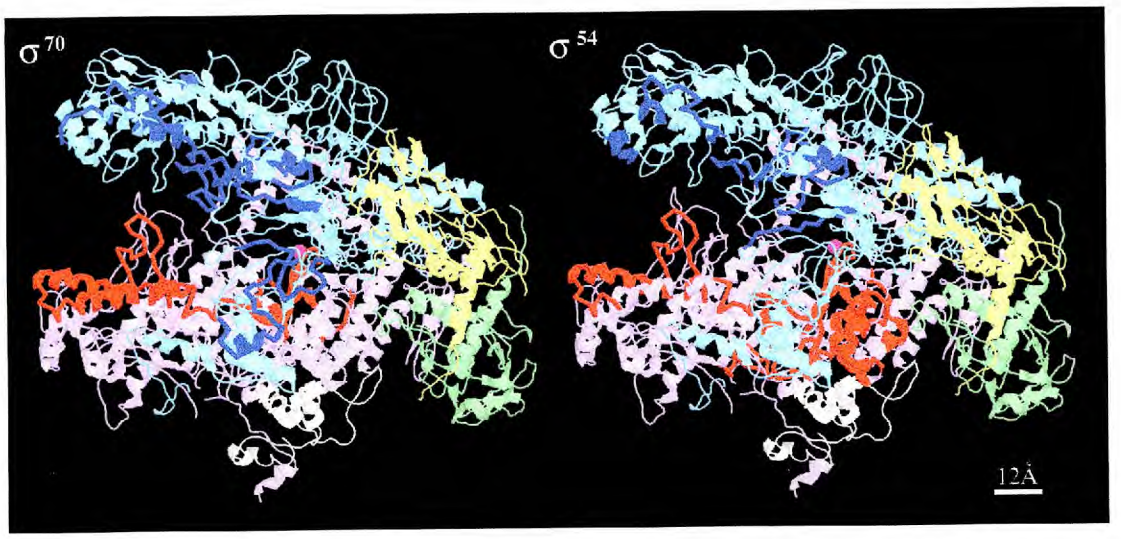


Figure 4E. RIBBONS diagrams of the three-dimensional structure of core RNAP from *T. aquaticus* (Zhang *et al.*, 1999). The β subunit is shown in pink and the β' subunit in blue. The two α subunits are shown in green and yellow. (**A**) σ^{70}- and (**B**) σ^{54}-interacting surfaces on the β and β' subunits are as indicated in dark red and blue, respectively.

4.5 Discussion

The tethered hydroxyl radical FeBABE footprinting methodology using FeBABE is a powerful tool to study protein-protein interactions that occur during transcription and has been successfully applied previously (Miyake *et al.*, 1998; Owens *et al.*, 1998). It is particularly well suited to the study of proximity relationships of a single site on one subunit to one or more other sites. Results can add considerably to our understanding of the overall architecture of large macromolecular complexes. We have used the FeBABE cleavage methodology to look at proximity relationships between functional domains of σ^{54} and the core RNAP subunits β and β' and to widen our understanding of this specialised form of bacterial enhancer dependent transcription.

Residue 36 in Region I. Our cleavage data for the Cys36 are consistent with the demonstrated interaction of Region I sequences with core RNAP (Gallegos & Buck, 1999; Casaz & Buck, 1999; 1997). Partial proteolysis studies on the σ^{54}-holoenzyme have shown

that an amino terminal region of σ^{54} spanning residues 36-100 is protected from proteolytic cleavage (Casaz & Buck, 1997) and free hydroxyl radical footprinting experiments have shown that core RNAP protects a region of σ^{54} comprising residues 36-140 from hydroxyl radical cleavage (Casaz & Buck, 1999). Even though Region I of σ^{54} is dispensable for the binding of both the core RNAP and DNA, it is essential in mediating response to activator proteins (Casaz *et al.*, 1999; Cannon *et al.*, 1995; Hsieh & Gralla, 1994; Hsieh *et al.*, 1994; Wong *et al.*, 1994; Sasse-Dwight & Gralla, 1990). The proximity of Region I E36 to core RNAP provides further physical evidence consistent with the view that Region I interacts with core RNAP to exert some of its effects, including the inhibition of polymerase isomerisation and the stabilisation of holoenzyme on melted DNA (Cannon *et al.*, 1999; Gallegos & Buck, 1999; Gallegos & Buck, 2000). It seems that each of these functions are achieved in part through interacting with core RNAP, but we cannot say precisely if proximity relationships revealed by the FeBABE data correlate to one or both roles for Region I. The interface of σ factors with core RNAP is now recognised as being functionally specialised, contributing to more than just anchoring core and σ together, and is likely to contribute to communicating information about promoter conformation to the core RNAP (Sharp *et al.*, 1999; Gross *et al.*, 1998).

Residue 336 in Region III. We previously concluded that an interaction between the amino and carboxyl sequences of σ^{54} in the holoenzyme is closely associated with inhibiting polymerase isomerisation (Chaney & Buck, 1999), consistent with the view that the σ^{54} Region I sequences may be in communication with other sequences in σ^{54} (Cannon *et al.*, 1995; Casaz & Buck, 1999) probably including the major carboxyl terminal DNA binding domain (Guo *et al.*, 1999; Guo & Gralla, 1997; Taylor *et al.*, 1996; Wong *et al.*, 1994). The recently discovered role of residue 336 in Region III of σ^{54} in maintaining closed complexes in a transcriptionally silent state in the absence of activation (Chaney & Buck, 1999) and the unambiguous similarity of Cys336 cleavage data to that of Cys36, together argue that residues around 336 contributes to the same interface with core RNAP. Our FeBABE cleavage data provide indirect structural evidence for an interaction between Regions I and III occurring via the core RNAP. However, the similarity in the cleavage patterns seen with Cys36 and Cys336 must also be viewed in the context that both mutants

(E36C and R336C) form the deregulated holoenzyme conformation, in which σ^{54} Region I and III may be proximal to each other hence providing similar cleavage patterns. Thus, we do not discount the possibility that Regions I and III may normally only interact with each other following polymerase isomerisation after activation, or at some earlier intermediate step en route to the open complex.

Residue 198 in Region III. Residue 198 seems to be able to establish two proximities to the β subunit (centred between regions B and D and between F and G) and β' subunit (between regions C and D and centred around region F). Interestingly, according to the *T. aquaticus* core RNAP structure the β proximities (as opposed to β' proximities) are well separated in space (Zhang *et al.*, 1999). Possibly binding of σ^{54} results in two different holoenzyme conformers, or movement in core RNAP upon the binding of σ^{54} brings the two sites closer together. Inspection of the native gel showing core RNAP binding of the FeBABE modified Cys198 σ^{54} supports the former suggestion. We also note that core RNAP binding of a small fragment of σ^{54} (120-215aa) resulted in two discrete complexes (Gallegos & Buck, 1999), consistent with the two binding modes suggested by the Cys198 FeBABE cleavage data. The 120-215aa fragment seems to contain the major high affinity determinant for core RNAP binding. A short σ^{70} similarity sequence of σ^{54} important for core RNAP binding in both is found between residues 175 and 189 of σ^{54} (Tintut & Gralla, 1995). Kinetic studies have shown that initial binding of σ^{54} to core RNAP is followed by a slower rearrangement of the holoenzyme (Scott *et al.*, unpublished data). The relationship of this conformational change to interactions made by sequences around Cys198 remains to be determined. Mutational analysis around position Cys198 has shown that the integrity of this σ^{54} sequence is important for holoenzyme stability on DNA that is conditional upon particular premelted DNA sequences (Pitt *et al.*, 2000). Communication of promoter structure to the core RNAP through a patch in σ^{54} within the 120-215aa fragment involving Cys198 seems likely. Intermolecular interactions within Region III that enhance DNA binding of σ^{54} are evident (Cannon *et al.*, 1997).

Residue 346 in Region III. C346 shows a similar proximity relationship to core RNAP as does residue 336, however substitutions in C346 do not result in deregulation of

RNAP isomerisation. The cutting by FeBABE at 346 was weaker than that with 336, indicating a greater distance to core RNAP and/or geometric constraint on positioning the FeBABE on core RNAP.

Residue 383 in Region III. The failure of the FeBABE derivative at position 383 to cleave core RNAP is consistent with a role in DNA binding, as proposed by Merrick & Chambers (1993), rather than in interacting with core RNAP. DNA interaction studies with FeBABE at position 383 might confirm such a role.

Overview. In conclusion, our Fe-BABE data provides a framework for interpretation of some of the genetic and biochemical data on σ^{54}. The most remarkable observation is the unambiguous similarity in the core RNAP surfaces proximal to the activator independent and activator dependent σ classes. This observation is further supported by small angle x-ray scattering data on σ^{54} which suggests that although σ^{54} and σ^{70} are unrelated by primary amino acid sequence, they both share a significant overall structural similarity (Svergun *et al.*, 2000 and Chapter 8), and protein footprinting studies (Traviglia *et al.*, 1999). It seems that the two different σs occupy similar positions in the core RNAP and use some overlapping points of interaction to contribute to the shared functionalities of the holoenzymes. With hindsight this is quite logical, although it could not have been reliably predicted *a priori*. For transcription by σ^{54}, the progression from the closed to the open promoter complex must occur along one or more pathways in which changing interactions between σ and DNA and σ and core RNAP occur to establish the alternate functional states of the holoenzyme. It seems doubtless that some of the proximity relationships we have detected will change upon activation of the σ^{54}-holoenzyme. Those associated with σ^{54} residues 36 and 336 seem most likely to change, given that these are associated with σ^{54} sequences needed to maintain the stable closed promoter complex and these when mutated lead to activator independent isomerisation and binding to premelted DNA. It seems reasonable to assume that one or more of the σ-core and σ-DNA interactions associated with maintaining the stable closed complex conformation must be changed by activator to initiate the holoenzyme isomerisation. Working out how activator and its nucleotide hydrolysis trigger these changes remains to be determined.

The proximity relationships we have determined with FeBABE modified σ^{54} proteins fall into two different classes with respect to the phenotypes of the σ^{54} single-cysteine mutants. The σ^{54} FeBABE derivatives at Cys36 and Cys336 show deregulated transcription whereas the FeBABE derivatives at Cys198 and Cys346 holoenzymes do not. Therefore it is formally possible that the proximity relationships deduced apply to two different conformers and therefore functional states of the holoenzyme. In contrast to the Cys198 and Cys346, Cys36 and Cys336 holoenzymes may have a conformation that could more resemble the conformation of the activated σ^{54}-holoenzyme in open promoter complexes. The extent to which the conformation of regulated and deregulated σ^{54} holoenzymes differ and how conformations change upon interacting with DNA remains to be determined. It is interesting that Cys198 is proximal to sequences in β' region G that are involved in maintenance of the active site of the RNAP.

CHAPTER FIVE

Regulatory Sequences in Sigma 54 Localise near the Start of DNA Melting.

5.1 Introduction

Recall that the Region III R336A mutant σ^{54} shares characteristics with some Region I mutants and the $\Delta I\sigma^{54}$ protein (Casaz *et al.*, 1999; Cannon *et al.*, 1999). Mutation of R336 (R336A) or removal of Region I ($\Delta I\sigma^{54}$) permits deregulated (activator independent) transcription initiation *in vitro* - probably due to disruption of a common or closely related, interaction required to prevent the RNAP holoenzyme from isomerisation in the absence of activation. Recent work with σ^{54} lacking Region I and carrying the Region III alteration R336A (ΔIR336A) suggested inter-related roles for Region I and Region III in open complex formation by the σ^{54}-holoenzyme (Ph.D. thesis, Matthew Chaney; Wigneshweraraj *et al.*, 2001). The DNA binding properties of ΔIR336A, $\Delta I\sigma^{54}$, R336A and the wild-type σ^{54} were compared on three types of DNA structure (homoduplex, early melted and late melted) representing possible intermediates along the DNA melting pathway from the closed complex to the open promoter complex. The main findings of that work are summarised in Table 5.1 (also see Chaney & Buck, 1999; Wigneshweraraj *et al.*, 2001), and lead to the suggestion that an interaction between the amino terminal Region I and the carboxyl terminal Region III is necessary for regulated transcription initiation by the σ^{54}-holoenzyme.

Assay	*Properties of the R336A, ΔIR336A & Δσ^{54} Proteins*
σ^{54}-DNA Interactions	R336A is defective for binding to the homoduplex and late melted DNA probes. Consistent with previous results using σ^{54} and ΔIσ^{54}, removal of Region I from the R336A mutant (ΔIR336A) resulted in 3-4 fold increase in DNA binding. In contrast to the σ^{54} and the ΔIσ^{54} the R336A mutant does not bind the early melted DNA probe. However, the ΔIR336A shows an affinity for the early melted DNA probe approaching that of ΔIσ^{54}, implying that removal of Region I allows recovery of early melted DNA binding, as also seen with the homoduplex and late melted DNA probes.

σ^{54}-holoenzyme DNA interactions	Like the $\Delta I\sigma^{54}$ and R336A holoenzymes, the ΔIR336A holoenzyme forms activator-independent, heparin stable complexes on the late melted DNA probe. In contrast, the wild-type σ^{54} requires activation for stable complex formation. In marked contrast to the $\Delta I\sigma^{54}$ and R336A holoenzymes, presenting Region I *in trans* to the ΔIR336A holoenzyme prior to DNA binding increases its heparin resistance on the late melted DNA probe. Addition of Region I *in trans* to the $\Delta I\sigma^{54}$ and R336A holoenzymes on the late melted DNA probe, in the absence of activation, confers heparin sensitivity. This suggests that the ΔIR336A forms a holoenzyme that is further (with respect to the $\Delta I\sigma^{54}$ and R336A holoenzymes) deregulated for interaction with late melted DNA probe. Region I supplied *in trans* cannot destabilise interactions of the ΔIR336A holoenzyme with the late melted DNA probe.
In vitro **activator-independent transcription**	RNAP-holoenzyme formed with the $\Delta I\sigma^{54}$ and R336A mutants initiate transcription independently of activation. Consistent with this the ΔIR336A holoenzyme was active for activator-independent transcription *in vitro*. Addition of Region I *in trans* to the reaction inhibits the activator-independence of the $\Delta I\sigma^{54}$ and R336A holoenzymes. Consistent with previous observation that $\Delta I\sigma^{54}$ and R336A holoenzyme become increasingly heparin sensitive on the late melted DNA probe when presented with Region I *in trans*, addition of Region I *in trans* to the ΔIR336A holoenzyme stimulated activator independent transcription. Since Region I *in trans* did not have any negative effect on the ΔIR336A holoenzyme complexed to the late melted DNA probe (see above), it appears that the sites in the ΔIR336A holoenzyme with which Region I should interact to inhibit activator independent transcription (and reduce heparin stability on the late melted DNA probe) are inaccessible or inactive in the ΔIR336A holoenzyme.

Table 5.1. Summary of DNA binding and transcription properties of the R336A, ΔIR336A, $\Delta\sigma^{54}$ (Chaney & Buck, 1999; Wigneshweraraj *et al.*, 2001).

The properties of the double mutant lacking R336 and Region I reinforce the idea of a positive role for Region I in open complex formation and show that Region I masks DNA binding activity. When supplied *in trans* Region I clearly stimulates bypass transcription by the ΔIR336A holoenzyme, an activity not evident in assays using $\Delta I\sigma^{54}$ or R336A holoenzymes. It seems that on transiently melted DNA the absence of R336 makes accessible a complex that Region I can stabilise so as to increase initiation. The presence of R336 may normally disfavour such positive effects of Region I under non-activating conditions because this arginine residue binds to and stabilises double stranded DNA. Furthermore, the solvent accessibility of Region III changes in the holoenzyme when Region I is removed (Casaz & Buck, 1999, 1997) and FeBABE tethered footprinting shows that Region I and Region III are proximal to similar surfaces on the core RNAP subunits β and β' (Wigneshweraraj *et al.*, 2000 & Chapter 4). Together with the binding

properties of the ΔIR336A (Wigneshweraraj *et al.*, 2001) these observations suggest that Region I contributes to the DNA binding properties of σ^{54}, possibly through an interaction with Region III (Chaney *et al.*, 2000; Casaz et al., 1999; Guo *et al.*, 1999).

To begin to unravel the physical basis of the multiple σ^{54}-DNA interactions that contribute to maintaining the stable σ^{54}-holoenzyme promoter complex, and which are believed to be directed through the -12 promoter region (Chaney *et al.*, 2000; Guo *et al.*, 1999) we examined using a tethered iron chelate footprinting method, the proximity relationships of Region I and Region III to promoter DNA. The FeBABE conjugated σ^{54} proteins were made using single-cysteine σ^{54} mutants (Wigneshweraraj *et al.*, 2000). The sites exploited for FeBABE conjugation were in Region I (E36) and carboxyl terminal DNA binding domain Region III (R336) within the DNA-crosslinking patch, on residues whose mutation deregulates the σ^{54}-holoenzyme *in vitro* (Wigneshweraraj *et al.*, 2000; Chaney & Buck, 1999). *S. meliloti* promoter-related DNA duplexes representing possible intermediates along the DNA melting pathway from the closed to the open promoter complex were used for the experiments (Figure 5A). The results reported below show that Region I and Region III FeBABE derivatives each lead to DNA cutting across the –12 promoter region. The cutting pattern changes under the influence of activator, and is also influenced by core RNAP and DNA structure. Together with related core RNAP proximity data (Wigneshweraraj *et al.*, 2000), the results lead to the clear conclusion that a localised structure contributed to by the -12 region promoter DNA and at least two different protein elements in σ^{54} exists within the closed complex. It appears that this structure is intimately linked to the activation target of the enhancer binding proteins, or may be included within it, and can help explain the overlapping properties of mutants in Region I and Region III.

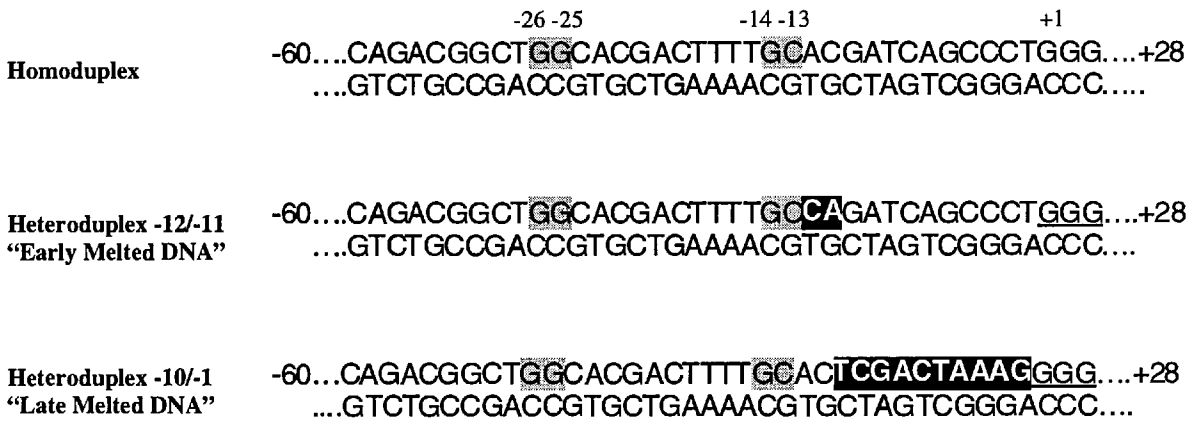


Figure 5A. *S. meliloti nifH* homoduplex, early melted and late melted DNA probes used for FeBABE footprinting (Cannon *et al.*, 1999; 2000). The main consensus bases of the promoter are shown in grey, and the bases altered to produce heteroduplex on black, on the upper strand.

5.2 FeBABE-tethered Proteins as Chemical Nucleases

The application of FeBABE modified σ-factors as chemical nucleases to map promoter DNA sequences proximal to the FeBABE attachment site has provided significant and valuable information on σ-promoter DNA interactions that occur during transcription initiation (Colland *et al.*, 1999; Bown *et al.*, 1999; Owens *et al.*, 1998; Murakami *et al.*, 1997a; Murakami *et al.*, 1997b). To date, libraries of single cysteine mutants have been constructed for σ^{70} and σ^{38}. Mapping of open promoter complexes using FeBABE tethered σ^{70} and σ^{38} subunits has indicated that σ^{70} and σ^{38} make contact with different regions along promoter DNA (Colland *et al.*, 1999; Owens *et al.*, 1998). Furthermore, the cleavage data supported previous genetic studies on the essential roles of the two hexanucleotide promoter sequences located at positions -35 and -10, respectively, and revealed hitherto unidentified promoter DNA contacts involving sequences other than those at -35 and -10. The comparison of promoter DNA cleavage data also suggested different modes of promoter DNA interaction by σ^{70} and σ^{38}. However, relatively little is known about the organisation of σ^{54} domains on promoter DNA that contribute to regulated transcription initiation.

When compared with footprinting by conventional methods (DNase I, hydroxyl radical, S1 nuclease etc.), which mainly reveal only the perimeters outlined by bound σ^{54} or σ^{54}-holoenzyme, the strength of cysteine tethered FeBABE footprinting is that it details the promoter sequences proximal to FeBABE conjugation sites within the σ^{54}-DNA or the σ^{54}-holoenzyme DNA closed and open complexes. Protein tethered FeBABE cleavage of DNA appears to occur through the generation of hydroxyl radicals ($OH^{\cdot}$) from the FeBABE, which then cleave DNA by removing a hydrogen atom from a deoxyribose sugar on the DNA backbone (Figure 5B). The mechanism of the DNA strand scission by FeBABE is not understood in detail, but a few key features have emerged. From the product analyses; namely the release of one base per cleavage event and the nature of the DNA polymer termini [5'-phosphoryl, 3'-phosphoryl, and 3'-(phosphoglycolic acid)].

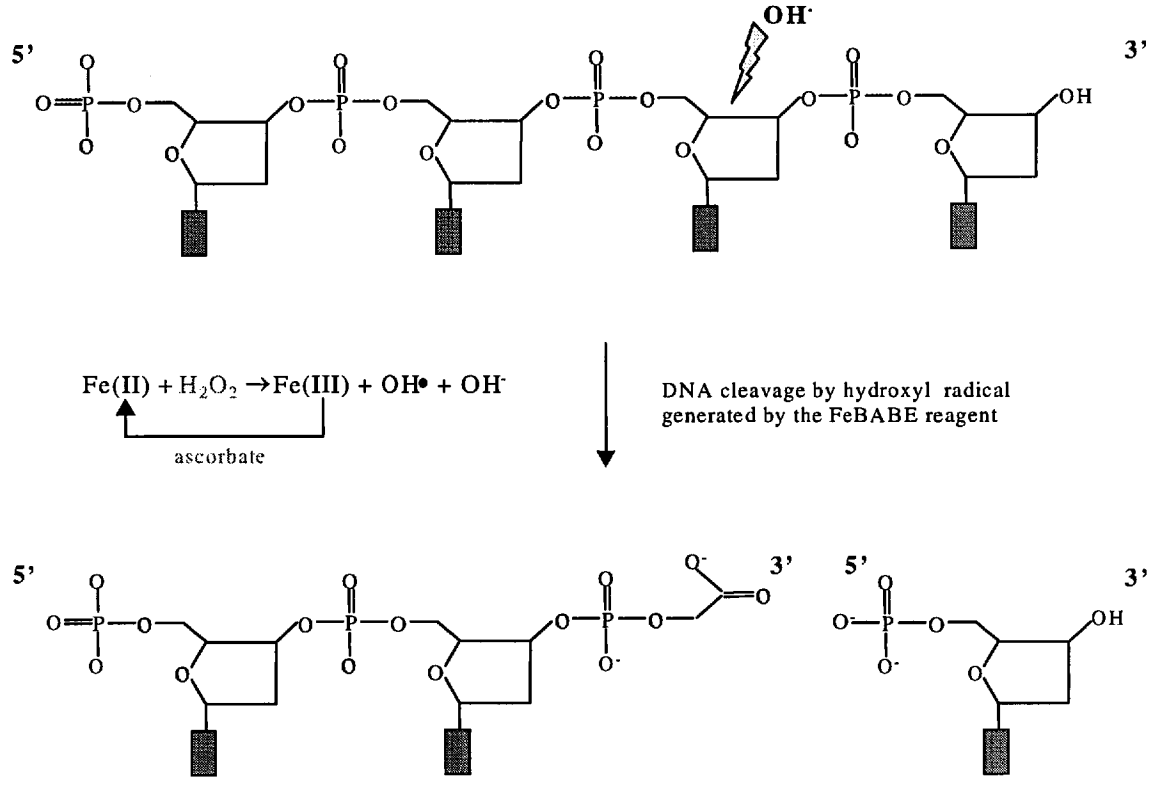


Figure 5B. DNA cleavage by hydroxyl radicals produced by the Fenton-Udenfriend reaction originating from the FeBABE reagents (see text for details)

FeBABE mediated oxidative degradation of DNA happens *via* a combination of Fenton & Udenfriend chemistry. As part of the Fenton reaction ferrous iron Fe(II) on the FeBABE reduces hydrogen peroxide to give hydroxyl radicals (**I**):

$$Fe(II) + H_2O_2 \rightarrow Fe(III) + OH\bullet + OH^- \qquad \textbf{(I)}$$

Udenfriend *et al.* showed that ascorbate acts to convert ferric iron Fe(III) (**I**) back to ferrous iron Fe(II) to continue the process (**II**):

$$Fe(II) + H_2O_2 \rightarrow Fe(III) + OH\bullet + OH^-$$

(ascorbate)

(II)

Since hydroxyl radical is exceedingly short lived and reactive, and attack sites (12Å from the FeBABE molecule) on the backbone of the DNA molecule, there is almost no sequence dependence in the cleavage reaction.

5.3 Region I and R336 proximal DNA sequences on homoduplex, early and late melted heteroduplex DNA structures

In this set of assays, we aimed to identify promoter DNA sequences proximal to Region I and R336 on homoduplex, early and late melted DNA structures bound by σ^{54} and its holoenzyme (Figure 5A). Because σ^{54}, in contrast to σ^{70} and σ^{70}-related sigma factors, can bind to promoter DNA in the absence of core RNAP *in vitro*, the FeBABE footprinting methodology allows identification of DNA sites proximal to the iron chelate

probe within σ^{54}-DNA and σ^{54}-holoenzyme-DNA complexes. We chose E36 in Region I because it is implicated in movements associated with activation (Casaz & Buck, 1997) and R336 in Region III as sites for FeBABE conjugation. The essential negative control reaction in the DNA cleavage assays involved a cysteine-free σ^{54} mutant (Cys(-)), which was subjected, together with the single cysteine mutants E36C (Cys36) and R336C (Cys336), to the FeBABE reagent (Wigneshweraraj *et al.*, 2000). Cleavage reactions conducted using the Cys(-) protein in parallel with the Cys36-FeBABE and Cys336-FeBABE resulted in no discernible cutting of the DNA, demonstrating that under the conditions used there is not artifactual DNA cleavage (see Figures 5C-5F and summarised in Figure 5G). Hence, we are confident that each DNA cut is attributable by the FeBABE specifically conjugated to either Cys36 or Cys336. Amongst the DNA cleavage experiments, the least strong cutting pattern was with Cys36-FeBABE on the early melted and late melted DNA probes where the background signal was highest. Nevertheless, quantitation of the gels showed that cutting of DNA with FeBABE conjugated proteins was clearly discernible above background. The results shown (Figures 5C, 5E and 5F) and summarised (Figure 5G) were confirmed with many replicate experiments. In general, we observed that the addition of low amounts of heparin (10μg/ml) to samples prior to initiating cleavage gave more discrete patterns of cutting, possibly through removing non-specific complexes, but the instability of the Cys36-FeBABE-DNA complex to heparin precluded use of this procedure in experiments shown in Figure 5C.

5.3.1 Promoter DNA sequences proximal to E36 (Region I) and R336 (Region III)

The binding of σ^{54} to its target promoter sequences in the absence of core RNAP allows identification of specific σ^{54}-DNA interactions. The influence of core RNAP upon the σ^{54}-DNA interaction was evaluated subsequently.

σ^{54}-Homoduplex DNA. FeBABE cleavage of homoduplex promoter DNA with Cys336-FeBABE σ^{54} reveals an R336 proximal interaction centred on position -7 (± 2), between positions –9 and -5 on the non-template strand (Figure 5C-A, *lane* 6) and position -8 (±2), between positions –10 and -6 on the template strand (Figure 5C-B, *lane* 6). We

analysed Figure 5C-A by densiometry. Results showed that there was a 50% increase in signal in the cleavage region when compared to the control lanes (2, 3, 7 and 8) in Figure 5C-A. No obvious cleavage of homoduplex DNA non-template strand is seen with Cys36-FeBABE σ^{54} implying that position 36 of Region I is not within 12Å of promoter DNA (Figure 5C-A, *lane* 4). Attempts to cut the template strand with the Cys36-FeBABE σ^{54} yielded no discernible cleavage products (Figure 5C-B, *lane* 4). Since gel shift assays showed that Cys36-FeBABE σ^{54} bound the DNA, we concluded that Region I is positioned away from the DNA in the initial σ^{54}-DNA complex.

σ^{54}-early melted DNA. Next we attempted to map the Region I (E36) and Region III (R336) interaction sites on an early melted promoter DNA probe. Weak binding of the R336 derivative to the early melted DNA precluded any cutting (Figure 5C-C, *lane* 8). In contrast to the homoduplex DNA (Figure 5C-A and 5C-B, *lane* 4), we detected cutting with Cys36-FeBABE. Cys36-FeBABE cleaves the early melted DNA probe at three distinct positions between –17 and -15 (±1), adjacent to the –14, -13 GC-element on the non-template strands (Figure 5C-C, *lane* 2), and between -18 and –15 (±1) on the template strand (Figure 5C-C, *lane* 10). The region of early melted DNA could help structure Region I for binding and cutting. Because residue 36 is approximately situated in the middle of Region I sequences important for σ^{54} function (Casaz *et al.*, 1999) and which are probably α-helical some Region I sequences could be involved in direct interactions with the consensus GC promoter doublet.

Addition of activator and hydrolysable nucleotide to the early melted DNA-σ^{54} complex causes σ^{54} to isomerise and further melt the DNA (to form the supershifted σ^{54}-DNA complex) (Cannon *et al.*, 2000; Gallegos *et al.*, 2000). Interestingly, we see a reduction in non-template DNA cutting between positions –17 and -15 (±1) by Cys36-FeBABE σ^{54} under conditions (i.e. in the presence of activator (PspFΔHTH) and hydrolysable nucleotide) that permit σ^{54} isomerisation (Figure 5C-C, compare *lanes* 2 and 4). Densiotmetry showed that when activated there is a 40% reduction in cutting efficiency of the early melted template strand by Cys36-FeBABE σ^{54} (data not shown and compare Figure 5C-C, *lanes* 11 and 12). We considered the possibility that activation may have

caused substantial dissociation of the early melted DNA-Cys36-FeBABE complex, thus resulting in the disappearance of the non-template cleavage profile. As shown, activation of the early melted DNA-Cys36-FeBABE complex produces an unstable, but supershifted σ^{54}-DNA complex (Figure 5C-D, compare *lanes* 4 and 5). However, the total amount of bound DNA remains constant. The reduction of DNA cutting between positions –17 (±1) and –15 (±1) by Cys36-FeBABE may reflect conformation changes in Region I that appear to position Region I (at least residue 36) away from promoter DNA sequences in response to activation (see also Casaz & Buck, 1997). The lack of detectable reduction of DNA cutting under non-activating conditions, i.e. in the presence of activator and GTPγS or in the absence of PspFΔHTH further supports this view (Figure 5C-C).

σ^{54}-late melted DNA probe. Cys336-FeBABE cuts the late melted DNA probe extensively between positions -10 and -3 (±2) on the non-template strand (Figure 5C-E, *lane* 6) and between -11 and -6 (±1) on the template strand (Figure 5C-F, *lane* 6). These cutting patterns are different to those seen on the homoduplex DNA, suggesting that σ^{54}-DNA relationships have changed by melting out sequences between -10 and -1. On the late melted DNA probe, Cys36-FeBABE weakly cleaves the non-template strand around and across the -14, -13 GC-element between positions -17 and –13 (Figure 5C-E, *lane* 4, marked with arrows). Densiotmetry of lane 4 and the control lanes (2, 3, 7 and 8 in Figure 5C-E) confirmed that the cutting by Cys36 did occur, but with very low efficiency. However, Cys36-FeBABE does not detectably cleave template strand DNA when bound to the late melted DNA probe (Figure 5C-F, compare *lanes* 3 and 4). Even though the efficiency of strand cleavage by Cys36-FeBABE on the late melted DNA probe is not strong, it appears that Region I is facing the –14, -13 GC element in the late melted DNA-σ^{54} complex and (at least residue 36) is orientated towards the non-template strand.

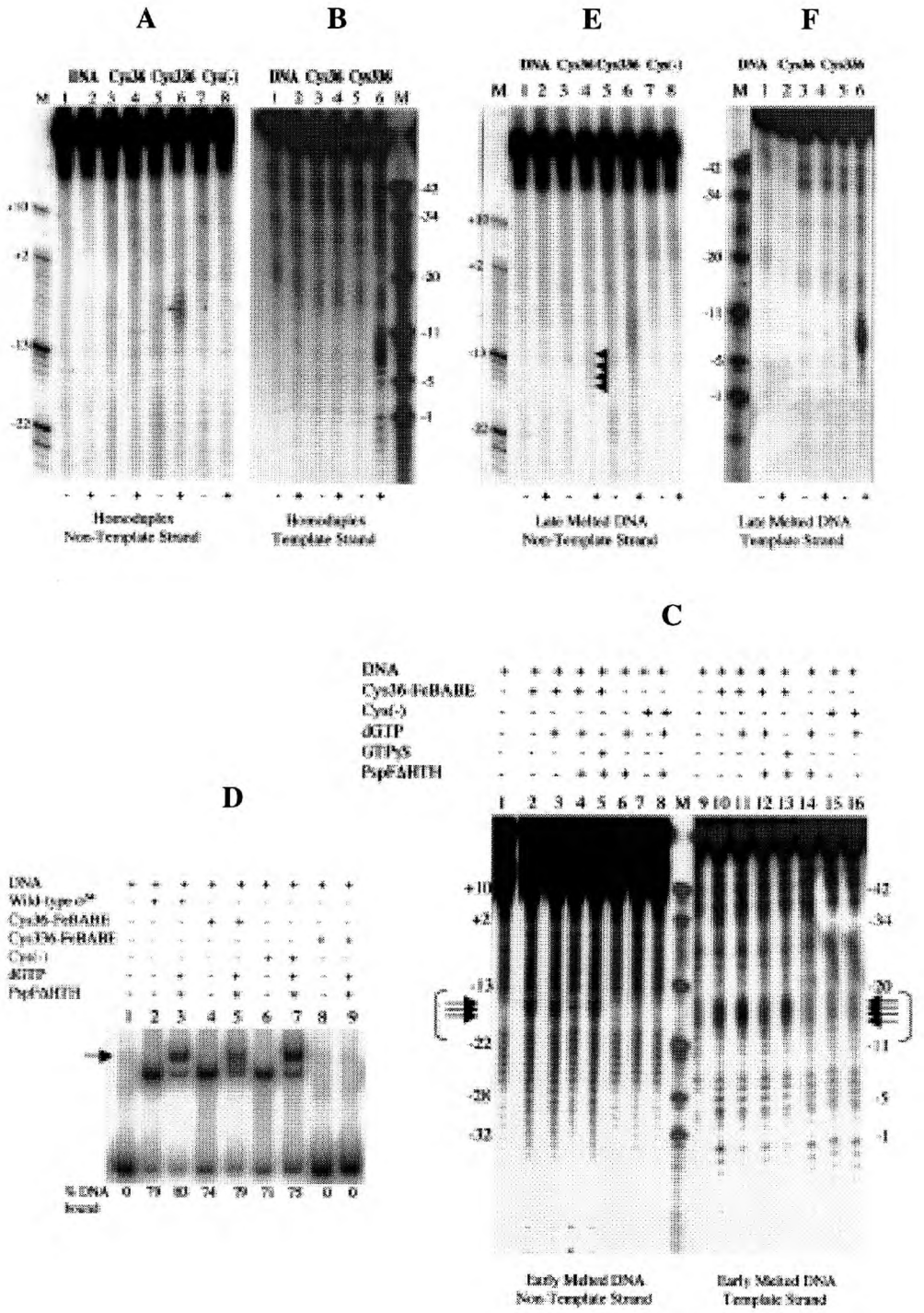

A
B
E
F
C
D
DNA
Cys36
Cys(-)
dGTP
GTPγS
PspF-ΔHTH
Homoduplex
Non-Template Strand
Homoduplex
Template Strand
Late Melted DNA
Non-Template Strand
Late Melted DNA
Template Strand
Early Melted DNA
Non-Template Strand
Early Melted DNA
Template Strand
% DNA bound

Figure 5C (previous page). *S. meliloti nifH* promoter DNA cleavage by FeBABE-modified σ^{54} proteins. Homoduplex (**A**) non-template strand and (**B**) template strand cleavage by Cys36-FeBABE & Cys336-FeBABE. (**C**) Early melted probe non-template strand (*lanes* 1-8) and template strand (*lanes* 9-16) DNA cleavage by Cys336-FeBABE. Arrows show the migration positios of the cleaved DNA band. (**D**) σ^{54}-isomerisation assay on the early melted DNA probe. (**E** and **F**) Late melted probe non-template strand and template strand DNA cleavage, respectively, by Cys36-FeBABE and Cys336-FeBABE. Reactions to which ascorbate and hydrogen peroxide were added are marked with +, control reactions to which no ascorbate and hydrogen peroxide were added are marked with -. Lanes M contain a mixture of end-labelled S. meliloti nifH promoter DNA fragments as molecular size markers. (A high quality reproduction of this figure is available in Wigneshweraraj *et al.*, 2001, which is attached at the back of this thesis)

5.3.2 Promoter DNA sequences proximal to Region I and R336 in the closed and open complex

Next, we investigated how the relationships between the three promoter DNA structures and Region I E36 and Region III R336 change in the transition from a closed to an open promoter complex. Holoenzymes were formed with 1:2 molar ratio of core RNAP to σ^{54} in order to ensure that all conjugated σ^{54} proteins were in the holoenzyme form. Holoenzymes were at 100nM concentration. Note that at 100nM, the sigma remaining free would not give significant binding and cutting. We conducted cleavage experiments on closed and activated holoenzyme complexes.

Homoduplex DNA cleavage. The Cys-336-FeBABE holoenzyme cuts homoduplex DNA strongly between positions -10 and -6 (±1) on the non-template strand (Figure 5D-A, *lane* 6), whereas the template strand is cut between positions –13 and -8 (Figure 5D-C, *lane* 6). Strong cleavage of the homoduplex DNA template strand is seen with Cys36-FeBABE holoenzyme and is centred around position –16 (±2), between –18 and –14 (Figure 5D-C, *lane* 4). In contrast, the non-template strand is weakly cleaved along the T-tract upstream of the GC-element between positions –16 and –11 (±2) (Figure 5D-A, *lane* 4). It seems that Cys336-FeBABE and Cys36-FeBABE cleavage positions of non-template DNA in the closed complex are close, consistent with previous observation that residues 36 and 336 are proximal to some of the same surfaces on the core RNAP subunits β and β' (Wigneshweraraj *et al.*, 2000). When PspFΔHTH and dGTP were added, no discernible changes were observed in the cutting patterns between homoduplex DNA-σ^{54}-holoenzyme

(compare Figures 5D-B and 5D-D) with 5D-A and 5D-C, respectively) consistent with the view that open complex formation on linear DNA is inefficient (Oguiza & Buck, 1997; Wedel & Kustu, 1995).

Early melted DNA cleavage. On the early melted DNA probe, Cys336-FeBABE holoenzyme produced strong cutting of the non-template strand between positions –18 and –10 (±1), whereas the template strand was cut moderately between positions –13 and -7 (±2). (Figure 5E-A and 5E-C, *lane* 6). Recall that binding by Cys336-FeBABE σ^{54} alone did not occur with the early melted DNA probe (Chaney & Buck, 1997 and see above). Core RNAP clearly stabilises the Cys336-FeBABE on the early melted DNA probe (Wigneshweraraj *et al.*, 2001). As in the homoduplex DNA-holoenzyme closed complex, Cys36-FeBABE cleavage of the non-template strand was along the promoter T-tract between positions –19 and –16 (±1), and between positions –19 and -15 on the template strand (Figure 5E-A, *lane* 4 and Figure 5E-C, *lane* 4). Strikingly, activation of the early melted DNA-holoenzyme complex reduces the efficiency of non-template strand cutting by Cys36-FeBABE (Figure 5E-B, *lane* 4, marked with arrows), whereas the cutting between positions –17 and -14 on the template strand (Figure 5E-C and 5E-D, *lanes* 4) remains unchanged. As deduced from the cutting patterns, the relationship between Cys336-FeBABE holoenzyme and the non-template early melted DNA probe remains unchanged under activating conditions (Figure 5E-A and 5E-B, *lanes* 6). In contrast, activation of the Cys336-FeBABE holoenzyme on the early melted DNA probe complex causes a reduction in the efficiency of cutting on the template strand (Figure 5E-C and 5E-D, *lanes* 6), comparable to the effect on Cys36-FeBABE cutting of the other strand. Consistent with other observations (Guo *et al.*, 2000; Casaz & Buck, 1997 and Figure 5C-C), these activator/NTP hydrolysis dependent changes in Cys36-FeBABE and Cys336-FeBABE DNA cleavage patterns provide evidence for activation-related conformational changes affecting Region I and Region III within the holoenzyme, leading to proximity and orientation changes in the relationships between the promoter DNA, Region I and Region III. Cleavage patterns produced by Cys336-FeBABE and Cys336-FeBABE holoenzyme did not change when the non-hydrolysable nucleotide GTPγS substituted dGTP in the activation assays (data not shown), confirming the significance of the above activation defects.

Late melted DNA cleavage. Because the late melted DNA structure is believed to represent a DNA conformation adopted by the promoter DNA following activation and isomerisation of σ^{54}-holoenzyme, we examined whether activation associated changes are still evident in holoenzyme complexes formed with the late melted DNA probe. If so, they would support the view that activator drives conformational changes within σ^{54} and that pre-melting the DNA does not bypass these (Cannon *et al.*, 1999; Wedel & Kustu 1995). Strong cleavage by Cys36-FeBABE holoenzyme occurs on the non-template strand of the late melted DNA probe between positions –16 and –11 (±2) (Figure 5F-A, *lane* 4), and moderate cleavage on the template strand centred around position –16 (between –18 and –14) (Figure 5F-C, *lane* 4). Consistent with our previous observation with early melted DNA probe (see above, Figure 5E-B), activation causes a decrease in non-template strand cleavage by Cys36-FeBABE between positions -16 and -11 (compare Figures 5F-A and 5F-B, *lanes* 4 and 2, respectively). Similarly, there is a slight reduction in the template strand cleavage by Cys36-FeBABE when activated, which was not observed with early melted DNA (Figure 5F-C and 5F-D, *lanes* 4). The relationship between Cys336-FeBABE and the template strand remains unchanged when activated as judged by the relatively constant cutting patterns between –12 and –4 (non-activated) and –14 and –5 (activated) in the Cys336-FeBABE holoenzyme late melted DNA complex (Figure 5F-C and 5F-D, *lanes* 6), whereas reduced cleavage was seen on early melted DNA. Interestingly, we note an increase in non-template strand cleavage between positions –12 and –3 by Cys336-FeBABE when activated (Figure 5C-A and 5D-B, compare *lanes* 6 and 4, respectively). It appears that conformational changes in response to activation have led to an increased interaction by Region III residue 336 with late melted non-template DNA strand. Clearly, these FeBABE cutting data support the view that changes in Region I and Region III brought about by activator are involved in the conformational changes that allow the holoenzyme to melt the promoter or to engage with melted DNA (Cannon *et al.*, 1999; Wedel & Kustu, 1995).

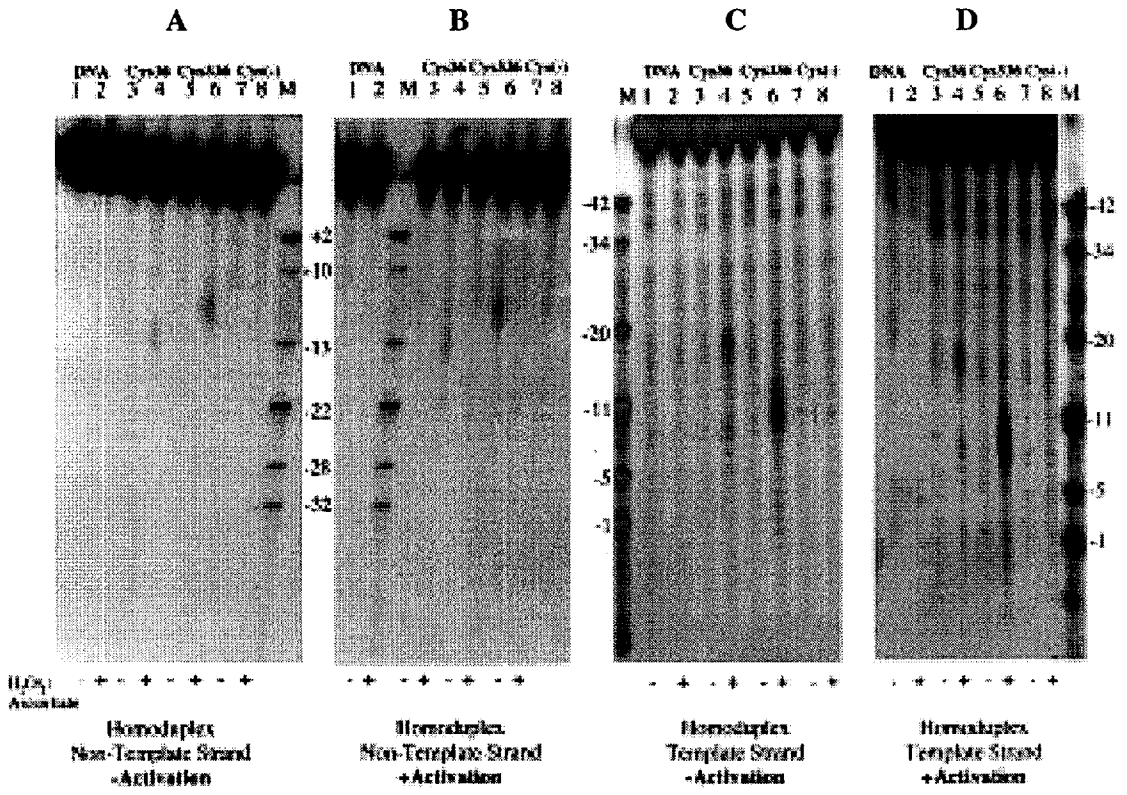


Figure 5D. *S. meliloti nifH* promoter DNA cleavage by RNAP holoenzymes containing FeBABE modified σ^{54} proteins. Homoduplex probe non-template and template strand DNA cleavage under non activating (**A** and **C**) , and activating, (**B** and **D**) conditions, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +, control reactions to which no ascrobate and hydrogen peroxide were added are marked with -. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as molecular size markers. (A high quality reproduction of this figure is available in Wigneshweraraj *et al.*, 2001, which is attached at the back of this thesis).

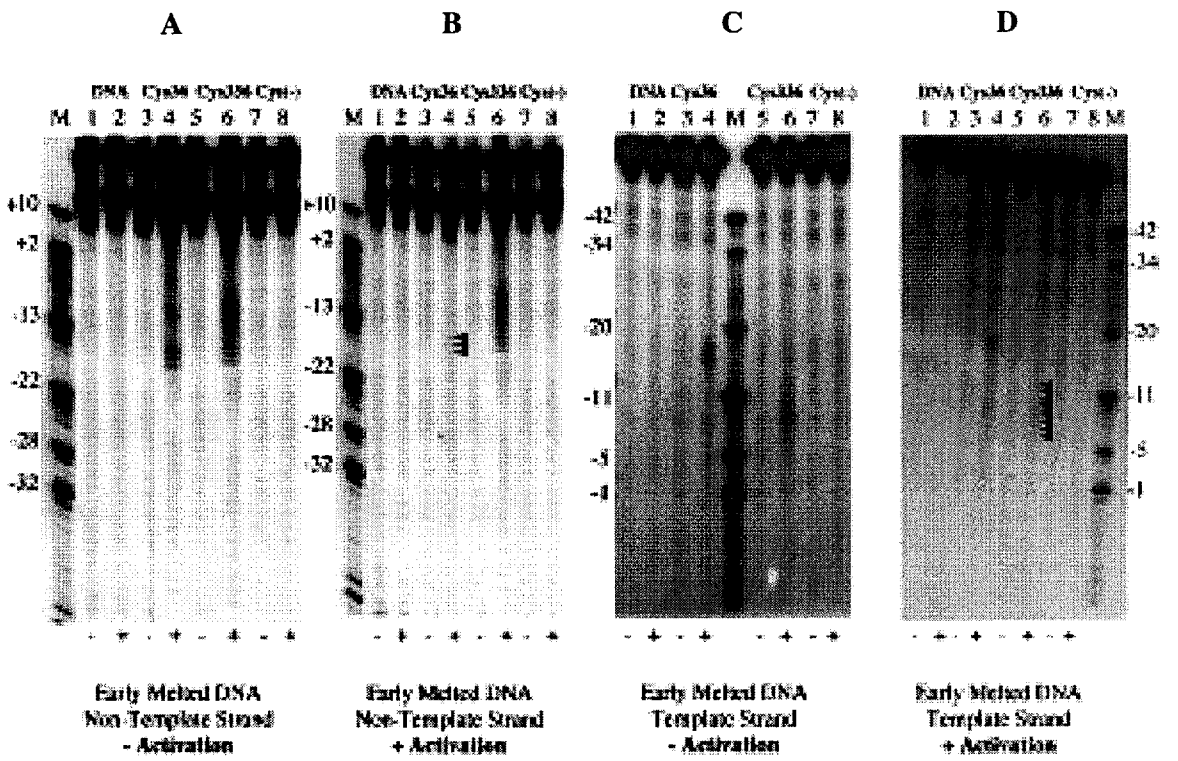


Figure 5D. *S. meliloti nifH* promoter DNA cleavage by RNAP holoenzymes containing FeBABE modified σ^{54} proteins. Early melted probe non-template and template strand DNA cleavage under non-activating (**A** and **C**), and activating (**B** and **D**) condition, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +, control reactions to which no ascorbate and hydrogen peroxide were added are marked with -. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as molecular size markers. (A high quality reproduction of this figure is available in Wigneshweraraj *et al.*, 2001, which is attached at the back of this thesis).

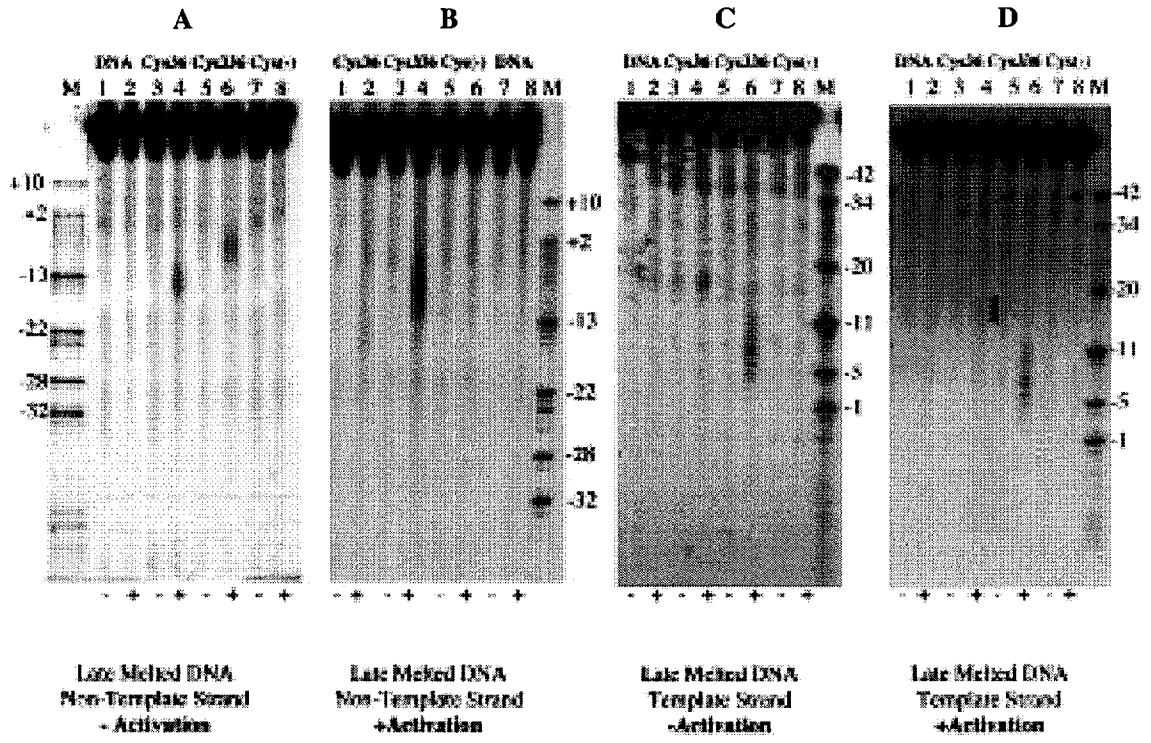
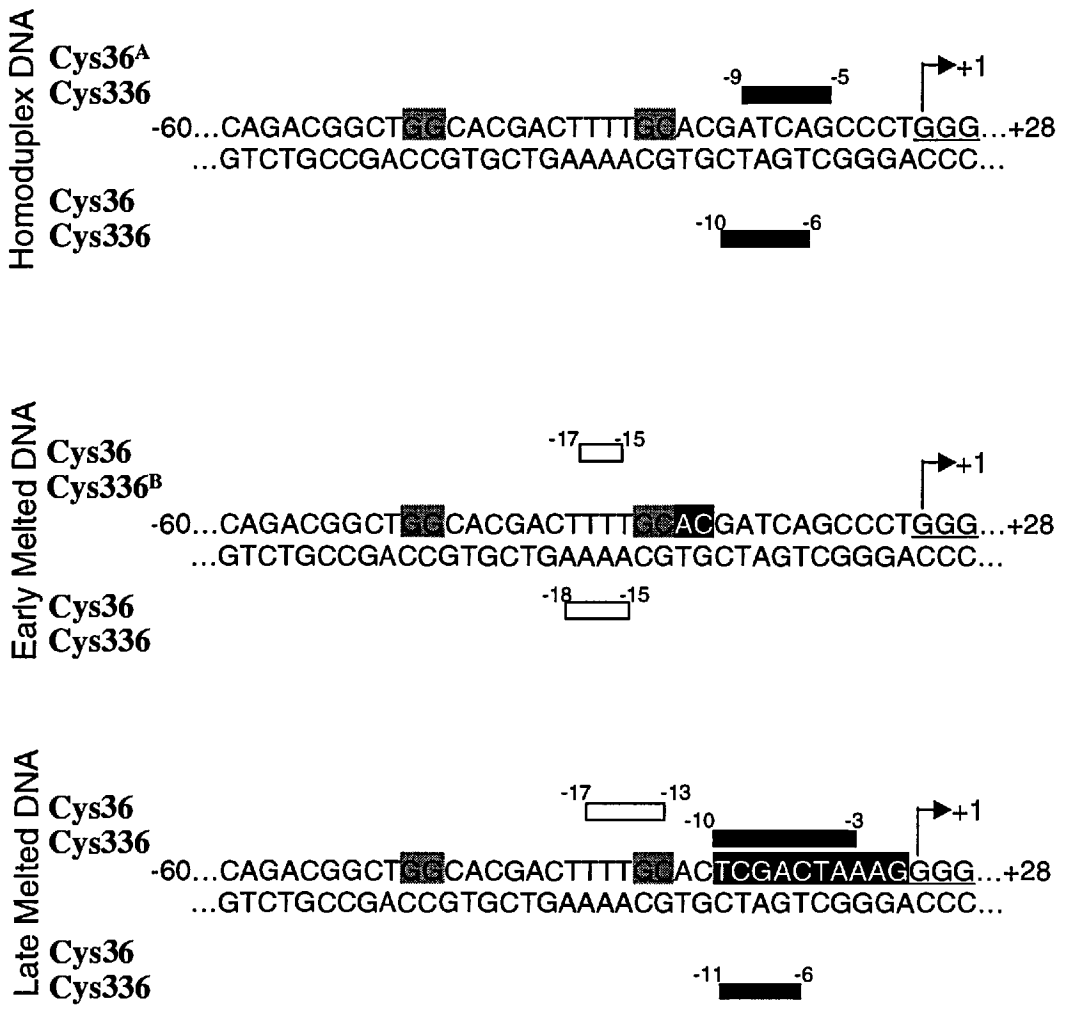


Figure 5F. *S. meliloti nifH* promoter DNA cleavage by RNAP holoenzyme containing FeBABE modified σ^{54} proteins. Late melted probe non-template and template strand DNA cleavage under non-activating (**A** and **C**) and activating (**B** and **D**) conditions, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +, control reactions to which no ascorbate and hydrogen peroxide were added are marked with -. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as molecular size markers. (A high quality reproduction of this figure is available in Wigneshweraraj *et al.*, 2001, which is attached at the back of this thesis).

(A) DNA Cleavage by σ^{54}



[A]Binding evident by gel mobility shift assays
[B]Binding not evident by gel mobility shift assays

Figure 5G. (continued on next page)

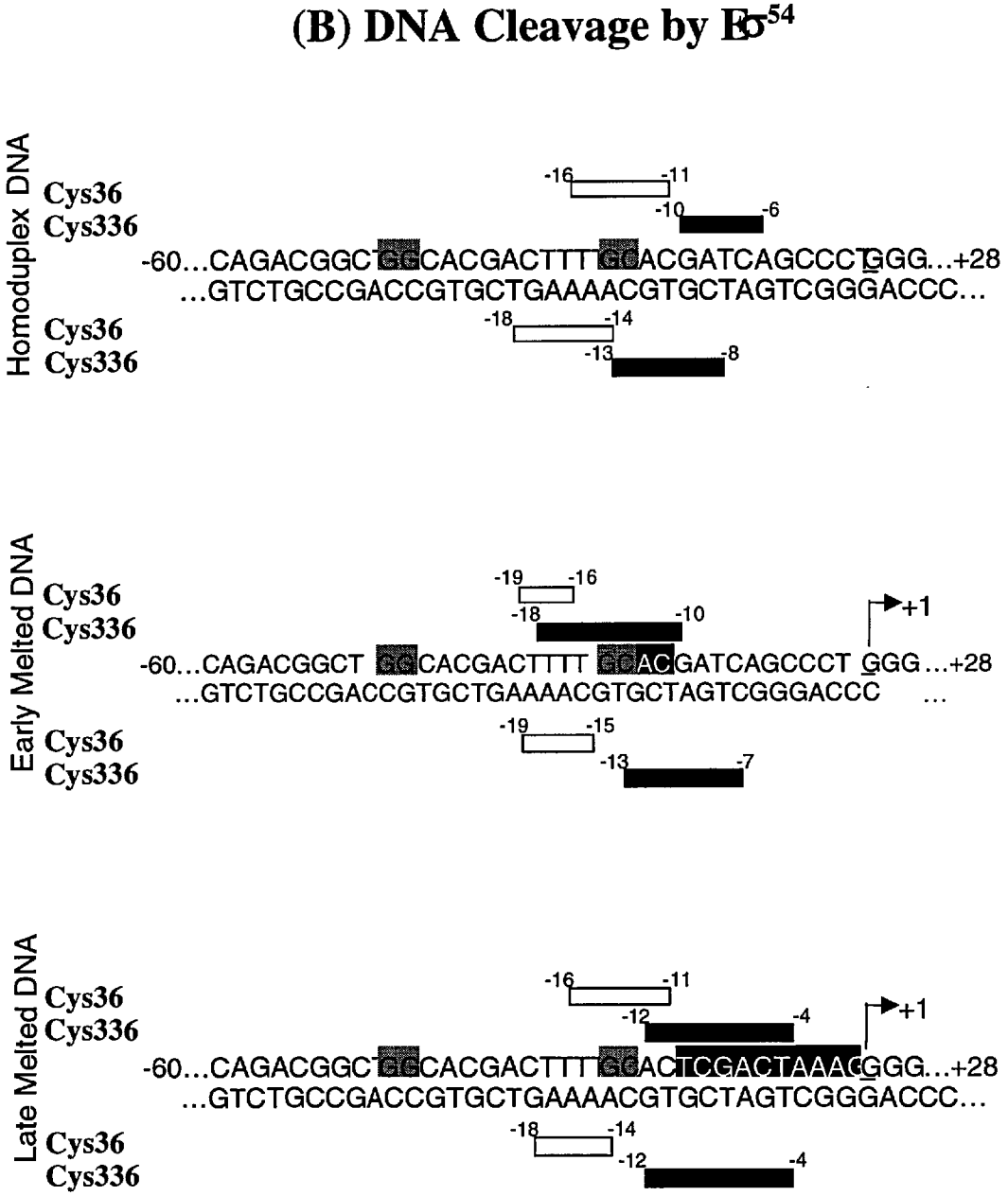


Figure 5G. (continued on next page)

(C) DNA Cleavage by Activated $E\sigma^{54}$

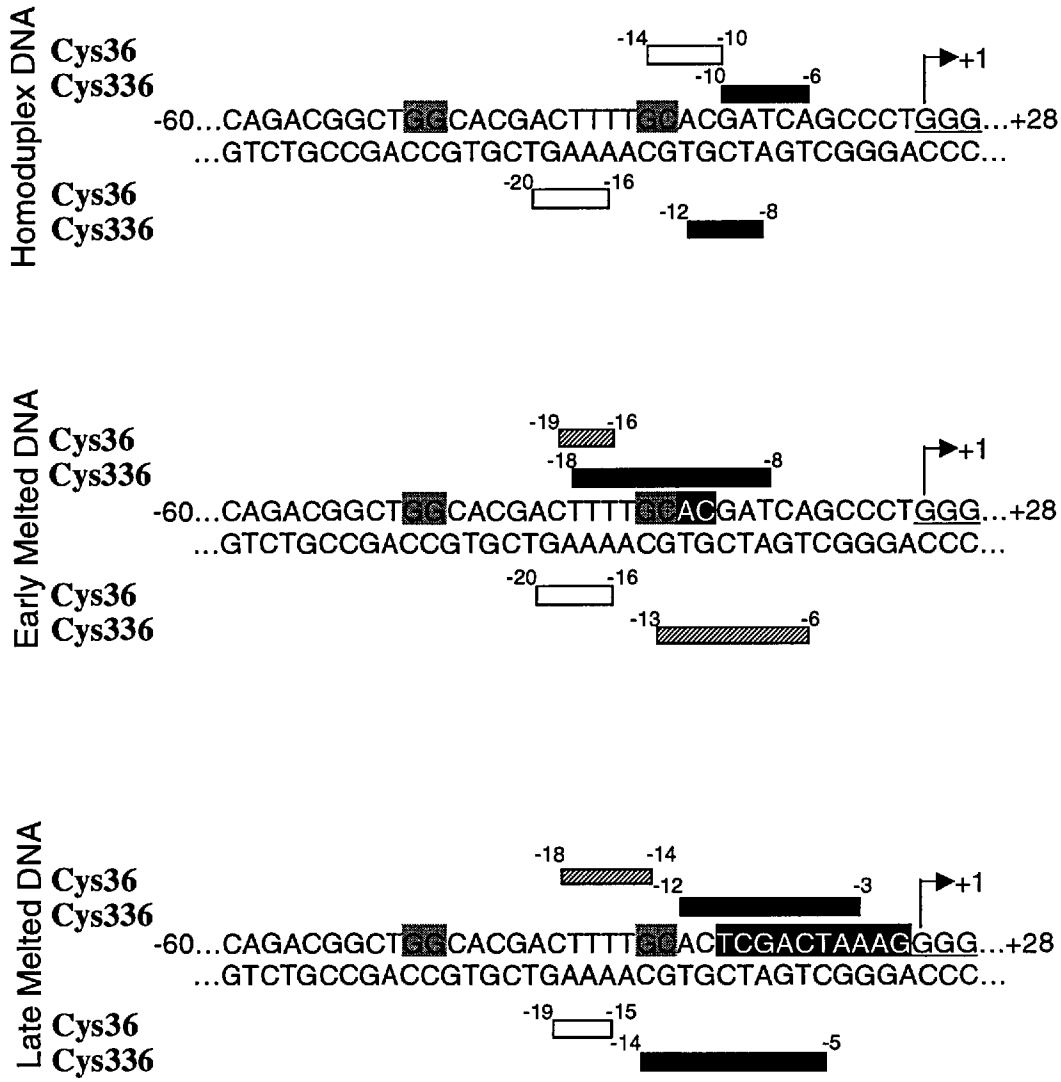


Figure 5G. Summary of *S. meliloti nifH* promoter cleavage results. Boxes above (non-template strand) and below (template strand) the DNA sequence indicate positions of cleavage by FeBABE-modified σ^{54} proteins. Open boxes show cutting by Cys36-FeBABE and filled boxes by the Cys336-FeBABE proteins. Shown are (**A**) σ^{54}-DNA complexes, (**B**) holoenzyme closed complexes and (**C**) holoenzyme complexes with activation on homoduplex, early melted and late melted DNA structures, respectively. In (C) the hatched box indicates the "repositioning" of Cys36 from the non-template DNA strand and Cys336 from the template DNA strand under activating conditions (see text for details).

5.4 Discussion & Conclusions

Results overview. Previous work examined the inter-relationship between Region I and a Region III DNA crosslinking patch in σ^{54} that, when mutated, displayed overlapping new functionalities of the σ^{54}-holoenzyme and changed DNA binding activities of σ^{54} (Wigneshweraraj *et al.*, 2001). Although the two sequences are separated by nearly 300 amino acids, they seem to occupy closely related positions in space with respect to the core RNAP (Wigneshweraraj *et al.*, 2000) and close to the GC consensus in the promoter DNA (this Chapter). This DNA region is associated with regulatable properties of the promoter, especially the dependence upon activator for open complex formation and DNA melting (Cannon *et al.*, 2000; Guo *et al.*, 1999; Wang *et al.*, 1999, 1997). Hence we suggest that parts of Region I and Region III form a centre in the holoenzyme which is closely associated with DNA melting. Changing relationships between σ^{54} and DNA were evident under activating conditions (Figure 5C-5F, summarised in Figure 5G), suggesting that movements in the centre occur and are probably needed for open promoter formation. Some of these movements seem to involve parts of σ^{54}, in particular Region I sequences, that direct σ^{54} to melted DNA structures downstream of the promoter GC element. The differences in DNA cutting patterns seen between the FeBABE modified free σ^{54} proteins and their holoenzymes implies some rearrangements in the σ^{54} DNA-interacting surfaces upon binding to core RNAP. Strikingly, either the addition of core RNAP or the use of early melted DNA is required to allow detectable non-template strand cutting by Cys36-FeBABE. This suggests that core RNAP and the distorted DNA structure created when holoenzyme binds homoduplex DNA can increase the interaction of Region I sequences with promoter DNA. This is consistent with differences seen between σ^{54} and σ^{54}-holoenzyme DNA –12 region footprints (Morris *et al.*, 1994).

The differences in the DNA cutting patterns seen with σ^{54} and σ^{54}-holoenzyme implies re-arrangements in the σ^{54} DNA-interacting surfaces upon binding to core RNAP, suggesting that the FeBABE is located in a slightly different place in the σ^{54} versus σ^{54}-holoenzyme. As summarised in Figure 5G, a region of promoter DNA between -18 and -5 are within cutting range (~12Å from the cysteine sulphur, ± 3-4Å diffusion distance of the

actual cutting species) of Region I residue 36 and Region III residue 336. Our cleavage data imply that the Cys36-FeBABE cleavage sites are, in general, located upstream of the consensus GC promoter element, whereas the Cys336-FeBABE cleavage sites are located downstream of the consensus GC element. This suggests they could form a clamp across the GC region. There is a certain degree of overlap in the DNA cleavages produced by Cys36-FeBABE and Cys336-FeBABE that occur over the GC element. A possible implication of the RegionI/R336-DNA proximity relationship in starting DNA melting is discussed below ("*open complex formation*")

Region I. It is now clear that part of Region I, including position 36, is located near position –15, close to the GC consensus element, under all conditions tested. Although this proximity relationship is suggestive of a contact between Region I and DNA, a role for Region I in structuring the DNA-contacting surfaces of σ^{54} is also plausible. Indeed, the two models need not to be mutually exlusive. Reduced proximity of residue 36 to the DNA was evident in at least a large fraction of the σ^{54} molecules nnder conditions of activaton, consistent with the view that Region I contributes to at least two different functional states of the holoenzyme, and controls their interconversion (Gallegos *et al.*, 1999). The FeBABE cleavage data imply that Region I moves away from the non-template strand upon activation and is positioned closer to the template strand sequences. The re-positioning of Region I is consistent with conformational changes that take place in response to activation (Casaz & Buck, 1997). The full significance of this movement remains to be elucidated. One possibility is that upon activation Region I switches its interactions to a new set that contributes to open complex formation - one such interaction could be in directing template strand DNA sequences towards the RNAP catalytic centre.

R336 and relationship to Region I. The DNA proximity relationships suggested for R336 indicate that it interacts with sequences just downstream of the GC element, close to the site of early melted DNA. Consistent with this, R336A binds probes with early melted DNA worse than homoduplex DNA. The known roles of arginine residues in DNA interactions suggests that R336 may well make a direct DNA contact. DNA melting within isomerised σ^{54}-DNA complexes is to –6 (Cannon *et al.*, 2000) suggesting that R336 interacts with DNA that exists in either the double stranded form, or is melted out. R336 may make a direct interaction with a promoter segment to be melted, in both double and

single stranded states. This suggests that sigma factor functions in DNA melting, as do conventional DNA helicases, through direct interaction with DNA that later melts.

Activation of the Cys336-FeBABE holoenzyme complex on early melted DNA probe causes a reduction in efficiency of template strand cutting, whereas activation leads to enhanced cutting of the non-template strand in the complex with late melted DNA. Cutting of non-template strand sequences by the Cys36-FeBABE seems to always be reduced under activating conditions on the early melted and late melted DNA probes. The FeBABE cutting data clearly provides evidence for movements in two σ^{54} –DNA relationships during transition from a closed complex to an open complex. Conformational changes in Region I upon activation have been observed (Casaz *et al.*, 1997).

Stable closed complex formation. Tight binding of σ^{54} to a distorted DNA structure, the fork junction, just downstream of the GC-element, is thought to hold the holoenzyme in a stable conformation and so maintain the closed complex (Cannon *et al.*, 2000; Gallegos & Buck, 2000; Guo *et al.*, 1999). The overlapping phenotypes of $\Delta I\sigma^{54}$ and the R336 mutants suggest that Region I could act in part through R336. Here, establishing the Region I negative effects of limiting unregulated DNA opening could be through the R336 dependent tight binding of σ^{54} to DNA downstream of the GC to restrict the spread of DNA melting. Since Region I is predicted to be α-helical it is also possible that parts of Region I extend to make DNA contact across the sequences that R336 is suggested to interact with.

Open complex formation. Activation of σ^{54} and its holoenzyme leads to DNA melting downstream of the early melted DNA (Tintut *et al.*, 1995; Popham *et al.*, 1989). Although it is unclear how activator achieves this in a nucleotide consuming reaction, it is evident from the altered DNA cutting patterns of the FeBABE modified σ^{54} proteins that changed interactions across the GC promoter region are occurring. These may serve to relieve the inhibition on polymerase isomerisation imposed by Region I and R336, and to establish a new set of interactions needed for open complex formation. The need for a concerted set of movements in σ^{54} is reinforced by the finding that pre-opened DNA does not bypass activator requirements for transcription (Cannon *et al.*, 1999; Wedel & Kustu, 1995) or changing cutting patterns (Figure 5D). Further, FeBABE derivatives of R336 and

E36 cut the core RNAP in a restricted set of positions, some of which are close to the polymerase active site (Wigneshweraraj *et al.*, 2000). This suggests that activation involves movements of σ^{54} within the holoenzyme associated with placing DNA correctly within the polymerase active centre. DNA placement may contribute to open complex formation and stability through core RNAP melted-DNA interactions (Brodolin *et al.*, 2000).

Strain forces can drive the DNA helix to adopt non-canonical structures. Results with intra-strand crosslinks support this view (van Aalten *et al.*, 1999). These can lead to new DNA structures that can very substantially increase DNA-protein interactions, possibly through compensating for the normal energetic costs of protein induced DNA deformation (van Aalten *et al.*, 1999). Further, torsional pressures applied to the DNA minor groove can be relieved by base pair rupture and base extrusion (van Aalten *et al.*, 1999). Clearly, parallels exist with the DNA strand separation events of open promoter complex formation. It is possible that local DNA distortion downstream of the GC element is created through strain induced by Region I and the element of structure that includes R336 interacting across the GC element. Footprinting experiments have suggested that σ^{54} makes minor groove interactions across promoter sequences (Cannon *et al.*, 1995; Morris *et al.*, 1994) and common properties of $\Delta I\sigma^{54}$ and R336A include lack of local DNA distortion in the closed complex and defects in binding the early melted DNA probe (Wigneshweraraj *et al.*, 2001; Chaney *et al.*, 2000; Gallegos & Buck, 2000; Gallegos *et al.*, 1999; Chaney & Buck, 1999). Strain may be used to start local DNA opening downstream of the GC element by a protein-driven destabilisation of double stranded DNA by a peeling-type mechanism.

Region I and Region III sequences form a regulatory centre. Region I appears to interact with the carboxyl terminal DNA-binding domain in Region III (Wigneshweraraj *et al.*, 2000; Casaz & Buck, 1999). The FeBABE cleavage data now shows that parts of Region I and Region III are close to the –12 promoter element, a DNA sequence known to influence the regulation of open complex formation (Guo *et al.*, 1999; Wang & Gralla, 1998) and to core RNAP subunits (Wigneshweraraj *et al.*, 2000). An advantage of this protein-DNA arrangement would be in co-ordinating concerted movements in σ^{54}, promoter DNA and core RNAP to orchestrate the formation of a stable open promoter complex upon activation. Region I may act as an organising centre that brings different σ^{54}

(and RNAP) and DNA components together. Region I and Region III change their relationship to DNA in response to activation, suggesting that the organising centre includes a target for the activator NTPase activity (Guo *et al.*, 2000; Cannon *et al.*, 2000). In σ^{70} mediated transcription initiation where activation is often by recruitment of the RNAP, determinants within the holoenzyme needed for activation often localise at or upstream of the start site distal –35 promoter element (see Introduction). In contrast, our work shows that for σ^{54}-holoenzyme, which is activated at the DNA melting step, activation determinants in the holoenzyme localise over the start site proximal –12 promoter region from where DNA melting originates (Cannon *et al.*, 2000; Guo *et al.*, 1999; Morris *et al.*, 1994). DNA looping could easily deliver enhancer bound activator to the centre formed by core RNAP, Region I and the DNA crosslinking patch of σ^{54}. Access of activator to the centre from sites on DNA proximal to the holoenzyme would be sterically hindered by RNAP, suggesting a mechanistic role for DNA looping in activation.

CHAPTER SIX

In Vitro Roles of Invariant Helix-Turn-Helix Motif Residue R383 in Sigma 54

6.1 Introduction

Promoter-specific DNA-binding activity of σ^{54} is central to formation of the $E\sigma^{54}$-promoter complex. DNA binding by σ^{54} appears complex and the interaction between σ^{54} and DNA is modulated by core RNAP (Casaz & Buck, 1999; Guo *et al.*, 1999). The specific DNA binding determinants of σ^{54} are located in the C-terminal Region III (residues 329-477 in *K. pneumoniae*), which includes a patch (residues 329-346) that UV-crosslinks to DNA, a putative helix-turn-helix (HTH) motif (residues 367-386) and a conserved patch (residues 454-463) of 10 amino acids termed the RpoN box which has been implicated to play a direct or indirect role in the recognition of both, -24 and -12 conserved promoter elements (Chaney *et al.*, 2000; Gallegos *et al.*, 1999; Guo & Gralla, 1997; Taylor *et al.*, 1996; Cannon *et al.*, 1995; Merrick & Chambers, 1992).

Region III residues 367-386 of σ^{54} are proposed to form a HTH DNA binding structure, and R383 in the recognition helix to interact with bases in the -12 promoter element, in particular with the consensus G of the GC promoter doublet (Merrick & Chambers, 1992; Figure 6A). Substitution of R383 with any other amino acid, except for lysine and to a lesser extent with histidine was suggested to result in an inactive protein, implying that the nature of the charge on this residue is an important criterion for σ^{54} function (Merrick & Chambers, 1992). The suppression of –12 promoter-down mutations in the *K. pneumoniae glnAp2* promoter by R383K *in vivo* is considered as evidence for a role of R383 in recognising the –12 promoter region. An extension of these conclusions was that the promoter interaction was direct, based largely on the view that the suggested bi-helical structure would make specific contacts to promoter DNA and the apparent suppression data might not be explained by indirect effects.

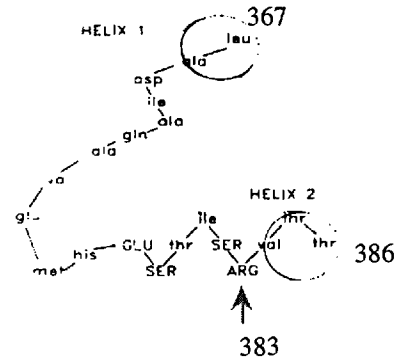


Figure 6A. The putative helix-turn-helix motif of *K. pneumoniae* σ^{54} (residues 367-386). The arginine in position 383 (R383) in the recognition helix (residues 378-386) is implicated in -12 promoter element interactions (Merrick & Chambers, 1992).

The GC promoter region of σ^{54}-dependent promoters in known to be a key DNA element in the network of interactions that keeps the polymerase silent and which establishes the activatable state of the polymerase (Cannon *et al.*, 2000; Wang *et al.*, 1999; Guo *et al.*, 1999 Wang & Gralla, 1998). Previously it was shown by tethered FeBABE footprinting and genetic studies that Region I, the σ^{54} UV-crosslinking patch and the -12 promoter region form a centre in the holoenzyme that contains protein and DNA determinants of activator responsiveness and DNA melting (Cannon *et al.*, 2000; Chaney *et al.*, 2000; Wigneshweraraj *et al.*, 2000; Wigneshweraraj *et al.*, 2001). Here the functionality of purified σ^{54} proteins altered at position 383 have been explored to determine if R383 is a part of the regulatory centre in the σ^{54}-holoenzyme. Results indicate that R383 is not a part of the centre, and that R383 may not establish a direct contact to DNA. Rather it seems that residue 346 is a part of the centre and is close to the promoter GC. However, it is clear that R383 contributes to DNA binding and for discriminating between bases at the G of the GC. It is also required for σ^{54} stability *in vivo*. We show that R383 contributes to maintaining stable promoter complexes in which limited one base DNA opening downstream of the -12 GC element has occurred. Unlike with the wild-type σ^{54}, holoenzymes assembled with the R383A or R383K mutants could not form activator independent heparin stable complexes on the heteroduplex *S. meliloti nifH* DNA mismatched next to the GC (Cannon *et al.*, 2000; Gallegos & Buck, 2000). Although R383

appears inessential for polymerase isomerisation, it appears to make a significant contribution to maintaining the holoenzyme in a stable complex when melting is initiating next to the GC element.

6.2 Results

6.2.1 Expression and Stability of the R383K and R383A σ^{54} Mutant Proteins

We constructed the σ^{54} mutants R383K and R383A to explore their activities *in vitro*. Denaturing gel analysis of whole-cell extracts from induced *E. coli* BL21 (pLysS) cultures revealed proteolysis of the overexpressed R383A protein (Figure 6B-A, *lane 2*). In contrast the R383K protein, harbouring the more conservative substitution, appeared to be more stable and migrated as a single band, as did the wild-type protein during denaturing gel electrophoresis (Figure 6B-A, *lane 4*). Overexpression of the R383A protein in different *E. coli* backgrounds and altering overexpression conditions (time & temperature) did not improve the stability of the R383A protein (data not shown). Since the truncated R383A co-purified with the full-length R383A protein during Ni affinity chromatography, the R383A appears proteolysed in its carboxyl terminal domain. These observations indicate R383 is structurally important, not readily predicted from the suggestion that R383 is solvent exposed (Merrick & Chambers, 1992). We therefore constructed the R383C mutant protein in a cysteine-free σ^{54} background to measure solvent accessibility (Wigneshweraraj *et al.*, 2000). The CPM reactivity of the R383C in its native state showed that R383C is indeed solvent accessible. Furthermore, the R383C protein was more stable upon overexpression and more active for transcription *in vitro* than is R383A and R383K (Wigneshweraraj *et al.*, 2000, data not shown). We therefore infer that R383 has a structural role related to the bulk of the side chain and that 383 is a surface accessible residue.

Previous *in vivo* studies led to the conclusion that R383A is unable to initiate transcription from the *E. coli glnAp2* promoter (Merrick & Chambers, 1992). The

instability of R383A we observed (Figure 6B-A) suggested that the apparent inactivity of the R383A protein could be due to proteolytic cleavage *in vivo* rather than solely a functional defect caused by the mutation. We therefore conducted *in vivo* promoter activation assays (promoter fusion ß-galactosidase assays) and Western blots on the R383A protein. Leaky expression of *rpoN* in pET28b+ allows the use of the overexpression plasmid in these assays (Casaz *et al.*, 1999). Consistent with the previous *in vivo* results (Merrick & Chambers, 1992), the R383A mutant did not support activation *in vivo* (data not shown). Analysis of whole-cell extracts prepared from *E. coli* TH1 cells under activating conditions containing the pSRW-R383A with anti-σ^{54} polyclonal antibodies failed to detect full-length σ^{54} protein (Figure 6B-B, *lane 3*); wild-type and R383K proteins were detected (Figure 6B-B, *lanes 1 & 4*). It appears that the activity of R383A may not be easily judged by *in vivo* activity assays. We therefore conducted a series of *in vitro* assays to explore the activity of the R383A and R383K mutants.

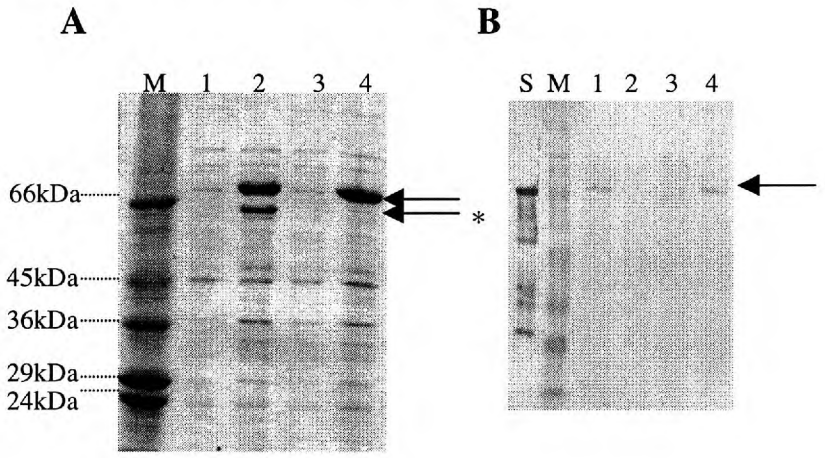


Figure 6B. Overexpression and *in vivo* stability of R383K and R383A mutant σ^{54} proteins. (**A**). Uninduced whole cell extracts from *E. coli* BL21 pLysS carrying pSRW-R383A (*lane 1*) and pSRW-R383K (*lane 3*), fractionated on SDS-PAGE. The upper arrow indicates the expression of the R383A (*lane 2*) and R383K (*lane 4*) in whole cell extracts after two hours of induction. * indicates the proteolysed R383A protein (*lane 2*) which was removed by heparin and Mono Q chromatography. The marker (*lane M*) is BroadRange SDS-7L (Sigma). (**B**). Whole cell extracts from *E. coli* strain TH1 carrying pMTHσ (*lane 1*), pET28b$^+$ (*lane 2*), pSRW-R383A (*lane 3*) and pSRW-R383K (*lane 4*) were probed with

anti-σ^{54} antibodies. The arrow indicates σ^{54}. The prestained marker M (*lane M*) is from BioRad (broad range) and *lane S* contains a partially purified sample of σ^{54}.

6.2.2 Interaction of R383K and R383A with the *E. coli* core RNAP

Native gel holoenzyme assembly assays were used to detect complexes forming between core RNAP and σ^{54} based on the different mobility of core versus holoenzyme. The wild-type σ^{54} shifted nearly all the core into the holoenzyme form at a 1:1 molar ratio of core to σ^{54} (Figure 6C, *lane 2*). Results showed that R383K has a slightly reduced affinity for core RNAP and shifted all the core to the holoenzyme form at a ratio of 1:2 (Figure 6C, *lane 6*). In contrast, R383A had a significantly reduced affinity, and compared to the wild-type σ^{54} forms holoenzyme with increased mobility on native gels. The R383A protein did not produce a characteristic σ^{54} band but was diffuse and slower running (Figure 6C, *lane 16*), in contrast to the R383K and wild-type proteins (Figure 6C, *lanes 14 & 15*). Previously we showed using Cys383-FeBABE tethered footprinting methods that R383 is not proximal to the core subunits β and β' (Wigneshweraraj *et al.*, 2000). We conclude that changing the invariant R383 to A has resulted in a conformational change that may not be localised and which results in significant changes in core RNAP binding and in the formation of holoenzyme with a different conformation. These observations further support a structural role for R383.

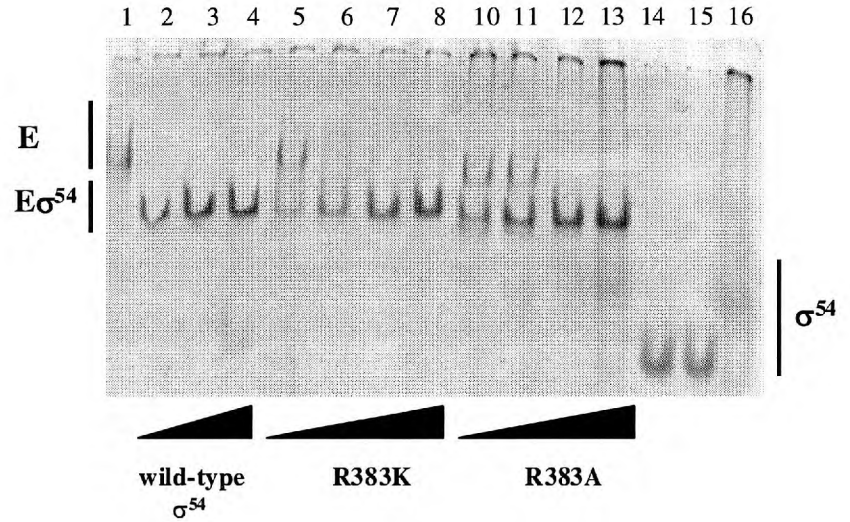


Figure 6C. Binding of σ^{54} to *E. coli* core RNAP. Native gel holoenzyme assembly assays were used to detect complexes forming between core RNAP and R383K and R383A, respectively. The formation of holoenzyme ($E\sigma^{54}$) was detected as the presence of a faster migrating species when compared with core (E, *lane 1*) alone. Titrations of core RNAP with σ^{54} were carried out using 250nM core RNAP and increasing concentrations of σ^{54} at ratios 1:1 (*lanes 2, 5 & 10*), 1:2 (*lanes 3, 6 & 11*), 1:4 (*lanes 4, 7 & 12*) and 1:8 (*lanes 8 & 13*). Free σ^{54} (at 2.5μM) proteins (*lanes 14* [wild-type], *15* [R383A] *& 16* [R383A]) are also shown.

6.2.3 DNA Binding Activities of the R383K and R383A Mutant σ^{54} Proteins and their Holoenzymes

The R383K mutant was suggested to show an altered DNA-binding preference for the promoter GC element (Merrick & Chambers, 1992). Using a gel mobility shift assay we compared the DNA binding activities of R383A, R383K and the wild-type σ^{54} proteins and their holoenzymes for the *E. coli glnHp2* (termed Hp2–13T) and a mutant derivative with a T to G substitution at position –13 (termed Hp2–13T→G) (Figure 6D-A and 6D-B, respectively). Previous *in vivo* and *in vitro* studies have shown that the *glnHp2* promoter with a G at –13 is a better substrate for σ^{54} holoenzyme function than one with a T at –13 (Claverie-Martin & Magasanik, 1992). Binding of σ^{54} confirmed the promoter with the G at position -13 (Hp2-13T→G) as the preferred template, to which σ^{54} has 7-fold increased binding (at 250nM) compared to the Hp2-13T promoter. The R383K mutant bound both promoter sequences similarly, but had a 7-8-fold reduced binding (at 1μM) to the Hp2-13T and Hp2–13T→G templates compared to the wild-type σ^{54} (Figure 6D-A). This observation, together with the inability of R383A to detectably bind either promoter probe,

even at higher protein concentrations (1.5μM and 2μM; Fig. 6D-A) establishes that R383 is important for DNA binding by σ^{54}.

By comparing the binding activities of the wild-type, R383A and R383K holoenzymes for the two promoter probes we determined that the wild-type holoenzyme (at 100nM) has a 4-5-fold higher binding to the Hp2-13T→G than the Hp2–13T promoter sequence (Figure 6D-B). Like the R383K σ^{54}, the R383K-holoenzyme, showed a similar binding preference to both promoter probes, being unable to distinguish them (Figure 6D-B), consistent with previous *in vivo* results (Merrick & Chambers, 1992). Clearly σ^{54}-core RNAP interactions are significant in determining promoter binding (compare Figure 6D-A & 6D-B) and it seems that σ^{54} dominates the promoter binding preference. Any binding preferences of the R383A-holoenzyme to Hp2-13T and Hp2-13T→G could not be determined due to the low binding activity of the R383A holoenzyme to both probes (data not shown). Overall, our data show that R383K and its holoenzyme bind the Hp2-13T→G and the Hp2-13T templates equally. In contrast, wild-type σ^{54} and its holoenzyme bind to the Hp2-13T→G probe better than to the Hp2-13T probe. R383 is significant for DNA binding by σ^{54} and is needed for preferential binding to Hp2-13G rather than the Hp2-13T probe.

To further explore the DNA binding properties of the mutant σ^{54} proteins and their holoenzymes we compared the binding activities of R383K and R383A to the *S. meliloti nifH* promoter DNA, a higher affinity binding site used for many σ^{54} activity measurements (see later). As shown (Figure 6D-C) R383K bound less *S. meliloti nifH* probe (3-fold reduced) compared to the wild-type σ^{54}. In contrast, the R383A mutant appeared very defective for DNA binding. Next, holoenzyme binding to the *S. meliloti nifH* promoter was assayed using saturating ratios of sigma to core RNAP. The wild-type holoenzyme shifted 70% of the *S. meliloti nifH* promoter probe DNA at 150nM, whereas mutant holoenzymes shifted 60% (R383K) and 45% (R383A) of the probe (Figure 6D-D). This observation contrasts with the behaviour of the R383A holoenzyme on the Hp2-13T→G and suggests sequences in the *S. meliloti nifH* promoter rescue promoter binding by R383A holoenzyme.

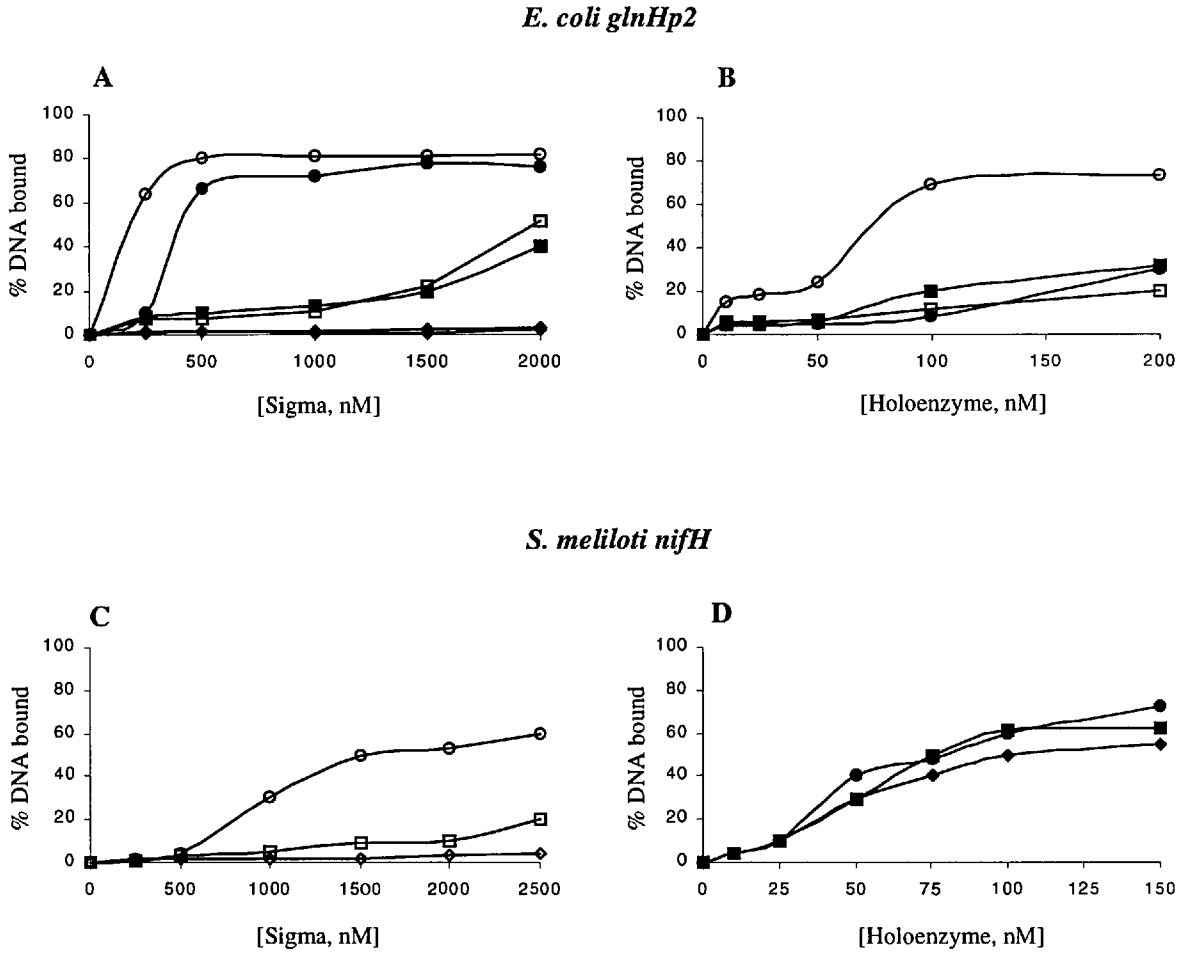


Figure 6D. Interactions of R383K and R383A with *E. coli glnHp2* and *S. meliloti nifH* promoter fragments. Gel mobility shift assays were used to detect the binding activities of the mutant sigmas and their holoenzymes to *E. coli glnHp2* promoter fragments: (**A**) Binding of wild-type σ^{54} (●), R383K (■) and R383A (◆) and (**B**) their holoenzymes to the Hp2-13T (closed symbols) and to the Hp2-13T→G (open symbols).(**C**) Binding of wild-type σ^{54} (○), R383K (□) and R383A (◇) and (**D**) their holoenzymes (same as in (**C**). but closed symbols) to the *S. meliloti nifH* promoter fragment.

6.2.4 *In vitro* Transcription Activity of the R383K and R383A Holoenzymes.

To begin to examine the consequences of altered DNA binding of R383K and R383A upon later steps in activation, we next examined the ability of R383K and R383A holoenzymes to support transcription *in vitro* from supercoiled plasmid pFC50 which contains the wild-type *E. coli glnHp2* promoter or the GC promoter region mutant derivatives of this promoter (Table 1). Changing the –13T to a C (pFC50-m33) or A (pFC50-m11) results in a strong promoter-down phenotype or a largely inactive mutant promoter, respectively (Claverie-Martin & Magasanik, 1992).

Promoter	Plasmid	Sequence
		-28 -13 +1
E. coli glnHp2	pFC50	ACT**GG**CACGATTTTT**TC**ATATATGTGAATGT
	pFC50-m12	ACT**GG**CACGATTTTT**GC**ATATATGTGAATGT
	pFC50-m33	ACT**GG**CACGATTTTT**CC**ATATATGTGAATGT
	pFC50-m11	ACT**GG**CACGATTTTT**AC**ATATATGTGAATGT
		-29 -14 +1
S. meliloti nifH	pMKC28	GCT**GG**CACGACTTTT**GC**ACGATCAGCCCTGGG
		-28 -13 +1
K. pneumoniae glnAp2		GTT**GG**CACAGATTTC**GC**TTTATATTTTTTAA

Table 6.1. *E.coli glnHp2* and *S. meliloti nifH* promoter sequences used for the *in vitro* transcription assays. The *E. coli glnHp2* (shown from -28 to +3) and *S. meliloti nifH* promoter sequences (shown from -29 to +3) used for the *in vitro* transcription assays. The -13GC (for the *glnHp2* promoters) and -14GC (for the *nifH* promoter) promoter elements are underlined and the consensus GG and GC are shown in bold. The *K. pneumoniae glnAp2* promoter sequence (from -28 to +3) is also given for comparison.

Initially the response of wild-type and mutant holoenzymes to saturating concentrations of the *E. coli* activator protein PspFΔHTH, which functions from solution, was tested. The ability of the holoenzymes to promote transcription at the *glnHp2*–13T, *glnHp2*–13T→G and *glnHp2*–13T→C was expressed as a percentage of wild-type holoenzyme activity at the *glnHp2*–13T→G promoter (Figure 6E-A). Assays were conducted with subsaturating amounts of holoenzyme to allow quantitative detection of

holoenzyme activities. Experiments were performed at least six times to enhance reliability. Results clearly showed that the R383A holoenzyme is significantly active and able to support transcription *in vitro* (40-50% of wild-type activity on the *glnHp2*–13T→G promoter; Figure 6E-A). Transcription by R383A is apparently greater than its holoenzyme promoter DNA binding. Formation of stable open complexes that do not rapidly decay to heparin sensitive closed complexes can explain this.

The R383K holoenzyme was ~20% less efficient in transcription than the wild-type holoenzyme at all the promoters tested (Figure 6E-A). This result differs from *in vivo* assays on the *K. pneumoniae glnAp2* which showed that the R383K holoenzyme transcribed better from –13T or –13C promoters than did wild-type σ^{54} (Merrick & Chambers, 1992). To explore the potential for suppression *in vitro*, we varied assay conditions. We failed to see any significant R383K specific suppression at the promoter-down mutant (*glnHp2*-13T & *glnHp2*-13T→C) sequences in the presence of nucleotides that facilitate the formation of initiated complexes prior to heparin challenge, at higher temperatures (37°C instead of 30°C) used to stimulate transient DNA melting or at higher PspFΔHTH concentrations (data not shown).

The possibility that the suppression of promoter-down *K. pneumoniae glnAp2* mutants by R383K seen *in vivo* could be either activator specific or require an enhancer-bound activator was considered. We used NtrC instead of PspFΔHTH and examined the *in vitro* transcription activity of the wild-type and mutant holoenzymes at the *glnHp2* promoters. The results showed that, when activated by NtrC, the R383K holoenzyme transcribed from the *glnHp2*-13T and *glnHp2*-13T→G promoters at levels consistent with our observation using *glnHp2*-13T→G and PspFΔHTH (compare Figure 6E-A and Figure 6E-B). Contrasting the PspFΔHTH results, a much lower activity of R383A holoenzyme at the *glnHp2*-13T→G or no activity at the *glnHp2*-13T and *glnHp2*–13T→C promoters, even at higher holoenzyme and NtrC concentrations (data not shown) was evident. This could be linked to an altered holoenzyme conformation (Figure 6C) and the reduced DNA-binding by the R383A holoenzyme (Figure 6D). The R383A holoenzyme could be

positioned on the DNA in a conformation unfavourable for activation by the more restricted enhancer bound NtrC but allowing activation by PspFΔHTH from solution.

In conclusion, our *in vitro* transcription results do not show the suppression of promoter-down phenotypes of the *E. coli glnHp2* promoter by R383K holoenzyme reported for *in vivo* for *K. pneumoniae glnAp2* promoter assays (Claverie-Martin & Magasanik, 1992). The *in vitro* activities of R383K and R383A holoenzymes argues that R383, at least as seen *in vitro*, is not absolutely required for productive transcription initiation by $E\sigma^{54}$. It seems R383 is not needed to allow preferential initiation of transcription in which the -13 base is G rather than T (Figure 6E). However, the R383A mutation severely reduces the ability of the holoenzyme to be activated by enhancer-bound NtrC (although not by free PspFΔHTH), especially when the –13 basepair is suboptimal.

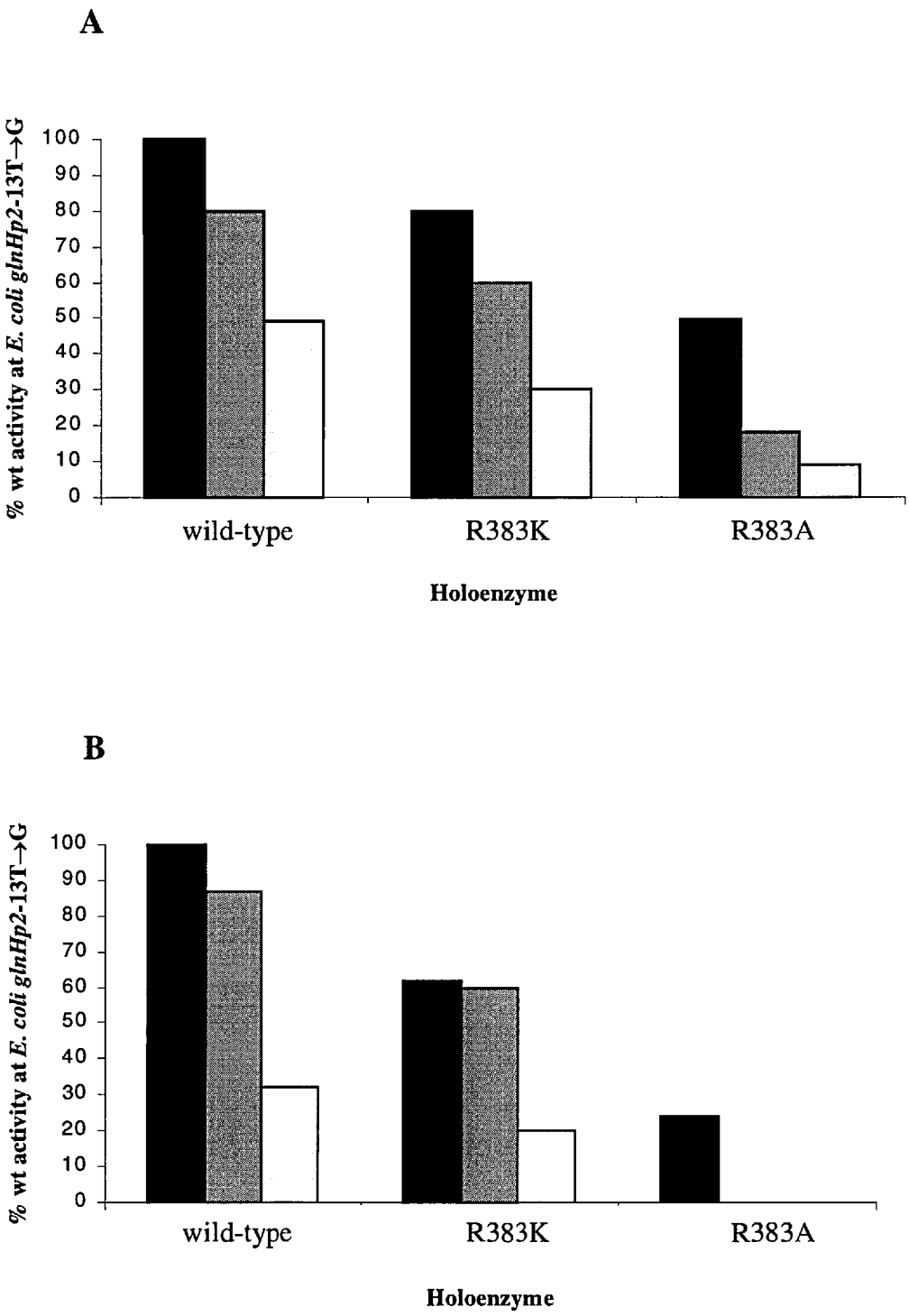


Figure 6E. *In vitro* activator dependent transcription assays on *E. coli glnHp2* wild-type and mutant promoter sequences. The sequence of the promoters used for transcription are as shown in Table 6.1. (**A**). PspFΔHTH activated transcription and (**B**) NtrC activated transcription on *glnHp2*-13T→G (pFC50-m12) (black bars), *glnHp2-13*T (pFC50) (grey bars) and *glnHp2-13*T→C

(pFC50-m33) (white bars). The transcription activities are expressed as a percentage of wild-type activity at the *glnHp2*-13T→G (pFC50-m12) promoter.

6.2.5 Activator-Independent Transcription Activity of the R383K and R383A Holoenzymes

Maintaining the transcriptionally silent state of $E\sigma^{54}$ in closed complexes depends upon the creation and interaction of σ^{54} with locally distorted promoter DNA downstream to the consensus -12 GC promoter element (Gallegos *et al.*, 1999; Morris *et al.*, 1994; Cannon *et al.*, 2000). σ^{54} proteins defective in the recognition of the –12 GC promoter element proximal DNA distortion are capable of increased activator independent transcription *in vitro*, so called bypass transcription (Guo *et al.*, 1999; Casaz *et al.*, 1999; Gallegos & Buck, 2000). We used the *in vitro* bypass assay to see whether holoenzymes formed with R383K and R383A were active for unregulated transcription from the *glnHp2*-13T→G promoter. We used R336A mutant σ^{54} as a positive control for the bypass transcription assay (Chaney & Buck, 1999) and PspFΔHTH for activator dependent transcript formation. As shown, no bypass transcription was observed with R383K or R383A (Figure 6F).

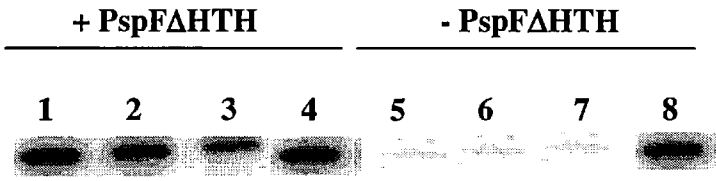


Figure 6F. *In vitro* activator independent transcription assays on the *E. coli glnHp2*-13T→G (pFC50-m12) promoter. Shown are activator dependent (*lanes 1-3*) and activator independent "bypass" transcription (*lanes 5-7*) for the wild-type σ^{54}, R383K and R383A, respectively. The R336A mutant σ^{54} (Chaney & Buck, 1999) was used as the positive control in both cases (*lane 4* for activator dependent and *lane 8* for activator independent reactions).

Additional assays from the *glnHp2* or pMKC28 carrying the *S. meliloti nifH* promoter (Claverie-Martin & Magasanik, 1992; Chaney & Buck, 1999) failed to give

unregulated transcription with the R383 mutants (data not shown). The failure to detect bypass transcription with R383K and R383A suggests tight binding of these mutant σ^{54} proteins to the early melted DNA formed in closed complexes, as seen in heteroduplex DNA-binding assays (Gallegos & Buck, 2000; Cannon *et al.*, 2000). Bypass transcription correlates with strong defects in the binding of σ^{54} to early melted DNA, a defect that is only weakly evident with the R383 mutants (see below). Thus, we infer that the R383A and R383K mutants appear intact in the generation and maintenance of locally melted promoter structures associated with the closed complex (Guo *et al.*, 1999; Gallegos & Buck, 1999). Further, chemical footprinting with *ortho*-copper phenanthroline of closed complexes formed with R383 mutant holoenzymes reveal a local distortion of promoter DNA 3' adjacent to the GC element as seen with the wild-type holoenzyme but lacking in bypass mutants (data not shown). Overall, R383 is not directly or indirectly associated with the inhibition of unregulated bypass transcription *in vitro*. By inference R383 does not closely interact with the elements of –12 promoter region that are involved in maintaining the stable conformation of the closed complex and limiting DNA opening prior to activation.

6.2.6 Interaction of R383K & R383A Mutant Proteins and Their Holoenzyme with Heteroduplex *S. meliloti nifH* Promoter DNA Probes

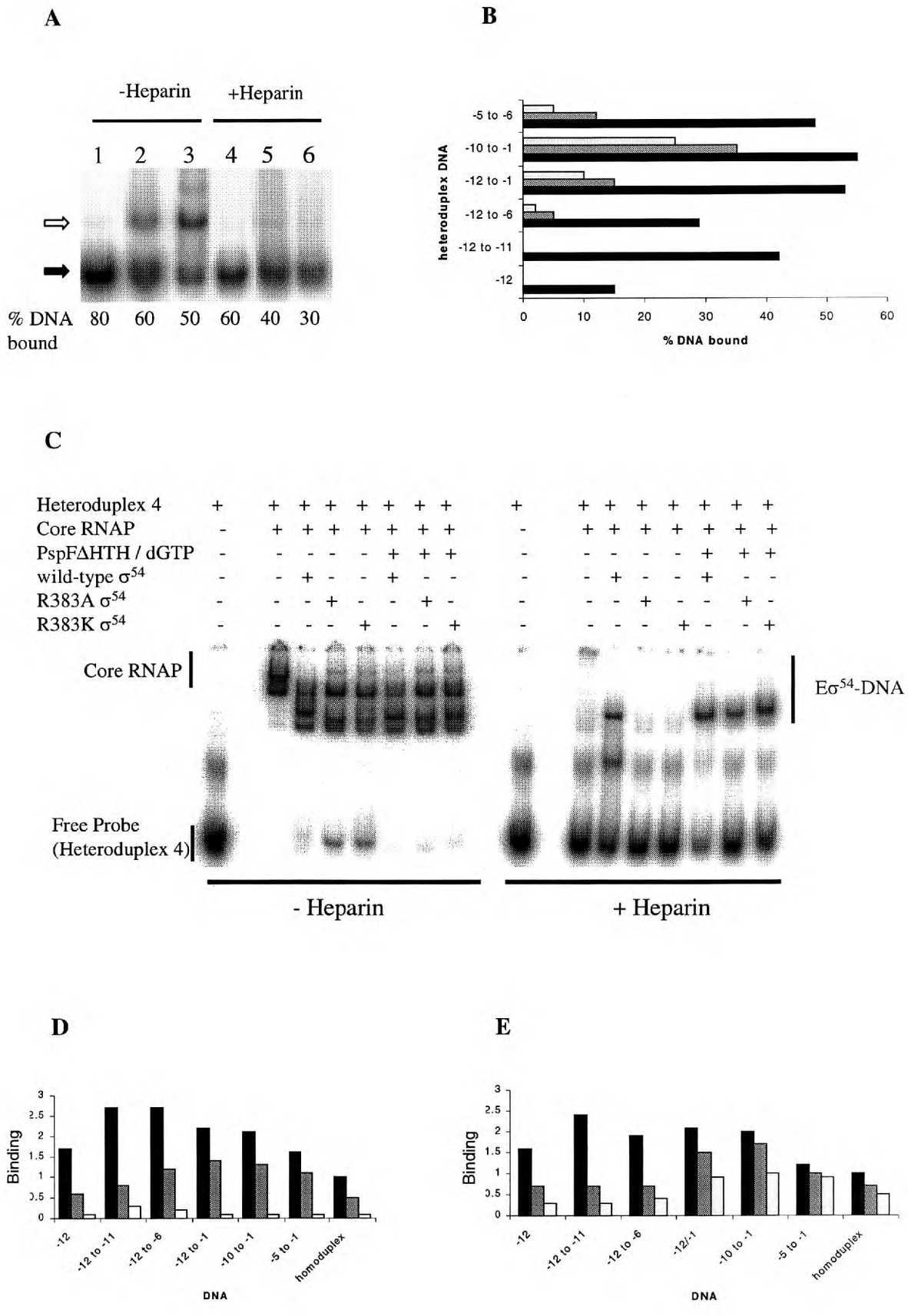
In the course of the *in vitro* transcription experiments we observed that R383K and R383A holoenzymes were essentially inactive for transcription at the *S. meliloti nifH* promoter even though promoter binding was sufficiently efficient to expect transcripts (Figure 6D and data not shown, respectively). The transcription inactivity of the mutant proteins at the *nifH* promoter, but not on the *glnHp2*-13T→G promoter with essentially the same sequences (Table 1) from –11 to +27 prompted us to further explore the properties of R383K and R383A mutant proteins. The *S. meliloti nifH* (from -11 to -1) is rich in G and C residues, whereas the *E. coli glnHp2* promoter is AT-rich in this region, which is melted in open complexes. This led us to consider that the R383K and R383A mutant holoenzymes might be defective in some aspect of DNA melting or single stranded DNA binding at the

S. meliloti nifH promoter. We used heteroduplex DNA that mimics the DNA at different stages of open complex formation to test this idea (Table 2). In marked contrast to the failure to transcribe from the *S. meliloti nifH* promoter (data not shown), both R383K and R383A mutant holoenzymes gave heparin stable activator and nucleotide hydrolysis dependent complexes on promoters with a region of heteroduplex from -10 to -1 (Table 2, heteroduplex 5) (Figure 6G-A). On this DNA structure the mismatched region includes the non-conserved sequence from -10 to -1 that interacts with σ^{54} within the closed complex (Cannon *et al.*, 1995) or with σ^{54}-holoenzyme in the open promoter complex (Popham *et al.*, 1989; Morret & Buck, 1989; Sasse-Dwight & Gralla, 1988). The ability of the wild-type, R383K and R383A holoenzymes to form activator and nucleotide hydrolysis dependent heparin stable complexes when bound to promoter DNA where sequences from –10 to –1 are heteroduplex (Table 6.2, heteroduplex 5) argues that the R383K and R383A holoenzyme are not *per se* defective in polymerase isomerisation at the *nifH* promoter. Also pre-opening from -10 to -1 appears to allow a range of activities with R383K and R383A similar to that seen with the *glnHp2* promoters in transcription assays. As expected from the *in vitro* activator independent transcription results, the wild-type holoenzyme, like the R383K and R383A holoenzymes, does not form heparin stable activator and nucleotide hydrolysis independent complexes on heteroduplex 5 (Cannon *et al.*, 2000, 1999; data not shown).

Heteroduplex		Sequence
1.	-12	-60....GCTGGCACGACTTTTGCCCGATCAGCCCTGGG....+28
2.	-12 to -11	-60....GCTGGCACGACTTTTGCCAGATCAGCCCTGGG....+28
3.	-12 to -6	-60....GCTGGCACGACTTTTGCCATCGACGCCCTGGG....+28
4.	-12 to -1	-60....GCTGGCACGACTTTTGCCATCGACTAAAGGGG....+28
5.	-10 to -1	-60....GCTGGCACGACTTTTGCACTCGACTAAAGGGG....+28
6.	-5 to -1	-60....GCTGGCACGACTTTTGCACGATCATAAAGGGG....+28
7.	wild-type	-60....GCTGGCACGACTTTTGCACGATCAGCCCTGGG ....+28
	glnHp2-13T→G(-12)	-60....ACTGGCACGATTTTTGCCTATATGTGAATGT....+28

Table 6.2. *S. meliloti nifH* and *E. coli glnHp2* DNA fragments used for the gel mobility shift assays. The *S. meliloti nifH* and *E. coli glnHp2*-13T→G (-12) heteroduplex promoter DNA fragments (from –60 to +28) used for the gel mobility shift assays. Shown are sequences from –26 to +3, where in bold are the σ^{54} consensus promoter sequences and boxed in black are mutant sequences introduced in the top strand to create the mismatch (Cannon *et al.*, 1999; Cannon *et al.*, 2000).

Heteroduplex with early melted sequences. Next we examined whether the R383A and R383K mutants were defective in interacting with DNA structures representing the early stages of DNA melting. Since the interaction of the wild-type holoenzyme with promoter DNA results in a structural distortion proximal to the GC element, believed to be the nucleation of DNA melting upon closed complex formation, we used heteroduplex promoter DNA fragments (Table 2) unpaired at –12 (Table 2, heteroduplex 1) and at –12/-11 (Table 2, heteroduplex 2) to mimic the structure believed to be involved in initial DNA opening (Gallegos & Buck, 2000; Cannon *et al.*, 2000).


A
-Heparin
+Heparin
1 2 3 4 5 6
% DNA bound 80 60 50 60 40 30
B
heteroduplex DNA
-5 to -6
-10 to -1
-12 to -1
-12 to -6
-12 to -11
-12
0 10 20 30 40 50 60
% DNA bound
C
Heteroduplex 4
Core RNAP
PspFΔHTH / dGTP
wild-type σ54
R383A σ54
R383K σ54
Core RNAP
Free Probe
(Heteroduplex 4)
Eσ54-DNA
- Heparin
+ Heparin
D
Binding
0 0.5 1 1.5 2 2.5 3
-12
-12 to -11
-12 to -6
-12 to -1
-10 to -1
-5 to -1
homoduplex
DNA
E
Binding
0 0.5 1 1.5 2 2.5 3
-12
-12 to -11
-12 to -6
-12/-1
-10 to -1
-5 to -1
homoduplex
DNA

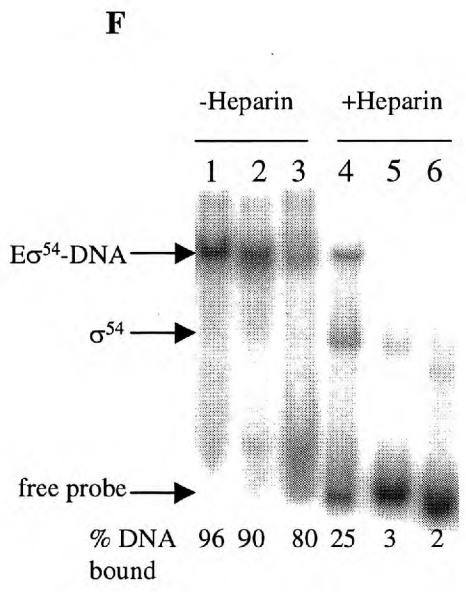


Figure 6G (from previous page). Binding of R383K and R383A to homo- and heteroduplex promoter DNA probes. (**A**). Activator and nucleotide dependent heparin resistant open complex formation on the *S. meliloti nifH* heteroduplex –10 to -1 (heteroduplex 5, Table 2) by wild-type σ^{54} (*lane 1*), R383K (*lane 2*) and R383A (*lane 3*) holoenzymes (100nM). Black arrows indicate the position of $E\sigma^{54}$-DNA complex; white arrows the position of free core RNAP. Shown are reactions without (*lanes 1-3*) and with (*lanes 4-6*) heparin (100μg/ml) challenge for five minutes prior loading. (**B**) Activator dependent, heparin resistant $E\sigma^{54}$-DNA complex formation on heteroduplexes (heteroduplexes 1-6, see Table 2). Shown are % DNA shifted with holoenzymes of wild-type σ^{54} (black bars), R383K (grey bars) and R383A (white bars).(**C**) Stability of wild-type and mutant $E\sigma^{54}$-DNA complexes on *S. meliloti nifH* heteroduplex –12 to –1 (heteroduplex 4) under non-activating and activating conditions in the presence and absence of heparin (100μg/ml), respectively (*lanes as indicated on figure)* (**D**) DNA binding activities of R383K (grey bars) and R383A (white bars) to homo- and heteroduplex DNA (see Table 2) expressed as a fraction of wild-type σ^{54} (black bars) binding. (**E**) As in (**D**) but for holoenzymes. (**F**) Activator and nucleotide hydrolysis independent, heparin stable $E\sigma^{54}$-DNA complex formation on the *E. coli glnHp2*-13T→G(-12) (see Table 2). Shown are reactions without (*lanes 1-3*) and with (*lanes 4-6*) heparin challenge.

These heteroduplexes allow the wild-type holoenzyme to form complexes that survive a heparin challenge independently of activator and nucleotide hydrolysis (Gallegos & Buck, 2000). As shown (Figure 6G-B), we were unable to form heparin stable complexes with the R383K and R383A holoenzymes on either of the heteroduplexes even under activating conditions. This defect correlates with the inability of the R383 mutants to

transcribe from the *nifH* promoter. However, when using heteroduplex promoter DNA where sequences from –12 to –6 (Table 2, heteroduplex 3) and –12 to –1 (Table 2, heteroduplex 4) are unpaired the R383K and R383A holoenzymes survived the heparin challenge, but only with activator and nucleotide hydrolysis (compare Figure 6G-B and Figure 6G-C). In contrast, the wild-type holoenzyme formed heparin stable complexes on heteroduplex 4 in the absence of activator and nucleotide as predicted from the results with heteroduplex 1 and 2, opened at –12 (Table 2, heteroduplex 1) and from -12 to –11 (Table 2, heteroduplex 2) (Gallegos & Buck, 2000; Cannon *et al.*, 2000; Cannon et al., 1999; Oguiza *et al.*, 1999). Although non-native structures near the –12 have the property of allowing the wild-type holoenzyme to form heparin stable complexes independently of activation, the R383K and R383A holoenzymes form heparin stable and activator and nucleotide hydrolysis dependent complexes with promoter probes containing –12 proximal melts (Table 2, heteroduplex 1 & 2) only when these heteroduplexes included further regions of heteroduplex proximal to the start site (Table 2, heteroduplex 3 & 4). These results show that R383 specifies an interaction in the closed complex associated with the stable complex formation by the holoenzyme when the –12 proximal sequences are melted. Other interactions that are required for acquiring heparin stable complex formation involving late melted sequences appear intact in the R383K and R383A mutants.

DNA binding activities on heteroduplexes. Next, we measured the DNA binding activities of the mutant proteins (Figure 6G-D) and their holoenzymes (Figure 6G-E) on *S. meliloti nifH* homoduplex and each of the heteroduplex DNA. Results are shown relative to the binding of the wild-type σ^{54} and its holoenzyme to the *S. meliloti nifH* homoduplex DNA. It is evident that R383K and especially its holoenzyme bound more (2 to 3-fold) of the heteroduplex DNA probes which contain start site proximal melts, whilst the wild-type σ^{54} and its holoenzyme prefer heteroduplex DNA with –12 proximal melts.

Results (Figure 6G) clearly implies a role for R383 in interactions within the closed complexes in which limited DNA opening next to the –12 element has occurred. To further explore this idea we used the *E. coli glnHp2*-13T→G promoter mismatched at position -11 (i.e. the *glnHp2*-13T→G equivalent to heteroduplex 1; Table 6.2). R383K had 80% of

wild-type transcription activity on this promoter *in vitro*. As shown (Figure 6G-F), whilst the wild-type holoenzyme was able to form activator and nucleotide hydrolysis independent heparin stable complexes on the *glnHp2* heteroduplex, the R383K holoenzyme did not. Activation enabled the R383K and to a lesser extent R383A holoenzymes to form some heparin stable complexes (< 10% DNA shifted) on this heteroduplex (data not shown). The ease of opening the *glnHp2* AT-rich sequence (from -11 to -1) may explain why activator allows acquisition of heparin stability, but why at *nifH* a longer segment of heteroduplex is needed also. Therefore, binding assays with two different σ^{54} dependent promoters are consistent with the view that R383 has a role in interactions within the initial closed complexes, after limited DNA opening next to the -12 element has already occurred. Unless DNA opening occurs easily, the defects associated with R383 dominate and few open complexes form.

6.2.7 Proximity of Residue 383 to Promoter DNA

To examine the physical proximity of R383 to promoter DNA, we constructed a σ^{54} with FeBABE located at 383. A single cysteine substitution at 383 was made, the naturally occurring cysteines of the *K. pneumoniae* σ^{54} at 198 and 346 having been replaced by alanine, to allow 383 specific conjugation of FeBABE. The Cys383-FeBABE σ^{54} retained DNA binding and transcription activity (see below and Wigneshweraraj *et al.*, 2000). The Cys383-FeBABE σ^{54} failed to produce detectable cutting of several different *S. meliloti nifH* promoter templates (homoduplex and heteroduplexes 2 and 5), in the presence or absence of core RNA polymerase and under activating conditions that allow open complex formation (see below and data not shown). We considered the possibility that Cys383-FeBABE and its holoenzyme might have dissociated from the promoter DNA under DNA cleavage conditions. However, binding assays conducted with the Cys383-FeBABE protein under DNA cleavage conditions show that 91 % of Cys383-FeBABE and 48% of Cys383-FeBABE-holoenzyme remain bound to DNA (Figure 6H-A). Repeated experiments failed to show DNA cutting by Cys383-FeBABE and its holoenzyme.

As one positive control σ^{54} with FeBABE conjugated to Cys346, at the edge of the DNA crosslinking patch of σ^{54}, produced cutting proximal to the GC element on the *S. meliloti nifH* homoduplex promoter DNA (Figure 6H-B). The putative HTH motif in σ^{54} is C-terminal to a patch of amino acids that UV-crosslinks to promoter DNA. In this patch, FeBABE conjugated to residue 336 cuts DNA downstream of the conserved GC promoter element (Wigneshweraraj *et al.*, 2001). As shown in Figure 6H-B, Cys346-FeBABE holoenzyme strongly cuts homoduplex promoter DNA between positions -14 and -7, mostly downstream of the GC element, but this cutting is upstream of that seen with the Cys336-FeBABE derivative (Wigneshweraraj *et al.*, 2001). The C-terminal to N-terminal orientation of the crosslinking patch is therefore 5' to 3' with respect to the template strand of the promoter DNA.

A

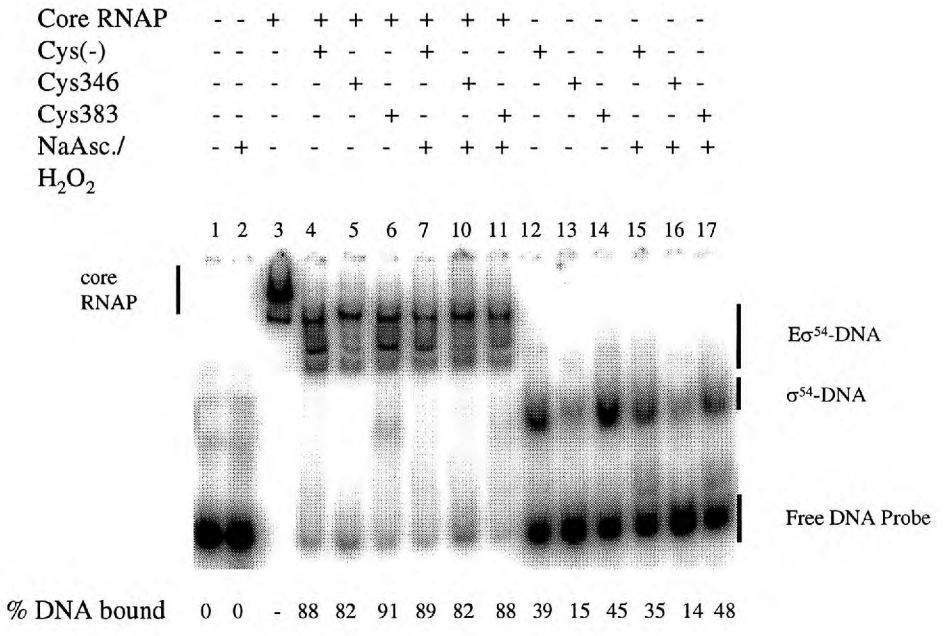


Figure 6H. (continued on next page)

B

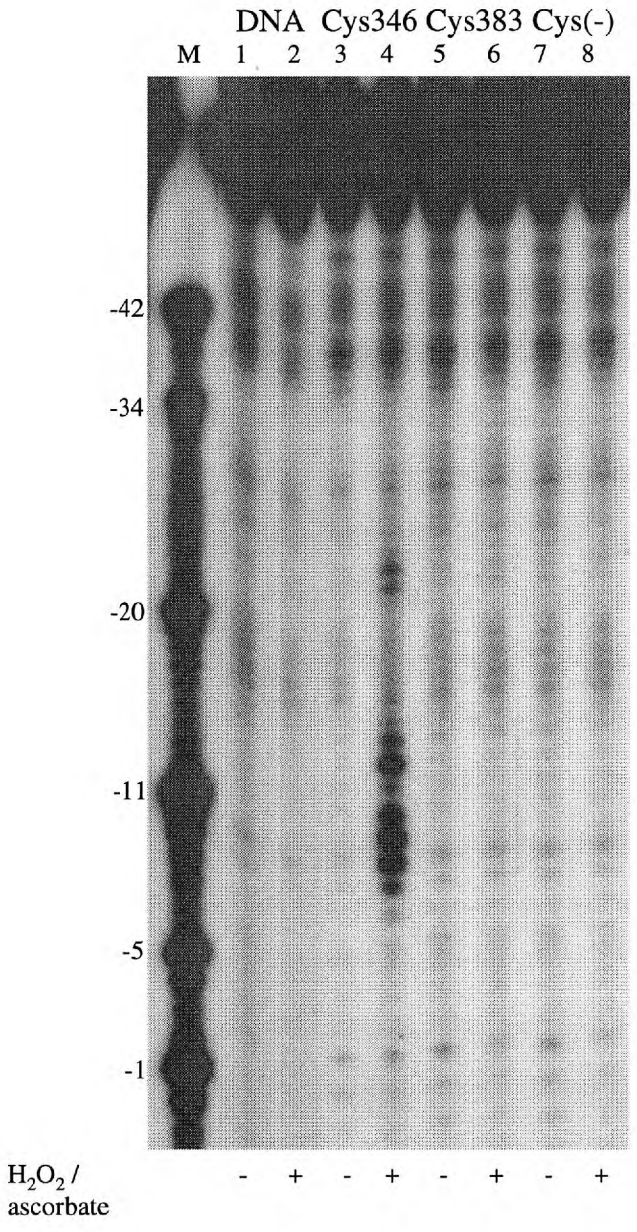


Figure 6H (from previous page). *S. meliloti nifH* promoter DNA template strand cleavage by RNAP holoenzyme containing FeBABE modified σ^{54} proteins. (**A**) Cys383-FeBABE σ^{54} and σ^{54}-holoenzyme binding to the *S. meliloti nifH* homoduplex probe under DNA cleavage conditions (*lanes as marked on figure*). (**B**) Homoduplex cleavage. Reactions to which hydrogen peroxide and ascorbate were added to initiate DNA cleavage are marked with '+', control reactions to which no ascorbate and hydrogen peroxide were added are marked with '-'. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as molecular weight markers.

6.3 Discussion

Specific recognition of promoter DNA by sigma factors has an essential role in locating the RNAP at the correct site for initiation. However, sigma factors are believed to have more functions than simple promoter recognition. In the case of σ^{54}, recent evidence has implicated it in the recognition of the DNA fork junction created when the DNA starts to melt and prevention of further progress with initiation in the absence of activator (Chaney *et al.*, 2000; Guo *et al.*, 1999; Gallegos *et al.*, 1999; Chaney & Buck, 1999). Protein footprints suggest that the DNA binding domain of σ^{54} is a part of the interface with core RNAP and properties of mutants indicate a role in generating polymerase isomerisation and facilitating promoter opening (Cannon *et al.*, 2000; Casaz & Buck, 1999; Chaney & Buck, 1999). Our results address the functions of invariant residue R383 of σ^{54}, previously implicated in interactions with the –12 promoter element.

Protein stability. R383 is clearly required for protein stability *in vivo*. The instability of R383A may be associated with an unfavourable change in structure directly increasing its proteolytic sensitivity. Alternatively, the reduced DNA-binding activity of R383A may indirectly increase proteolysis through changing the intracellular location of σ^{54}. The suggestion from *in vivo* promoter activation assays that R383 is essential for σ^{54} function may be incorrect as purified R383A does support transcription from the *E. coli glnHp2* promoter *in vitro*.

Core RNAP binding. The part of σ^{54} strongly footprinted by core is centred on residue 397, and in the absence of Region I ($\Delta I\sigma^{54}$) much of the 325-440 sequence is protected by core RNA polymerase (Casaz & Buck, 1999). The R383A had reduced core RNA polymerase binding and formed a holoenzyme with altered mobility on native gels, suggesting that R383 contributes to core RNAP binding. Possibly, the core interface of σ^{54} is altered in the R383A mutant, potentially in a manner that involves Region I sequence (see later). Interestingly, Cys383-tethered FeBABE footprinting of core failed to show any proximity of residue 383 (at least within 12Å) to the core subunits β and β'

(Wigneshweraraj *et al.*, 2000). Assuming that this is not an artefact of the R383C mutation or FeBABE attachment, it is suggests either than R383 contacts α, or that the defects in core RNAP binding by the R383 mutants are indirect. The R383K mutant was less disrupted for core binding, consistent with the conservative nature of the mutation.

DNA binding. Gel mobility shift assays showed that R383A was very defective in DNA binding, R383K less so. Whereas wild-type σ^{54} shows preferential binding to *E. coli glnHp2* with consensus G (as against T) in the –13 position, R383K abolishes this ability to discriminate (as judged by DNA binding). The fact that R383K does not show better binding than wild-type σ^{54} to the –13T promoter is also interesting, because R383 has been reported to transcribe more efficiently than wild-type σ^{54} from a similar promoter sequence *in vivo* (mutant *K. pneumoniae glnAp2*) (Merrick & Chambers, 1992). Clearly, the increased transcription reported may not simply correlate with increased DNA binding of σ^{54}. R383K might therefore influence other steps to allow increased transcription *in vivo*. The R383K holoenzyme did not footprint the *glnAp2*-13G→T promoter *in vivo*, but the wild-type σ^{54} did, consistent with the view (developed below) that promoter occupancy may not be dominant for the increased transcription observed (Merrick & Chambers, 1992). It was striking that defects in *in vitro* transcription at *E. coli glnHp2* with the R383K and R383A mutants were less than the defects *in vitro* sigma and holoenzyme DNA binding. We interpret this to mean that promoter occupancy is not reduced to a point so as to severely limit transcription *in vitro*. It is plausible that tight binding of holoenzyme to the promoter increases a transition barrier for open complex formation. The R383K and R383A mutants may reduce this barrier, compensating for reduced promoter occupancy. This favourable effect might contribute to the elevated activities observed with R383K *in vivo*, and the good level of transcription detected *in vitro*. It may also partly compensate for the defect in forming stable complexes with early melted DNA (discussed below).

Interactions with heteroduplex DNA. Results from transcription and assays with heteroduplexes suggested that slow opening of the DNA at the *S. meliloti nifH* promoter might mean that the closed complex or an activator dependent isomerised holoenzyme formed with R383K dissociates prior to full strand opening. In contrast, more frequent

opening of *E. coli glnHp2* promoter or stable opening as in heteroduplexes may explain why these templates do support stable open complex formation with R383K. Even so, these complexes do decay more rapidly than those forming with wild-type σ^{54} (data not shown), suggesting R383 contributes to DNA binding within the open complex. However, R383 is not essential for transcription, at least *in vitro*. Our ability to recover activator dependent stable complex formation using premelted *S. meliloti nifH* DNA templates with heteroduplex –10 to –1 sequences suggests the mutant holoenzymes are limited at some promoters in steps leading to full DNA melting. Compared to DNA templates with start site proximal melts, DNA templates with melted sequences proximal to the –12 promoter element were poor DNA binding sites for the R383K and R383A mutants. Such –12 proximal heteroduplex templates are believed to mimic the DNA distortion seen in the close complex that forms with the wild-type holoenzyme and then allow heparin stable complexes to form without activation. The failure of the R383A and R383K mutants to bind well or make stable complexes on the heteroduplex promoter fragments with –12 proximal melts suggests that the R383 mutant proteins are directly or indirectly defective for interactions with such structures. With the mutant proteins the activator did not allow the use of heteroduplex promoters with –12 proximal melts for efficient stable complex formation on both the *S. meliloti nifH* and *E. coli glnHp2* promoters, but in the context of the *S. meliloti nifH* promoter the activator allowed the formation of stable complexes on promoter sequences which were further opened to include the –1 residue. Overall, the results suggest that DNA melting from –10 to –1 stabilises the promoter complexes that form with the R383K and R383A mutant holoenzymes, and that the initial melting at –12 is unfavourable for stable $E\sigma^{54}$- DNA binding when R383 is A or K. Rapid melting at the AT-rich *E. coli glnHp2* may allow stable complexes to form, but slow melting at the GC-rich *S. meliloti nifH* may result in reduced stable complex formation.

Region I mutants of σ^{54} have defects in binding to early melted DNA structures (Gallegos & Buck, 2000). Indirect effects of substitutions at position 383 upon the function of Region I which is required for tight binding to heteroduplex promoter with –12 proximal

melts is a possibility, suggested by protein-protein and protein-DNA footprints (Wigneshweraraj *et al.*, 2000; Casaz & Buck, 1999, Wigneshweraraj *et al.*, 2001).

Bypass transcription. Interaction of σ^{54} with the –12 GC promoter element appears important for maintaining the transcriptionally silent state of the holoenzyme. Mutations that change the sequences adjacent to the GC or certain amino acids in σ^{54} that interact with it allow transcription independently of activator (Chaney *et al.*, 2000; Wang *et al.*, 1999; Wang & Gralla, 1998). The σ^{54} DNA-binding domain mutant R336A gives strong bypass transcription (Chaney & Buck, 1999). The R383K or R383A did not, despite supporting levels of activated transcription and binding to -10 to -1 heteroduplex DNA (heteroduplex 5) that suggested a bypass activity might be readily detected. If R383 interacts with the –12 GC promoter sequence, it would appear not to be an interaction contributing to maintaining the silent state of the holoenzyme prior to activation, contrasting properties of the R336A and Region I mutants.

DNA proximities. Residue 383 was suggested to be within a HTH motif, expected to establish direct contacts with DNA, and thought to interact with the -12 region of the promoter (Merrick & Chambers, 1992). Although our data show substitutions at 383 do influence DNA binding, other data suggest a simple direct interaction of R383 with DNA may not be occurring. Cleavage of the promoter DNA by Cys383-FeBABE was not evident. Closed complex promoter DNA cleavage by the FeBABE derivative of σ^{54} in the UV-crosslinking patch (Cys336-FeBABE) is centred around position -9 (±1) (Wigneshweraraj *et al.*, 2001) and DNA cutting by the Cys346-FeBABE derivative is centred further upstream at position -11 (±1) (Figure 6H). Given that the UV-crosslinking patch is α-helical in structure, the FeBABE cleavage data suggests that UV-crosslinking patch (N-terminal to the HTH motif) aligns with or is inclined towards the promoter DNA template strand in closed complexes. The lack of any discernible cleavage by Cys383-FeBABE suggests that either residue 383 is not involved in a direct DNA contact, or some structural consequences of making substitutions at 383 do not allow detection using the FeBABE methodology. However, R383C was active for transcription *in vitro*, and more so than R383A (Wigneshweraraj *et al.*, 2001).

Summary. Overall our data are consistent with a requirement for R383 to distinguish between T or G at -13, and an overall reduced DNA binding when substituted by K or A (Merrick & Chambers, 1992). It is possible some of the overall loss in DNA-binding activity has a basis in an altered protein structure rather than the simple loss of a DNA-interacting side chain, a view supported by the *in vivo* instability of the R383A. Although instability is unexpected of substitution by alanine in a surface exposed residue in an α-helix, there are suggestions from secondary structure predictions that the R383 may exist within a non-helical structure. The data presented in Figures 6A to 6G suggest that the interpretation placed on the *in vivo* activation data may need reconsideration and suggest that R383 has a previously unexpected role in the stability of σ^{54}.

In the absence of additional structural or proximity data, any suggestion that the HTH motif lies within a fold that localises sufficiently near the –12 region to contact DNA is speculative. A colinear arrangement, with the promoter non-template strand aligned 5'-3' (i.e. –24, -12, start) beside σ^{54} Region III running C-terminal to N-terminal (i.e. RpoN box, HTH and crosslink-patch) is feasable, but unproven. Further, the clear involvement of the σ^{54} Region I sequences in promoter binding has shown that determinants outside of the DNA-binding domain make critical contributions to the DNA binding function of σ^{54} (Gallegos & Buck, 2000; Guo *et al.*, 1999; Casaz *et al.*, 1999; Wigneshweraraj *et al.*, 2001).

6.4 Conclusions

Specialisation of function across the σ^{54} DNA binding domain is evident. Residues F402, F403, F354 and F355 appear to be associated with interactions needed for efficient polymerase isomerisation (Oguiza *et al.*, 1999; Oguiza & Buck, 1999), R383 with forming stable initially melted DNA complexes, and R336 with maintaining the inhibited silent state of the polymerase (Gallegos *et al.*, 1999; Chaney & Buck, 1999). We note functional similarities between the putative α-helix in which C346 and R336 in σ^{54} lie and the helix

14 of the *E. coli* σ^{70}. Both helices interact with promoter sequences that include recognition sequences (Chaney *et al.*, 2000; Gross *et al.*, 1998, Malhotra *et al.*, 1996). They also contain determinants for binding locally single stranded DNA structures from which melting originates and spreads (Fenton *et al.*, 2000; Chaney & Buck, 1999; Wigneshweraraj *et al.*, 2001). These and other considerations lead to the view that a series of linked interactions that involve σ^{54}-DNA interaction and σ^{54}-core interfaces are required to change in order to allow the polymerase to progress to the open complex. It seems the HTH motif of σ^{54} contributes as a structural element rather than as a major direct DNA contacting surface. However, changes in σ^{54} structure must occur during activation and different functional states may exist to give rise to very transient complexes that escaped our analysis of Cys383-FeBABE proximities to DNA. Current data suggest that the centre formed by Region I, the DNA crosslinking patch of σ^{54} and the -12 promoter region does not include the HTH motif as an element of proximity (Chaney *et al.*, 2000; Wigneshweraraj *et al.*, 2000; Wigneshweraraj *et al.*, 2001).

CHAPTER SEVEN

Correlating Protein Footprinting with Mutational Analysis - Identifying Potential Targets in Sigma 54 for FeBABE Conjugation

7.1 Introduction

Previous work by Casaz & Buck (1997) showed that σ^{54} undergoes specific alterations in protease sensitivity at different stages on the pathway leading from free σ^{54} to one in an engaged transcription complex. The three main findings of this study were:

1. Two distinct regions in σ^{54} were protected from proteolytic cleavage upon interaction with core RNAP: (1) a region in the amino terminus from residues 36-100, consistent with recent observations that amino terminal Region I has low affinity for core RNAP and that residue 36 is in close proximity to the core RNAP subunits β and β' (Gallegos & Buck, 1999; Wigneshweraraj *et al.*, 2000); and (2) a region in the carboxyl terminus spanning residues 325-334, further supporting recent σ^{54}-tethered FeBABE footprinting data which showed residue 336 is close to the catalytic centre of the core RNAP (Wigneshweraraj *et al.*, 2000). However, the protected regions largely lie outside the minimal core binding fragment (residues 70-215) and mainly within regions involved in response to activation (residues 1-50 and 329-346) and near or within the DNA binding domain (residues 329-477), therefore it can be implied that the protection is due to conformational changes in σ^{54} that occur upon interaction with core RNAP (Chaney & Buck, 1999; Chaney *et al.*, 2000; Cannon *et al.*, 1995; Wigneshweraraj *et al.*, 2001).
2. In the closed complex, three sites in σ^{54} located in Region III were protected from proteolytic cleavage: E410, E414 and E431 which lie within the minimal DNA binding domain of σ^{54} (residues 329-477) and between the putative helix-turn-helix motif (residues 367-387) and the RpoN box (residues 454-463) - two elements implicated in promoter DNA recognition (Taylor *et al.*, 1996; Merrick & Chambers, 1992). Thus, it was concluded

that E410, E414 and E431 were probably protected from protease attack by their proximity to promoter DNA in the closed complex.

3. During open complex formation enhanced cleavage was seen at residue E36 in Region I which was absent in closed complexes. This supports the role σ^{54} Region I has in activator response and is consistent with recent FeBABE footprinting data which suggest conformational changes in Region I in response to activation (Wigneshweraraj *et al.*, 2001).

In summary (Figure 7A), the work by Casaz & Buck (1997) showed that surface-exposed regions of σ^{54} undergo conformational changes as a result of interaction with core RNAP and during the transition from the closed complex to the transcription competent open complex. Amino acid residues in Regions I and III were identified which could be either in direct contact or proximal to the core RNAP or promoter DNA. The authors concluded that the changes in protease sensitivity of σ^{54} are due to the interactions σ^{54} makes (with core RNAP, promoter DNA and the activator protein), which result in localised conformational changes within σ^{54} that are required for productive transcription initiation by σ^{54}.

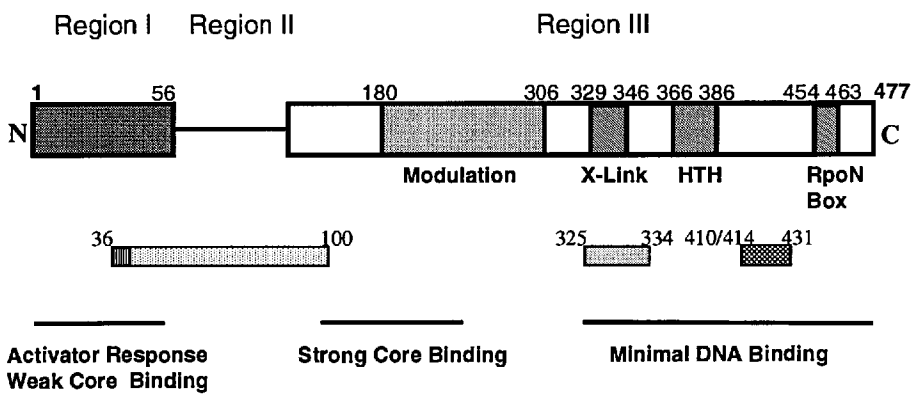


Figure 7A. Summary of protein-footprinting results for σ^{54} (modified from Casaz & Buck, 1997). Coded as follows: dotted bars - core RNAP specific protection, checked bars- closed complex specific protection, and vertically striped bars - cleavage enhanced in open complexes.

In this work, surface exposed residues in Region III of σ^{54} (F318, E410, E414 and E431) that were protected or are proximal to residues protected from protease cleavage by core RNAP and by promoter DNA in the closed complex, and hence may be involved in specific interactions with core RNAP or DNA, were mutated to alanine residues (Table 7.1). Region I residue E36 was substituted with glycine in an attempt to increase conformational flexibility of Region I around position 36, in the anticipation that this mutation could have an altered response to activation compared to the E36C mutation. Recall from before, that the E36C mutant σ^{54} (on the cysteine free *rpoN* background) was able to initiate transcription in the absence of activation *in vitro via* an unstable open complex (Wigneshweraraj *et al.*, 2000). Similarly, residue E325 that was protected from protease attack by core RNAP was substituted with glycine. Preliminary functionalities of the mutant σ^{54} proteins were determined by *in vivo* and purified proteins were used in *in vitro* assays in order to identify potential target residues for future FeBABE footprinting studies.

Residue	IN σ^{54} REGION	CHANGED TO
E36	I	G
F318	III	A
E325	III	G
E410	III	A
E414	III	A
E431	III	A
E414/E410	III	A/A

Table 7.1. σ^{54} residues substituted with either alanine (A) or glycine (G) residues

7.2. *In vivo* Activity of σ^{54} Mutants

β-galactosidase assays were used to assess the *in vivo* activity of the mutants. *K. pneumoniae* UNF2792, which lacks functional σ^{54}, was transformed with mutant *rpoN* containing low copy number vectors carrying mutant *rpoN* alleles, and with pMB1 which contains the β-galactosidase gene under control of the σ^{54}-dependent *K. pneumoniae nifH* promoter (Cannon & Buck, 1992), while the activator NifA is chromosomally expressed under nitrogen limiting conditions. Figure 7B shows the *in vivo* β-galactosidase activity measured for each mutant. With the exception of the E325G and E410A mutants, all others are apparently inactive (< 5% of wild-type) for *in vivo* promoter activation.

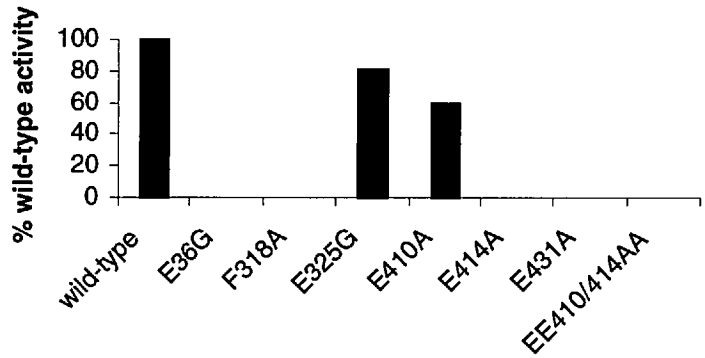


Figure 7B. *In vivo* activity of σ^{54}-mutants. ß-galactosidase assays were performed with strain UNF2792/pMB1. The results of one experiment with four independent colonies for each mutant are presented.

One explanation for the lack of *in vivo* activity of the σ^{54} mutants could be protein instability. Thus, the levels of mutant σ^{54} expressed in UNF2792 cells under the same conditions as used in the activation assays were measured by immunoblotting. Strikingly, the mutants F318A, E414A and EE410/414AA, that were inactive for promoter activation *in vivo* were not stably expressed (Figure 7C). However, the E36G and E431A mutant was found to be expressed to near wild-type levels (Figure 7C, compare lanes 1 and 2, and 1

and 7, respectively). Thus, for the E36G and E431A mutants protein instability appears to be not the cause for the lack of *in vivo* activity. The lower level of detection of σ^{54} in immunoblotting experiments is 2-3ng, which corresponds to roughly 200-300 molecules of σ^{54} per cell (in *E. coli* K12 cell measured by Jishage *et al.*, 1996). Therefore, we conducted *in vivo* promoter activation assays using pMB210.1 which harbours the β-galactosidase gene under the control of the *S. meliloti nifH* promoter. In contrast to the *K. pneumoniae nifH* promoter (in pMB1), the *S. meliloti nifH* is a strong binding site for σ^{54} (Cannon & Buck, 1992). Consistent with previous observations (see above), we failed to detect β-galactosidase activity when using pMB210.1 with the F318A, E414A and EE410/414AA σ^{54} mutants (data not shown). Interestingly, the *S. typhimurium* σ^{54} harbouring the F318S mutation was shown to be stably expressed and to have near wild-type levels of *in vivo* activity at the *glnAp2* promoter in *S. typhimurium* TRH107 cells (Kelly and Hoover, 1999). Overall, the data suggests that low levels of mutant protein expression and stability *in vivo* explain the inability of the F318A, E414A and EE410/414AA mutants to support transcription at the *K. pneumoniae* and *S. meliloti nifH* promoters. Overexpression studies showed that the unstable mutant could be produced (section 7.3). To further confirm the instability of the F318A and E414A mutants *in vivo*, the overexpression plasmids harbouring the F318A and E414A were transformed into *E. coli* TH1 cells (*rpoN* mutants). Immunoblots of cell extract prepared from *E. coli* TH1 cells under *in vivo* promoter activation conditions (see above) showed that F318A and E414A were not stably expressed (data not shown).

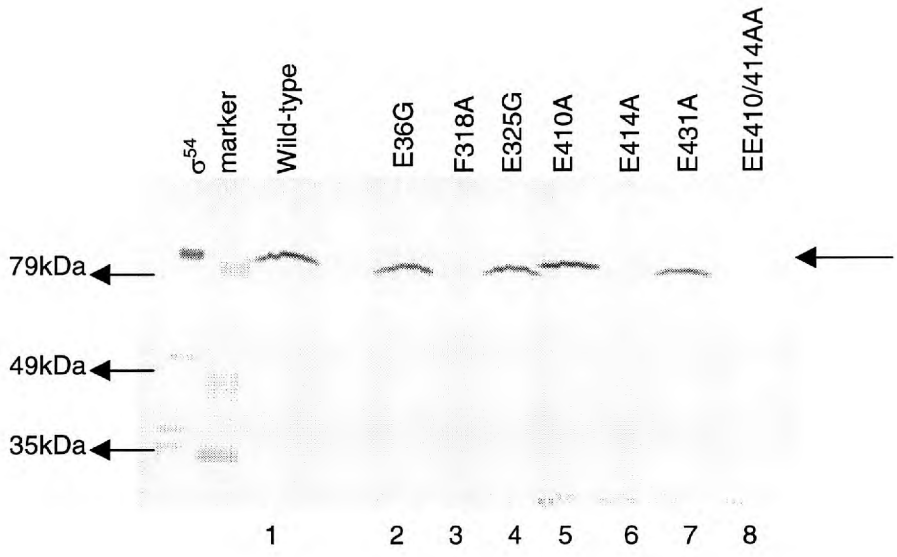


Figure 7C. *In vivo* stability of mutant proteins. Immunoblots of lysates of UNF2792 cells containing plasmids carrying mutant *rpoN* genes. Extracts from equal numbers of cells were loaded in each lane. A prestained marker (Biorad Broad Range) and purified σ^{54} were used to as molecular size references. (See text for details).

7.3 Core RNAP Binding Properties of Purified σ^{54} Mutants

To characterise holoenzymes formed with the mutant σ^{54} proteins, purified proteins were assayed for their ability to bind to core RNAP using a native gel assembly assay (Cannon *et al.*, 1995; Gallegos & Buck, 1999). As shown in Figure 7D, E36G, F318A, E410A, E414A, E431A and the double mutant EE410/414AA bound the core RNAP, as judged by the depletion of the bands corresponding the core RNAP and the appearance of the faster migrating band corresponding to the holoenzyme ($E\sigma^{54}$). The mutants E325G (Figure 7D-B) and to a much lower extent the E36G (Figure 7D-A) showed a clearly reduced affinity to the core RNAP, consistent with the demonstration that Region I and carboxyl terminal residues 325-334 are close to core RNAP in the holoenzyme (Wigneshweraraj *et al.*, 2000; Casaz & Buck, 1997). Both, E325G and E36G did not migrate cleanly into the native gel, and no free fast-running σ^{54} protein was detected in the

lower part of the gel, contrasting with the behaviour of the wild-type protein (Figure 7D, compare lane 6 in panel A with lanes 11 and 6 in panels A and B, respectively) and possibly suggesting aggregation of the purified proteins. The E325G and E36G mutants were needed at a ratio of 8:1 (σ^{54} to core RNAP) to saturate the core RNAP and may reflect a reduced availability of E325G and E36G due to aggregation (Figure 7D, compare lanes 2-5 in panel A with lanes 7-10 and lanes 2-5 in panels A and B, respectively). Thus, in the view that residues E36 and E325 lie outside the high affinity core RNAP binding determinant (residues 120-215) in σ^{54}, but within sequences that are proximal to the core RNAP in the holoenzyme (Wigneshweraraj *et al.*, 2000; Gallegos & Buck, 1999; Casaz & Buck, 1997) we conclude that changing E36 and E325 to glycine has produced σ^{54} derivatives with conformations that either inhibit the initial contacts with core enzyme, or subsequent structural rearrangements required for stable core binding.

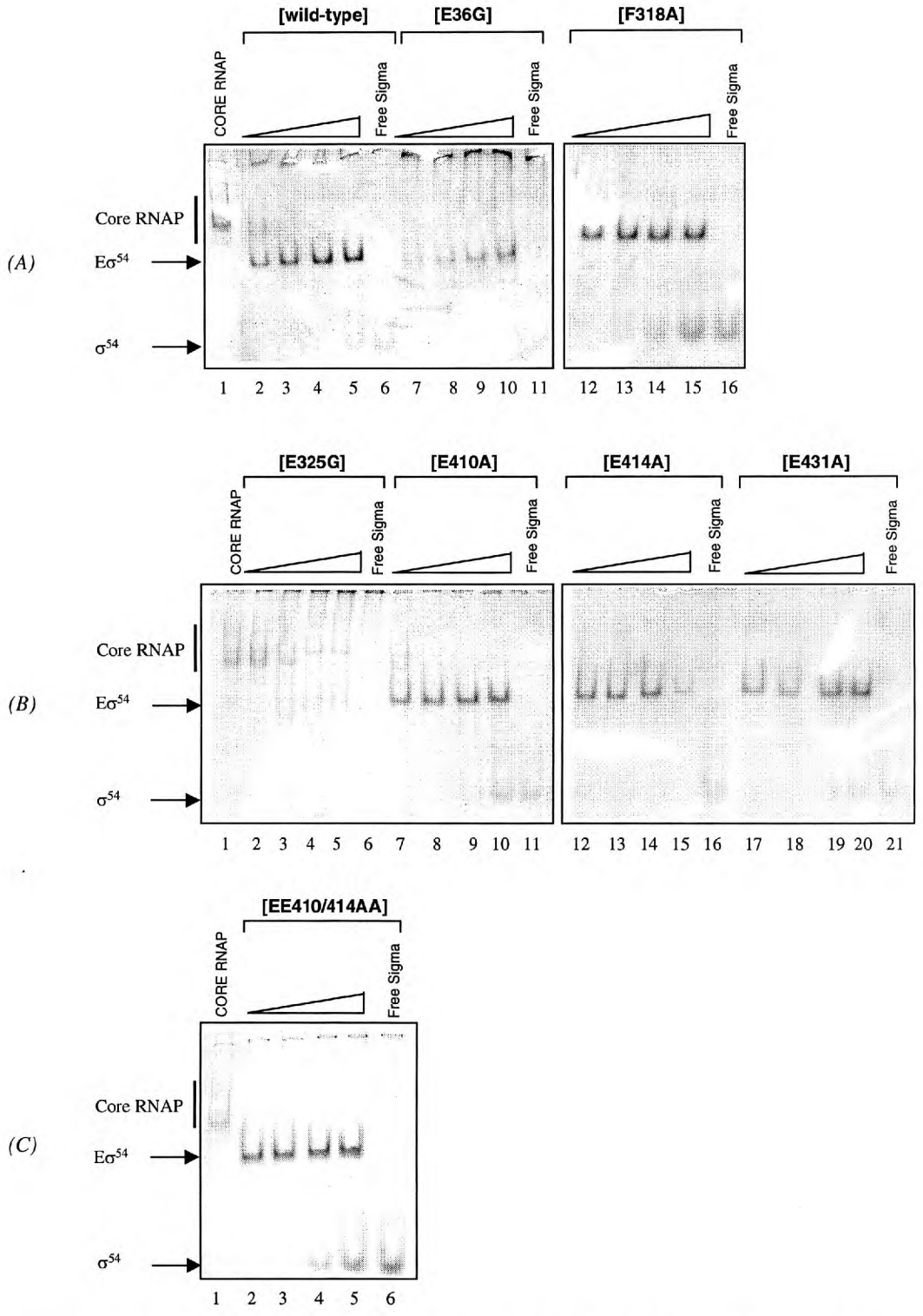


Figure 7D. Holoenzyme gel assembly assay with σ^{54} mutants. A constant amount of core RNAP (250nM) was combined with increasing amounts (at molar ratios of 1:1, 1:2, 1:4 and 1:8) of σ^{54}. As indicated, the assay depends upon the native mobility difference between core RNAP alone,

holoenzyme ($E\sigma^{54}$) complexes and free σ^{54}. When increasing amounts of σ^{54} are added to core RNAP, the bands corresponding to the core RNAP diminish and complexes with higher mobility corresponding to $E\sigma^{54}$ form. The gel was stained with Coomassie Blue stain.

7.4 Interaction of Mutant σ^{54} Proteins with Promoter DNA

To explore the promoter DNA binding properties of the σ^{54} mutants, we conducted DNA binding assays. In addition to the double stranded homoduplex *S. meliloti nifH* DNA we used different DNA templates to represent DNA structures in which the nucleation of DNA melting has occurred in the closed complex (early melted DNA) or where the extended melting in the open complex is evident (late melted DNA), (see Chapter 5).

Homoduplex DNA Binding. Using a native gel mobility shift assay, we examined binding of the mutant σ^{54} proteins to double stranded promoter DNA.

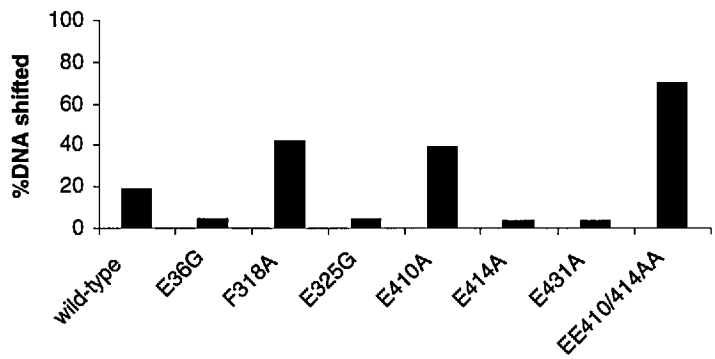


Figure 7E. Binding of σ^{54} mutants (1μM) to the *S. meliloti nifH* homoduplex DNA probe (16nM).

Results in Figure 7E show that E36G, E325G, E414A and E431A have at least a 4-fold lower binding activity to the homoduplex promoter DNA compared to the wild-type σ^{54}. Residues E325, E414 and E431 are within the minimal DNA binding domain in σ^{54} (residues 329-477) and E414 and E431 are located between the putative helix-turn-helix motif (residues 367-386) and the RpoN box (residues 454-463) (Taylor *et al.*, 1996; Cannon *et al.*, 1995; Merrick & Chambers, 1992). Residue 36 tethered FeBABE

footprinting of promoter DNA showed that residue 36 is close to the consensus GC promoter element. Therefore, given that residues E414 and E431 are also protected in closed complexes by DNA from protease cleavage (Casaz & Buck, 1997), it is possible that glutamate residues at positions 36, 325, 414 and 431 make sequence specific contact with promoter DNA. Strikingly, mutating E414 to A in the E410 mutant apparently rescues homoduplex DNA binding (Figure 7E). Further, the binding assays reveal that the double mutant EE414/410AA has a 4-fold increased binding activity than the wild-type protein to homoduplex promoter DNA. Strikingly, an increased binding activity is also apparent with the F318A and E410A mutant proteins. σ^{54} titration assays show that EE410/414AA, and to a lesser extent F318A and E410A, saturate homoduplex promoter DNA binding at a much earlier point in the binding curve when compared to the wild-type protein (Figure 7F).

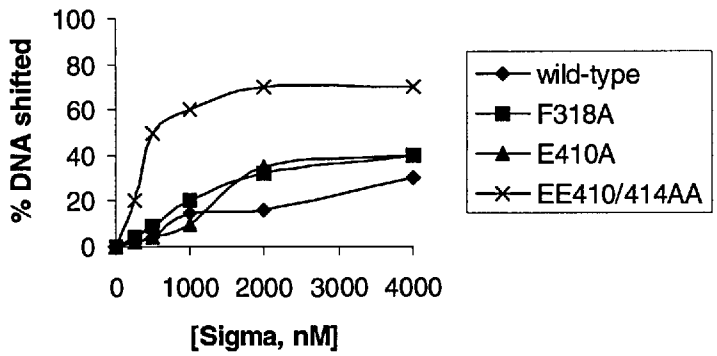


Figure 7F. Binding of mutant σ^{54} proteins to the *S. meliloti nifH* homoduplex promoter DNA probe. The promoter DNA (16nM(was incubated with increasing concentrations of σ^{54} as indicated in the figure.

Thus, it appears that mutating surface exposed residues F318, E410A and E414 (only in combination with E410A) unmasks extra DNA binding activity in σ^{54}. DNA binding assays with Region I deleted σ^{54} ($I\Delta\sigma^{54}$) revealed that this mutant σ^{54} had increased binding to promoter DNA than the wild-type σ^{54} (Cannon *et al.*, 1999). In the recent view that amino terminal Region I and carboxyl terminal Region III interact in σ^{54} (Wigneshweraraj *et al.*, 2001) it is possible that mutating surface exposed residues in

Region III of σ^{54} repositions Region I in such a way that extra DNA binding activity is revealed. However, whether this extra binding is promoter specific interaction of σ^{54} with DNA remains to be elucidated.

Early Melted DNA Binding. The early melted DNA probe represents the locally distorted structure adopted by the promoter DNA when σ^{54}-holoenzyme binds to form closed complexes. This distortion downstream of the consensus GC is believed to be the nucleation of DNA melting and involves determinants in Region I and Region III of σ^{54} suggesting σ^{54}-DNA interactions in DNA melting (Morris *et al.*, 1994). The formation and recognition of the distorted DNA structure next to the consensus GC element is possibly a signature for regulated (activator dependent) transcription initiation by the σ^{54}-holoenzyme and appears to be closely associated with the tight binding of σ^{54} to the early melted DNA structure (Gallegos & Buck, Guo *et al.*, 1999). σ^{54} is a DNA binding protein that occupies certain consensus promoters in the absence of the core RNAP, although higher concentrations of σ^{54} are required for DNA binding (Buck & Cannon, 1992). The wild-type σ^{54} protein shows a 3-fold higher affinity for the early melted DNA probe than for the homoduplex, binding about 60% of the early melted DNA probe at 1μM (Figure 7G) (Gallegos & Buck, 2000; Cannon *et al.*, 1999).

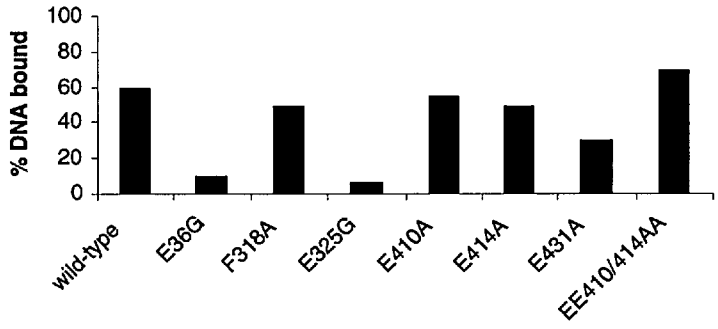


Figure 7G. Binding of σ^{54} mutants (1μM) to the *S. meliloti nifH* early melted DNA probe (16nM).

F318A, E410A and EE410/414AA mutant σ^{54} proteins, which showed higher binding to the homoduplex DNA probe than the wild-type σ^{54}, apparently bind the early melted DNA

probe with near wild-type activity (Figure 7G). The σ^{54} mutants (E36G, E325G and E431A) that were very defective for homoduplex DNA binding (Figure 7E) show slightly increased binding to the early melted probe (Figure 7G). Strikingly, we note that E414A, which bound the homoduplex promoter poorly, has a near wild-type level of binding activity to the early melted DNA probe. Overall, our early melted DNA binding assays with the σ^{54} mutants suggest that the Region I residue E36 and Region III residue E325 in σ^{54} are important for binding to both homoduplex and early melted DNA probes. The enhanced (with respect to wild-type σ^{54}) homoduplex DNA binding by F318A, E410A and EE410/414AA is not evident on the early melted DNA probe even at lower protein concentrations (data not shown). This suggests that determinants of the 6-7 fold increased binding activity of σ^{54} to the early melted promoter probe dominates binding patterns (Cannon *et al.*, 2000; Gallegos & Buck, 2000). Residue E414 appears important for binding to homoduplex DNA (Figure 7E), however redundant for early melted DNA binding (Figure 7G)

σ^{54}- isomerisation on early melted DNA probe. σ^{54} bound to the early melted DNA probe responds to activator in an nucleotide hydrolysis dependent manner to form a new supershifted DNA complex in which σ^{54} has isomerised (Cannon *et al.*, 2000). The isomerised complex is characterised by an extended DNase I footprint in the direction of transcription (by about five base pairs) (Cannon *et al.*, 2000). With the σ^{54} mutant and the early melted DNA, we attempted to form the isomerised σ^{54}-DNA (ssσ-DNA) species using the activator protein PspFΔHTH and hydrolysable nucleotide dGTP (Figure 7H).

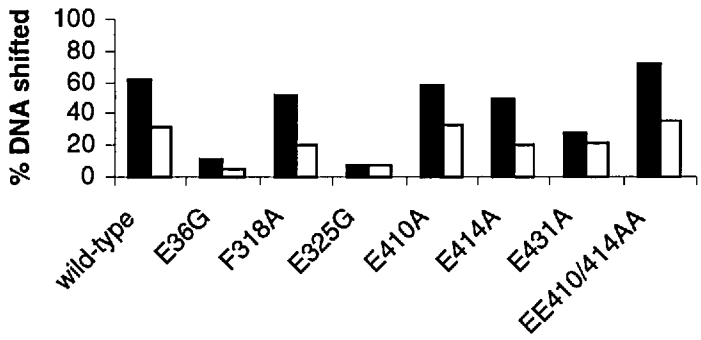


Figure 7H. Isomerisation of σ^{54} mutants bound to the early melted DNA probe (black bars) to form the supershifted complex (ssσ - see text for details) (white bars) in response to activation and nucleotide hydrolysis.

For each mutant, a comparison of the fraction of initially bound DNA converted (in an activator dependent and nucleotide-requiring reaction) to the new slower running ssσ-DNA species was made. As shown (Table 7.2), the efficiency of conversion to the isomerised state was 34-66%, compared to ~60% for the wild-type, suggesting that the mutant σ^{54} proteins are largely functional for isomerisation. Hence, residues E36 in Region I and Region III residues F318, E325, E410, E414 and E431 seem dispensable for σ^{54} interaction with activator, at least within the scope of the isomerisation assay.

Protein	Binding to Early Melted DNA Probe (%)	Isomerisation (%)
wild-type	62	59
E36G	12	34
F318A	52	37
E325G	5	52
E410A	58	53
E414A	50	53
E431A	28	66
EE410/414AA	72	53

Table 7.2. Supershift assay with mutant σ^{54} proteins. Gel shift assays were conducted with 1μM σ^{54} proteins and 16nM DNA, bound and unbound DNA were separated and the percentage of bound DNA quantified by phosphorimager analysis. Data for the early melted DNA binding come from Figure 7H. Isomerisation is calculated as a percentage of initially bound DNA converted in a

activator and nucleotide requiring reaction, to the supershifted DNA complex (ssσ-DNA complex). Note that isomerisation is not a measure of total isomerised σ^{54} complex formed.

Late melted DNA binding. The late melted DNA structure (heteroduplex from –10 to –1) includes the non-conserved sequence from –10 to –5 that interacts with σ^{54} within the open complex (Cannon *et al.*, 1999; Cannon *et al.*, 1995). σ^{54} binding to the late melted DNA probe is thought to reflect σ^{54}-promoter DNA interactions that occur in the open complex.

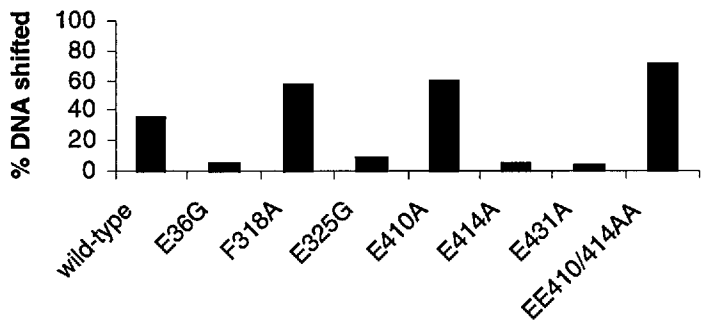


Figure 7I. Binding of σ^{54} mutants (1μM) to the *S. meliloti nifH* late melted DNA probe (16nM).

Figure 7I shows that the mutant σ^{54} proteins bind the late melted DNA probe with a pattern similar to their binding to the homoduplex DNA (Figure 7E). Consistent with homoduplex DNA binding, F318A, E410A and EE410/414AA have 1.5-2 fold increased binding to the late melted DNA probe. The mutants that were defective for homoduplex DNA binding (E36G, E325G, E414A and E431A) show an equal (E36G) or more severe defect with respect to the late melted DNA probe.

7.5 Interaction of Mutant Holoenzymes with Promoter DNA

Homoduplex DNA binding. We compared the DNA binding activities of holoenzymes (formed with 1:2 ratio of core RNAP to σ^{54}) for the homoduplex promoter

probe. E36G and E325G had reduced affinity to the core RNAP (Figure 7D). This might be offered as the reason why the holoenzymes formed with these proteins are defective for binding to the homoduplex promoter probe (Figure 7J).

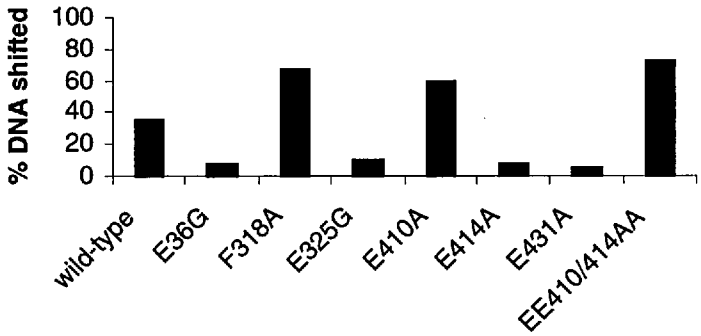


Figure 7J. Mutant σ^{54}-holoenzyme (100nM) binding to the *S. meliloti nifH* homoduplex promoter DNA probe (16nM).

However, holoenzymes formed with E36G and E325G using a 1:4 or 1:6 ratio of core RNAP to mutant σ^{54} also failed to bind the homoduplex promoter (data not shown). Moreover, the defects shown by the E36G, E325G, E414A and E431A holoenzymes, as well as the increased binding using the F318A, E410A and double mutant holoenzymes (Figure 7J) are quiet similar in degree to those seen in the absence of the core (Figure 7E). This indicates that the promoter affinity of the holoenzyme is contributed by the associated sigma subunit. The absence of evidence of an additional defect attributable to core binding problems in the E36G and E325G cases might, if significant, imply that the presence of promoter DNA assists core binding by these proteins.

Early melted DNA binding. The wild-type holoenzyme binds to the early melted DNA probe to form an heparin resistant complex independent of activation and associated nucleotide hydrolysis (Cannon *et al.*, 1999). Heparin resistance is a property of open complexes that form with the σ^{54}-holoenzyme in a reaction normally requiring activator and nucleotide hydrolysis (Wedel & Kustu, 1995). Figure 7K shows that holoenzymes formed with the mutant σ^{54} proteins E36G, E410A, E414, E431A and EE410/414AA form heparin stable complexes on the early melted DNA probe suggesting that determinants required for the acquisition of heparin stability are largely intact in the mutant σ^{54} proteins.

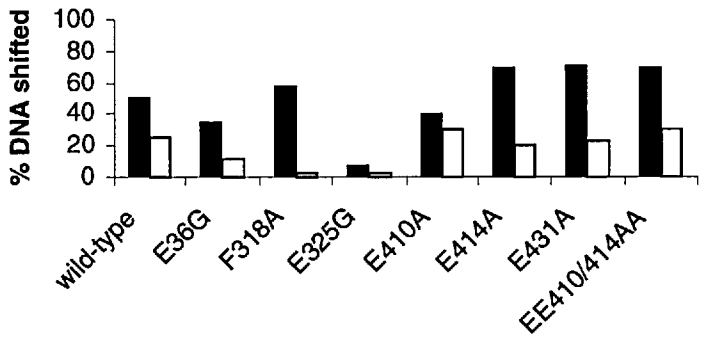


Figure 7K. Mutant σ^{54}-holoenzyme (100nM) binding to the *S. meliloti nifH* early melted promoter DNA probe (16nM) in the absence (black bars) and presence (white bars) of 100μg/ml of heparin.

Region III residues E414 and E431 appear not essential for holoenzyme binding to the early melted DNA probe, consistent with previous data (see above) in which free E414A and E431A σ^{54} mutants bind poorly to the homoduplex promoter, they bind considerably better (relative to wild-type) to the early melted DNA probe (compare Figures 7K, 7E and 7G). Holoenzymes containing the E325G mutant σ^{54} are defective for binding to the early melted DNA probe, a defect probably associated with the promoter DNA (Figure 7G) and/or core RNAP binding (Figure 7D) of the E325G mutant. Strikingly, even though the holoenzyme containing the F318A mutant σ^{54} binds the early melted DNA probe with wild-type levels of activity, it fails to form heparin stable complexes on the early melted DNA probe (Figure 7K). Previous work has shown that the formation of heparin unstable complexes on early melted DNA correlated strongly with poor binding of the mutant σ^{54} to the early melted DNA probe. However, as shown in Figure 7G, the F318A mutant σ^{54} apparently binds the early melted DNA probe with near wild-type levels of affinity and responds to activator in an nucleotide-dependent reaction to form the supershifted σ^{54}-DNA complex (ssσ). Heparin resistance was not acquired by the F318A holoenzyme on the early melted DNA when activator and hydrolysable nucleotide were present (data not shown). Since the σ^{54}-core RNAP complex is very heparin sensitive, the heparin instability of the F318A-holoenzyme on early melted DNA probe may involve altered σ^{54}-core RNAP interactions (Gallegos & Buck, 1999). Consistent with this view,

F318A-early melted DNA complexes were detected in the heparin challenge assay, hence reflecting the dissociation of core RNAP from F318A σ^{54} (data not shown). Thus, it seems that the F318A holoenzyme is easily disrupted by heparin. The surface exposed residue F318 is close to a region (residue 325-334) protected by protease attack by the core RNAP and hence may be involved in interactions with the core RNAP. Given that the F318A mutant σ^{54} binds to core RNAP with wild-type levels of affinity in the absence of DNA, it appears that the defects in F318A may reflect problems with changing interactions with core RNAP when particular DNA probes are bound by sigma.

Late melted DNA binding. Efficient formation of heparin stable complexes on the late melted DNA probe requires activator and a hydrolysable nucleotide if the σ^{54}-holoenzyme is used, indicating that a conformation change in the holoenzyme must occur to allow heparin stable interactions with late melted DNA (Cannon *et al.*, 1999; Wedel & Kustu, 1995). Initially, we measured the activator independent binding of the mutant holoenzymes to the late melted DNA probe. Results in Figure 7L (black bars) show that all but one of the holoenzymes of all the mutant σ^{54} proteins bind the late melted DNA probe with similar to wild-type levels of affinity.

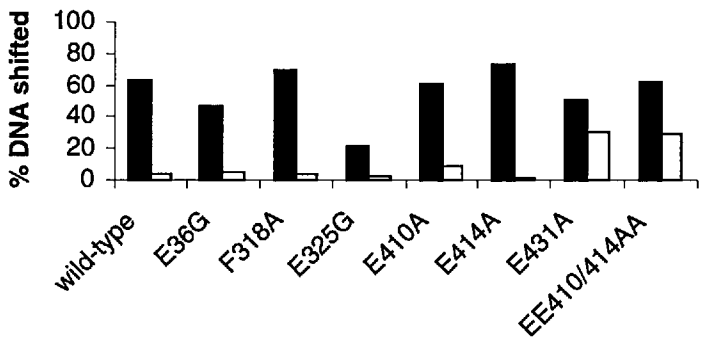


Figure 7L. Activator-independent binding of mutant σ^{54}-holoenzymes (100nM) to the *S. meliloti nifH* late melted promoter DNA probe (16nM) in the absence (black bars) and presence (white bars) of 100μg/ml of heparin.

The reduced binding activity of E325G-holoenzyme is consistent with its defects in binding to core and to the other DNA probes testes (above) . However, the poor homoduplex binding activity of the E414A and E431A holoenzymes (Figure 7J) is not echoed when they bind to late melted DNA. Taken together with their strong binding to early melted probe, this suggests that residues E414 and E431 are important in the initial stages of closed complex formation, but not later. Next, we assayed for heparin stable complex formation on the late melted DNA probe with the wild-type and the mutant σ^{54} holoenzymes. As shown in Figure 7L (white bars), only the mutant holoenzymes harbouring E431A and EE410/414AA σ^{54} formed activator independent heparin resistance complexes with the late melted DNA probe. Heparin resistance, independent of activation, is normally a property of mutant σ^{54}s assembling holoenzymes deregulated for activation (activator bypass mutant σ^{54}) (Gallegos & Buck, 2000). It appears that substitution at residues E431 and EE410/414AA confer upon the holoenzyme an ability to interact with the late melted DNA in an activator independent manner to give heparin resistant complexes. It is thus possible that the mutant holoenzymes (E431A and EE410/E414AA holoenzymes) and the wild-type σ^{54} holoenzyme represent two distinct functional states. Interestingly, we also note that the holoenzymes containing single mutant σ^{54} proteins only harbouring the E410A or E414A mutation are not able to form heparin stable complexes on the late melted DNA probe independent of activation (Figure 7L). Thus, it appears that both E residues at 410 and 414 have to be replaced by alanine to confer heparin resistance to the holoenzyme, perhaps by reducing the overall negative charge of the σ^{54} surfaces in the holoenzyme required for DNA interaction and increased binding to the negatively charged late melted DNA probe occurs. Further, some σ^{54} DNA-binding domain mutant holoenzymes have been described that do show a strong property of heparin stable interaction with the late melted DNA probe independent of activator and associated nucleotide hydrolysis (Chaney *et al.*, 2000; Chaney & Buck, 1999), but that are in the DNA crosslinking patch (residues 329-346) of σ^{54}. Overall, the activator independent and heparin resistant complex formation by E431A and EE410/414AA holoenzymes predicts activator bypass transcription by the holoenzymes harbouring these mutant σ^{54} proteins.

Next, we measured whether the number of heparin stable complexes formed with the late melted DNA could be increased by activation. Results in Figure 7M show an activator-dependent increase in the number of heparin stable holoenzyme complexes forming with the late melted DNA probe, suggesting that all σ^{54} mutant holoenzymes are functional in responding to the activator protein. We note that the mutant holoenzymes containing the E431A and EE410/414AA respond to activation as evidenced by the increased number of heparin stable complexes formed in the activation assays (compare Figure 7L and 7M).

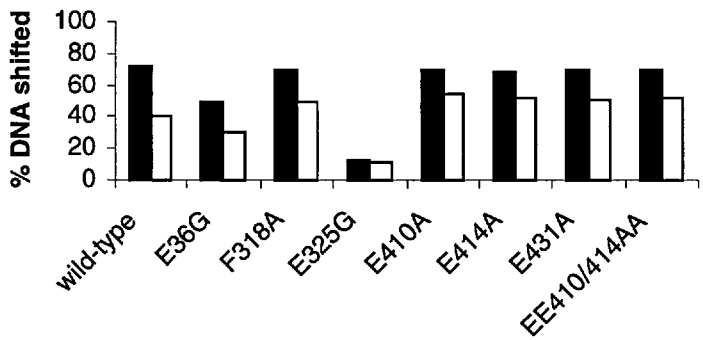


Figure 7M. Binding of activated mutant σ^{54}-holoenzymes (100nM) to the *S. meliloti nifH* late melted promoter DNA probe (16nM) in the absence (black bars) and presence (white bars) of 100μg/ml of heparin.

7.6 *In vitro* Transcription Assays

Activator-dependent assays. To correlate the *in vitro* DNA binding data, *in vitro* transcription assays were conducted from a supercoiled *S. meliloti nifH* promoter template (pMKC28, Chaney & Buck, 1999). Holoenzymes and template DNA were incubated with ATP to allow initiated complexes to form. Stimulation of RNA synthesis by activator protein PspFΔHTH was assayed. Single-round RNA synthesis from stable initiated complexes was then allowed to proceed by the addition of the remaining three NTPs together with heparin. As shown in Figure 7N all mutant holoenzymes were active for *in vitro* transcription (40-70% wild-type holoenzyme activity). The reduced activity (39%

wild-type holoenzyme activity) of the holoenzyme containing the E325G mutant σ^{54} may be due to the core binding and/or promoter binding defects of this holoenzyme. Consistent with this is the observation, that the E325G holoenzyme only formed 10-20% activator dependent heparin stable complexes on the late melted DNA probe (Figure 7M). Holoenzymes containing the E410A, E414A, E431A and EE410/414AA mutant σ^{54} proteins formed more activator dependent heparin stable complexes on the late melted promoter probe compared to the wild-type holoenzyme (Figure 7M). Interestingly, this pattern of activity is not reflected in the *in vitro* transcription assays conducted with 100nM supercoiled DNA. Possibly, the holoenzymes containing the E410A, E414A, E431A and EE410/414AA mutant σ^{54} proteins are defective in some later stages in the transcription process and that the E410A, E414A, E431A and EE410/414AA mutants markedly impair the holoenzyme in steps leading to changes in RNAP necessary for the productive interactions necessary for efficient transcript formation. Additional *in vitro* transcription assays with non-saturating levels of holoenzyme and under conditions that facilitate open complex formation (i.e. higher temperatures) on the *S. meliloti nifH* and other σ^{54}-dependent promoters will elucidate more on the *in vitro* transcription activity of these mutants. Figure 7N also makes it clear that transcription is greatly stimulated by activation, in every case. Thus, the expectation of activator-independent transcription by the E431A and EE410/414AA enzymes has not been confirmed.

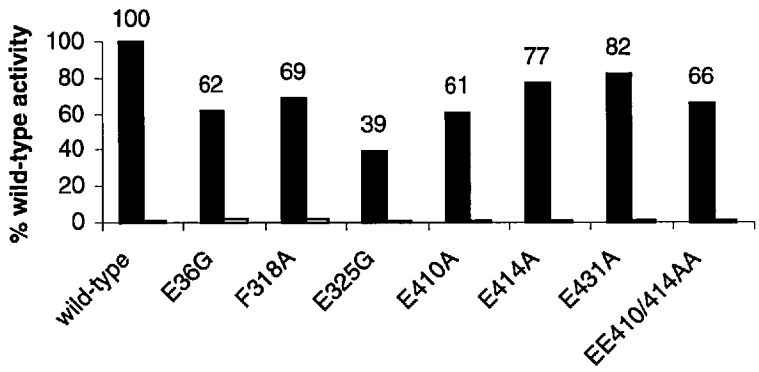


Figure 7N. PspFΔHTH activated (black bars) and inactivated (white bars) *in vitro* transcription on supercoiled *S. meliloti nifH* promoter template.

Activator-independent assays. Preliminary activator independent transcription assays reveal that E36G and F318A are active for some activator independent (bypass) transcription initiation *in vitro* (Figure 7O, *lanes 2 & 3*, respectively). Region I in σ^{54} has a negative function in disallowing transcription initiation by the holoenzyme, demonstrated by Region I mutants that allow increased activator-independent transcription *in vitro* (Casaz *et al.*, 1999; Cannon *et al.*, 1999; Wang *et al.*, 1997). Thus, it is possible that a glycine (this work) or cysteine residue (Wigneshweraraj *et al.*, 2000) at position 36 allows Region I to lose an interaction that allows productive transcription initiation *in vitro* independent of activation. Region III mutants R336A, R336C, K338A and R342A (Wang & Gralla, 2000; Chaney *et al.*, 2000; Wigneshweraraj *et al.*, 2000; Chaney & Buck, 1999) exist that confer their holoenzymes with the activator-independent phenotype, suggesting Region I may function in part through Region III. In the recent view that amino terminal Region I and carboxyl terminal Region III interact in σ^{54} (Wigneshweraraj *et al.*, 2001 & chapter 5) it is possible that mutating surface exposed aromatic residues in Region III of σ^{54}, such as R336A (Chaney *et al.*, 1999), R336C (Wigneshweraraj *et al.*, 2000), R342A (Chaney *et al.*, 2000) and F318A (this work) "repositions" Region I to permit activator independent initiation of transcription *in vitro*.

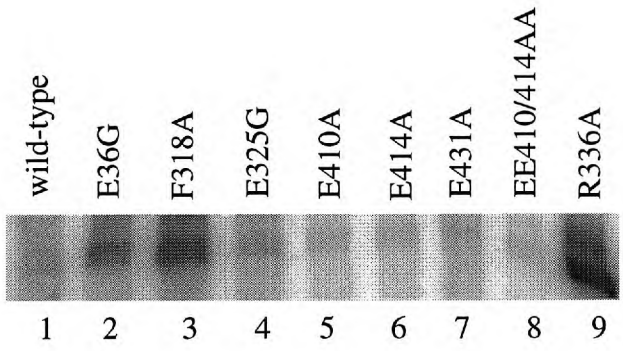


Figure 7O. Activator-independent *in vitro* transcription assay at the supercoiled *S. meliloti nifH* promoter

7.7 Discussion

The σ^{54} holoenzyme is distinctive in its ability to remain transcriptionally silent. Productive transcription initiation by the σ^{54}-holoenzyme only occurs in response to an interaction of the σ^{54}-holoenzyme with an enhancer binding activator protein in an nucleotide hydrolysis dependent manner. Conformational adjustments must occur in the σ^{54}-holoenzyme closed complex to relieve the inhibited state, to allow RNAP isomerisation in response to activation and stable DNA melting. σ^{54} undergoes specific alterations in protease sensitivity at different steps on the pathway leading from free σ^{54} to one engaged in a transcription complex (Casaz & Buck, 1997). Surface exposed residues in Region III in σ^{54} that were protected from protease attack by the core RNAP in the holoenzyme or by promoter DNA in the closed complex were substituted with alanine (F318, E410, E414, E431) or glycine (E325) to seek a correlation between the protein footprinting and the functional importance of the amino acid. Region I residue 36 that becomes sensitive to protease attack in initiated complexes was also substituted with glycine. The *in vivo* and *in vitro* activities of the altered σ^{54} proteins were then studied. The work should also help indentify new potentially significant surface exposed residues for FeBABE footprinting work.

Residue E36. FeBABE tethered footprinting experiments revealed that residue 36 locates close to the catalytic centre of the core RNAP in the holoenzyme and the GC promoter region in closed promoter complexes (Wigneshweraraj *et al.* 2000; Wigneshweraraj *et al.*, 2001). Thus, residue 36 was suggested to be part of the regulatory centre in the σ^{54}-holoenzyme closed complex that possibly contains a target for the activator protein (Wigneshweraraj *et al.*, 2001). Residue 36 is located in Region I of σ^{54} which has a prime function in activator response (Casaz *et al.*, 1999; Cannon *et al.*, 1999; Wang *et al.*, 1995), but is dispensable for core RNAP and DNA binding. Region I sequences have a low affinity for the core RNAP (Gallegos & Buck, 1999) and a Region I deleted or some Region I mutant σ^{54} proteins show activator-independent transcription activity *in vitro* (Gallegos *et al.*, 1999; Wang *et al.*, 1995). Here the *in vivo* and *in vitro* activity of σ^{54} protein containing the E36G mutation was examined. A glycine substitution likely confers increased conformational flexibility of Region I around position 36. Thus, the massively reduced ability of the holoenzyme containing the E36G mutant to initiate transcription *in vivo*, but to 62% of wild-type activity *in vitro* could be due to the altered conformation of the E36G mutant σ^{54} in the holoenzyme. In the *in vivo* assays, activation is dependent on the enhancer-bound NifA protein, whereas *in vitro* it is dependent on PspFΔHTH acting from solution. Thus, the DNA bound NifA protein could be restricted in accessing the mutant activation target in the E36G mutant holoenzyme, than its counterpart PspFΔHTH which activates unrestricted from solution. Equally, the reduced affinity to the core RNAP and the apparent DNA binding defects of the purified E36G mutant could contribute to holoenzyme and/or closed complex stability *in vivo*. The E36G mutation also confers a partially activator-independent initiation (bypass) phenotype on the holoenzyme, like E36C (Chapter 3) and other Region I mutations already mentioned. The E36C mutation is less severe in its effects than E36G; in particular the E36C proteins shows no core-binding defefect, and in holoenzyme form it retains 20% of the wild-type transcription activity *in vivo*. Overall, the phenotype of the E36G mutant argues for a significant role of this residue in Region I conformation and its interaction with core RNAP and promoter DNA as indicated by residue 36 tethered FeBABE footprinting studies (Wigneshweraraj *et al.*, 2000; Wigneshweraraj *et al.*, 2001).

Residue F318. Even though this residue lies outside the minimal DNA binding domain (residues 329-477), an alanine substitution at F318 (F318A) has significant consequence on the DNA binding function of σ^{54}. The F318A mutant σ^{54} shows an increased binding activity to promoter DNA compared to the wild-type protein. It appears that mutating F318 has unmasked extra DNA binding in σ^{54}. More interestingly, is the preliminary observation that holoenzymes containing the F318A mutant σ^{54} are active for activator independent transcription initiation. A function in activator responsiveness of aromatic amino residues in Region III of σ^{54} is not uncommon (Wang & Gralla, 2000; Chaney *et al.*, 2000; Chaney and Buck, 1999) and with the associated involvement of aromatic residues in single-stranded DNA interactions (Parkinson *et al.*, 1996; Bocharev *et al.*, 1997), it is possible that the F318 protein contributes to single-stranded DNA binding required for regulated (activated) transcription initiation. Thus, the surface exposed residue F318 presents a suitable target for FeBABE conjugation studies and Cys318 tethered FeBABE footprinting will reveal positions on the promoter DNA proximal to F318 in σ^{54}-DNA and within the closed and open promoter complexes. The data arising from this work suggest that residue F318 is part of the regulatory centre in σ^{54} that comprises the –12 promoter region, the UV-crosslinking patch and Region I (Wigneshweraraj *et al.*, 2001). Clearly, there are differences between the F318A holoenzyme and holoenzyme containing the R336A activator-bypass mutant σ^{54} (Chaney *et al.*, 1999; Wigneshweraraj *et al.*, 2001). In contrast to the R336A holoenzyme, the F318A σ^{54} binds early melted DNA with wild-type levels of activity and its holoenzyme is sensitive to heparin on the late melted DNA probe in the absence of activation. Further functional assays are underway to dissect the differences between the carboxyl terminal mutants F318A and R336A. In light of recent evidence (Wang & Gralla, 2000), it appears that F318A represents a different class of activator bypass σ^{54} mutant, contrasting the R336A and K338A.

Residue E325. The mutant σ^{54} protein harbouring the E325G mutation is very defective for holoenzyme formation and DNA binding. This phenotype is reflected in the reduced *in vitro* transcription activity of the holoenzyme harbouring the E325G mutant σ^{54}. E325 is positioned outside the minimal core binding sequence in σ^{54} (residues 120-215).

Previously, Cys336 and Cys346 conjugated FeBABE footprinting work (Wigneshweraraj *et al.*, 2000; Wigneshweraraj *et al.*, 2001) indicate that E325 is probably located close to the core RNAP subunits and proximal to positions downstream of the GC-region on promoter DNA.

Residues E410, E414 and E431. These residues were protected from protease attack in the closed complex. These closed complex-specific protections can also be interpreted as conformational changes, but because the protected residues lie within the DNA-binding domain, it is also possible that protection is probably due to their proximity to promoter DNA (Casaz and Buck, 1997). Further, the phenotypes of mutant σ^{54} proteins with alanine substitutions at E410, E414 and E431, respectively, suggest a direct role of these residues in promoter DNA interaction. Immunoblots showed that the mutants E414A and EE410/414AA were not stably produced *in vivo*, whereas the E410A had a near wild-type level of stability, suggesting that E414 probably has a structural role in σ^{54} . In contrast, even though E431A mutant was stably expressed, it was inactive for promoter activation *in vivo*. The reduced double-stranded DNA binding activity of the E431A mutant σ^{54} and its holoenzyme could be a reason for the lack of *in vivo* activity. Strikingly, like the F318A mutant, the E410A and EE410/414AA mutants showed increased DNA binding activity with respect to the wild-type protein. Binding assays with Region I deleted σ^{54} showed that the Region I deleted protein had an increased DNA binding activity compared to the full-length wild-type protein (Cannon *et al.*, 1999; Wigneshweraraj *et al.*, 2001). Therefore, it is possible that mutating surface exposed glutamate (and phenylalanine) residues in Region III of σ^{54} positions Region I in such a way that extra DNA binding activity is unmasked – especially in the context of the recent view that Region I and Region III interact in σ^{54} (Wigneshweraraj *et al.*, 2001). However, the increased promoter DNA binding activity of the mutants (E410A, EE410/414A and F318A) compared to the wild-type protein is not reflected in *in vitro* transcription activity. This implies that the mutant holoenzymes are defect in one or many steps after closed complex formation. Also, the increased DNA binding activity seen with the E410A EE410/414AA and F318A mutant σ^{54} proteins may impair efficient promoter clearance by the mutant holoenzymes once transcription is initiated. Interestingly, in an analysis of

amino acid-DNA interactions in crystal structures, it appears that glutamate residues have a distinct preference for cytosine residues (Mandel-Gutfreund *et al.*, 1995). Given that the σ^{54} binding site consensus is GC-rich, TGGCA-N_5-TTGC (Barrios *et al.*, 1999), it may be possible that, glutamate residues that are protected from protease attack in the closed complex (E410, E414 and E431) as well as the core RNAP protected E325 make specific contacts with promoter cytosine residues. FeBABE tethered to positions 325, 410, 414 and 431 will provide more insight into the roles of these residues in DNA binding.

7.8 Conclusions

We have characterised the phenotypes of mutations affecting surface exposed residues in σ^{54} that were protected against protease attack either by the core RNAP (E325) or by their proximity to promoter DNA (F318, E410, E414, E431) in the closed complex. The results presented here argue for a positive role of these residues in σ^{54} function. These residues provide a suitable conjugation site for FeBABE and the resulting footprinting studies (on both, core RNAP and DNA) will (1) reveal whether these residues are part of the regulatory centre in σ^{54} that has been shown to include Region I, the X-link patch, core RNAP catalytic centre and the promoter GC element (Wigneshweraraj *et al.*, 2001), and (2) will help further elucidate the orientation of σ^{54} on DNA in closed and open promoter complexes (Wigneshweraraj *et al.*, 2001).

CHAPTER EIGHT

Low Resolution Structure of Sigma 54 Revealed by Small Angle X-ray Scattering

8.1 Introduction

σ^{54} has a modular domain organisation (Cannon *et al.*, 1997 & 1995; Wong *et al.*, 1994). Deletion analysis of the σ^{54} protein has shown σ^{54} comprises of distinct biochemically active parts which can be expressed as separate, functionally active fragments (Gallegos & Buck, 1999; Cannon *et al.*, 1997, 1995). The promoter DNA binding domain is located in the carboxyl terminal part of the protein (Region III) and the amino terminal domain (Region I) is predominantly involved in activation. The central domain (Region II and amino terminal parts of Region III) contains the major core RNAP binding determinants.

The σ^{54}-holoenzyme binds to promoters to form a closed complex that is silent for transcription initiation. Conversion of the closed complex to a transcriptionally productive open promoter complex is dependent upon activation by enhancer binding activator proteins in a nucleotide hydrolysis dependent manner and involves an as yet poorly defined interaction of the activator with the σ^{54}-holoenzyme in the closed complex. To date, an extensive body of evidence suggests that σ^{54} is a target for the activator and Region I plays a role in the activator responsiveness of the σ^{54}-holoenzyme (Cannon *et al.*, 2000; Casaz *et al.*, 1999; Casaz & Buck, 1997, 1999; Hsieh & Gralla, 1994; Hsieh *et al.*, 1994; Wong *et al.*, 1994). However, other studies have shown Region I to constitute a part of the σ^{54} interface with core RNAP (Wigneshweraraj *et al.*, 2000; Gallegos *et al.*, 1999) and Region I to be proximal to the consensus GC promoter element of σ^{54}-dependent promoters (Wigneshweraraj *et al.*, 2001). Further, one of the main findings of the work presented in this thesis is that σ^{54} Regions I and III interact to form a regulatory centre in the closed

promoter complex, which includes the consensus GC promoter element located –12 base pairs from the transcription start site (Wigneshweraraj *et al.*, 2001). This regulatory centre possibly constitutes a target with which the activator protein has to interact with to initiate open complex formation in an nucleotide hydrolysis dependent manner.

Clearly, σ^{54} contributes to several functions of the holoenzyme required for enhancer dependent transcription. Although discrete activities are known to reside within individual domains in σ^{54}, the domains appear to interact with each other for the full function of the σ^{54}. The availability of structural information on the tertiary domain organisation of σ^{54} will not only confirm or disprove our previous suggestions but will provide us with a valuable tool for interpreting FeBABE footprinting and mutational analysis data. The objectives of the present work are to evaluate the envelope shape and tertiary domain organisation of the *K. pneumoniae* σ^{54} by using small angle x-ray solution scattering (SAXS) – a method yielding low resolution structural information at nearly physiological conditions. For this purpose, the σ^{54} fragments comprising residues 70-324, 1-306 and 117-477 were used as amino terminal, internal and carboxyl terminal containing fragments representing various combinations of domains (Figure 8A). SAXS measurements on the σ^{54} proteins were done in collaboration with Dr Dmitri Svergun at the European Molecular Biology Laboratories (EMBL) at Hamburg (Germany) using the synchrotron radiation facilities of the Deutsches Elekronen Synchrotron (DESY).

8.2 Small Angle X-ray Scattering (SAXS)

Structure determination of biological macromolecules using high resolution methods such as macromolecular crystallography and nuclear magnetic resonance (NMR) requires a specific set of conditions: it is often difficult to grow crystals of high molecular weight (MW) assemblies that are suitable for diffraction, and the application of NMR is fundamentally limited to small (MW < 30kDa) monodispersed proteins. In contrast, SAXS

as a method for structure determination of biological macromolecules can yield low-resolution information in a broad range of conditions and particle sizes.

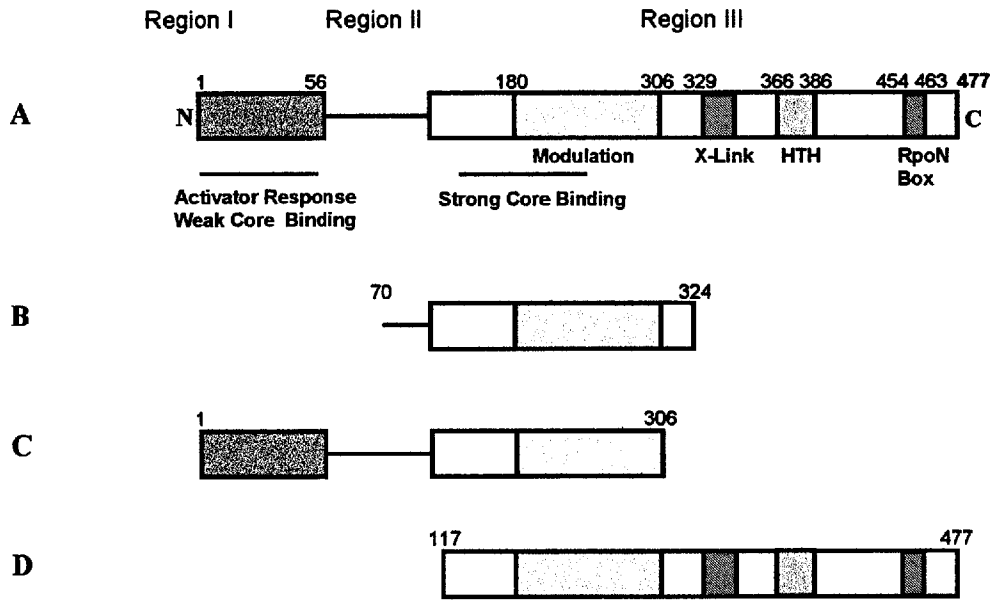


Figure 8A. Schematic diagram of σ^{54} fragments used for the small angle X-ray scattering measurements: full-length σ^{54} (**A**) and fragments consisting of amino acid residues 70-324 (**B**), 1-306 (**C**) and 117-477 (**D**).

SAXS is a useful technique for obtaining structural information from macromolecules and has been extensively used for determining the size, shape and molecular weight of proteins and nucleic acids. It permits analysis of biological macromolecules and their complexes in nearly physiological environments and direct study of structural responses to changes in external conditions. Applications of SAXS to biological systems include studies on envelope shape of proteins (Bu *et al.*, 1998; Shilton *et al.*, 1998), conformational changes and protein folding/unfolding (Flanagan *et al.*, 1992), and ligand induced oligomerisation and enzyme-substrate interactions (Tuzikov *et al.*, 1996; Lemmon *et al.*, 1997).

The mathematical models and theories used in SAXS studies on biological macromolecules and applied for shape determination from SAXS data are complex and

clearly beyond the scope of this thesis work, and thus are not included in this chapter. However, a brief theoretical introduction to SAXS and shape determination of macromolecules from SAXS data, together with a list of accessible and useful references are provided in Appendix D for the interested reader.

8.3 Sample Preparation for SAXS

The full-length σ^{54} and the fragments 70-324aa, 1-306aa and 117-477aa were constructed and purified as amino terminal 6His-tagged proteins. In order to reduce background scattering and to obtain protein samples which were >99% pure and monodispersed, several purification steps were applied to each protein preparation (see Materials & Methods). In an attempt to minimise aggregation and precipitation and increase the monodispersity of the protein samples during purification, storage and SAXS measurements, all proteins were prepared in the presence of 75mM sodium thiocyanate – a chaotropic agent used to help minimise aggregation. The preparation of the 70-324aa fragment is shown in Figure 8B as an example for the purification steps used and the purity of the protein sample used for SAXS.

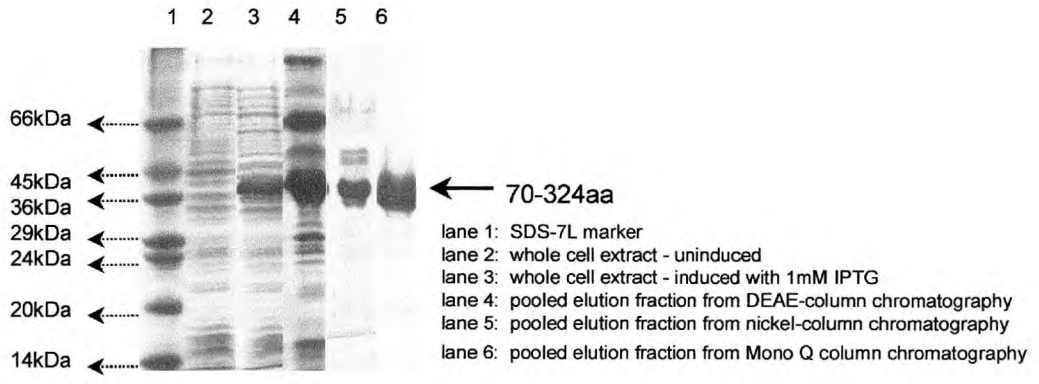


Figure 8B. Purification of the 70-324aa σ^{54} fragment for SAXS measurements. Shown is a Coomassie blue stained 12.5% SDS-polyacrylamide gel.

8.4 SAXS Measurements on Purified Sigma 54 Proteins

For the full-length σ^{54} and the 70-324aa fragment, the SAXS measurements at protein concentrations 3-20mg/ml displayed no significant concentration effects, and the scattering values were in good agreement with that from a reference sample (bovine serum albumin, Sigma) indicating that the samples were not aggregated and were monodispersed (data not shown). However, the scattering data for the 1-306aa and 117-477aa fragments showed that the background scattering was too high and the protein samples were strongly aggregated. Attempts to improve the quality of sample preparation by increasing the sodium thiocyanate concentration and reducing the storage and handling times of the samples failed to usefully improve the scattering data for the 1-306aa and 117-477aa σ^{54} fragments.

The experimental scattering parameters of the full-length σ^{54} and the 70-324aa fragment are shown in Table 8.1 (Svergun *et al.*, 2000). The molecular masses of the solutes indicate that both σ^{54} and the 70-324aa fragment are monomeric in solution, in agreement with the results of analytical centrifugation (D. Scott & J Hoggett, personal communication). The experimental values of D_{max} (maximum diameter) and R_g (radius of gyration) suggest that the two macromolecules, and especially σ^{54} are rather anisometric. It implies that σ^{54} is a very elongated protein and that the 30kDa 70-324aa fragment, although having a 2nm smaller D_{max}, is still rather elongated.

Parameter	σ^{54}	70-324aa
Mass (kDa)	58±6	32±4
D_{max} (nm)	13.0±0.5	11.0±0.5
R_g (nm)	3.43±0.06	3.05±0.05
Resolution (nm)[1]	2.8	2.4

[1]of the restored envelope models.

Table 8.1. Experimental scattering parameters of σ^{54} and the 70-324aa fragment.

8.5 Low resolution models of σ^{54} and the 70-324aa fragment

The low resolution models of σ^{54} and the 70-324aa fragment were restored from the experimental data using two *ab initio* procedures (Svergun *et al.*, 2000 and references within Appendix D). The restored envelope shapes and dummy atom models of σ^{54} and the 70-324aa fragment are presented in Figure 8C (panels A and B, respectively). The envelope shape of the 70-324aa fragment can be unambiguously positioned inside that of σ^{54}. Moreover, a boomerang-like shape of the core RNAP binding 70-324aa fragment of σ^{54} correlates well with the gross structure of the core RNAP and DNA binding fragment of the *E. coli* σ^{70} protein (the atomic model of the latter is superimposed with the low resolution models in Figure 8C; Malhotra *et al.*, 1996).

The low resolution solution structure for σ^{54} determined in this work provides the first initial model for a protein that functions to convert *E. coli* RNAP into an enhancer dependent enzyme. Clearly developed structures are present that probably represent either a discrete individual domain or alternatively domains that are directly interacting within the tertiary structure. The amino terminal Region I is required for transcription silencing and activator responsiveness and the carboxyl terminal parts of Region III DNA binding domain are thought to interact (Cannon *et al.*, 1999; Casaz & Buck, 1999; Wigneshweraraj *et al.*, 2001). Comparison of the structure of the 70-324aa fragment that lacks both of these domains (Figure 8A) is consistent with such an interaction, since two (possibly more) lobes are clearly absent (Figure 8C). Possibly , both Region I and the carboxyl parts of Region III contribute to this lobe and do directly interact with each other. Consistent with this view (further developed below) is the finding that Region I and parts of Region III interact to form a regulatory centre over the consensus GC promoter element (Wigneshweraraj *et al.*, 2001).

σ^{54} footprints promoter DNA from positions –35 to –5 with respect to the transcription start site (Buck & Cannon, 1992). Major points of contact are within four bases, namely TGCA, centred at –12 (the GC-element) and seven bases CTGGCAC centred at –24 within the consensus sequence of σ^{54}-dependent promoters.

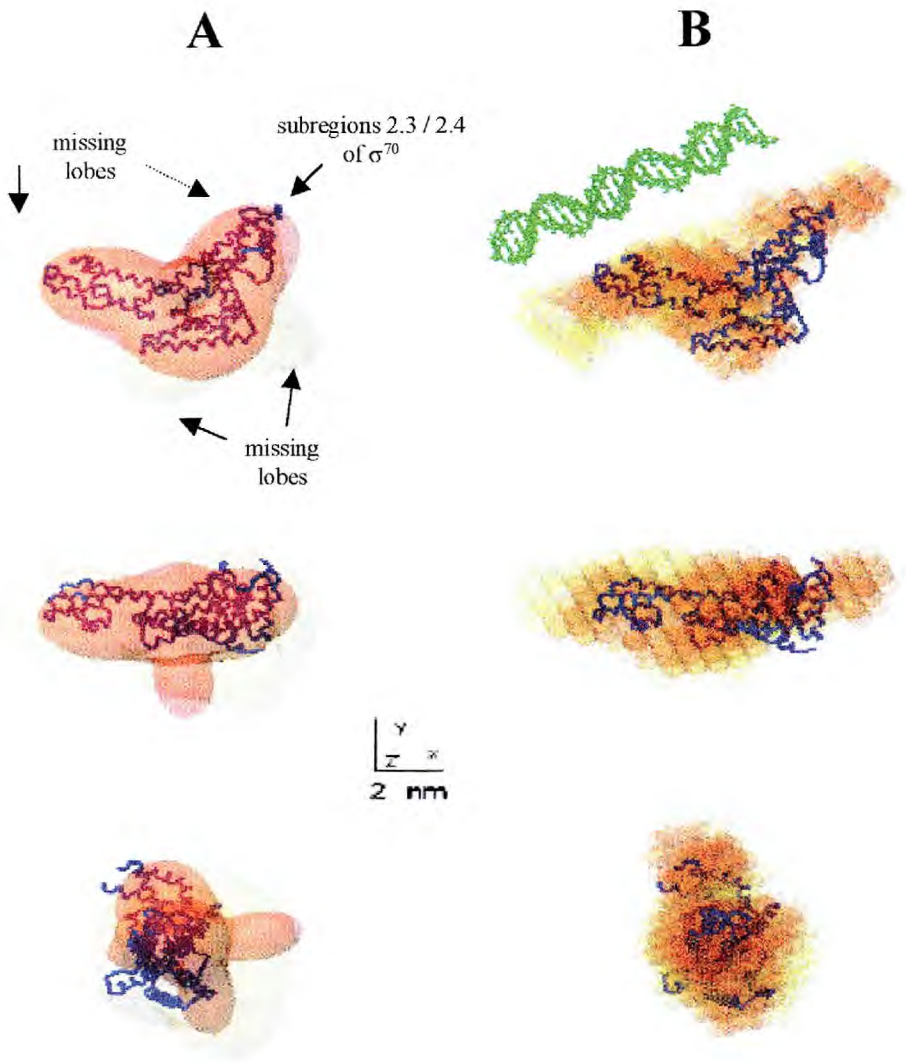


Figure 8C. Low resolution and dummy atom models of σ^{54}. Panel (**A**). Low resolution envelope models of σ^{54} (pale yellow) and the 70-324aa fragment (red). The missing lobes are indicated (see text); the helices that correspond to σ^{70} subregions 2.3 (involved in –10 promoter element interaction) and 2.4 are shown (see Introduction). Panel (**B**). Dummy atom models (spheres of radius 0.45nm). The middle and bottom rows are rotated counterclockwise by 90°C around the x and y axes, respectively. The models in Panel (**A**) and (**B**) are superimposed with the crystal structure of the σ^{70} fragment from *E. coli* RNAP (blue). A 28bp B-DNA fragment comprising the σ^{54} binding site (green) is shown above the potential DNA binding surfaces (Panel (**B**), top row).

Superimposition of the low resolution dummy atom models with a 28bp B-DNA fragment (Figure 8C, panel B) summarises a relationship between σ^{54}, and the 70-324aa fragment and the promoter DNA that can account for the known extent of σ^{54}-promoter

DNA interactions. The 70-324aa fragment lacks the Region III DNA binding determinants and Region I sequences that also contribute to DNA recognition (Wigneshweraraj *et al.*, 2001). It therefore seems possible that the lobe missing from the 70-324aa fragments (see above), but present in the full-length σ^{54} model includes Region I and parts of Region III DNA binding determinants that direct promoter binding. Further, the upper surface of the σ^{54} depicted in Figure 8C (panel B) further distinguishes the full-length σ^{54} from the 70-324aa fragment and may contribute to additional DNA contacting structures. The 70-324aa fragment does not bind DNA. Thus, together, the missing lobes and the upper surface can provide DNA contacting surfaces of approximately the necessary extent to fully interact with promoter DNA.

Another remarkable observation is that the fragment of σ^{70} for which the atomic structure has been determined (Malhotra *et al.*, 1996) has a similar shape and size compared with the 70-324aa fragment of σ^{54} (Figure 8C, panel A); this structural information further supports our previous finding that the two classes of sigma factor, the enhancer dependent σ^{54} and enhancer independent σ^{70}, both occupy similar positions in the core RNAP and use some overlapping points of interaction to contribute to the shared functionalities of the holoenzymes (Wigneshweraraj *et al.*, 2000; Datwyler *et al.*, 2000).

8.6 Conclusions

The structural study of proteins is a necessary complement to biochemical and genetic analysis of a system. Understanding the mechanisms of transcription initiation is dependent upon all three approaches. The low resolution structure of σ^{54} shows a well developed anisometric shape and is the first initial structural model for this class of σ factor. Similarity to a part of σ^{70} that binds core RNAP and interacts with the –10 promoter element is somewhat suggestive. Now that crystallography and cryoelectron microscopy are beginning to reveal the structures of multisubunit RNA polymerases (Finn *et al.*, 2000; Mooney & Landick, 1999), one can begin to evaluate the functional significance of this

similarity. These structural approaches together with the use of further tethered iron chelate methods for examining proximity relationships within transcription complexes should help determine the organisation of σ^{54}-RNAP holoenzymes.

OVERVIEW

In bacteria, the process of transcription of DNA to produce RNA relies upon a large multisubunit RNA polymerase (RNAP) containing two dissociable components: the core enzyme (E), which catalyses phophodiester bond formation, and one of a family of sigma (σ) subunits, transcription factors which determine promoter specificity. The core RNAP of *Escherichia coli*, which is the best-studied enzyme of its class, contains five polypeptide subunits: β, β', ω and a dimer of α. Binding of one of the several specificity σ factors converts the core enzyme to an holoenzyme form, which can recognise promoters. Of the seven *E. coli* σ factors, σ^{54} is in a class of its own. As discussed in the Introduction, members of the σ^{70}-class and σ^{54} bind to the common core RNAP but form holoenzymes with markedly different functional properties. The enhancer-dependent σ^{54}-holoenzyme relies on the ATPase activity of enhancer bound activator proteins for transcription initiation, whereas the enhancer-independent σ^{70} types of holoenzyme forms closed complexes that spontaneously isomerise open the promoter DNA and allow initiation.

Using the FeBABE tethered footprinting methodology we have shown that several of the surfaces proximal to σ^{54} and σ^{70} are common to both sigmas. This observation was further supported by small-angle x-ray scattering data on σ^{54}, which suggest that although σ^{54} and σ^{70} are unrelated by primary amino acid sequence, they may share a significant overall structural similarity. A significant outcome of the FeBABE protein footprint work in unraveling the enhancer dependence of the σ^{54}-holoenzyme in transcription initiation is in revealing the location of two regulatory residues of σ^{54} (Region I residue 36 and Region III residue 336) which are proximal to the catalytic centre of the core RNAP. σ^{54} mutants at residues 36 and 336, even though separated by 300 amino acids, share the same functional phenotype. It seems doubtless that some of the proximity relationships we have detected using FeBABE will change upon closed complex formation and activation of the σ^{54}-holoenzyme.

Those associated with residues 36 and 336 seem most likely to change, given that these residues are associated with σ^{54} sequences needed to maintain the stable closed promoter complex and that these, when mutated, lead to activator-independent isomerisation of the holoenzyme. Future work will target these issues and will aim to characterise the σ^{54}-core RNAP interface in the closed and open complexes.

σ^{54} must somehow manage to keep the promoter bound holoenzyme silent by preventing the required melting of promoter DNA until signalled by the ATPase activity of the activator protein. Region I of σ^{54} has been implicated in being a key element in the transcription silencing event and in preventing ATPase independent DNA melting at promoters (Cannon *et al.*, 2000; Casaz *et al.*, 2000; Gallegos *et al.*, 2000; Wang *et al.*, 1995). It has been proposed to form and bind to a repressive fork junction structure created by transient unpairing of base pairs adjacent (downstream) to the consensus –12 GC promoter element and to interact with single-strand DNA binding elements located in the carboxyl terminal Region III (Guo *et al.*, 1999) (Figure 1). The activator and its ATPase activity is important for triggering changes in the way how σ^{54} binds promoter DNA and for its action requires σ^{54} Region I. To fully engage the transcription bubble for initiation, holoenzymes (both σ^{54} and σ^{70}) need to establish a stable connected interaction between the fork junction at the upstream edge of the bubble and the melted downstream DNA. The σ^{54}-holoenzyme lacks a detectable interaction with nontemplate strand adjacent to the fork junction at –12 (Guo *et al.*, 1999) (Figure 1). This interaction is a critical feature of the σ^{70}-holoenzyme that distinguishes the two holoenzymes. Recently it was revealed that the interaction of the activator and its associated ATP binding activity help the σ^{54}-holoenzyme to establish the missing nontemplate strand interaction (Guo *et al.*, 2000). The interaction of the activator with Region I of σ^{54} is said to unmask single-strand DNA binding activity of σ^{54} causing full stable DNA opening to occur, leading to transcription (Guo *et al.*, 2000).

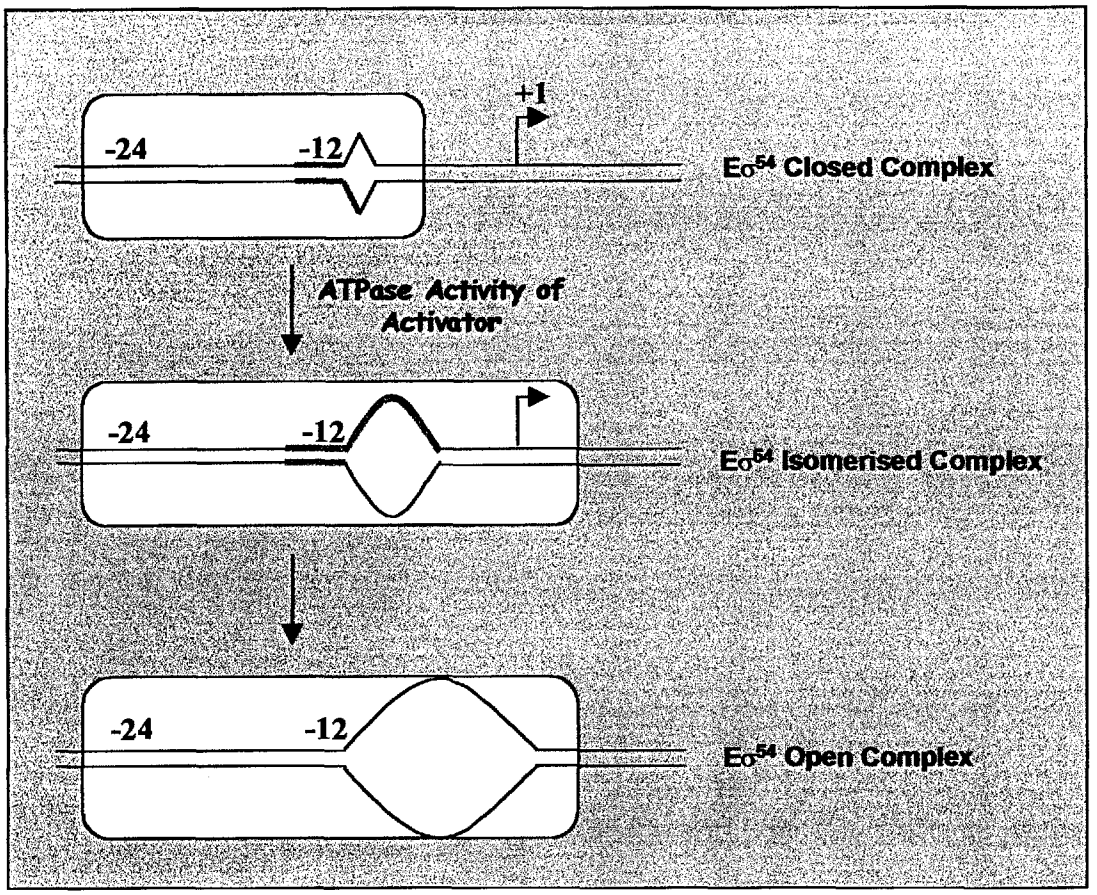


Figure 1. Model for the role of σ^{54} in transducing the signal for transcription activation (adapted from Gralla, 2000). Shown in green is the σ^{54}-holoenzyme and in dark green DNA sequences interacted by σ^{54} during the progression from the closed complex to an open one (see text for detals).

Using σ^{54} derivatives with FeBABE tethered at positions at 36 or 336, respectively, we have provided, for the first time, physical evidence for the model outlined in Figure 1. The DNA cleavage data presented show that parts of Region I and Region III are close to the –12 promoter element known to strongly influence regulation of open complex formation (Guo *et al.*, 1999). An advantage of these protein-DNA arrangements would be in co-ordinating concerted movements in σ^{54}, promoter DNA and core RNAP. We propose a model in which Region I and Region III forms an organising centre that brings the key σ^{54}, core RNAP and DNA components together (Figure 2).

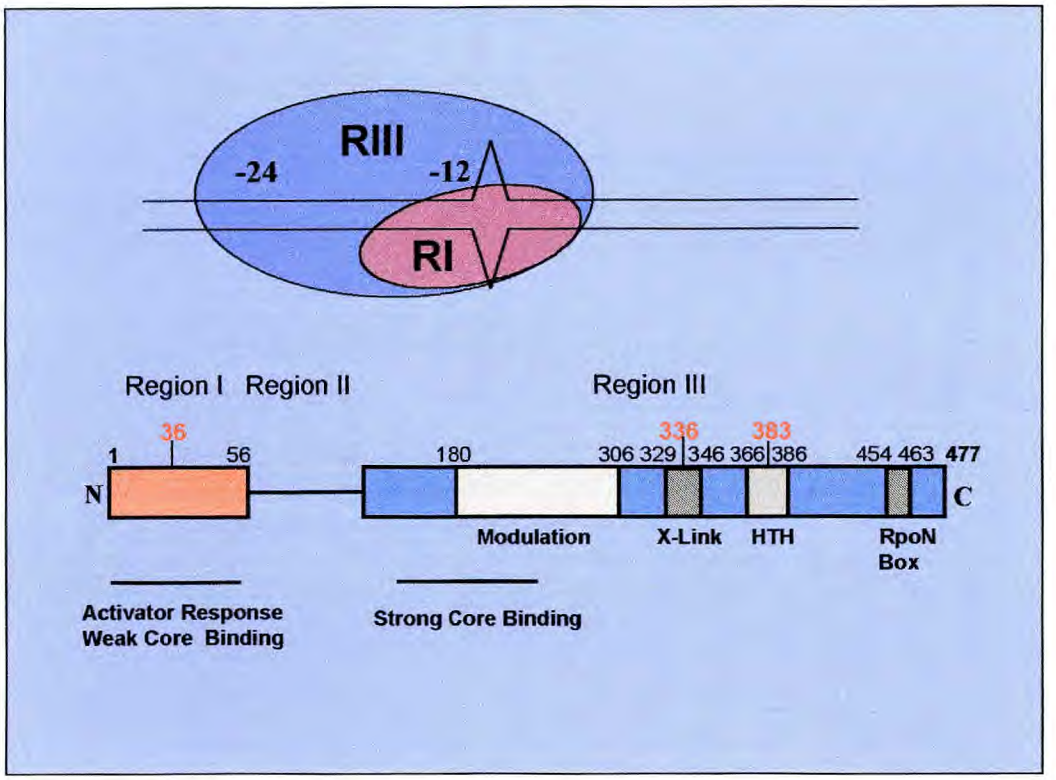


Figure 2. Regulatory sequences in σ^{54} localise over the start of DNA melting. Residues which were modified with FeBABE reagent are shown in red over the domain organisation of σ^{54}.

Region I changes its relationship to DNA in response to activation suggesting that the organising centre includes a target for the activator (Guo *et al.*, 2000; Cannon *et al.*, 2000). Indeed, the activator interacts with Region I (Chaney *et al.*, in preparation) probably to enforce unmasking of nontemplate single-strand DNA binding activity of σ^{54} that is proposed to reside in Region III (Kelly *et al.*, 2000; Guo *et al.*, 2000). Interestingly, residue 336 is located within this proposed single-strand DNA binding sequence and the σ^{54} derivative with FeBABE at position 336 shows increased cleavage of nontemplate single strand DNA upon activation. Figure 1 also highlights an important distinguishing feature of σ^{54} and σ^{70} holoenzymes: σ^{70}-holoenzymes prefer binding to the nontemplate strand whereas, in the silent state prior activation, σ^{54}-holoenzymes prefer binding to the template strand. Since extensive interactions along the nontemplate strand are an essential requirement in the promoter melting process leading to transcription initiation (Guo *et al.*, 2000; Fenton *et al.*, 2000; Guo

et al., 1999; Wang *et al.*, 1997), this activity is only granted to the σ^{54}-holoenzyme by the ATPase activity of the activator protein.

Further, we have shown that a DNA interacting residue (R383) within the proposed HTH structure of σ^{54} is not part of this regulatory centre and additional DNA cleavage data show that a colinear arrangement, C-terminal to N-terminal (5'-3' non-template strand) of the RpoN box, the HTH motif and theUV-crosslinking patch is possible, but unproven. Our data suggest that residue 383 in the putative HTH motif has a function in σ^{54} stability and is required for maintaining the holoenzyme in a stable complex when melting is initiating next to the GC element.

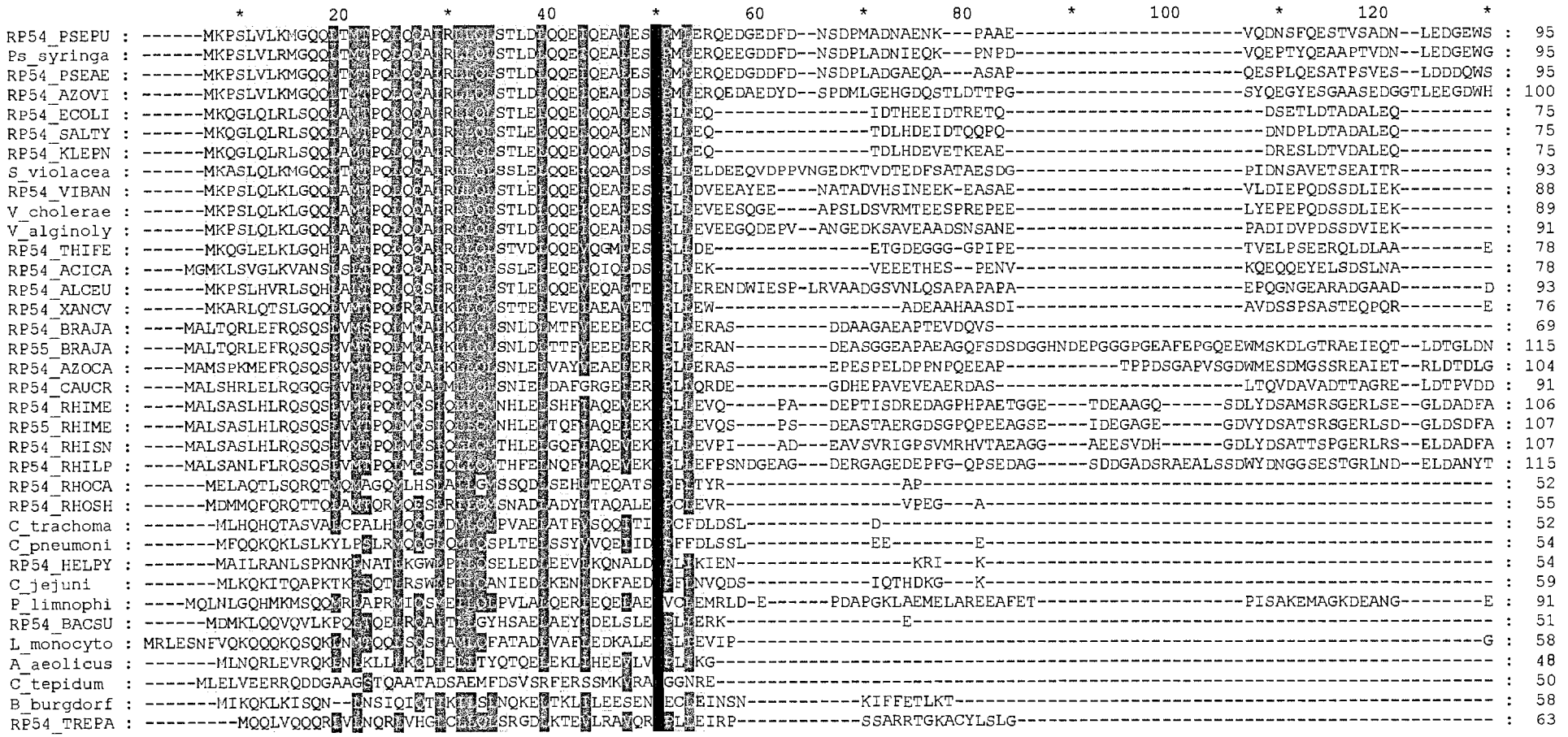
The σ^{54} promoter recognition sequence includes short elements at nucleotides –12 and –24 with extensive conservation in between these two. It is proposed that the –24 element makes a greater contribution to promoter binding. The –24 recognition is presumed to occur via Region III, but remains unproven. We have demonstrated the use of FeBABE as a tool in determining protein-protein and protein-DNA relationships in σ^{54} transcription initiation in the absence of any high resolution structural information. Current work involving a "FeBABE scanning" approach of the Region III is underway with the aim of identifying the –24 recognition element and the orientation of σ^{54} on promoter DNA. Some of the σ^{54} mutants characterised in the final parts of this thesis work show altered DNA binding properties and may prove to be suitable and interesting targets for FeBABE work. Strikingly, the DNA binding and activator-independent transcription properties of the F318A σ^{54} mutant suggest that residue F318 to be part of the regulatory centre.

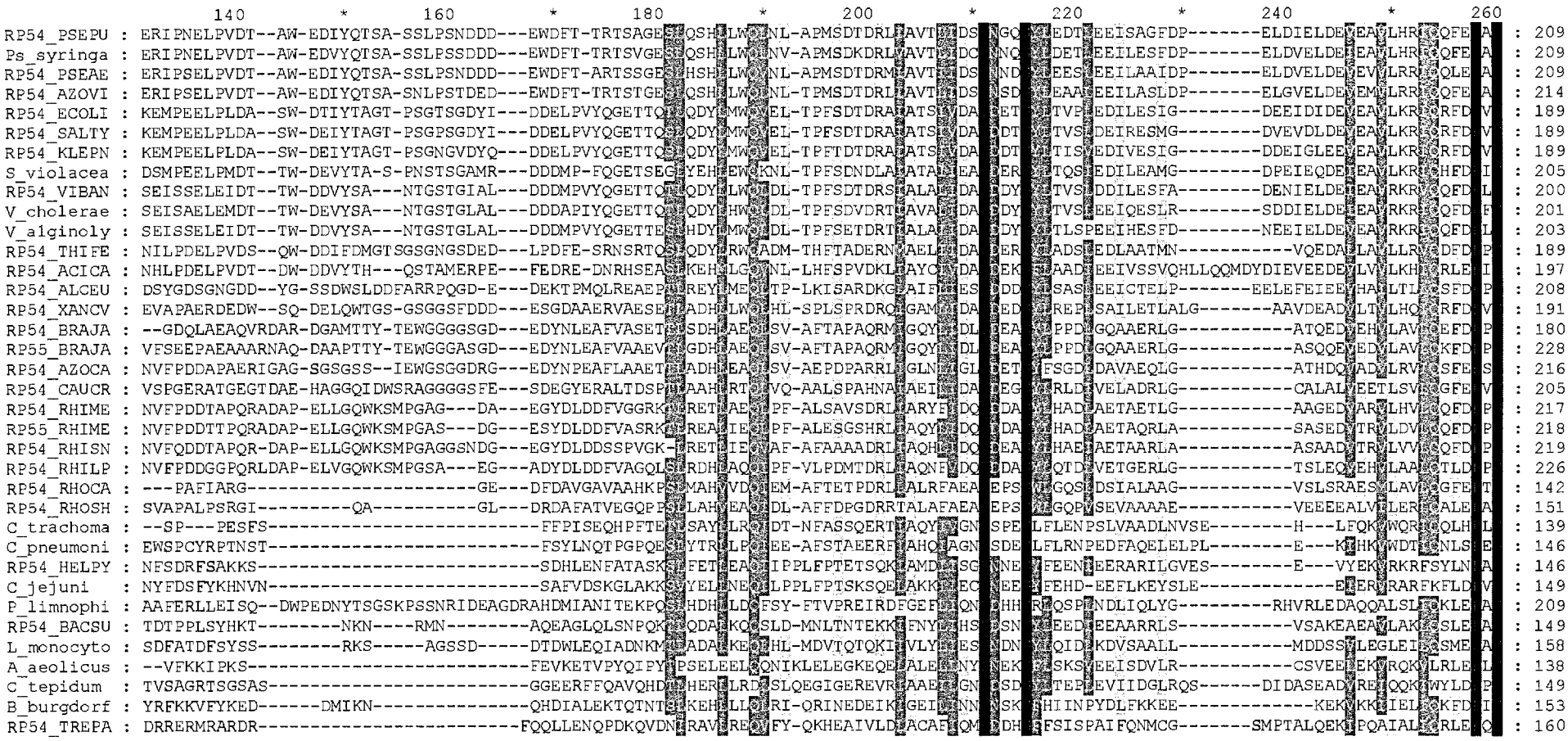
The uniqueness of transcription initiation by σ^{54} is its dependence on the nucleotide-hydrolysing activator proteins, but relatively little is known about the activator-σ^{54} interface – mainly owing to the transient nature of this interaction. Emerging data from our laboratory show that activator-bound σ^{54} complex can be trapped by using a transition state analogue of ATP ($ADP:AlF_4$) bound activator proteins (Chaney *et al.*, in preparation). FeBABE footprinting studies, using $ADP:AlF_4$ bound activator in combination with σ^{54}-FeBABE should provide us

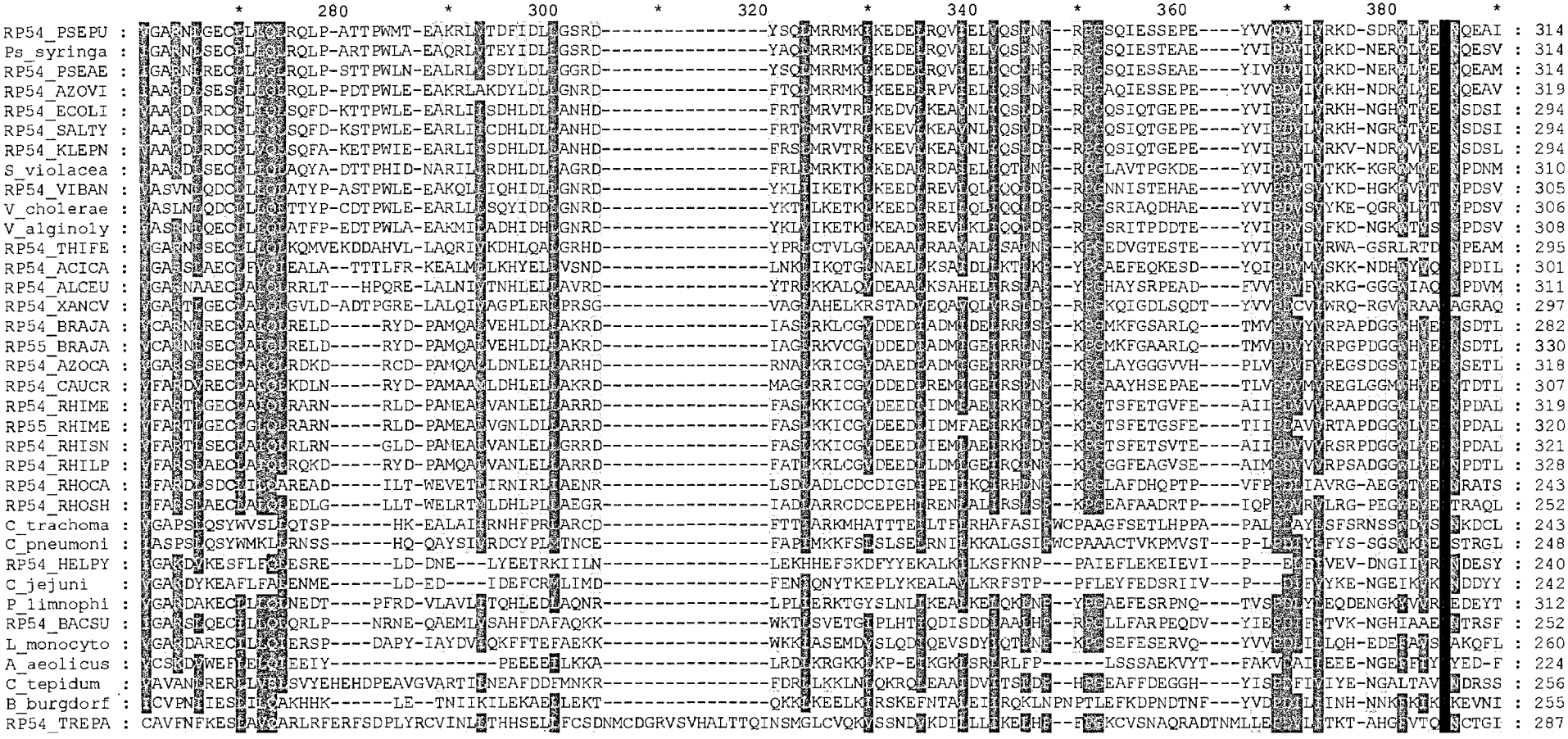
valuable insights into how this unique mechanism of bacterial transcription is orchestrated by the RNA polymerase subunit - σ^{54} and ATP binding and hydrolysis.

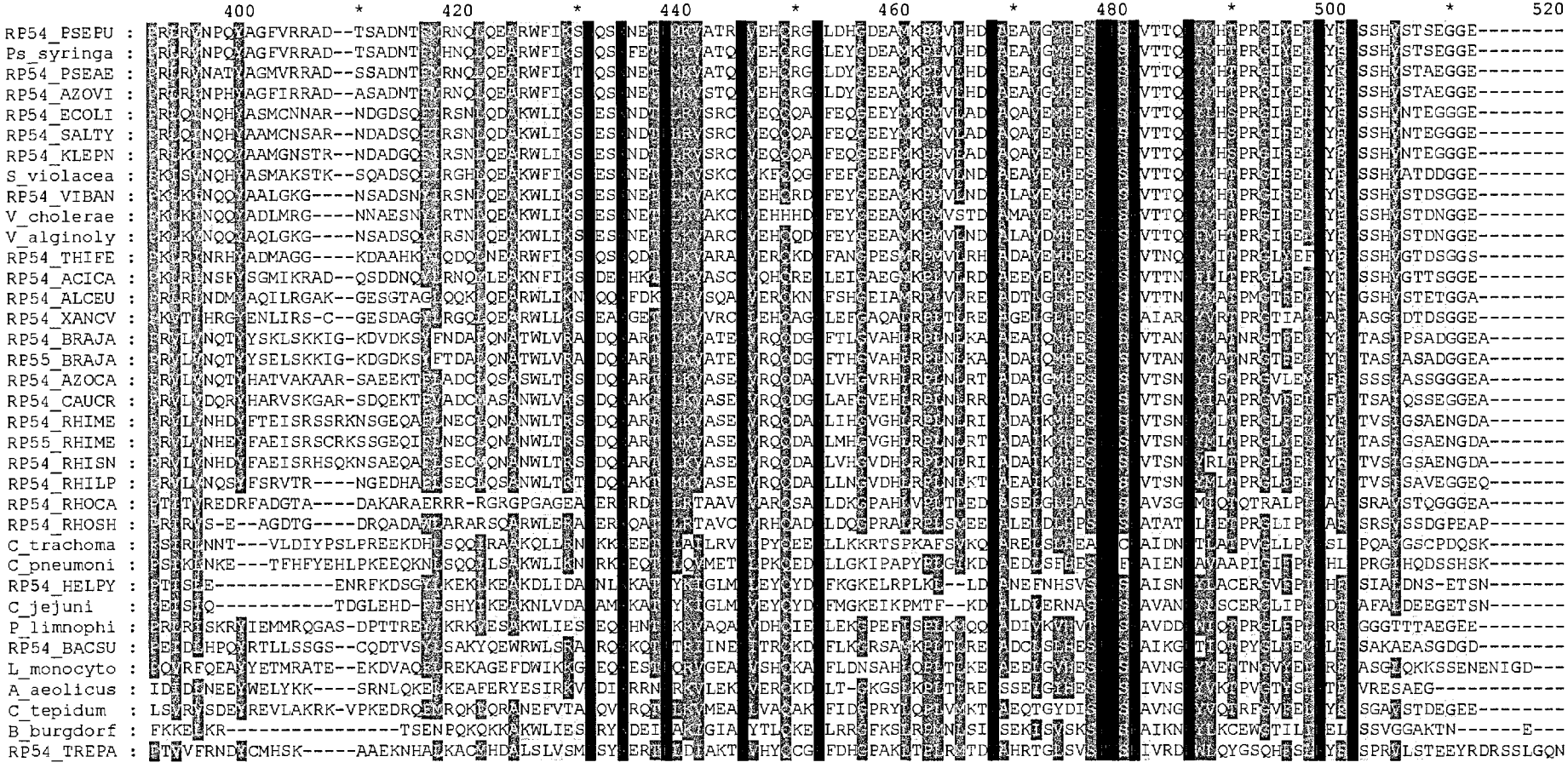
APPENDIX A: Sigma 54 Protein Sequence Alignment

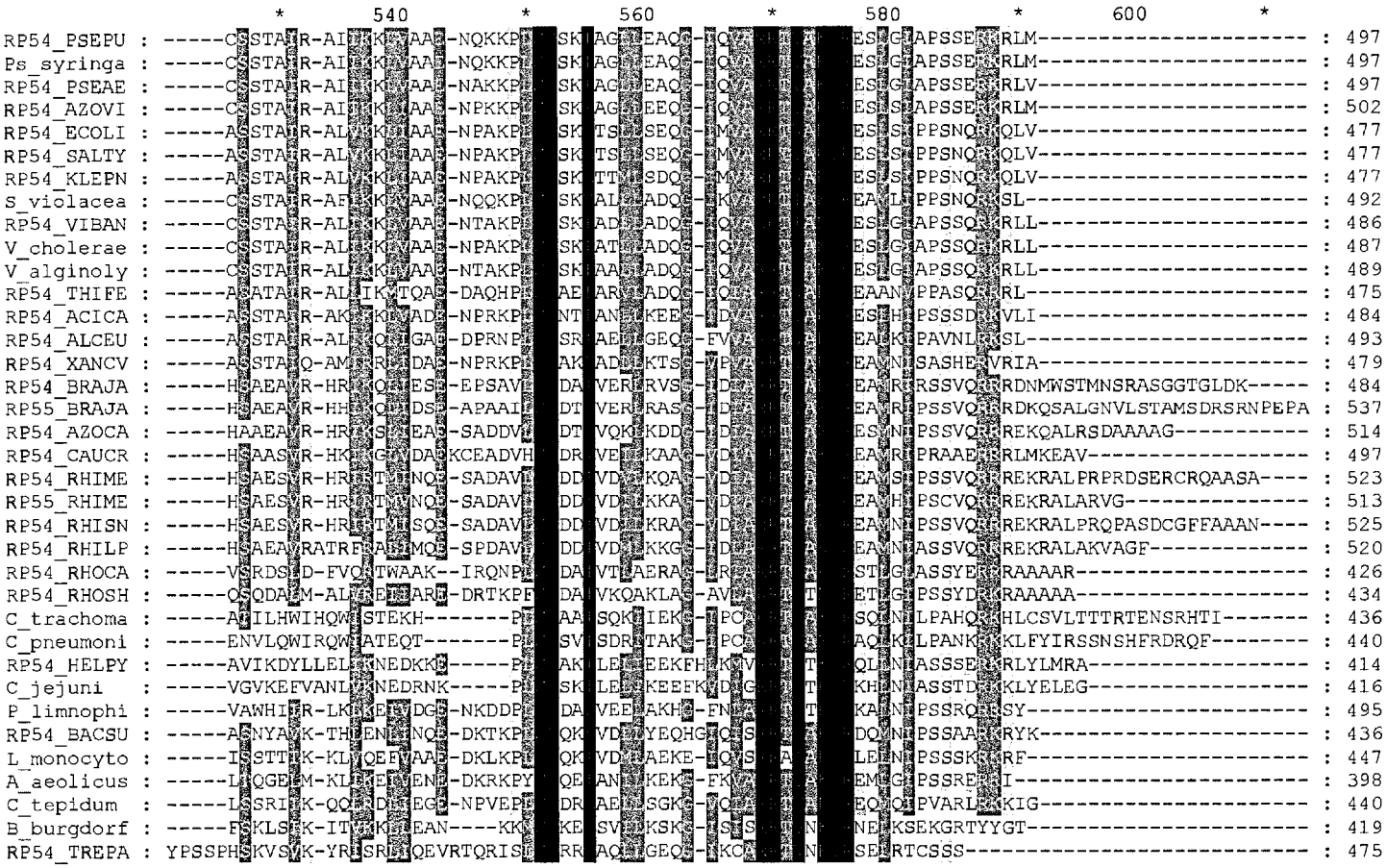
Highlighted in black are sequences with complete homology; in dark grey






* 280 * 300 * 320 * 340 * 360 * 380 *
RP54_PSEPU : 314
Ps_syringa : 314
RP54_PSEAE : 314
RP54_AZOVI : 319
RP54_ECOLI : 294
RP54_SALTY : 294
RP54_KLEPN : 294
S_violacea : 310
RP54_VIBAN : 305
V_cholerae : 306
V_alginoly : 308
RP54_THIFE : 295
RP54_ACICA : 301
RP54_ALCEU : 311
RP54_XANCV : 297
RP54_BRAJA : 282
RP55_BRAJA : 330
RP54_AZOCA : 318
RP54_CAUCR : 307
RP54_RHIME : 319
RP55_RHIME : 320
RP54_RHISN : 321
RP54_RHILP : 328
RP54_RHOCA : 243
RP54_RHOSH : 252
C_trachoma : 243
C_pneumoni : 248
RP54_HELPY : 240
C_jejuni : 242
P_limnophi : 312
RP54_BACSU : 252
L_monocyto : 260
A_aeolicus : 224
C_tepidum : 256
B_burgdorf : 255
RP54_TREPA : 287


400 * 420 * 440 * 460 * 480 * 500 * 520
RP54_PSEPU
Ps_syringa
RP54_PSEAE
RP54_AZOVI
RP54_ECOLI
RP54_SALTY
RP54_KLEPN
S_violacea
RP54_VIBAN
V_cholerae
V_alginoly
RP54_THIFE
RP54_ACICA
RP54_ALCEU
RP54_XANCV
RP54_BRAJA
RP55_BRAJA
RP54_AZOCA
RP54_CAUCR
RP54_RHIME
RP55_RHIME
RP54_RHISN
RP54_RHILP
RP54_RHOCA
RP54_RHOSH
C_trachoma
C_pneumoni
RP54_HELPY
C_jejuni
P_limnophi
RP54_BACSU
L_monocyto
A_aeolicus
C_tepidum
B_burgdorf
RP54_TREPA



APPENDIX B: *E. coli* σ^{54} Activators

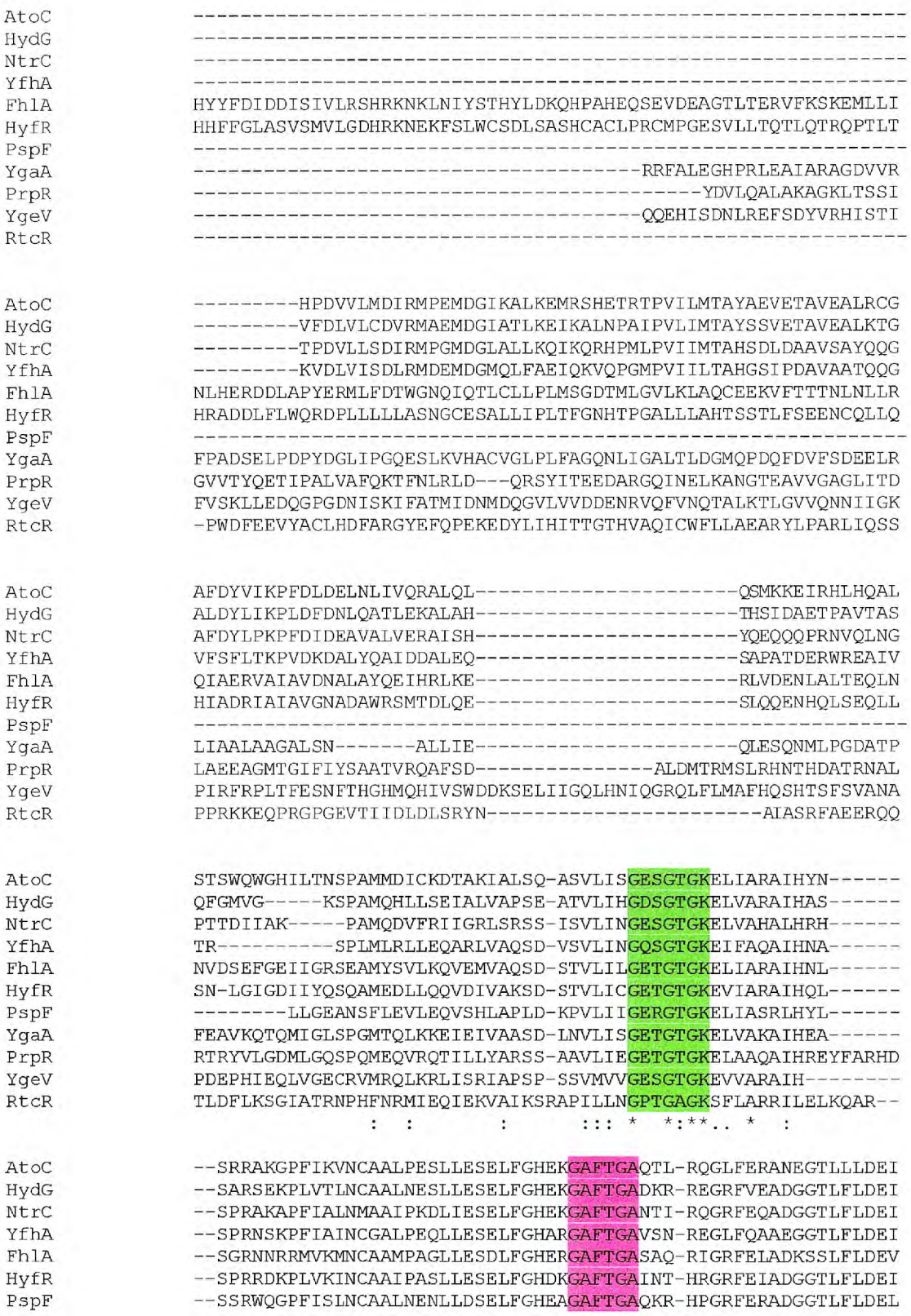
Shown below is the protein sequence alignment of *E. coli* σ^{54} activators. The nucleotide binding motif (Walker A motif) is highlighted in green and the σ^{54} interaction sequence is shown in pink (see introduction for further details).

```
AtoC    ------------------------------------------------------------
HydG    ------------------------------------------------------------
NtrC    ------------------------------------------------------------
YfhA    ------------------------------------------------------------
FhlA    --MSYTPMSDLGQQGLFDITRTLLQQPDLASLCEALSQLVKRS-----------------
HyfR    --MFAPPQGITIEAVNGMLAERLAQKHGKASLLRAFIPLPPPF-----------------
PspF    ------------------------------------------------------------
YgaA    ------MSSQNDYQSNGRYDNIYSQIDSEAKMSFSVDVLANIA-----------------
PrpR    ---MAHPPRLNDDKPVIWTVSVTRLFELFRDISLEFDHLANIT-----------------
YgeV    MELATTQSVLMQIQPTIQRFARMLASVLQLEVEIVDENLCRVAGTGAYGKFLGRQLSGNS
RtcR    ----------MRKTVAFGFVGTVLDYAGRGSQRWSKWRP---------------------

AtoC    ----------MTAINRILIVDDEDNVRRMLSTAFALQGFETHCANNGRTALHLFADI---
HydG    ---------MTHDNIDILVVDDDISHCTILQALLRGWGYNVALANSGRQALEQVREQ---
NtrC    ----------MQRGIVWVVDDDSSIRWVLERALAGAGLTCTTFENGAEVLEALASK---
YfhA    ---------MSHKPAHLLLVDDDPGLLKLLGLRLTSEGYSVVTAESGAEGLRVLNRE---
FhlA    -------ALADNAAIVLWQAQTQRASYYASREKDTPIKYEDETVLAHGPVRSILSRPDTL
HyfR    -------SPVQLIELHVLKSNFYYRYHDDGSDVTATTEYQGEMVDYSRHAVLLGSSG---
PspF    -----------------MANFIMAEYKDN------------------------------
YgaA    -------IELQRGIGHQDRFQRLITTLRQVLECDASALLRYDSRQFIPLAIDGLAKDVLG
PrpR    -----PIQLGFEKAVTYIRKKLANERCDAIIAAGSNGAY-----LKSRLSVPVILIKPSG
YgeV    RLLRHVLETKTEKVVTQSRFDPLCEGCDSKENCREKAFLGTPVILQDRCVGVISLIAVTH
RtcR    -----TLCLCQQESLVIDRLELLHDARSRSLFETLKRDIASVSPETEVVSVEIELHN---

AtoC    ------------------------------------------------------------
HydG    ------------------------------------------------------------
NtrC    ------------------------------------------------------------
YfhA    ------------------------------------------------------------
FhlA    HCSYEEFCETWPQLDAGGLYPKFGHYCLMPLAAEGHIFGGCEFIRYDDRPWSEKEFNRLQ
HyfR    ------------------------------------MAELRFIRTHGSRFTSQDCTLFN
PspF    ------------------------------------------------------------
YgaA    ------------------------------------------------------------
PrpR    ------------------------------------------------------------
YgeV    E-----------------------------------------------------------
RtcR    ------------------------------------------------------------

AtoC    ------------------------------------------------------------
HydG    ------------------------------------------------------------
NtrC    ------------------------------------------------------------
YfhA    ------------------------------------------------------------
FhlA    TFTQIVSVVTEQIQSRVVNNVDYELLCRERDNFRILVAITNAVLSRLDMDELVSEVAKEI
HyfR    WLARIITPVLQSWLNDEEQQVALRLLEKDRDHHRVLVDITNAVLSHLDLDDLIADVAREI
PspF    ------------------------------------------------------------
YgaA    ------------------------------------------------------------
PrpR    ------------------------------------------------------------
YgeV    ------------------------------------------------------------
RtcR    ------------------------------------------------------------
```


AtoC
HydG
NtrC
YfhA
FhlA
HyfR
PspF
YgaA
PrpR
YgeV
RtcR


YgaA --SPRAVNPLVYLNCAALPESVAESELFGHVKGAFTGAISN-RSGKFEMADNGTLFLDEI
PrpR ARQGKKSHPFVAVNCGAIAESLLEAELFGYEEGAFTGSRRGGRAGLFEIAHGGTLFLDEI
YgeV KLSGRRNKPFIAINCAAIPEQLLESELFGYVKGAFTGASANGKTGLIQAANTGTLFLDEI
RtcR ---HQFSGAFVEVNCATLRGDTAMSTLFGHVKGAFTGARES-REGLLRSANGGMLFLDEI
AtoC GEMPLVLQAKLLRILQEREFERIGGHQTIKVDIRIIAATNRDLQAMVKEGTFREDLFYRL
HydG GDISPMMQVRLLRAIQEREVQRVGSNQIISVDVRLIAATHRDLAAEVNAGRFRQDLYYRL
NtrC GDMPLDVQTRLLRVLADGQFYRVGGYAPVKVDVRIIAATHQNLEQRVQEGKFREDLFHRL
YfhA GDMPAPLQVKLLRVLQERKVRPLGSNRDIDINVRIISATHRDLPKAMARGEFREDLYYRL
FhlA GDMPLELQPKLLRVLQEQEFERLGSNKIIQTDVRLIAATNRDLKKMVADREFRSDLYYRL
HyfR GDLPLELQPKLLRVLQEREIERLGGSRTIPVNVRVIAATNRDLWQMVEDRQFRSDLFYRL
PspF ATAPMMVQEKLLRVIEYGELERVGGSQPLQVNVRLVCATNADLPAMVNEGTFRADLLDRL
YgaA GELSLALQAKLLRVLQYGDIQRVGDDRCLRVDVRVLAATNRDLREEVLAGRFRADLFHRL
PrpR GEMPLPLQTRLLRVLEEKEVTRVGGHQPVPVDVRVISATHCNLEEDMQQGRFRRDLFYRL
YgeV GDMPLMLQAKLLRAIEAREILPIGASSPIQVDIRIISATNQNLAQFIAEGKFREDLFYRL
RtcR GELGADEQAMLLKAIEEKTFYPFGSDRQVSSDFQLIAGTVRDLRQLVAEGKFREDLYARI
AtoC NVIHLILPPLRDRREDISLLANHFLQKFSSENQR-DIIDIDP----MAMSLLTAWSWPGN
HydG NVVAIEVPSLRQRREDIPLLAGHFLQRFAERNRK-AVKGFTP----QAMDLLIHYDWPGN
NtrC NVIRVHLPPLRERREDIPRLARHFLQVAARELGV-EAKLLHP----ETEAALTRLAWPGN
YfhA NVVSLKIPALAERTEDIPLLANHLLRQAAERHKP-FVRAFST----DAMKRLMTASWPGN
FhlA NVFPIHLPPLRERPEDIPLLAKAFTFKIARRLGR-NIDSIPA----ETLRTLSNMEWPGN
HyfR NVFPLELPPLRDRPEDIPLLAKHFTQKMARHMNR-AIDAIPT----EALRQLMSWDWPGN
PspF AFDVVQLPPLRERESDIMLMAEYFAIQMCREIKLPLFPGFTE----RARETLLNYRWPGN
YgaA SVFPLSVPPLRERGDDVILLAGYFCEQCRLRQGL-SRVVLSA----GARNLLQHYSFPGN
PrpR SILRLQLPPLRERVADILPLAESFLKVSLAALSAPFSAALRQG-LQASETVLLHYDWPGN
YgeV NVIPITLPPLRERQEDIELLVHYFLHLHTRRLGSVYPGIAPD-----VVEILRKHRWPGN
RtcR NLWTFTLPGLRQRQEDIEPNLDYEVERHASLTGDSVRFNTEARRAWLAFATSPQATWRGN
AtoC IRELSNVIERAVVMNSGPIIFSEDLPPQIRQPVCNAGEVKTAPVGER----------
HydG IRELENAVERAVVLLTGEYISERELPLAIAS--------TPIPLGQS----------
NtrC VRQLENTCRWLTVMAAGQEVLIQDLPGELFESTVAESTSQMQPDSWATLLAQWADRALRS
YfhA VRQLVNVIEQCVALTSSPVISDALVEQALEGENTALP----------
FhlA VRELENVIERAVLLTRGNVLQLSL----------PDIVLPEPETPPAAT---------
HyfR VRELENVIERAVLLTRGNSLNLHLNVRQSRLLPTLNEDSALRSSMAQLLHP---------
PspF IRELKNVVERSVYRHGTSDYPLDDIIIDPFKRRPPEDAIAVSETTSLPTLP---------
YgaA VRELEHAIHRAVVLAR-ATRSGDEVILEAQHFAFPEVTLPTPEVAAVPVVK---------
PrpR IRELRNMMERLALFLSVEPTPDLTPQFMQLLLP----------ELARES-------
YgeV LRELSNLMEYLVNVVPSGEVIDSTLLPPNLLNNGTTEQSDVTEVSEAHLSLDD-------
RtcR FRELSASVTRMATFATSGRITLDVVEDEINRLRYNWQESRPSALTALLGAEAEN---IDL
AtoC ----NLKEEIKRVEKRIIMEVLEQQEGNRTRTALMLGISRRALMYKLQEYGIDPADV---
HydG ----QDIQPLVEVEKEVILAALEKTGGNKTEAARQLGITRKTLLAKLSR----------
NtrC GHQNLLSEAQPELERTLLTTALRHTQGHKQEAARLLGWGRNTLTRKLKELGME-------
YfhA ----TFVEARNQFELNYLRKLLQITKGNVTHAARMAGRNRTEFYKLLSRHELDANDFKE-
FhlA --VVALEGEDEYQLIVRVLKETNGVVAGPKGAAQRLGLKRTTLLSRMKRLGIDKSALI--
HyfR --TTPENDEEERQRIVQVLRETNGIVAGPRGAATRLGMKRTTLLSRMQRLGISVREVL--
PspF --LDLREFQMQQEKELLQLSLQQGKYNQKR-AAELLGLTYHQFRALLKKHQI--------
YgaA --QNLREATEAFQRETIRQALAQN-HHNWAACARMLETDVANLHRLAKRLGLKD------
PrpR -----AKTPAPRLLTPQQALEKFN--GDKTAAANYLGISRTTFWRRLKS----------
YgeV -----AGGTALEEMEKQMIREALSRHNSKKQVADELGIGIATLYRKIKKYELLNT-----
RtcR FDRMQLEHVIAICRQAKSLSAAGRQLFDVSRQGKASVNDADRLRKYLARFGLTWEAVQDQ
AtoC ----
HydG ----
NtrC ----

YfhA	----
FhlA	----
HyfR	----
PspF	----
YgaA	----
PrpR	----
YgeV	----
RtcR	HSSS

APPENDIX C: Primers & Oligonucleotides

Listed below are the sequences of oligonucleotides used for PCR, sequencing and mutagenesis in the work described in chapters 3-7. All sequences are given in 5' to 3' direction. Unless otherwise stated, all oligonucleotides were purchased from Sigma Genosys Europe Ltd. as HPLC purified products. (* indicates the complementary strand sequence).

1. Mutagenesis oligonucleotides

E36G Mutagenesis of *K. pneumoniae rpoN*

cgttagaactccagcaaggcctccagcaggcgctg
cagcgcctgctggaggccttgctggagttctaacg*

Cys36 (E36C) Mutagenesis of *K. pneumoniae rpoN (cys-)*

ctgtccacgttagaactccagcaatgcctccagcaggcgctggacagcaac
gttgctgtccagcgcctgctggaggcattgctggagttctaacgtggacag*

Cys198 (C346A) Mutagenesis of *K. pneumoniae rpoN*

ctgcgcgtcagccgcgccatcgtcgaacagcag
ctgctgttcgacgatggcgcggctgacgcgcag*

F318A Mutagenesis of *K. pneumoniae rpoN*

caatgacgctgacggccaggctatccgtagcaacctgcag
ctgcaggttgctagcgatagcctggccgtcagcgtcattg*

E325A Mutagenesis of *K. pneumoniae rpoN*

cgtagcaacctgcaggcagcgcgctggctgatc
gatcagccagcgcgctgcctgcaggttgctacg*

Cys346 (C198A) Mutagenesis of *K. pneumoniae rpoN*

gcggcaaaagatttgcgtgatgccctgctggttcagctttcacag
ctgtgaaagctgaaccagcagggcatcacgcaaatcttttgccgc*

Cys336 (R336C) Mutagenesis of *K. pneumoniae rpoN(cys-)*

ctgatcaagagcctggagagctgcaacgacaccctgctgcgcgtc
gacgcgcagcagggtgtcgttgcagctctccaggctcttgatcag*

Cys383 (R383C) Mutagenesis of *K. pneumoniae rpoN(cys-)*

gaaatgcatgaatccactatttcatgcgttactacgcaaaaatacctgcac
gtgcaggtatttttgcgtagtaacgcatgaaatagtggattcatgcatttc*

R383A Mutagenesis of *K. pneumoniae rpoN*

gccgtcgaaatgcatgaatccactatttcagccgttactacgcaaaaatacctgcacagtcca
tggactgtgcaggtatttttgcgtagtaacggctgaaagagtggattcatgcatttcgacggc*

R383K Mutagenesis of *K. pneumoniae rpoN*

gccgtcgaaatgcatgaatccactatttcaaaggttactacgcaaaaatacctgcacagtcca
tggactgtgcaggtatttttgcgtagtaacctttgaaagagtggattcatgcatttcgacggc*

E410A Mutagenesis of *K. pneumoniae rpoN*

gccatgtgaacaccgcaggcggcggcgaagc
gcttcgccgccgcctgcggtgttgacatggc*

E414A Mutagenesis of *K. pneumoniae rpoN*

ccgaaggcggcggcgcagcatcgtcgacggc
gccgtcgacgatgctgcgccgccgccttcgg*

EE410/414AA Mutagenesis of *K. pneumoniae rpoN*

gccatgtgaacaccgcaggcggcggcgcagcatcgtcgacggc
gccgtcgacgatgctgcgccgccgcctgcggtgttcacatggc*

E431A Mutagenesis of *K. pneumoniae rpoN*

gaaattaatcgccgcggcaaaccccgcgaaaccattg
caatggtttcgcggggtttgccgcggcgattaatttc*

2. Sequencing oligonucleotides

T7 primer

taatacgactcactatagg

T7T primer

gctagttattgctcagcgg*

MP64

gcgatcctcggcctctttgg*

N1780

actgctccagtttcgt*

RpoN117

catatgcaaagcctgcaggattacctgat

RpoN120

catatgctggcaggattacctgatgtgg

3. PCR oligonucleotides

glnHp2-53

cgcaccagattggtgccc

glnHp2-35

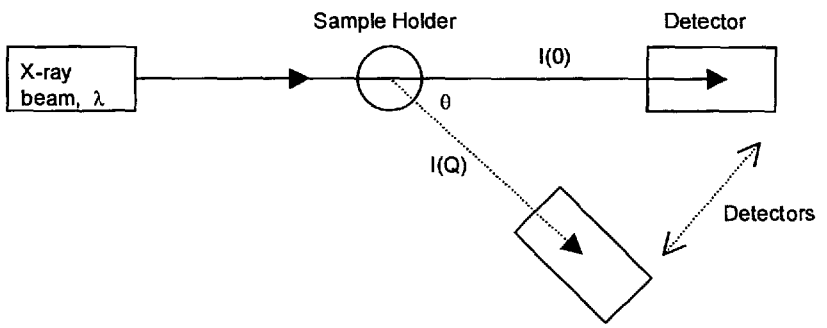
gggtcagagcagccagtg*

APPENDIX D: Small Angle X-ray Scattering (SAXS)

SAXS – Data Collection

In the small angle x-ray scattering (SAXS) experiments, an intense x-ray beam is focused on a protein sample in solution. The sample's electrons scatter a portion of the incident flux and the scattered x-ray intensity is measured as a function of scattering angle. From the scattering data, the following can be inferred: the protein's overall size (radius of gyration), overall shape (pair distribution function), and association state (monomer, dimer, high order oligomer etc.).

The SAXS measurements were performed on beamline Doris III of DESY (Deutsches Elektronen Synchrotron) at the EMBL Outstation in Hamburg. A monochromator is used for energy selection of the synchrotron radiation. The protein sample solution sits in a sample cell with mica windows allowing transmission of the x-ray (λ = 0.15nm) beam through the sample. The radiation which scatters from the sample is detected typically 2m downstream by a detector:



SAXS Data Analysis

1. Radius of Gyration

Interpreting SAXS data can be a very difficult task unless one is very lucky and the sample fits one of the many idealised models that have been developed over the years. The primary data from the experiment are contained in the function I(Q), the intensity of the scattered radiation, where

$$Q = \frac{4\pi \sin(\theta)}{\lambda}$$

where 2θ is the angle between the transmitted and scattered beams and λ is the wavelength.

I(Q) is often measured at several concentrations of protein, and then corrected for instrumental factors and extrapolated to zero before use. At very small angles of scattering of the incident influx I(0), the shape of the scattering in the so-called "Guinier" region can be used to give us an idea of the radius of gyration. At small angles, the scattering intensity follows the Guinier approximation:

$$I(Q) = I(0)\exp(\frac{-Q^2 R_g{}^2}{3})$$

where R_g is the radius of gyration and I(0) the intensity of scattering at zero angle scattering. Taking the natural log of the Guinier approximation, the following relationship results:

$$Ln(I(Q) = Ln(I(0) - \frac{4\pi^2 Q^2 Rg^2}{3}$$

The natural log of the Guinier approximation shows that there is a linear relationship between Ln(I(Q)) and Q^2. From the Guinier plot (Ln(I(Q)) against Q^2), the data can be fitted to a straight line and the radius of gyration (R_g) can be extracted from the slope of $4\pi^2R_g^2/3$ of the plot.

2. Conformational & Association State

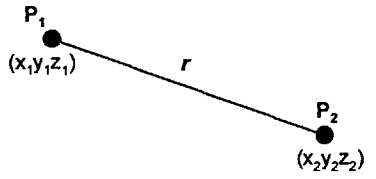
At smaller I(Q) values, the molecular weight of the protein specimen can be estimated from the value I(Q→0), that is when scattering at 0 angle (forward scattering), I(0) is coincident with the direct x-ray beam and is proportional to the molecular weight.

The forward scattering amplitude (I(0)) yields information on the association state of the protein sample. I(0) is proportional to the number (n) of scatterers in solution and proportional to the square of the scatterer's volume (V). For a completely dimerised sample, n will decrease by a factor of two and V will increase by a factor of two. The net result will be a factor of 2 increase in I(0).

3. Pair Distribution Function

At larger values of Q, I(Q) values can be used to extract dimensional information by computer analysis on the protein specimen. For this purpose, the Fourier transformation of I(Q), (a signal form understood by the computer algorithm) conventionally termed as P(r), is often used. The pair distribution function P(r) measures the relative number of electron pairs with a given distance separation. P(r) which provides shape information can be obtained from the scattering pattern. P(r) is a distance pair-distribution function – a function which analyses the distribution of distances between pairs of points (scattering

elements, n) in a molecule, that is, it is a function on two scattering elements in the molecule, say P_1 and P_2, defined by the coordinates (x_1,y_1,z_1) and (x_2,y_2,z_2), respectively:



The value of r, the distance between P_1 and P_2 can be determined from P(r) from the following relationship by computer analysis.

$$P(r) = 8\pi R \int_0^\infty I(Q)Q^2 \sin(2\pi QR)dQ$$

Therefore, the co-ordinates of several scattering elements in the molecule can be found, which can be used to determine the low resolution envelope shape of the protein sample. Thus, for the *ab initio* shape determination two computer alogorithms, GNOM and SASHA have been developed by our collaborators at EMBL Hamburg, Dr. Dimitri Svergun and Dr Michel Koch.

4. Useful References

Trewhella J, Gallagher SC, Krueger JK, Zhao J (1998). Neutron and X-ray solution scattering provide insights into biomolecular structure and function. Sci Prog., vol. 81, pp. 101-122

Bras W, Ryan AJ., (1998). Sample environments and techniques combined with small angle X-ray scattering. Adv Colloid Interface Sci., vol. 75., pp. 1-43.

Perkins SJ, Ashton AW, Boehm MK, Chamberlain D (1998). Molecular structures from low angle X-ray and neutron scattering studies. Int J Biol Macromol., vol. 22, pp. 1-16.

Trewhella J. (1997). Insights into biomolecular function from small-angle scattering. Curr Opin Struct Biol.,vol. 7, pp.702-708.

PUBLICATIONS ARISING FROM THIS WORK

The work described in Chapters 3 and 4:

Siva R Wigneshweraraj, Nobuyuki Fujita, Akira Ishihama and Martin Buck (2000). "Conservation of sigma-core RNA polymerase proximity relationships between the enhancer-independent and enhancer-dependent sigma classes". *EMBO Journal*, 19, pp. 3038-3048.

The work described in Chapter 5:

Siva R Wigneshweraraj, Matthew K Chaney, Akira Ishihama and Martin Buck (2001). "Regulatory Sequences in Sigma 54 Localise Near the Start of DNA Melting". *Journal of Molecular Biology*, 306, pp. 681-701.

The work described in Chapter 6:

Siva R Wigneshweraraj, Akira Ishihama and Martin Buck (2001). "*In vitro* roles of invariant helix-turn-helix motif residue R383 in σ^{54}". *Nucleic Acids Research*, 29, pp. 1163-1174.

The work described in Chapter 7:

Siva R Wigneshweraraj, Paul Casaz and Martin Buck "Correlating Protein Footprinting with Mutational Analysis in the enhancer-dependent RNA polymerase specificity factor sigma 54". Submitted.

The work described in Chapter 8:

Dimitri I Svergun, Marc Malfois, Michel H Koch, **Siva R Wigneshweraraj** and Martin Buck (2000). "Low Resolution Structure fo the σ^{54} Transcription Factor Revealed by X-ray Solution Scattering". *Journal of Biological Chemistry*, 275, pp. 4210-4214.

The EMBO Journal Vol. 19 No. 12 pp. 3038–3048, 2000

Conservation of sigma–core RNA polymerase proximity relationships between the enhancer-independent and enhancer-dependent sigma classes

Siva R.Wigneshweraraj, Nobuyuki Fujita[1], Akira Ishihama[1] and Martin Buck[2]

Imperial College of Science, Technology and Medicine, Department of Biology, Sir Alexander Fleming Building, Imperial College Road, London SW7 2AZ, UK and [1]Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411, Japan

[2]Corresponding author
e-mail: m.buck@ic.ac.uk

Two distinct classes of RNA polymerase sigma factors (σ) exist in bacteria and are largely unrelated in primary amino acid sequence and their modes of transcription activation. Using tethered iron chelate (Fe-BABE) derivatives of the enhancer-dependent σ^{54}, we mapped several sites of proximity to the β and β′ subunits of the core RNA polymerase. Remarkably, most sites localized to those previously identified as close to the enhancer-independent σ^{70} and σ^{38}. This indicates a common use of sets of sequences in core for interacting with the two σ classes. Some sites chosen in σ^{54} for modification with Fe-BABE were positions, which when mutated, deregulate the σ^{54}–holoenzyme and allow activator-independent initiation and holoenzyme isomerization. We infer that these sites in σ^{54} may be involved in interactions with the core that contribute to maintenance of alternative states of the holoenzyme needed for either the stable closed promoter complex conformation or the isomerized holoenzyme conformation associated with the open promoter complex. One site of σ^{54} proximity to the core is apparently not evident with σ^{70}, and may represent a specialized interaction.

Keywords: enhancers/σ factors/protein footprinting/RNA polymerase

Introduction

The *Escherichia coli* RNA polymerase is a multisubunit enzyme that exists in two major forms with respect to its subunit composition. The core enzyme ($\alpha_2\beta\beta'$; E) is capable of processive transcription elongation followed by termination, but is unable to initiate transcription specifically unless it is in its σ-bound holoenzyme form (Eσ). The σ subunit confers promoter-specific DNA binding and transcription initiation capabilities on the RNA polymerase. Based on structural and functional criteria, the seven different σ factors (σ^{70}, σ^{54}, σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{18}) identified in *E.coli* are grouped into two classes. Most σ factors, with the exception of σ^{54}, belong to the σ^{70} class, the major σ factor in *E.coli* that is involved in expression of most genes during exponential growth. σ^{54} (σ^{N}) represents a unique class that differs functionally from the σ^{70} class. The presence of σ^{54} has been well documented in several proteobacteria as well as in the Gram-positive *Bacillus subtilis* (Debarbouille *et al.*, 1991) and in *Planctomyces limnophilus* (Leary *et al.*, 1998). σ^{54} directs its holoenzyme to recognize and bind promoters with conserved –12/–24 regions located between positions –11 and –26 relative to the transcription start. These can be considered the functional analogue of the –10/–35 core promoter elements recognized by σ factors belonging to the σ^{70} class. In *E.coli*, both σ^{70}, the major σ factor for growth-related genes, and σ^{54} are always present, but other minor σ^{70} class members are induced under certain stress conditions (Jishage *et al.*, 1997).

Although σ^{54} binds the same core RNA polymerase as the σ^{70} class members, each σ class confers different properties on the holoenzymes formed. Eσ^{54} causes transcription to assume several features reminiscent of higher eukaryotic mechanisms (Sasse-Dwight *et al.*, 1990). In marked contrast to the σ^{70}–holoenzyme, which often initiates transcription in the absence of transcription activators, the σ^{54}–holoenzyme remains in a closed promoter complex in the absence of activator protein. Isomerization of this closed complex to a transcriptionally competent open complex depends upon the activity of an activator protein bound to a DNA sequence with enhancer-like properties coupled to γ–β bond nucleoside triphosphate hydrolysis by the activator protein (Weiss *et al.*, 1991).

The functional domain organization of σ^{54} is complex and appears to differ significantly from that of the members of the σ^{70} class. In particular, overall DNA binding by σ^{70} is modulated by N-terminal sequences, whereas N-terminal sequences in σ^{54} are required for activation. Extensive deletion and mutational analyses of σ^{54} have allowed functions to be assigned to different regions of the protein (Merrick *et al.*, 1993; Wong *et al.*, 1994; Cannon *et al.*, 1995; Gallegos and Buck, 1999). The N-terminal region I of σ^{54} is regulatory and closely implicated in polymerase isomerization and DNA melting (Cannon *et al.*, 1999; Gallegos *et al.*, 1999; Guo *et al.*, 1999). It plays a central role in mediating the response to activator proteins and in binding early melted DNA structures (Gallegos and Buck, 2000; W.Cannon, M.T.Gallegos and M.Buck, submitted), but is dispensable for core and overall DNA binding (Cannon *et al.*, 1999). Region II is acidic and not conserved. Region III includes a major core RNA polymerase-binding determinant and sequences that directly contact DNA and that enhance σ^{54}–DNA binding (Merrick and Chambers, 1992; Cannon *et al.*, 1997; Gallegos and Buck, 1999) (Figure 1A). Emerging roles for region III appear to be in maintaining the closed promoter complex in a transcriptionally silent state and in generating polymerase isomerization upon activation (Chaney and Buck, 1999).

Published by Oxford University Press

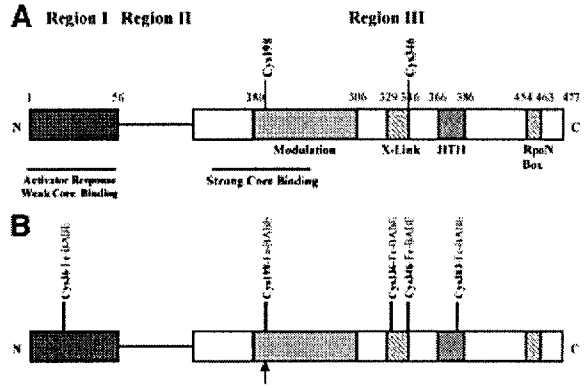


Fig. 1. (**A**) Schematic representation of *Klebsiella pneumoniae* σ^{54} domain organization. The locations of the native cysteine residues are indicated. (**B**) Sites where a single cysteine was introduced and conjugated with Fe-BABE. The arrow indicates that conjugation was performed under denaturing conditions.

The core RNA polymerase-binding interface of σ^{54} comprises at least two functionally distinct sequences: a 95 amino acid residue sequence (120–215) within region III, which binds the core strongly, and a second sequence within region I that binds the core more weakly (Gallegos and Buck, 1999). Comparison of hydroxyl radical-mediated cleavage profiles of wild-type and region I-deleted (ΔR1σ^{54}) proteins indicated that the overall solvent-exposed surface of σ^{54} is not greatly altered by deletion of region I (Casaz and Buck, 1999). Residues protected from hydroxyl radical-mediated cleavage in the holoenzyme made with ΔR1σ^{54} largely coincided with residues protected specifically by DNA in the wild-type holoenzyme, implying a role for region I in establishing an appropriate conformation of the DNA-binding domain in the wild-type holoenzyme (Casaz and Buck, 1999). Thus, it is possible that the interaction of region I with core RNA polymerase is important for controlling DNA-binding domain conformation.

To date, little is known about the location of σ^{54}-binding sites on the core RNA polymerase. Recently, we showed by small-angle X-ray scattering studies that core RNA polymerase-binding fragments of σ^{54} (amino acids 70–324) and the crystal structure of core RNA polymerase-binding σ^{70} fragment (amino acids 114–448) have a similar envelope shape, leading to the suggestion that both σ classes might be located in similar places on the core (Svergun *et al.*, 2000). However, the sequence differences that exist between σ^{54} and σ^{70} must also account for the markedly different properties of the two holoenzymes. The affinities of σ^{54} and σ^{70} for the core RNA polymerase measured *in vitro* are similar (Gallegos and Buck, 1999). Both σ factors could use the same binding site on core RNA polymerase, have overlapping sites or utilize additional independent binding sites. Recent protein–protein interaction and footprinting studies have revealed that the β' subunit of core RNA polymerase provides the major binding interaction for σ^{70} in the holoenzyme, while the β subunit contributes a further binding interaction (Arthur and Burgess, 1998; Owens *et al.*, 1998; Katayama *et al.*, 2000). In the present work, we have used tethered nucleophile-mediated footprinting to extend our understanding of the σ^{54}–core RNA polymerase interface. Using (*p*-bromoacetamidobenzyl)-EDTA Fe (Fe-BABE) derivatives of σ^{54}, we mapped several sites of proximity to the β and β' subunits of the core RNA polymerase. Remarkably, most sites localized to those previously identified as close to enhancer-independent σ^{70} (Owens *et al.*, 1998) and σ^{38} (A.Kolb, personal communication). This indicates use of the same or overlapping sequences in core RNA polymerase for interaction with the two σ classes. One site of σ^{54} proximity to the β' subunit apparently is not evident with σ^{70}, and may represent a specialized interaction of σ^{54} with core RNA polymerase.

Results

Location of Fe-BABE conjugation targets on σ^{54}

The bifunctional chelating agent Fe-BABE covalently conjugates to the free sulfhydryl groups of cysteine residues. After the addition of ascorbate and peroxide, the Fe-BABE generates nucleophiles, which cleave polypeptide chains in proximity (~12 Å) to the chelate, independently of the amino acid sequence involved (Rana and Meares, 1991; Ishihama, 2000). Using information from extensive mutational analysis of σ^{54}, we selected five sites (E36, C198, R336, C346 and R383, described below) of potential interest for tethering Fe-BABE. Figure 1B shows the schematic representation of the primary structure of σ^{54}, in which the positions of introduced cysteine residues are indicated.

Cys198 and Cys346. The σ^{54} from *Klebsiella pneumoniae* contains two highly conserved cysteine residues at amino acid positions 198 and 346. Cys198 is located in a 95 amino acid sequence (120–215), which binds core RNA polymerase strongly and is likely to include most of the fold critical for forming stable holoenzyme (Gallegos and Buck, 1999). Cys346 borders a DNA-interacting patch in σ^{54} (Cannon *et al.*, 1994) and is strongly protected by the core in hydroxyl radical cleavage experiments when region I is absent (Casaz and Buck, 1999). By changing Cys198 and Cys346 to alanine, we constructed two mutant σ^{54} proteins with single cysteines at positions 198 (C346A) and 346 (C198A). The Cys(–)*rpoN* construct was used to purify the cysteine-free double mutant (C198A and C346A) protein, which was used as the negative control protein in the cleavage assays. The Cys(–)σ^{54} protein was found to maintain 90% of its transcriptional activity (see later), implying that substitution of the conserved Cys198 and Cys346 with alanine does not significantly affect the conformation of the mutant protein.

Cys36. In the next step, single cysteine residues were introduced into Cys(–)σ^{54}. Region I of σ^{54} has multiple roles in regulating σ^{54} function: (i) inhibiting polymerase isomerization in the absence of activation and directing fork junction DNA binding (Wang *et al.*, 1995; Syed and Gralla, 1998; Cannon *et al.*, 1999; Guo *et al.*, 1999); and (ii) stimulating initiation of open complex formation in response to activation (Sasse-Dwight and Gralla, 1990; Syed and Gralla, 1998; Casaz *et al.*, 1999; Gallegos and Buck, 2000). The low affinity of region I sequences for

core RNA polymerase and its protection from proteolysis by core (Casaz and Buck, 1997, 1999; Gallegos and Buck, 1999) prompted us to introduce a single cysteine residue into region I in the Cys(–)σ^{54} protein. The increased protease sensitivity of residue E36 detected upon open complex formation provided us with a potential target for introducing a cysteine that could be in close proximity to the core RNA polymerase subunits in closed complexes (Casaz and Buck, 1997).

Cys336. Recently, we showed that a single amino acid substitution, R336A, in the region III DNA-binding domain of σ^{54} allows holoenzyme to isomerize and transcribe without activator (Chaney and Buck, 1999), a phenotype previously only associated with region I mutations (Hsieh and Gralla, 1994; Hsieh *et al.*, 1994; Syed and Gralla, 1998; Casaz *et al.*, 1999). We justify our choice of R336 for introducing a single cysteine residue by suggesting that R336 is a part of a network of interactions necessary for maintaining the transcriptionally silent state of the holoenzyme and thus might interact directly or indirectly with core RNA polymerase subunits.

Cys383. The highly conserved arginine at position 383 is located within the second helix of the proposed helix–turn–helix motif in region III of σ^{54}, suggested to make direct contact with the conserved –12 promoter sequences (Coppard and Merrick, 1991; Merrick and Chambers, 1992). Recent evidence shows conserved arginine residues in the region III DNA-binding domain of σ^{54} prevent polymerase isomerization (Chaney and Buck, 1999). Furthermore, the –12 promoter sequence is part of a molecular switch that has to be thrown by the action of the activator for transcription to proceed (Guo *et al.*, 1999) through a network of communication possibly involving R383, other σ^{54} parts and/or the core subunits. We therefore targeted R383 to introduce a single cysteine residue.

Single cysteine σ^{54} mutants are active for transcription

Next, the activity of each single cysteine mutant and the Cys(–)σ^{54} protein was compared with the activity of wild-type σ^{54} in an *in vitro* activator-dependent single round transcription assay using the *E.coli glnHp2*-m12 promoter as the template (Claverie-Martin and Magasanik, 1992). Assays were conducted with subsaturating amounts of holoenzyme to allow quantitative detection of holoenzyme activities. The mutant proteins were found to exhibit >50% of the activity of wild-type σ^{54} (Figure 2A). We therefore conclude that a replacement of all endogenous cysteines with alanine residues and the introduction of cysteine residues at alternative positions (see above) had no gross negative effect on the function of the protein. Also, because transcriptional activity of σ^{54} involves numerous unique stereospecific interactions, the preservation of transcriptional activity in the single cysteine mutants and the Cys(–)σ^{54} protein suggests that the conformation of the mutant proteins important for function is not altered significantly.

Region I sequences and R336 are involved in preventing unregulated polymerase isomerization (Syed *et al.*, 1997; Syed and Gralla, 1998; Casaz *et al.*, 1999; Chaney and Buck, 1999). We therefore tested whether the single cysteine σ^{54} mutants were active for unregulated transcription (also called 'activator bypass' transcription), i.e. whether they allowed polymerase isomerization in the absence of activator proteins. The protocol for these assays begins by incubating the holoenzyme with GTP, CTP and UTP. The first six bases transcribed from the *E.coli glnHp2*-m12 template are UGUCAC (+1 to +6), so transcripts can be initiated under these conditions if an activator-independent open complex forms. ATP and [α-^{32}P]UTP are then added, together with heparin, to destroy residual closed complexes and unstable open complexes, and allow elongation of any initiated transcripts. E36C and R336C showed activator-independent transcription activity (Figure 2B). The lack of unregulated transcription activity from the Cys(–)σ^{54} protein suggests that the bypass phenotype of E36C and R336C is attributable to the substitutions at amino acid residues at positions E36 and R336, and not because of the replacement of the two endogenous cysteines at positions C198 and C346. Consistent with previous observations by us (Casaz *et al.*, 1999; Chaney and Buck, 1999) and others (Hsieh and Gralla, 1994; Hsieh *et al.*, 1994; Syed and Gralla, 1998), E36 and R336 in σ^{54} have a function in inhibiting polymerase isomerization in the absence of activation.

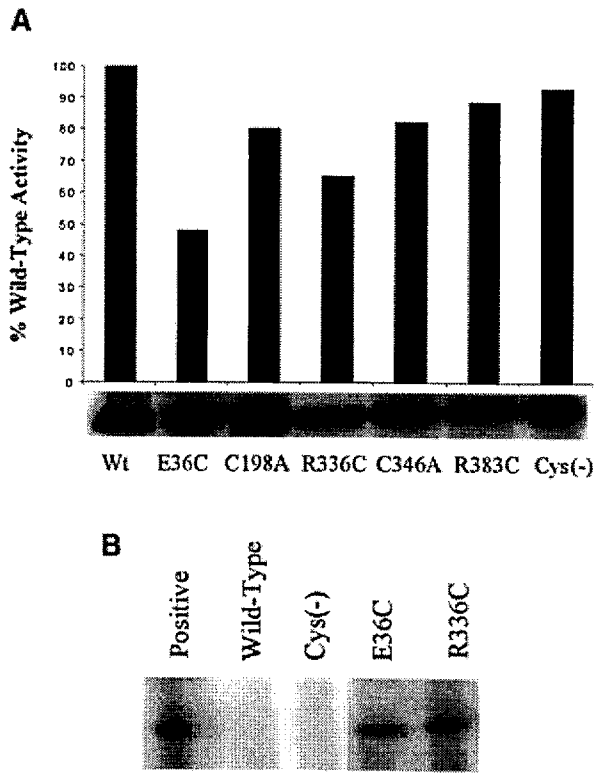


Fig. 2. *In vitro* transcription activity of σ^{54} single cysteine mutants on supercoiled *E.coli glnHp2*-m12 promoter. (**A**) Activator-dependent (using PspFΔHTH) and (**B**) activator-independent transcription activity of σ^{54} single cysteine mutants and Cys(–)σ^{54} relative to the wild type. Δ(21–27)σ^{54} (Gallegos and Buck, 2000) was used as the positive control (Positive) in the activator-independent assays.

Derivatization of single cysteine σ^{54} mutants with the chemical protease Fe-BABE

Each single cysteine σ^{54} mutant and Cys(–)σ^{54} were derivatized with the chemical cleavage reagent Fe-BABE under native conditions and the conjugation yield was determined fluorometrically by the CPM test (*N*-[4-[7-(diethylamino)-4-methylcoumarin-3-yl]phenyl]maleimide) (Greiner *et al.*, 1997) from the fluorescence difference between the conjugated and unconjugated proteins. The conjugation yield for the single cysteine mutants E36C, R336C, C198A and R383C was estimated to be >40% (Table I). The mutant σ^{54} harbouring the native single cysteine at position C198 (the C346A mutation) proved difficult to conjugate under native conditions (conjugation yield of 20%), implying that C198 is probably slightly buried or located in a hydrophobic region in the folded state because the hydrophobic reagent CPM reacted with high yield (data not shown) while the hydrophilic molecule Fe-BABE did not. When conjugation with C198 was performed under denaturing conditions, the yield increased to 50%. The Cys(–)σ^{54} protein failed to react detectably with CPM and Fe-BABE under native and denaturing conditions, directly implying that under the conditions used, Fe-BABE conjugation has only occurred at the free sulfhydryl side groups of cysteine residues.

The Fe-BABE-derivatized single cysteine mutants of σ^{54} are active for holoenzyme formation and transcription

We investigated the properties of the derivatized single cysteine mutants of σ^{54} proteins *in vitro* to determine whether core RNA polymerase binding and transcription were affected. Native gel holoenzyme assembly assays were used to detect complex formation between core RNA polymerase and σ^{54}. The holoenzyme migrates more slowly than the core enzyme. With the exception of the conjugated R336C mutant, all derivatized σ^{54} mutant proteins bound core RNA polymerase with the same apparent affinity as the non-conjugated wild-type σ^{54} (Figure 3A). Residue 336 lies outside the minimal core-binding region defined by deletion mutagenesis (Gallegos and Buck, 1999), but adjacent to a region in σ^{54} (amino acids 290–310) that is protected by core RNA polymerase from hydroxyl radical-mediated cleavage (Casaz and Buck, 1999) and within a sequence (amino acids 325–440) strongly protected by core RNA polymerase when region I is removed (Casaz and Buck, 1999). Because the unconjugated R336C binds core RNA polymerase with an affinity similar to that of wild-type σ^{54}, we conclude that the presence of the 490 Da Fe-BABE molecule at position R336 has caused the R336C mutant to undergo a conformational change that diminishes binding to the core RNA polymerase. We also note that conjugated Cys198 binding to core RNA polymerase results in two different holoenzyme conformers, one of which has a slower native gel mobility (Figure 3A, lane 10, marked with an arrow).

To determine the effect of Fe-BABE conjugation on the transcriptional activity of each single cysteine σ^{54} mutant, single-round transcription assays were also performed after conjugation with the Fe-BABE probe. Assays were conducted with subsaturating amounts of holoenzyme to allow detection of activity differences. After correction for the presence of unconjugated σ^{54} proteins, the conjugates showed >40% transcription activity relative to wild-type σ^{54}, with the exception of the R336C mutant, which has a markedly reduced transcription activity (10%) following conjugation (Figure 3B), consistent with our previous observation that the conjugated R336C mutant has reduced affinity for the core. Nevertheless, the conjugated R336C σ^{54} forms a holoenzyme that is active for a significant level (10%) of activated transcription.

We used the activator bypass transcription assay on the conjugated E36C and R336C mutants as a further indicator of conformational preservation of these proteins following conjugation. We reasoned that if the E36C and R336C mutants retained bypass activity following Fe-BABE conjugation, then the presence of the Fe-BABE molecule at these positions did not greatly affect the conformation of the conjugated proteins. Both mutants E36C and R336C retained their bypass activity following conjugation with Fe-BABE (Figure 3C). In conclusion, the core RNA polymerase-binding assays, the activator-dependent and the bypass transcription assays altogether suggest that Fe-BABE conjugation did not grossly affect the overall conformation and structural integrity of the mutant σ^{54} proteins.

Table I. Fe-BABE conjugation efficiency

Residue	Mutation	Conjugation efficiency (%)
36	E36C	53
198	C346A	46[a]
336	R336C	51
346	C198A	47
383	R383C	76
Cys(–)	C198A/C346A	3/5[a]

[a]Denotes conjugation under denaturing conditions.

Enhancer-dependent σ^{54}- and enhancer-independent σ^{70}-interacting sites on β and β′ core subunits are conserved

Holoenzymes were formed with conjugated and unconjugated single cysteine mutants of σ^{54} and cleavage reactions were carried out essentially as previously described (Owens *et al.*, 1998). In our approach, we assigned the cleavage sites on the β and β′ subunits by comparing the relative migration of the Fe-BABE cleavage products with sequence-specific markers generated by cleaving β and β′ subunits at either cysteine or methionine residues [using 2-nitro-5-thiocyanobenzoate (NTCB) and cyanogen bromide (CNBr), respectively]. The cleavage sites for each β and β′ marker fragment were assigned by calculating the molecular weight of each fragment with respect to the locations of cysteine or methionine residues in the amino acid sequence of β and β′ subunits (Figure 4A) using a plot of log molecular weight versus the relative migration distance of a known set of marker proteins (data not shown).

Initially, we showed that self-cleavage of σ^{54} did not occur (data not shown), indicating that cutting of β and β′ was attributed to binding of intact σ^{54} proteins. Ascorbate-

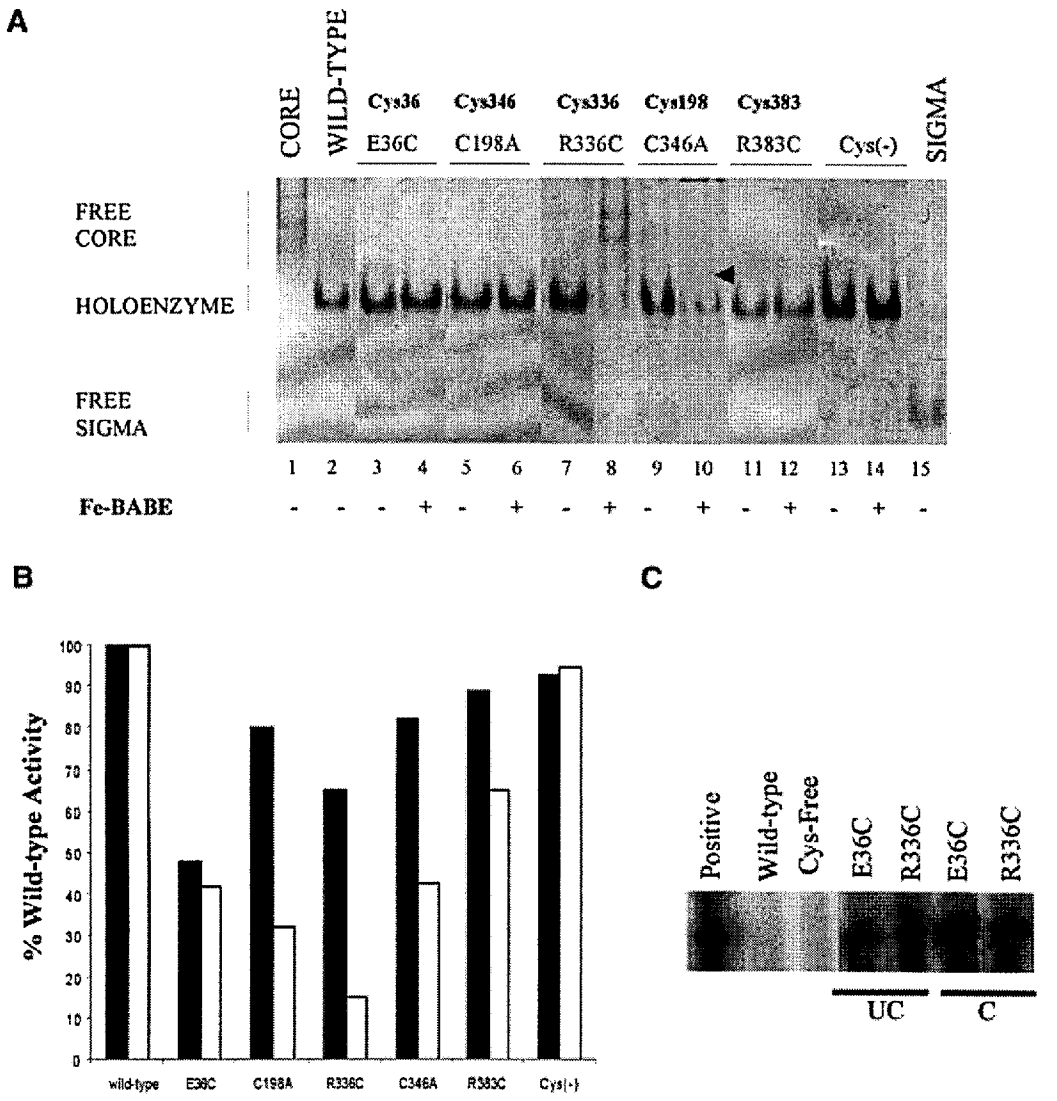


Fig. 3. (**A**) Formation of holoenzyme by the unconjugated (–) and Fe-BABE-conjugated (+) σ^{54} single cysteine mutants. Migration positions of free core RNA polymerase, holoenzyme and free σ protein are indicated. (**B**) *In vitro* transcription activity as in Figure 2. The activities of the σ^{54} single cysteine mutants before (black bars) and after (white bars) Fe-BABE conjugation are compared. The values for conjugated proteins are corrected for the presence of unconjugated proteins. (**C**) Activator bypass activity of the E36C and R336C σ^{54} proteins after Fe-BABE conjugation (C); the activator bypass activity of unconjugated proteins (UC) is shown for comparison.

and hydrogen peroxide-treated and untreated reactions of each single cysteine mutant σ^{54}–holoenyzme were separated, blotted and visualized by immunostaining with affinity-purified β and β′ N-terminus-specific antibodies (Figure 4B and C). The lack of cleavage of free core RNA polymerase and the holoenzyme formed with Cys(–)σ^{54} (Figure 4B and C, controls) protein serves as an essential control experiment, demonstrating that under the conditions used there is no artefactual cutting of the core RNA polymerase subunits. Hence, we believe that each cut (Figure 4B and C) is attributed to a specific Fe-BABE form of σ^{54} created by derivatizing a single cysteine residue.

Cys198 and Cys346. Initially, we looked for the β and β′ interactions made by the σ^{54} single cysteine mutants with Fe-BABE conjugated to the endogenous cysteines at positions Cys198 (C346A) and Cys346 (C198A), respectively. Cys198 cleaves β and β′ at multiple sites. On the β subunit, strong cleavage by Cys198 is seen within the Rif1 cluster (between conserved β regions C and D) and in conserved region G (Figure 4B, Cys198, lane 4). A few weak cleavages are also seen within the DR1 region. Moderate cleavage by Cys198 of the β′ subunit occurs in conserved regions B and C and within region G (Figure 4C, Cys198, lane 2). In contrast, Cys346 only moderately cleaves β and β′ at single sites; these map proximal to the conserved region D on the β subunit and within conserved regions C and D on the β′ subunit (Cys346, lane 6, Figure 4B and C, respectively). These results are consistent with previous observations: Cys198 is located within a sequence

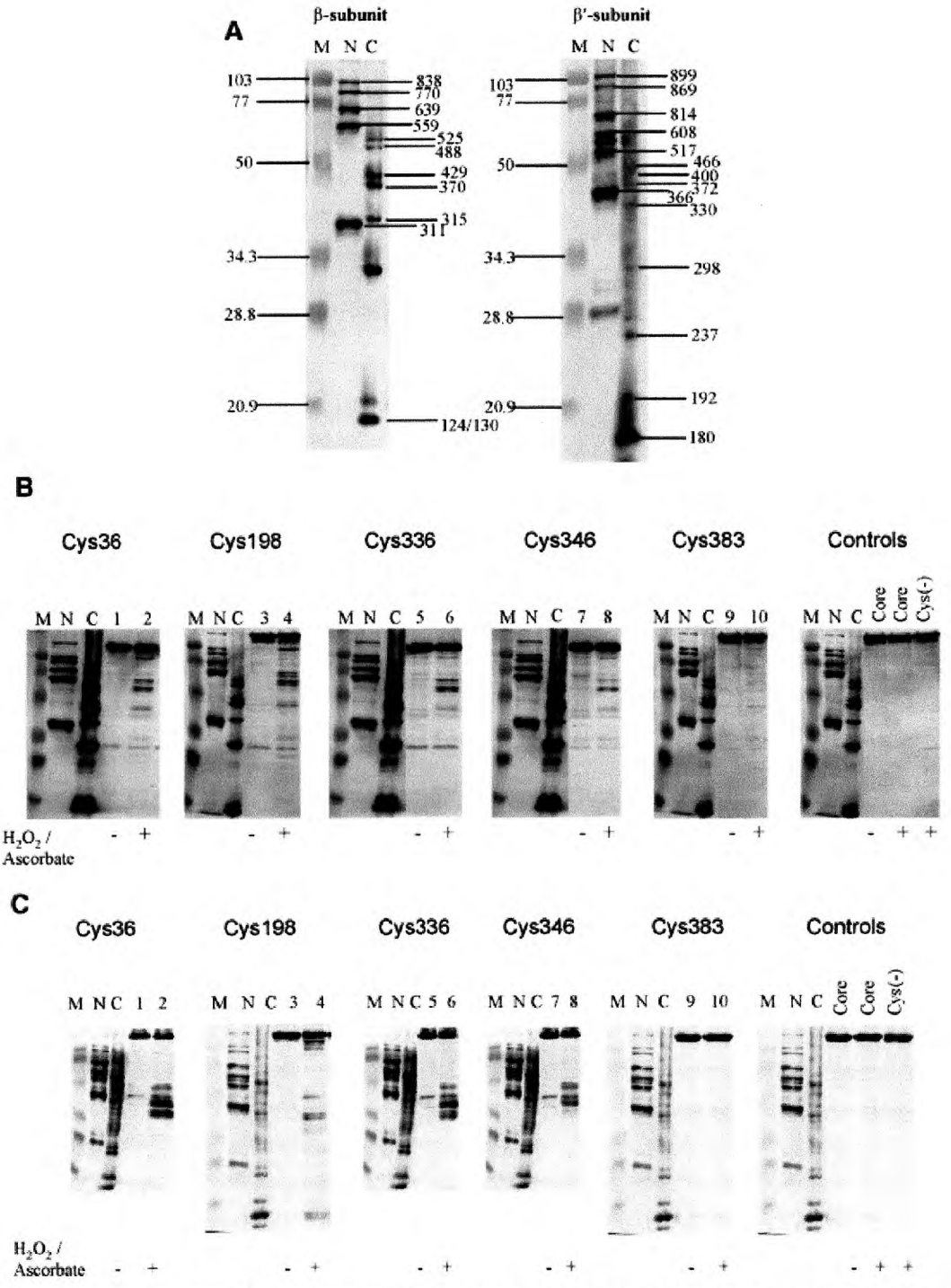


Fig. 4. (**A**) Immunoblot of markers generated by treating core RNA polymerase with NTCB (N) and CNBr (C), respectively, using anti-β or anti-β′ N-terminus-specific antibodies. Pre-stained molecular weight markers (M; Bio-Rad, Low Range SDS–PAGE Standards) were used as reference molecular weight markers. Holoenzyme cutting by σ^{54} single cysteine–Fe-BABE conjugates. Immunostained blots of SDS–PAGE detecting either (**B**) β or (**C**) β′ subunit fragments by using affinity-purified N-terminal antibodies. All reactions were either treated (+) or untreated (–) with ascorbate and hydrogen peroxide.

(amino acids 120–215) in region III of σ^{54} that has a strong affinity for the core RNA polymerase (Gallegos and Buck, 1999), while Cys346 is located within a region (amino acids 325–340) that is weakly protected by core RNA polymerase during hydroxyl radical cleavage (Casaz and Buck, 1999), consistent with the differences in the β and β′ cleavage patterns generated by the Cys198- and Cys346-conjugated σ^{54} proteins.

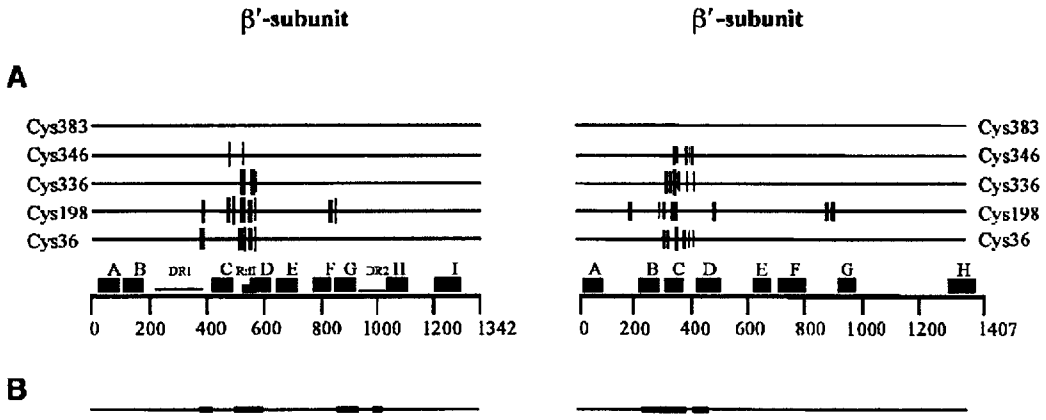


Fig. 5. (A) Summary of the cleavage data in Figure 4 showing the proximities of σ^{54} residues to surfaces on the β and β′ subunits. The thickness of the lines indicates the cleavage efficiency observed on immunoblots. (B) Surfaces (shaded boxes) proximal to σ^{70} as determined by Fe-BABE cleavage studies (Owens *et al.*, 1998). Conserved regions in β and β′ are indicated along the horizontal axes.

Cys 36 and Cys336. Next, we used the activator-independent single cysteine mutants of σ^{54} with Fe-BABE conjugated to Cys36 (E36C) and Cys336 (R336C) in the Cys(–)σ^{54} background to map their interaction sites on the β and β′ subunits of core RNA polymerase. Cys36 is in region I of σ^{54}, which functions to keep the closed complexes in a transcriptionally silent state in the absence of activation (Gallegos *et al.*, 1999) and has been shown to have weak affinity for core RNA polymerase (Gallegos and Buck, 1999). The core RNA polymerase cleavage by Cys36 shows that region I interacts strongly with β and β′ subunits. Cleavage sites map to within the Rif1 cluster and the conserved region D on β and within conserved region C on β′ (Cys36, lane 2, Figure 4B and C). Cys336 is located within a sequence in σ^{54} that lies outside the minimal DNA-binding domain, but nevertheless cross-links to DNA (Chaney and Buck, 1999; M.K.Chaney, M.S.Pitt and M.Buck, submitted). Remarkably, the cleavage data indicate that Cys336 interaction sites on both β and β′ subunits unambiguously overlap with those of Cys36. Given the shared activator-independent transcription phenotypes of the E36C and R336C σ^{54} mutants, we rationalize this observation by suggesting that E36 and R336 belong to an interface between σ^{54} and core RNA polymerase that acts to prevent polymerase isomerization in the absence of activators, and propose that the activator probably functions in changing the interface.

Cys383. Attempts to cleave core RNA polymerase by Cys383 (R383C)-conjugated σ^{54} revealed no discernible cleavage of either the β or β′ subunit (Cys383, lane 10, Figure 4B and C, respectively). In σ^{54}, Cys383 is proposed to fall within the recognition helix (helix 2) of the putative helix–turn–helix motif and to contribute to the recognition of the –12 promoter element by the σ^{54}–holoenzyme (Coppard and Merrick, 1991; Merrick and Chambers, 1992). We can conclude that Cys383 is not located proximal (at least not within a radius of 12 Å) to the core RNA polymerase subunits β and β′ and may be facing promoter DNA rather than core.

A comparison with σ^{70}–core RNA polymerase proximity. Because binding of σ^{54} to core RNA polymerase results in the formation of a holoenzyme that has very distinct functional properties compared with the σ^{70}–holoenzyme, we compared our cleavage data with those obtained for σ^{70}–holoenzyme cleavage by Fe-BABE-modified σ^{70} (Owens *et al.*, 1998; Figure 5). Notably, cleavage by Cys198 is within the β′ subunit conserved region G, whereas an interaction of σ^{70} with this region has not been observed. Several lines of evidence show that the N-terminal portion of the β′ subunit is involved in the specific binding of the σ^{70} subunit (Luo *et al.*, 1996; Arthur *et al.*, 1998; Owens *et al.*, 1998; Katayama *et al.*, 2000). The interaction made by Cys198 within conserved β′ region G may well be σ^{54} specific. Interestingly, a part of the sequence in β′ involved in chelating Mg^{2+} to the active centre of RNA polymerase is in conserved region G (Zaychikov *et al.*, 1996). We note that Cys198 is also proximal to the β′ rudder, a region distant from the active centre of the RNA polymerase (Zhang *et al.*, 1999; Figure 5A), suggesting significant conformational changes occurring in β′ following σ binding. Remarkably, other surfaces of the core RNA polymerase proximal to the two classes of σ factors appear to overlap greatly (residues 480–600 and around residue 900 ± 10 in the β subunit, and between residues 200 and 350 ± 20 in the β′ subunit). Some are proximal to the parts of the RNA polymerase catalytic centre, contributed by β conserved regions D and H and β′ conserved regions D and H (Mustaev *et al.*, 1997). Locating the proximity sites of σ^{54} and σ^{70} within the recently resolved crystal structure of *Thermus aquaticus* core RNA polymerase (Zhang *et al.*, 1999) serves to illustrate further the degree of overlap between the core surfaces that are proximal to the two classes of σ factors (Figure 6).

Discussion

The tethered hydroxyl radical footprinting methodology using Fe-BABE is a powerful tool to study protein–protein interactions that occur during transcription, and has been applied successfully previously (Miyake *et al.*, 1998; Owens *et al.*, 1998). It is particularly well suited to the study of proximity relationships of a single site on one

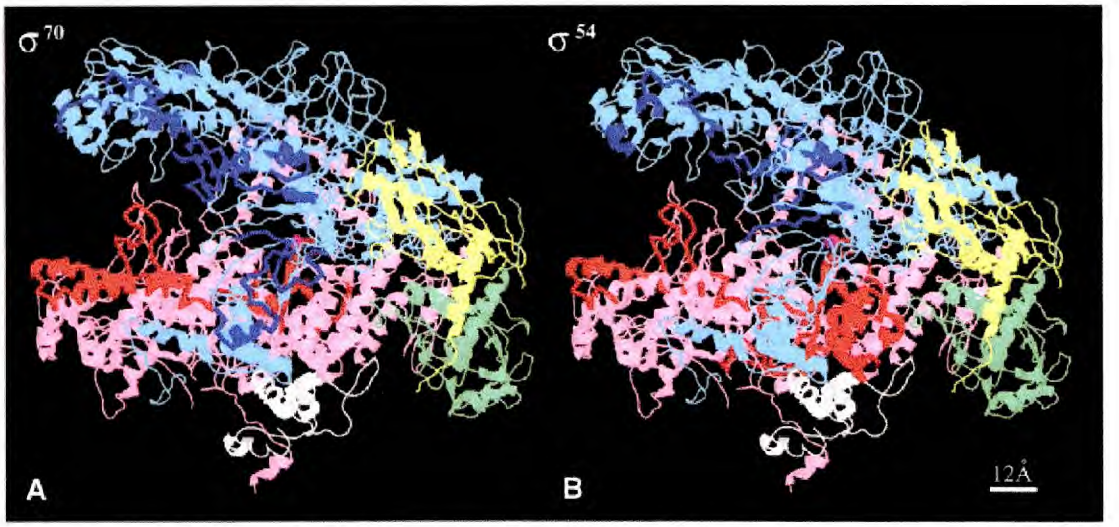


Fig. 6. RIBBONS diagrams of the three-dimensional structure of core RNA polymerase from *T.aquaticus* (Zhang *et al.*, 1999). The β subunit is shown in pink and the β′ subunit in blue. The two α subunits are shown in green and yellow. (**A**) σ^{70}- and (**B**) σ^{54}-interacting surfaces on the β and β′ subunits are indicated in blue and red, respectively.

subunit to one or more other sites. The results can add considerably to our understanding of the overall architecture of large macromolecular complexes. We have used the Fe-BABE cleavage methodology to look at proximity relationships between functional domains of σ^{54} and the core RNA polymerase subunits β and β′, and to widen our understanding of this specialized form of bacterial enhancer-dependent transcription.

Residue 36 in region I

Our cleavage data for Cys36 are consistent with the demonstrated interaction of region I sequences with core RNA polymerase (Casaz and Buck, 1997, 1999; Gallegos and Buck, 1999). Partial proteolysis studies on the σ^{54}–holoenzyme have shown that an N-terminal region of σ^{54} spanning residues 36–100 is protected from proteolytic cleavage (Casaz and Buck, 1997) and free hydroxyl radical footprinting experiments have shown that core RNA polymerase protects a region of σ^{54} comprising residues 36–140 from hydroxyl radical cleavage (Casaz and Buck, 1999). Even though region I of σ^{54} is dispensable for the binding of both the core RNA polymerase and DNA, it is essential in mediating the response to activator proteins (Sasse-Dwight and Gralla, 1990; Hsieh and Gralla, 1994; Hsieh *et al.*, 1994; Wong *et al.*, 1994; Cannon *et al.*, 1995; Casaz *et al.*, 1999). The proximity of region I E36 to core RNA polymerase provides further physical evidence consistent with the view that region I interacts with core RNA polymerase to exert some of its effects, including the inhibition of polymerase isomerization and the stabilization of holoenzyme on melted DNA (Cannon *et al.*, 1999; Gallegos and Buck, 1999, 2000). It seems that each of these functions is achieved in part through interacting with core RNA polymerase, but we cannot say precisely whether proximity relationships revealed by the Fe-BABE data correlate with one or both roles of region I. The interface of σ factors with core RNA polymerase is now recognized as being functionally specialized, contributing to more than just anchoring core and σ together, and is likely to contribute to communicating information about promoter conformation to the core RNA polymerase (Gross *et al.*, 1998; Sharp *et al.*, 1999).

Residue 336 in region III

We previously concluded that an interaction between the N- and C-terminal sequences of σ^{54} in the holoenzyme is closely associated with inhibiting polymerase isomerization (Chaney and Buck, 1999), consistent with the view that the σ^{54} region I sequences may be in communication with other sequences in σ^{54} (Cannon *et al.*, 1995; Casaz and Buck, 1999), probably including the major C-terminal DNA-binding domain (Wong *et al.*, 1994; Taylor *et al.*, 1996; Guo and Gralla, 1997; Guo *et al.*, 1999). The recently discovered role of residue 336 in region III of σ^{54} in maintaining closed complexes in a transcriptionally silent state in the absence of activation (Chaney and Buck, 1999), and the unambiguous similarity of Cys336 cleavage data to those of Cys36, together argue that residues around 336 contribute to the same interface with core RNA polymerase. Our Fe-BABE cleavage data provide indirect structural evidence for an interaction between regions I and III occurring via the core RNA polymerase. However, the similarity in the cleavage patterns seen with Cys36 and Cys336 must also be viewed in the context that both mutants (E36C and R336C) form the deregulated holoenzyme conformation, in which σ^{54} regions I and III may be proximal to each other, hence providing similar cleavage patterns. Thus, we do not discount the possibility that regions I and III may normally only interact with each other following polymerase isomerization after activation, or at some earlier intermediate step en route to the open complex.

Residue 198 in region III

Residue 198 seems to be able to establish two proximities to the β subunit. According to the *T.aquaticus* core RNA polymerase structure, these are well separated in space (Zhang *et al.*, 1999). Possibly, binding of σ^{54} results in two different holoenzyme conformers, or movement in core RNA polymerase upon the binding of σ^{54} brings the two sites closer together. Inspection of the native gel showing core RNA polymerase binding of the Fe-BABE-modified Cys198 σ^{54} supports the former suggestion. We also note that core RNA polymerase binding of a small fragment of σ^{54} (amino acids 120–215) resulted in two discrete complexes (Gallegos and Buck, 1999), consistent with the two binding modes suggested by the Cys198 Fe-BABE cleavage data. The 120–215 amino acid fragment seems to contain the major high affinity determinant for core RNA polymerase binding. A short σ^{70} similarity sequence of σ^{54} important for core RNA polymerase binding is found between residues 175 and 189 (Tintut and Gralla, 1995). Kinetic studies have shown that initial binding of σ^{54} to core RNA polymerase is followed by a slower rearrangement of the holoenzyme (D.J.Scott, A.L.Ferguson, M.T.Gallegos, M.S.Pitt and J.G.Hoggett, unpublished data). The relationship of this conformational change to interactions made by sequences around Cys198 remains to be determined. Mutational analysis around position Cys198 has shown that the integrity of this σ^{54} sequence is important for a holoenzyme stability that is conditional upon particular pre-melted DNA sequences (M.S.Pitt, M.T.Gallegos and M.Buck, in preparation). Communication of promoter structure with the core RNA polymerase through a patch in σ^{54} within the 120–215 fragment involving Cys198 seems likely. Intermolecular interactions within region III that enhance DNA binding of σ^{54} are evident (Cannon *et al.*, 1997).

Residue 346 in region III

C346 shows a similar proximity relationship to core RNA polymerase as does residue 336; however, substitutions in C346 do not result in deregulation of RNA polymerase isomerization. The cutting by Fe-BABE at 346 was weaker than that at 336, indicating a greater distance to core RNA polymerase and/or geometric constraint on positioning the Fe-BABE on core RNA polymerase.

Residue 383 in region III

The failure of the Fe-BABE derivative at position 383 to cleave core RNA polymerase is consistent with a role in DNA binding *per se* rather than in interacting with core RNA polymerase. DNA interaction studies with Fe-BABE at position 383 might reveal such a role.

Overview

In conclusion, our Fe-BABE data provide a framework for interpretation of some of the genetic and biochemical data on σ^{54}. The most remarkable observation is the unambiguous similarity in the core RNA polymerase surfaces proximal to the activator-independent and activator-dependent σ classes. This observation is supported further by small-angle X-ray scattering data on σ^{54}, which suggest that although σ^{54} and σ^{70} are unrelated by primary amino acid sequence, they both share a significant overall structural similarity (Svergun *et al.*, 2000), and by protein footprinting studies (Traviglia *et al.*, 1999). Possibly, the two different σ classes occupy similar positions in the core RNA polymerase and use some overlapping points of interaction to contribute to the shared functionalities of the holoenzymes. For transcription by σ^{54}, the progression from the closed to the open promoter complex must occur along one or more pathways in which changing interactions between σ and DNA and σ and core RNA polymerase occur to establish the alternative functional states of the holoenzyme. It seems doubtless that some of the proximity relationships we have detected will change upon activation of the σ^{54}–holoenzyme. Those associated with σ^{54} residues 36 and 336 seem most likely to change, given that these are associated with σ^{54} sequences needed to maintain the stable closed promoter complex and that these, when mutated, lead to activator-independent isomerization and binding to pre-melted DNA. It seems reasonable to assume that one or more of the σ–core and σ–DNA interactions associated with maintaining the stable closed complex conformation must be changed by activator to initiate the holoenzyme isomerization. How activator and its nucleotide hydrolysis trigger these changes remains to be determined.

The proximity relationships we have determined with Fe-BABE-modified σ^{54} proteins fall into two different classes with respect to the phenotypes of the σ^{54} single cysteine mutants. The σ^{54} Fe-BABE derivatives at Cys36 and Cys336 show deregulated transcription, whereas the Fe-BABE derivatives at Cys198 and Cys346 holoenzymes do not. Therefore, it is formally possible that the proximity relationships deduced apply to two different conformers and therefore to two functional states of the holoenzyme. In contrast to the Cys198 and Cys346 holoenzymes, Cys36 and Cys336 holoenzymes may have a conformation that could resemble more the conformation of the activated σ^{54}–holoenzyme in open promoter complexes. The extent to which the conformations of regulated and deregulated σ^{54} holoenzymes differ and how conformations change upon interacting with DNA remains to be determined. It is interesting that Cys198 is proximal to sequences in the β′ region G that are involved in maintenance of the active site of the RNA polymerase.

Materials and methods

Site-directed mutagenesis

The pET28b(+) (Novagen)-based plasmid pMTHσ^{N} (Gallegos and Buck, 1999), which directs the synthesis of the N-terminal His_6-tagged σ^{54} from *K.pneumoniae*, was used as the template for construction of single cysteine mutants of σ^{54} using the Quickchange mutagenesis kit (Stratagene). Briefly, mutated DNA was synthesized by *Pfu* DNA polymerase in a reaction that contains pMTHσ^{N} and a large molar excess of complementary mutagenic oligonucleotides. Following 30 cycles of heating at 95°C for 1 min, 50°C for 1 min and 72°C for 14 min, the reactions were treated with *DpnI* to remove the parental DNA. The resulting reaction mix was used to transform *E.coli* JM109 cells. Mutant clones were verified by DNA sequencing. Initially, the two single cysteine σ^{54} mutants, C198A (pSRW198) and C346A (pSRW346), were constructed by replacing cysteine residues at positions Cys198 and Cys346 with alanine. An *NdeI–BamHI* fragment from pSRW198 carrying the C198A mutation and a *BamHI–HindIII* fragment from pSRW346 carrying C346A were cloned into pET28b(+) (Novagen) to create pSRWCys(–). For the introduction of new cysteine residues at positions 36, 336 and 383 in σ^{54}, pSRWCys(–) was used as the template plasmid in the site-directed mutagenesis reactions to create pSRW36, pSRW336 and

pSRW383. The DNA sequences of each single cysteine σ^{54} mutant construct and that of cys(–)*rpoN* were confirmed by DNA sequencing.

Protein overproduction and purification

The plasmids directing the overproduction of Cys(–)σ^{54} and the single cysteine mutants were overexpressed in *E.coli* B834(DE3) cells. Proteins were purified by nickel affinity chromatography and eluted with an imidazole gradient essentially as described previously for wild-type σ^{54} (Gallegos and Buck, 1999). All purified proteins were dialysed against TGED buffer [10 mM Tris–HCl pH 8.0), 5% (v/v) glycerol, 0.1 mM EDTA, 1 mM dithiothreitol (DTT)] to which 50% (v/v) glycerol and 50 mM NaCl were added, and stored at –70°C (long-term storage) or –20°C (short-term storage). The activator protein PspFΔHTH used in the *in vitro* transcription assays was purified and stored essentially as described (Jovanovic *et al.*, 1996). Core RNA polymerase was purified from *E.coli* W3350 cells (Fujita *et al.*, 1987; Kusano *et al.*, 1996).

Conjugation of σ^{54} single cysteine mutants with Fe-BABE

With the exception of C346A, each purified σ^{54} protein was dialysed overnight at 4°C against conjugation buffer [10 mM MOPS pH 8.0, 0.2 M NaCl, 5% (v/v) glycerol, 2 mM EDTA]. Conjugation reactions were initiated by adding a 10-fold molar excess of Fe-BABE (300 μM) (Dojindo Chemicals, Japan) to each σ^{54} mutant (20 μM) at pH 8.0. After 1 h incubation at 37°C, excess Fe-BABE was removed by overnight dialysis against storage buffer (without DTT) at 4°C. The concentration of free cysteine sulfhydryl groups on the conjugated and unconjugated σ^{54} mutants and Cys(–)σ^{54} was determined fluorometrically by the CPM test (Parvari *et al.*, 1983) and the conjugation yield was calculated as described previously (Greiner *et al.*, 1997). The C346A preparation in storage buffer was dialysed overnight at 4°C against conjugation buffer containing 6 M deionized urea and treated essentially as described above.

Core RNA polymerase-binding assays

Escherichia coli core RNA polymerase (250 nM) and different amounts of unconjugated and conjugated mutant σ^{54} proteins were mixed together in Tris–NaCl buffer [40 mM Tris–HCl pH 8.0, 10% (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT, 100 mM NaCl] and incubated at 30°C for 10 min, followed by the addition of glycerol bromophenol blue loading dye. Samples were loaded onto native 4.5% polyacrylamide Bio-Rad Mini-Protean II gels and run at 50 V for 2 h at room temperature in Tris–glycine buffer (25 mM Tris, 200 mM glycine). Proteins were visualized by Coomassie Blue staining.

In vitro transcription assays

Supercoiled pFC50-m12 (Claverie-Martin and Magasanik, 1992) containing the *E.coli glnHp2*-m12 promoter was used as the template for the *in vitro* transcription assays. Activator-dependent transcription was performed essentially as described previously (Chaney and Buck, 1999), except that 30 nM σ^{54} holoenzyme, 10 nM template DNA (30 nM core RNA polymerase:120 nM σ^{54}) were used. For activation, 4 μM PspFΔHTH was added with 4 mM ATP to the final reaction volume of 10 μl. The reactions were incubated for 20 min to allow open complexes to form. The remaining rNTPs (CTP and GTP at 0.1 mM and UTP at 0.05 mM), 1.5 μCi of [α-^{32}P]UTP and heparin (100 μg/ml) were added and incubated for a further 20 min at 30°C. The reactions were stopped with 4 μl of formamide loading buffer and, after heating for 5 min at 95°C, 7 μl of the samples were loaded on a 6% denaturing sequencing gel. The dried gel was analysed on a phosphoimager. The activator-independent transcription assay was performed essentially as described above, but in the absence of the PspFΔHTH activator protein. The *E.coli glnHp2*-m12 start sequence is TGTCAC (+1 to +6). To allow activator-independent initiated complex formation, 4 mM GTP, 0.1 mM UTP and 0.1 mM CTP were incubated with the holoenzyme prior to challenge with heparin and the addition of 0.1 mM ATP and 1.5 μCi of [α-^{32}P]UTP.

Core RNA polymerase cleavage by Fe-BABE-conjugated σ^{54}

The *E.coli* RNA polymerase holoenzymes were prepared by incubating core RNA polymerase with Fe-BABE-conjugated σ^{54} mutants and Cys(–)σ^{54} (1:1 molar ratio) at 30°C for 10 min in cleavage buffer [10 mM MOPS pH 8.0, 10 mM $MgCl_2$, 10% (v/v) glycerol, 200 mM NaCl, 2 mM EDTA]. Cleavage reactions were initiated by the rapid sequential addition of freshly prepared ascorbate (5 mM final) and hydrogen peroxide (5 mM final) and were allowed to proceed for 2 min at 30°C. Reactions were stopped with 5× SDS sample buffer [62.5 mM Tris–HCl pH 8.2, 2% (w/v) SDS, 5% (v/v) 2-mercaptoethanol, 10% (v/v) glycerol, 25 mM EDTA, 0.02% (w/v) bromophenol blue] and immediately frozen in liquid nitrogen and stored at –70°C. The cleaved fragments were separated, blotted, and visualized by immunostaining with affinity-purified subunit β and β′ N-terminus-specific antibodies essentially as described previously (Greiner *et al.*, 1996).

Assignment of cleavage sites on β or β′ subunits

The β and β′ subunits were cleaved at either cysteine or methionine residues using the sequence-specific chemical reagents NTCB and CNBr, respectively. NTCB cleavage was performed as described previously (Jacobson *et al.*, 1973). For the NTCB cleavage, core RNA polymerase was buffer exchanged into unfolding buffer (0.1 M MOPS pH 8.5, 8 M urea) and incubated for 10 min at 37°C. Cleavage was initiated by adding a 5-fold molar excess of NTCB (in 0.1 M MOPS buffer) over total sulfhydryl groups. After overnight incubation at 37°C, the cleavage reactions were stopped with 5× SDS sample buffer [125 mM Tris–HCl pH 6.8, 4% (w/w) SDS, 20% (v/v) glycerol, 10% (v/v) 2-mercapto-ethanol, 0.004% (w/v) bromophenol blue] and analysed by SDS–PAGE and immunostaining with β and β′ N-terminus-specific antibodies. CNBr cleavage was performed essentially as in Grachev *et al.* (1989). Initially, core RNA polymerase was treated with 10% SDS and incubated at 37°C for 30 min. For denaturation, 1 M HCl and 1 M CNBr in acetonitrile were then added and incubated at room temperature for 6 h. The reaction was stopped with 5× SDS sample buffer and analysed as described above. The cleavage sites for each NTCB- and CNBr-generated fragment of β and β′ were assigned by using a third-order polynomial fit of log molecular weight versus the migration distance on SDS–PAGE of a known set of marker fragments and by β and β′ amino acid sequence analysis.

Acknowledgements

We thank Annie Kolb for communicating data prior to publication, members of the M.B. and A.I. laboratories for their valuable suggestions and contributions to the project, and Seth Darst for making available the co-ordinates of the *T.aquaticus* core RNA polymerase used in Figure 6. Work in M.B.'s laboratory was supported by a CEC Biotechnology grant BIO4-CT97-2143 and Biotechnology and Biological Sciences Research project grant, in A.I.'s laboratory by a CREST grant of Japan Science and Technology Corporation, and Grants-in-aid from the Ministry of Education, Science and Culture of Japan. A postgraduate studentship to S.R.W. was from the LEA of Karlsruhe, Germany.

References

Arthur,T.M. and Burgess,R.R. (1998) Localization of a σ^{70} binding site on the N-terminus of the *Escherichia coli* RNA polymerase β′ subunit. *J. Biol. Chem.*, **273**, 31381–31387.

Cannon,W., Claverie-Martin,F., Austin,S. and Buck,M. (1994) Identification of a DNA-contacting surface in the transcription factor σ-54. *Mol. Microbiol.*, **11**, 227–236.

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

Cannon,W., Chaney,M.K., Wang,X. and Buck,M. (1997) Two domains within σ^{N} (σ^{54}) cooperate for DNA binding. *Proc. Natl Acad. Sci. USA*, **94**, 5006–5011.

Cannon,W., Gallegos,M.T., Casaz,P. and Buck,M. (1999) Amino-terminal sequences of σ^{N} (σ^{54}) inhibit RNA polymerase isomerization. *Genes Dev.*, **13**, 357–370.

Casaz,P. and Buck,M. (1997) Probing the assembly of transcription initiation complexes through changes in σ^{N} protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.

Casaz,P. and Buck,M. (1999) Region I modifies DNA-binding domain conformation of σ^{54} within the holoenzyme. *J. Mol. Biol.*, **285**, 507–514.

Casaz,P., Gallegos,M.T. and Buck,M. (1999) Systematic analysis of σ^{54} N-terminal sequences identifies regions involved in positive and negative regulation of transcription. *J. Mol. Biol.*, **292**, 229–239.

Chaney,M.K. and Buck,M. (1999) The σ^{54} DNA-binding domain includes a determinant of enhancer responsiveness. *Mol. Microbiol.*, **33**, 1200–1209.

Claverie-Martin,F. and Magasanik,B. (1992) Positive and negative effects of DNA bending on activation of transcription from a distant site. *J. Mol. Biol.*, **227**, 996–1008.

Coppard,J.R. and Merrick,M.J. (1991) Cassette mutagenesis implicates a

helix–turn–helix motif in promoter recognition by the novel RNA polymerase σ factor σ^{54}. *Mol. Microbiol.*, **5**, 1309–1317.

Debarbouille,M., Martin-Verstraete,I., Kunst,F. and Rapoport,G. (1991) The *Bacillus subtilis sigL* gene encodes an equivalent of σ^{54} from gram-negative bacteria. *Proc. Natl Acad. Sci. USA*, **88**, 9092–9096.

Fujita,N., Nomura,T. and Ishihama,A. (1987) Promoter selectivity of *Escherichia coli* RNA polymerase. Purification and properties of holoenzyme containing the heat-shock σ subunit. *J. Biol. Chem.*, **262**, 1855–1859.

Gallegos,M.T. and Buck,M. (1999) Sequences in σ^{N} determining holoenzyme formation and properties. *J. Mol. Biol.*, **288**, 539–553.

Gallegos,M.T. and Buck,M. (2000) Sequences in σ^{54} region I required for binding to early melted DNA and their involvement in σ–DNA interactions. *J. Mol. Biol.*, **297**, 849–859.

Gallegos,M.T., Cannon,W. and Buck,M. (1999) Functions of the σ^{54} region I *in trans* and implications for transcription activation. *J. Biol. Chem.*, **274**, 25285–25290.

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

Greiner,D.P., Hughes,K.A., Gunasekera,A.H. and Meares,C.F. (1996) Binding of the σ^{70} protein to the core subunits of *Escherichia coli* RNA polymerase, studied by iron-EDTA protein footprinting. *Proc. Natl Acad. Sci. USA*, **93**, 71–75.

Greiner,D.P., Miyake,R., Moran,J.K., Jones,A.D., Negishi,T., Ishihama,A. and Meares,C.F. (1997) Synthesis of the protein cutting reagent iron (*S*)-1(*p*-bromoacetamidobenzyl)-ethylenediaminetetraacetate and conjugation to cysteine side chains. *Bioconj. Chem.*, **8**, 44–48.

Gross,C.A., Chan,C., Dombroski,A., Gruber,T., Sharp,M., Tupy,J. and Young,B. (1998) The functional and regulatory roles of σ factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 141–155.

Guo,Y. and Gralla,J.D. (1997) DNA-binding determinants of σ^{54} as deduced from libraries of mutations. *J. Bacteriol.*, **179**, 1239–1245.

Guo,Y., Wang,L. and Gralla,J.D. (1999) A fork junction DNA-binding switch that controls promoter melting by the bacterial enhancer-dependent σ factor. *EMBO J.*, **18**, 3736–3745.

Hsieh,M. and Gralla,J.D. (1994) Analysis of the N-terminal leucine heptad and hexad repeats of σ^{54}. *J. Mol. Biol.*, **239**, 15–24.

Hsieh,M., Tintut,Y. and Gralla,J.D. (1994) Functional roles for the glutamines within the glutamine-rich region of the transcription factor σ^{54}. *J. Biol. Chem.*, **269**, 373–378.

Ishihama,A. (2000) Molecular anatomy of RNA polymerase using protein-conjugated metal probes with nuclease and protease activities. *Chem. Commun.*, in press.

Jacobson,G.R., Schaffer,M.H., Stark,G.R. and Vanaman,T.C. (1973) Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues. *J. Biol. Chem.*, **248**, 6583–6591.

Jishage,M. and Ishihama,A. (1997) Variation in RNA polymerase σ subunit composition within different stocks of *Escherichia coli* W3110. *J. Bacteriol.*, **179**, 959–963.

Jovanovic,G., Weiner,L. and Model,P. (1996) Identification, nucleotide sequence and characterization of PspF, the transcriptional activator of the *Escherichia coli* stress-induced *psp* operon. *J. Bacteriol.*, **178**, 1936–1945.

Katayama,A., Fujita,N. and Ishihama,A. (2000) Mapping of subunit–subunit contact surfaces on the β′ subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.*, **275**, 3583–3592.

Kusano,S., Ding,Q., Fujita,N. and Ishihama,A. (1996) Promoter selectivity of *Escherichia coli* RNA polymerase E σ^{70} and E σ^{38} holoenzymes. Effect of DNA supercoiling. *J. Biol. Chem.*, **271**, 1998–2004.

Leary,B.A., Ward-Rainey,N. and Hoover,T.R. (1998) Cloning and characterization of *Planctomyces limnophilus rpoN*: complementation of a *Salmonella typhimurium rpoN* mutant strain. *Gene*, **221**, 151–157.

Luo,J., Sharif,K.A., Jin,R., Fujita,N., Ishihama,A. and Krakow,J.S. (1996) Molecular anatomy of the β′ subunit of the *E.coli* RNA polymerase: identification of regions involved in polymerase assembly. *Genes Cells*, **1**, 819–827.

Merrick,M.J. (1993) In a class of its own—the RNA polymerase σ factor σ^{54} (σ^{N}). *Mol. Microbiol.*, **10**, 903–909.

Merrick,M. and Chambers,S. (1992) The helix–turn–helix motif of σ^{54} is involved in recognition of the –13 promoter region. *J. Bacteriol.*, **174**, 7221–7226.

Miyake,R., Murakami,K., Owens,J.T., Greiner,D.P., Ozoline,O.N., Ishihama,A. and Meares,C.F. (1998) Dimeric association of *Escherichia coli* RNA polymerase α subunits, studied by cleavage of single-cysteine α subunits conjugated to iron-(*S*)-1-[*p*-(bromoacetamido)benzyl]ethylenediaminetetraacetate. *Biochemistry*, **37**, 1344–1349.

Mustaev,A., Kozlov,M., Markovtsov,V., Zaychikov,E., Denissova,L. and Goldfarb,A. (1997) Modular organization of the catalytic center of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **94**, 6641–6645.

Owens,J.T., Miyake,R., Murakami,K., Chmura,A.J., Fujita,N., Ishihama,A. and Meares,C.F. (1998) Mapping the σ^{70} subunit contact sites on *Escherichia coli* RNA polymerase with a σ^{70}-conjugated chemical protease. *Proc. Natl Acad. Sci. USA*, **95**, 6021–6026.

Parvari,R., Pecht,I. and Soreq,H. (1983) A microfluorometric assay for cholinesterases, suitable for multiple kinetic determinations of picomoles of released thiocholine. *Anal. Biochem.*, **133**, 450–456.

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.

Sasse-Dwight,S. and Gralla,J.D. (1990) Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor σ^{54}. *Cell*, **62**, 945–954.

Sharp,M.M., Chan,C.L., Lu,C.Z., Marr,M.T., Nechaev,S., Merritt,E.W., Severinov,K., Roberts,J.W. and Gross,C.A. (1999) The interface of σ with core RNA polymerase is extensive, conserved and functionally specialized. *Genes Dev.*, **13**, 3015–3026.

Svergun,D.I., Malfois,M., Koch,M.H.J., Wigneshweraraj,S.R. and Buck,M. (2000) Low resolution structure of the σ^{54} transcription factor revealed by X-ray solution scattering. *J. Biol. Chem.*, **275**, 4210–4214.

Syed,A. and Gralla,J.D. (1997) Isolation and properties of enhancer-bypass mutants of σ^{54}. *Mol. Microbiol.*, **23**, 987–995.

Syed,A. and Gralla,J.D. (1998) Identification of an N-terminal region of σ^{54} required for enhancer responsiveness. *J. Bacteriol.*, **180**, 5619–5625.

Taylor,M., Butler,R., Chambers,S., Casimiro,M., Badii,F. and Merrick, M. (1996) The RpoN-box motif of the RNA polymerase σ factor σ^{N} plays a role in promoter recognition. *Mol. Microbiol.*, **22**, 1045–1054.

Tintut,Y. and Gralla,J.D. (1995) PCR mutagenesis identifies a polymerase-binding sequence of σ^{54} that includes a σ^{70} homology region. *J. Bacteriol.*, **177**, 5818–5825.

Traviglia,S.L., Datwyler,S.A., Yan,D., Ishihama,A. and Meares,C.F. (1999) Targeted protein footprinting: where different transcription factors bind to RNA polymerase. *Biochemistry*, **38**, 15774–15778.

Wang,J.T., Syed,A., Hsieh,M. and Gralla,J.D. (1995) Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH_2-terminal leucine patch in σ^{54}. *Science*, **270**, 992–994.

Wang,J.T., Syed,A. and Gralla,J.D. (1997) Multiple pathways to bypass the enhancer requirement of σ^{54} RNA polymerase: roles for DNA and protein determinants. *Proc. Natl Acad. Sci. USA*, **94**, 9538–9543.

Weiss,D.S., Batut,J., Klose,K.E., Keener,J. and Kustu,S. (1991) The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell*, **67**, 155–167.

Wong,C., Tintut,Y. and Gralla,J.D. (1994) The domain structure of σ^{54} as determined by analysis of a set of deletion mutants. *J. Mol. Biol.*, **236**, 81–90.

Zaychikov,E. *et al.* (1996) Mapping of catalytic residues in the RNA polymerase active center. *Science*, **273**, 107–109.

Zhang,G., Campbell,E.A., Minakhin,L., Richter,C., Severinov,K. and Darst,S.A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell*, **98**, 811–824.

Received March 24, 2000; revised and accepted April 17, 2000

Reprinted from *J. Mol. Biol.* (2001) **306**, 681–701

JMB

Regulatory Sequences in Sigma 54 Localise Near the Start of DNA Melting

Siva R. Wigneshweraraj, Matthew K. Chaney, Akira Ishihama and Martin Buck

doi:10.1006/jmbi.2000.4393 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 306, 681–701

JMB

Regulatory Sequences in Sigma 54 Localise Near the Start of DNA Melting

Siva R. Wigneshweraraj[1], Matthew K. Chaney[1], Akira Ishihama[2] and Martin Buck[1]*

[1]*Imperial College of Science Technology and Medicine Department of Biology, Imperial College Road, London SW7 2AZ, UK*

[2]*Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411, Japan*

**Corresponding author*

Transcription initiation by the enhancer-dependent σ^{54} RNA polymerase holoenzyme is positively regulated after promoter binding. The promoter DNA melting process is subject to activation by an enhancer-bound activator protein with nucleoside triphosphate hydrolysis activity. Tethered iron chelate probes attached to amino and carboxyl-terminal domains of σ^{54} were used to map σ^{54}-DNA interaction sites. The two domains localise to form a centre over the −12 promoter region. The use of deletion mutants of σ^{54} suggests that amino-terminal and carboxyl-terminal sequences are both needed for the centre to function. Upon activation, the relationship between the centre and promoter DNA changes. We suggest that the activator re-organises the centre to favour stable open complex formation through adjustments in σ^{54}-DNA contact and σ^{54} conformation. The centre is close to the active site of the RNA polymerase and includes σ^{54} regulatory sequences needed for DNA melting upon activation. This contrasts systems where activators recruit RNA polymerase to promoter DNA, and the protein and DNA determinants required for activation localise away from promoter sequences closely associated with the start of DNA melting.

Keywords: DNA cleavage; DNA melting; σ^{54}-holoenzyme; RNA polymerase; regulatory centre

Introduction

Transcription initiation in bacteria, an intricate multistep process, is orchestrated by the multimeric DNA-dependent RNA polymerase (RNAP) enzyme consisting of the core subunits, $\alpha_2\beta\beta'$, and a σ factor subunit. In *Escherichia coli*, the RNAP σ subunit plays an essential role in promoter recognition and open complex formation. The seven σ factors identified in *E. coli* are categorised into two classes on the basis of sequence similarity and transcription mechanism. The σ^{70}-class includes the major σ factor, σ^{70}, that is responsible for most of the genes expressed during exponential cell growth, and the structurally and functionally related alternative σ factors (σ^{38}, σ^{32}, σ^{28}, σ^{24} and σ^{18}) that are responsible for expressing genes under various environmental conditions (reviewed by Ishihama, 1997). The second class is uniquely represented by the σ^{54} factor, which differs functionally from the σ^{70} class. The σ^{54}-holoenzyme binds promoters with the consensus YTGG-CACGrNNNTTGCW (Barrios *et al.*, 1999) containing conserved sequences (underlined) centred around −24 and −12 , respectively, from the transcription start site, in contrast to the −10 and −35 sequence recognition elements of the σ^{70}-holoenzyme family. σ^{54} endows its RNAP holoenzyme with a distinctive mechanism of transcription initiation (reviewed by Buck *et al.*, 2000). Isomerisation of the σ^{54}-holoenzyme-DNA closed complex to a transcription-competent open complex requires nucleoside triphosphate hydrolysis by an enhancer-bound activator protein and is mediated by direct activator-closed complex interactions (Cannon *et al.*, 2000; Kelly & Hoover, 2000; Lee & Hoover, 1995; Rippe *et al.*, 1997; Su *et al.* 1990; Wedel *et al.*, 1990). Since the rate-limiting step in transcription initiation by σ^{54}-holoenzyme is the isomerisation of the closed complex to an open one, the major function of σ^{54}-dependent activator

S.R.W. and M.K.C. contributed equally to this work.

Abbreviations used: RNAP, RNA polymerase; NTP, nucleoside triphosphate; FeBABE, iron [S]-1-[*p*-bromoacetamidobenzyl] ethylenediaminetetraacetate.

E-mail address of the corresponding author: m.buck@ic.ac.uk

0022-2836/01/040681–21 $35.00/0

proteins is not that of recruitment but to catalyse post-binding steps.

The σ^{54} protein can be divided into three regions (Regions I-III). The amino-terminal Region I is a regulatory domain, playing a central role in mediating response to activator proteins and consequent isomerisation of the σ^{54}-holoenzyme closed complex. Region II is an acidic linker connecting Region I and Region III, which contains sequences required for core RNAP binding (residues 120 to 215 in *Klebsiella pneumoniae*) and determinants for DNA interactions, including the helix-turn-helix motif (residues 366 to 386) and a patch that UV-crosslinks to promoter DNA (residues 329 to 346) (Chaney *et al.*, 2000; Gallegos & Buck, 1999; Chaney & Buck, 1999; Cannon *et al.*, 1994; Merrick & Chambers, 1992). The amino-terminal Region I is bifunctional: (i) it is required for inhibiting RNAP isomerisation and initiation in the absence of activation (Kelly & Hoover, 2000; Cannon *et al.*, 1999; Syed & Gralla, 1998; Wang *et al.*, 1995); and (ii) for aiding RNAP isomerisation and stable open complex formation upon activation (Casaz *et al.*, 1999; Syed & Gralla, 1998; Sasse-Dwight & Gralla, 1990). Holoenzymes containing σ^{54} without Region I ($\Delta I\sigma^{54}$) are active for so-called bypass transcription, in which the $\Delta I\sigma^{54}$ holoenzyme is able to isomerise and initiate transcription in the absence of activator proteins and nucleotide hydrolysis (Cannon *et al.*, 1999; Wang *et al.*, 1995, 1997). Open complexes formed with the $\Delta I\sigma^{54}$ holoenzyme are unstable (Wang *et al.*, 1995, 1997) and Region I mutants are unable to bind well to certain DNA structures thought to mimic early DNA melting in the DNA opening pathway (Gallegos & Buck, 2000; Guo *et al.*, 1999). The solvent accessibility of Region III changes in the holoenzyme when Region I is removed (Casaz & Buck, 1997, 1999). These observations suggest that Region I contributes to the DNA binding properties of σ^{54} possibly through an interaction with Region III (Guo *et al.*, 1999, 2000; Cannon *et al.*, 2000; Gallegos & Buck, 2000).

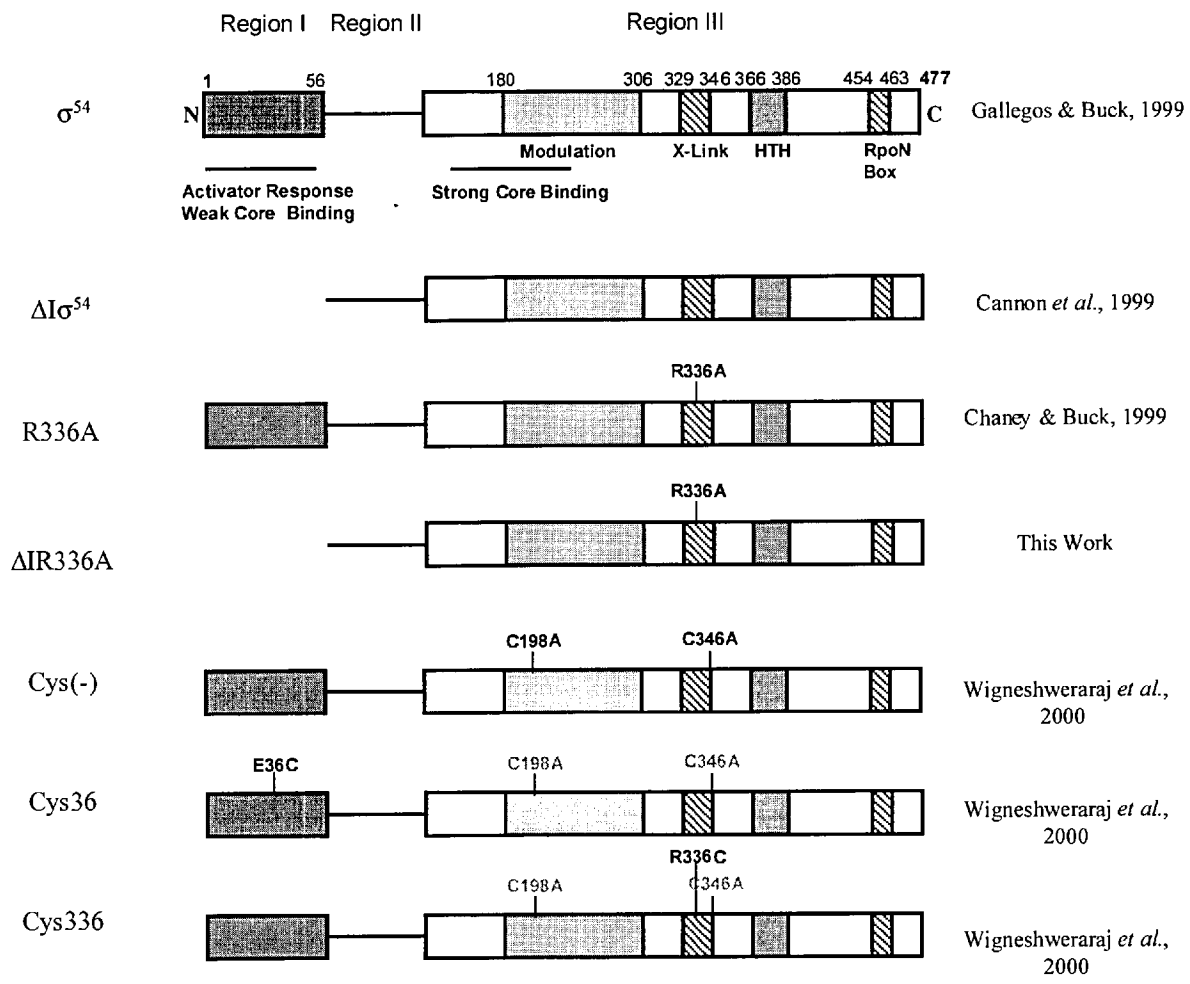
The Region III R336A mutant σ^{54} protein shares characteristics with some Region I mutants and the $\Delta I\sigma^{54}$ protein (Gallegos & Buck, 2000; Chaney & Buck, 1999; Cannon *et al.*, 1995, 1999; Casaz *et al.*, 1999; Hsieh & Gralla, 1994; Morris *et al.*, 1994; Wong *et al.*, 1994; Sasse-Dwight & Gralla, 1990). Mutating R336 (R336A) or removing Region I deregulates transcription *in vitro*, possibly due to disruption of a common or closely related interaction required to keep the RNAP holoenzyme silent for isomerisation in the absence of activation. Loss of stabilising interactions that normally restrict the conformation of the holoenzyme appear to allow unregulated isomerisation of the σ^{54} holoenzyme. However, R336A remains largely dependent upon activator to form stable open promoter complexes (Chaney & Buck, 1999).

Here, we constructed a Region I-deleted σ^{54} protein harbouring the R336A mutation (ΔIR336A) to further investigate the inter-related roles of Region I and Region III (Figure 1(a)). To begin to unravel the physical basis of the multiple σ^{54}-DNA interactions that contribute to maintaining the stable σ^{54}-holoenzyme promoter complex, and which are believed to be directed through the −12 promoter region (Chaney *et al.*, 2000; Guo *et al.*, 1999), we examined the proximity relationships of Region I and Region III to promoter DNA using a tethered iron chelate footprinting method. The FeBABE (iron [S]-1-[*p*-bromoacetamidobenzyl] ethylenediaminetetraacetate) conjugated σ^{54} proteins were made using single-cysteine σ^{54} mutants (Wigneshweraraj *et al.*, 2000). Sites for FeBABE conjugation were in the Region I (E36) and carboxyl-terminal, DNA-binding domain Region III (R336), at locations that partially deregulate the σ^{54}-holoenzyme *in vitro* but still retain activator responsiveness for stable open complex formation (Wigneshweraraj *et al.*, 2000; Chaney & Buck, 1999). *Sinorhizobium meliloti* promoter DNA complexes representing possible intermediates along the DNA melting pathway from a closed to the open promoter complex were analysed (Figure 1(b)). The results show that Region I and Region III FeBABE derivatives each lead to DNA cutting across the −12 promoter region. Cutting changes in an activator-dependent manner, and is influenced by core RNAP and DNA structure. Our results, together with related core RNAP proximity data (Wigneshweraraj *et al.*, 2000), led to the clear conclusion that a localised structure contributed to by the −12 region promoter DNA and at least two different structural elements of σ^{54} exists within a closed complex. It appears that this structure is intimately linked to the activation target of the enhancer binding proteins, or may be included within it, and can help explain the overlapping properties of mutants in Region I and Region III.

Results

Region I and R336 proximal DNA sequences on homoduplex, early and late melted heteroduplex DNA structures

To explore the relationship between Region I and Region III we attempted to identify promoter DNA sequences proximal to Region I and R336. In addition to double-stranded homoduplex DNA on which the closed complex forms, we used different DNA templates (Figure 1(b)) to represent structures formed when the nucleation of DNA melting has occurred in the closed complex (early melted DNA) or when extended melting is evident, as in the open complex (late melted DNA). The use of different template types allows formation of complexes that reflect the state of the DNA in initial, closed and open complexes to be examined, under non-activating and activating conditions. The cysteine-tethered chemical nuclease (FeBABE) allows the identification of DNA sites proximal to probe conjugation sites through the generation of local hydroxyl radicals that cleave the DNA



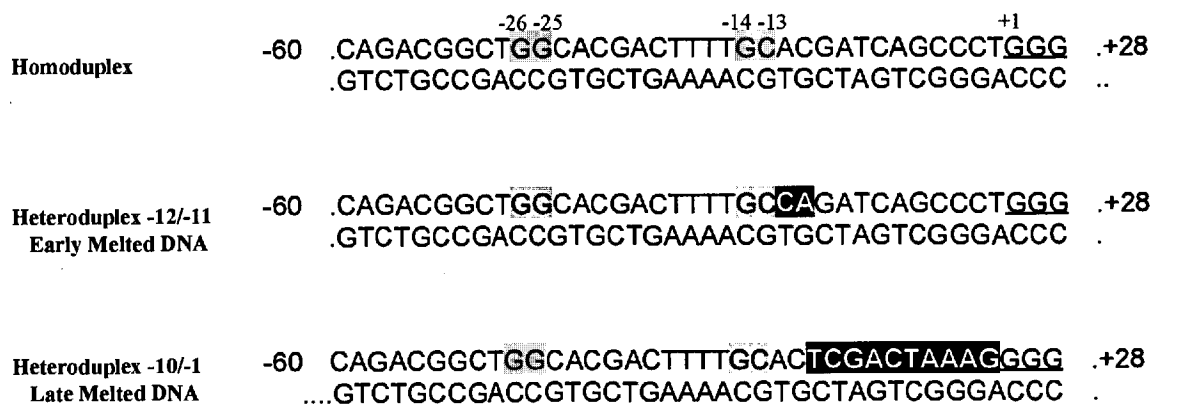


Figure 1. (a) *K. pneumoniae* σ^{54} proteins used in this work. (b) *S. meliloti nifH* homo- and heteroduplex promoter DNA probes used for the gel mobility shift and DNA cleavage assays. The consensus GG and GC of the σ^{54}-binding sites are highlighted and the mismatched regions are boxed in black.

(Ishihama, 2000; Bown *et al.*, 1999, 2000; Owens *et al.*, 1998; Murakami *et al.*, 1997a,b). Because σ^{54}, in contrast to σ^{70} and σ^{70}-related sigma factors, can bind to promoter DNA in the absence of core RNAP *in vitro*, the FeBABE footprinting methodology allows identification of DNA sites proximal

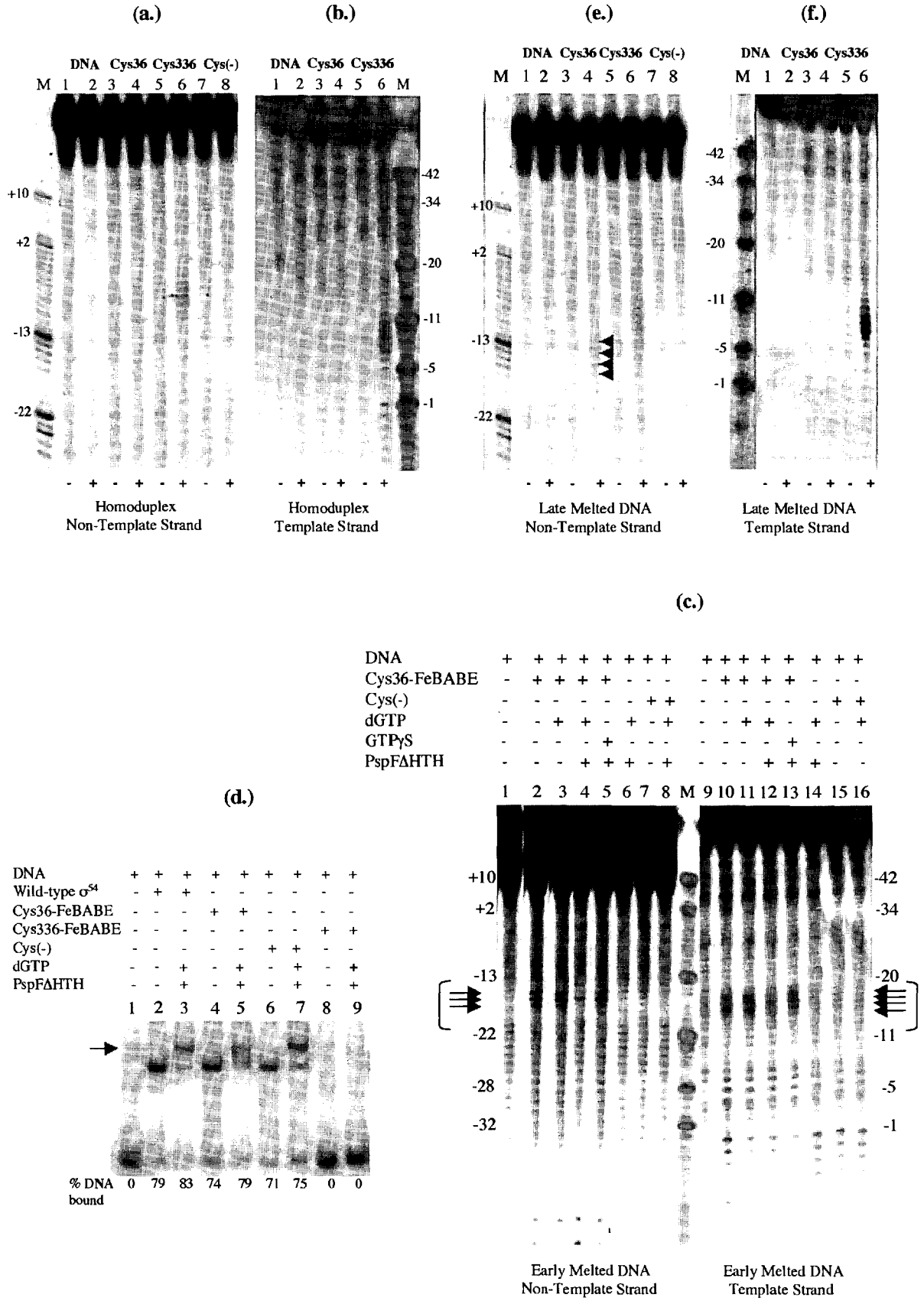


Figure 2 *(legend opposite)*

to the iron chelate probe within σ^{54}-DNA and σ^{54}-holoenzyme-DNA complexes (reviewed by Ishihama, 2000). For conjugation we chose E36 in Region I, because it is implicated in movements associated with activation (Casaz & Buck, 1997) and R336 in Region III (see Introduction). The single-cysteine mutants Cys36 (E36C) and Cys336 (R336C) remain dependent on activation for stable open promoter complex formation, demonstrating that productive interactions occur between the activator and the closed complex formed by these mutants (Wigneshweraraj *et al.*, 2000; Chaney *et al.*, 1999). The essential negative control reaction in the DNA cleavage assays involved a cysteine-free σ^{54} mutant (Cys(−)), which was subjected, together with the single cysteine mutants Cys36 and Cys336, to the FeBABE reagent (Wigneshweraraj *et al.*, 2000). Cleavage reactions conducted using the Cys(−) protein in parallel with the Cys36-FeBABE and Cys336-FeBABE resulted in no discernible cutting of the DNA, demonstrating that under the conditions used there is no artifactual DNA cleavage (see Figures 2-5 and summarised in Figure 6). Hence, we are confident that each DNA cut is attributed to the FeBABE specifically conjugated to either Cys36 or Cys336. Amongst the DNA cleavage experiments, the weakest cutting pattern was with Cys36-FeBABE σ^{54} on the early melted and late melted DNA probes where the background signal was highest. Nevertheless, quantification of the gels showed that cutting of DNA with FeBABE-conjugated proteins was clearly discernible above background. The results shown (see Figures 2-5) and summarised (see Figure 6) were confirmed with many replicate experiments. In general, we observed that the addition of small amounts of heparin (10μg/ml) to samples prior to initiating cleavage gave more discrete patterns of cutting, possibly through removing non-specific complexes, but the instability of the Cys36-FeBABE-DNA complex to heparin precluded use of this procedure in the experiments shown in Figure 2(c).

Promoter DNA sequences proximal to E36 (Region I) and R336 (Region III)

The binding of σ^{54} to its target promoter sequences in the absence of core RNAP allows identification of specific σ^{54}-DNA interactions. The influence of core RNAP upon the σ^{54}-DNA interaction were evaluated subsequently.

σ^{54}-homoduplex DNA. FeBABE cleavage of homoduplex promoter DNA with Cys336-FeBABE σ^{54} reveals an R336 proximal interaction centred around position −7 (±2), between positions −9 and −5 on the non-template strand (Figure 2(a), lane 6) and position −8 (±2), between positions −10 and −6 on the template strand (Figure 2(b), lane 6). We analysed by densitometry the gel shown in Figure 2(a). The results showed that there was a 50% increase in signal in the cleavage region when compared with the control lanes (2, 3, 7 and 8) in Figure 2(a). No obvious cleavage of homoduplex DNA non-template strand is seen with Cys36-FeBABE σ^{54} implying that position 36 of Region I is not within 12 Å of promoter DNA (Figure 2(a), lane 4). Attempts to cut the template strand with the Cys36-FeBABE σ^{54} yielded no discernible cleavage products (Figure 2(b), lane 4). Since gel shift assays showed that Cys36-FeBABE σ^{54} bound the homoduplex DNA probe (data not shown), we concluded that Region I is positioned mainly away from the DNA in the initial σ^{54}-DNA complex.

σ^{54}-early melted DNA. Chemical footprinting of the σ^{54}-holoenzyme closed complex reveals a hypersensitive site downstream of the consensus GC element, where the DNA appears to be locally melted out (Morris *et al.*, 1994). This nucleation of melting results in the formation of a fork junction structure downstream of the GC element, which σ^{54} binds tightly to (Cannon *et al.*, 2000; Gallegos & Buck, 2000; Guo *et al.*, 1999). Next, we attempted to map the Region I (E36) and Region III (R336) interaction sites on an early melted promoter DNA probe. Weak binding of the R336 derivative to the early melted DNA precluded detection of any cutting (Figure 2(d) and see below). In contrast to the homoduplex DNA (Figure 2(a) and (b), lanes 4), we detected cutting of the early melted DNA probe with Cys36-FeBABE. The region of early melted DNA could help structure Region I for binding and cutting. Cys36-FeBABE cleaves the early melted DNA probe at three distinct positions; between −17 and −15 (±1), adjacent to the −14, −13 GC element on the non-template strand (Figure 2(c), lane 2), and between −18 and −15 (±1) on the template strand (Figure 2(c), lane 10). Residue 36 is situated approximately in the middle of Region I sequences important for σ^{54} function (Casaz *et al.*, 1999) and which are probably α-helical, suggesting that some Region I sequences

Figure 2. (a) *S. meliloti nifH* promoter DNA cleavage by FeBABE-modified σ^{54} proteins. Homoduplex (a) non-template and (b) template strand cleavage by Cys36-FeBABE and Cys336-FeBABE. (c) Early melted probe non-template (lanes 1-8) and template strand (lanes 9-16) DNA cleavage by Cys36. An arrow shows the migration position of the cleaved DNA bands. (d) σ^{54}-isomerisation assay on the early melted DNA probe. Late melted probe (e) non-template and (f) template strand DNA cleavage by Cys36-FeBABE and Cys336-FeBABE. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +, control reactions to which no ascorbate or hydrogen peroxide were added are marked with −. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as a molecular mass marker.

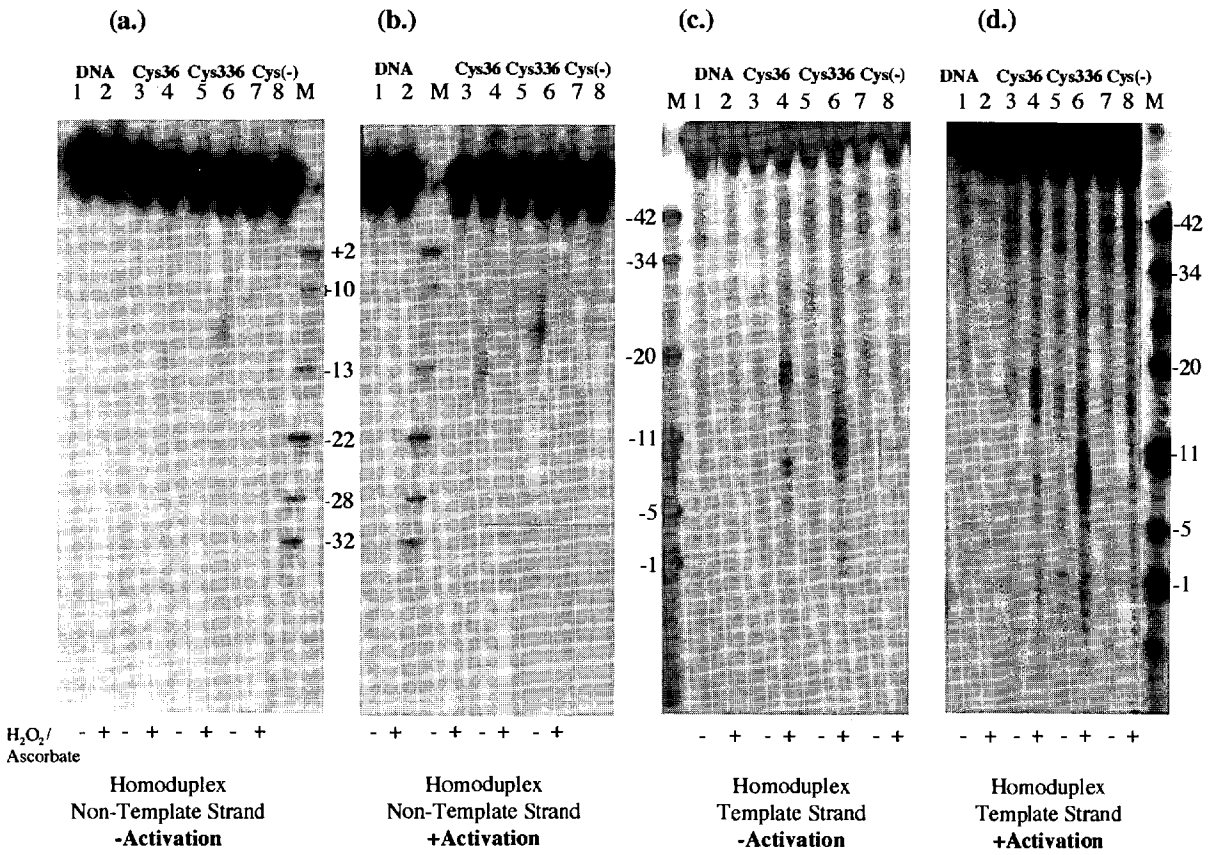


Figure 3. *S. meliloti nifH* promoter DNA cleavage by RNAP holoenzymes containing FeBABE-modified σ^{54} proteins. Homoduplex probe non-template strand and template strand DNA cleavage under ((a) and (c)) non-activating, and ((b) and (d)) activating conditions, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +; control reactions to which no ascorbate or hydrogen peroxide were added are marked with –. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as a molecular mass marker.

could interact with the consensus GC promoter doublet.

Addition of activator and hydrolysable nucleotide to the early melted DNA-σ^{54} complex causes σ^{54} to isomerise and further melt out the DNA (Cannon *et al.*, 2000; Gallegos & Buck, 2000). Interestingly, we see a reduction in non-template DNA cutting between positions −17 and −15 (±1) by Cys36-FeBABE σ^{54} under conditions that permit σ^{54} isomerisation (i.e. in the presence of activator and hydrolysable nucleotide) (Figure 2(c), compare lanes 2 and 4). Densitometry showed that when activated there is a 40% reduction in cutting efficiency of the early melted template strand by Cys36-FeBABE σ^{54} (data not shown, and compare Figure 2(c), lanes 11 and 12). We considered the possibility that activation may have caused substantial dissociation of the early melted DNA-Cys36-FeBABE complex, thus resulting in the disappearance of the cleavage profile. As shown, activation of the early melted DNA-Cys36-FeBABE complex produces an unstable, but supershifted σ^{54}-DNA complex (Figure 2(d), compare lanes 4 and 5). However, the total amount of bound DNA remains constant. The reduction of DNA cutting between positions −17 (±1) and −15 (±1) by Cys36-FeBABE may reflect conformation changes in Region I that appear to position Region I (at least residue 36) away from promoter DNA sequences in response to activation (see also Casaz & Buck, 1997) The failure to see a reduction of DNA cutting under non-activating conditions, i.e. in the presence of activator and GTP(S or in the absence of PspF(HTH further supports this view (Figure 2(c)).

σ^{54}-late melted DNA probe. On this DNA structure the mismatched region includes the non-conserved sequence from −10 to −5 that interacts with σ^{54} within closed complexes (Cannon *et al.*, 1995) or with σ^{54}-holoenzyme in the open promoter complexes (Popham *et al.*, 1991; Morett & Buck, 1989; Sasse-Dwight & Gralla 1988). Binding of σ^{54} to the late melted DNA structure is thought to reflect interactions that occur in the open complex. Cys336-FeBABE cuts the late melted DNA probe between positions −10 and −3 (±2) on the non-template strand (Figure 2(e), lane 6) and between

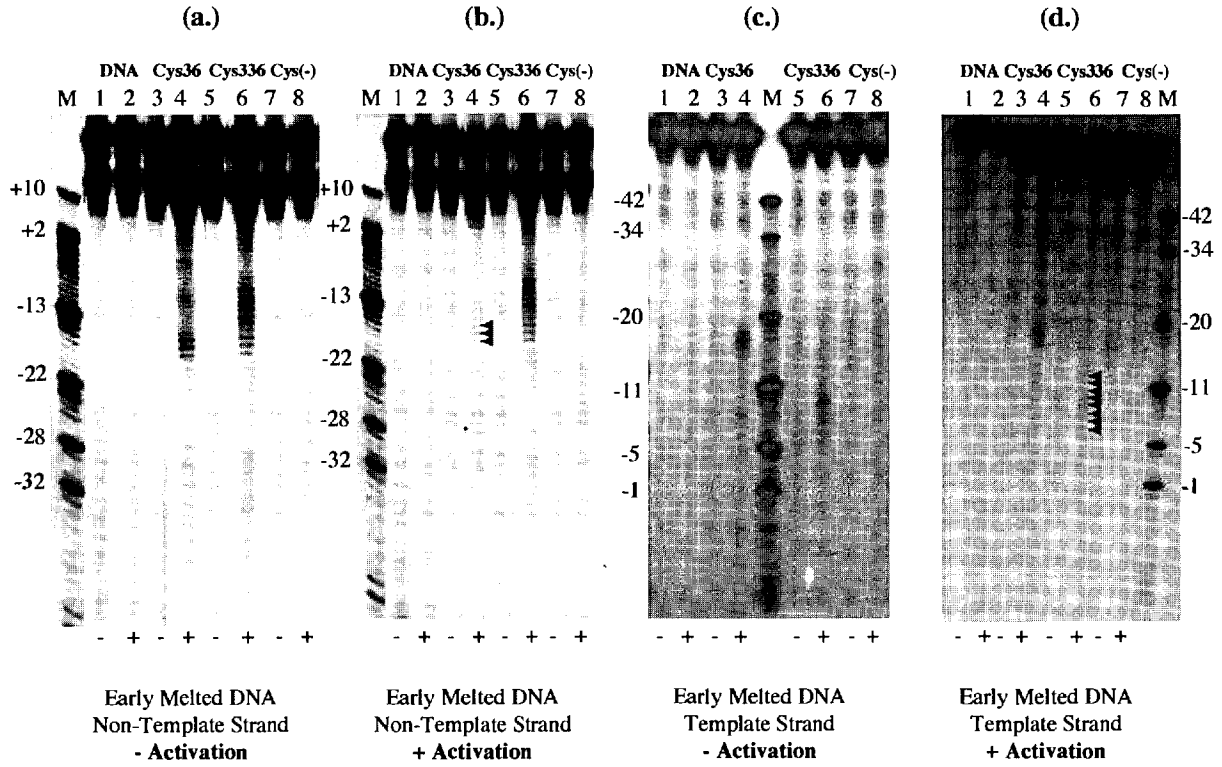


Figure 4. *S. meliloti nifH* early melted promoter DNA cleavage by RNAP holoenzymes containing FeBABE-modified σ^{54} proteins. Early melted probe non-template strand and template strand DNA cleavage under ((a) and (c)) non-activating, and ((b) and (d)) activating conditions, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +; control reactions to which no ascorbate or hydrogen peroxide were added are marked with −. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as a molecular mass marker.

−11 and −6 (±1) on the template strand (Figure 2(f), lane 6). These cutting patterns are different to those seen on the homoduplex DNA, suggesting that σ^{54}-DNA relationships have changed by melting out sequences between −10 and −1. On the late melted DNA probe, Cys36-FeBABE weakly cleaves the non-template strand around and across the −14, −13 GC element between positions −17 and −13 (Figure 2(e), lane 4, marked with arrows). Densiometry of the control lanes (Figure 2(e) lanes 2, 3, 7 and 8) showed that the cutting by Cys36 did occur, but with very low efficiency. However, Cys36-FeBABE fails to cleave template strand DNA detectably when bound to the late melted DNA probe (Figure 2(f), compare lanes 3 and 4). Even though the efficiency of strand cleavage by Cys36-FeBABE on the late melted DNA probe is not strong, it appears that Region I is facing the −14, −13 GC element in the late melted DNA-σ^{54} complex and (at least residue 36) is orientated towards the non-template strand.

Promoter DNA sequences proximal to Region I and R336 in the closed and open complex

We investigated how the relationship between the three promoter DNA structures and Region I E36 and Region III R336 change in the transition from closed to an open promoter complex. Although the Cys36 and Cys336 mutants are deregulated for DNA melting on supercoiled DNA, they are not deregulated on linear DNA and form activatable closed complexes on homoduplex DNA (Chaney & Buck, 1999). Holoenzymes were formed with 1:2 molar ratio of core RNAP: σ^{54} in order to ensure that all conjugated σ^{54} proteins were in the holoenzyme form. Holoenzymes were at 100 nM concentrations where the free 100 nM sigma would not give efficient binding and cutting. To investigate whether the σ^{54}-DNA relationship changes upon binding to core RNAP and activation of the closed complex, we conducted cleavage experiments on closed and activated holoenzyme complexes.

Homoduplex DNA. The Cys-336-FeBABE holoenzyme cuts the homoduplex DNA probe between positions −10 and −6 (±1) on the non-template strand (Figure 3(a), lane 6), whereas the template strand is cut strongly between positions −13 and −8 (Figure 3(c), lane 6). Intense cleavage of homoduplex DNA template strand is seen with Cys36-FeBABE holoenzyme and is centred around position −16 (±2), between −18 and −14 (Figure 3(c), lane 4). In contrast, the non-template strand is cleaved weakly along the T-tract upstream of the

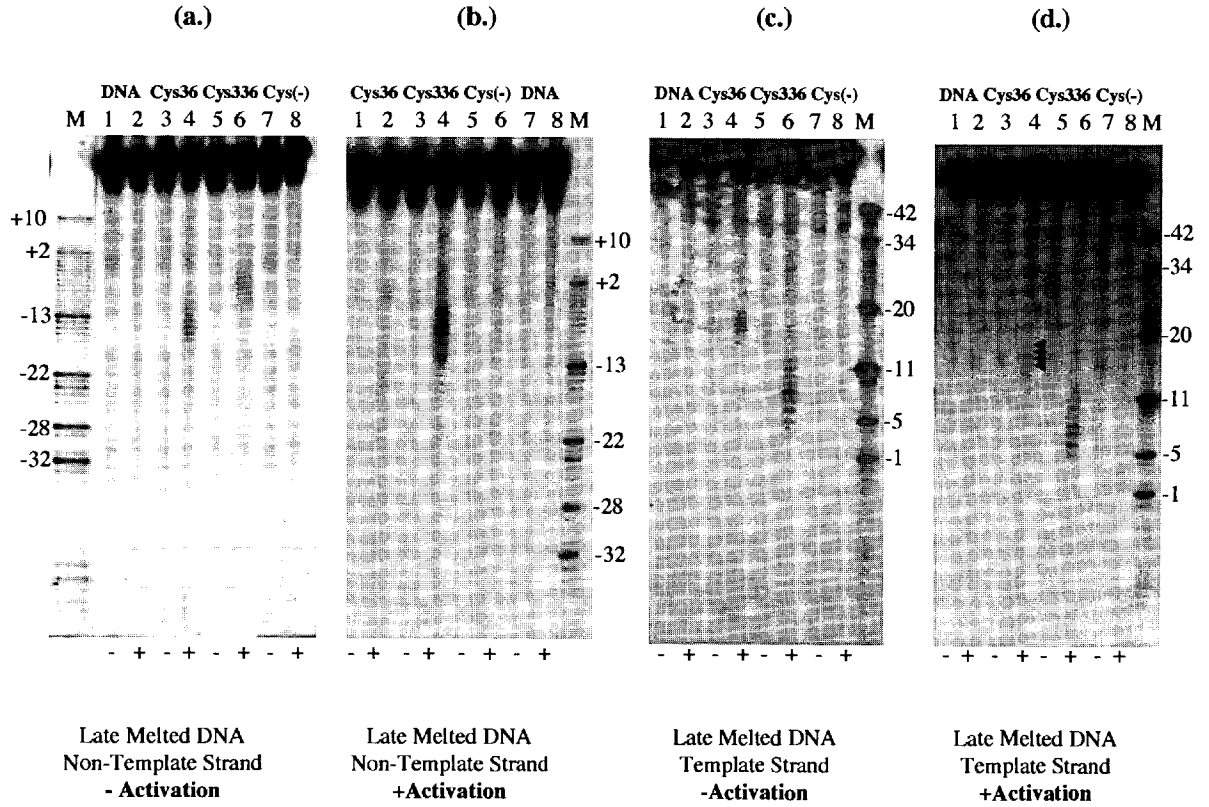


Figure 5. *S. meliloti nifH* late melted promoter DNA cleavage by RNAP holoenzymes containing FeBABE modified σ^{54} proteins. Late melted probe non-template strand and template strand DNA cleavage under ((a) and (c)) non-activating, and ((b) and (d)) activating conditions, respectively. Reactions to which ascorbate and hydrogen peroxide were added to initiate DNA cleavage are marked with +; control reactions to which no ascorbate or hydrogen peroxide were added are marked with −. Lane M contains a mixture of end-labelled *S. meliloti nifH* promoter DNA fragments as a molecular mass marker.

GC element between positions −16 and −11 (±2) (Figure 3(a), lane 4). It seems that Cys336-FeBABE and Cys36-FeBABE cleavage positions of non-template DNA in the closed complex are very close, consistent with previous observation that residues 36 and 336 are proximal to some of the same surfaces on the core RNAP subunits β and β′ (Wigneshweraraj *et al.*, 2000). When PspFΔHTH and dGTP were added, no discernible changes were observed in the cutting patterns between homoduplex DNA-σ^{54}-holoenzyme (compare Figure 3(b) and (d) with (a) and (c), respectively), consistent with the view that open complex formation on linear DNA is inefficient (Oguiza & Buck, 1997; Wedel & Kustu, 1995; and our unpublished data).

Early melted DNA. On the early melted DNA probe, Cys336-FeBABE holoenzyme produced strong cutting of the non-template strand between positions −18 and −10 (±1), whereas the template strand was cut moderately between positions −13 and −7 (±2). (Figure 4(a) and 4(c), lane 6). Recall that binding Cys336-FeBABE σ^{54} alone did not occur with the early melted DNA probe (Figure 2(c) and (d)). Core RNAP clearly stabilises the Cys336-FeBABE on the early melted DNA probe (see Figure 4 and below). As in the homoduplex DNA-holoenzyme closed complex, Cys36-FeBABE cleavage of non-template strand was along the promoter T-tract between positions −19 and −16 (±1), and between positions −19 and −15 on the template strand. It is striking that activation of the early melted DNA-holoenzyme complex reduces the efficiency of non-template strand cutting by Cys36-FeBABE (Figure 4(b), lane 4, marked with arrows), whereas the cutting between positions −17 and −14 on the template strand (Figure 4(c) and (d), lanes 4) remains unchanged. As deduced from the constant cutting patterns, the relationship between Cys336-FeBABE holoenzyme and the non-template early melted DNA probe remains unchanged even under activating conditions (Figure 4(a) and (b), lanes 6). In contrast, activation of the Cys336-FeBABE holoenzyme on the early melted DNA probe complex causes a reduction in the efficiency of cutting on the template strand (Figure 4(c) and (d), lanes 6). Consistent with other observations (Guo *et al.*, 2000; Casaz & Buck, 1997 and Figure 2(c)), activation-dependent changes in Cys36-FeBABE and Cys336-FeBABE-DNA cleavage provide evidence for conformational changes in

Region I and Region III within the holoenzyme that are associated with activation leading to proximity and orientation changes between the promoter DNA, Region I and Region III. Cleavage patterns by Cys36-FeBABE and Cys336-FeBABE holoenzyme did not change when the non-hydrolysable nucleotide GTPγS substituted dGTP in the activation assays (data not shown).

Late melted DNA cleavage. Since the late melted DNA structure is believed to represent a DNA conformation adopted by the promoter DNA following activation and isomerisation of σ^{54}-holoenzyme, we examined whether activation dependent changes were evident in holoenzyme complexes with the late melted DNA probe. If so, they would support the view that activator drives conformational changes within σ^{54} and that pre-melting the DNA does not bypass the action of the activator (Cannon *et al.*, 1999; Wedel & Kustu 1995). Strong cleavage by Cys36-FeBABE holoenzyme occurs on the non-template strand of the late melted DNA probe between positions −16 and −11 (±2) (Figure 5(a), lane 4), and moderate cleavage on the template strand centred around position −16 (between −18 and −14) (Figure 5(c), lane 4). Consistent with our previous observation (see above, Figure 4(b)), activation causes a decrease in non-template strand cleavage by Cys36-FeBABE between positions −16 and −11 (compare Figure 5(a) and (b), lanes 4 and 2, respectively). Similarly, there is a slight reduction in the template strand cleavage by Cys36-FeBABE when activated (Figure 5(c) and (d), lanes 4). The relationship between Cys336-FeBABE and template strand remains unchanged when activated as judged by the relatively constant cutting patterns between positions −12 and −4 (non-activated) and −14 and −5 (activated) in the Cys336-FeBABE holoenzyme late melted DNA complex (Figure 5(c) and (d), lanes 6). It is interesting that there is an increase in non-template strand cleavage between positions −12 and −3 by Cys336-FeBABE when activated (Figure 5(a) and (b), compare lanes 6 and 4, respectively). It appears that conformational changes in response to activation have led to an increased interaction of Region III residue 336 with late melted non-template DNA strand. It is clear that these FeBABE cutting data support the view that changes in Region I and Region III brought about by activator are involved in the conformational changes that allow the holoenzyme to engage with melted DNA (Guo *et al.*, 2000; Cannon *et al.*, 1999; Wedel & Kustu, 1995). The cleavage results are summarised in Figure 6.

R336A and ΔIR336A mutants have altered binding properties to late and early melted promoter DNA probes

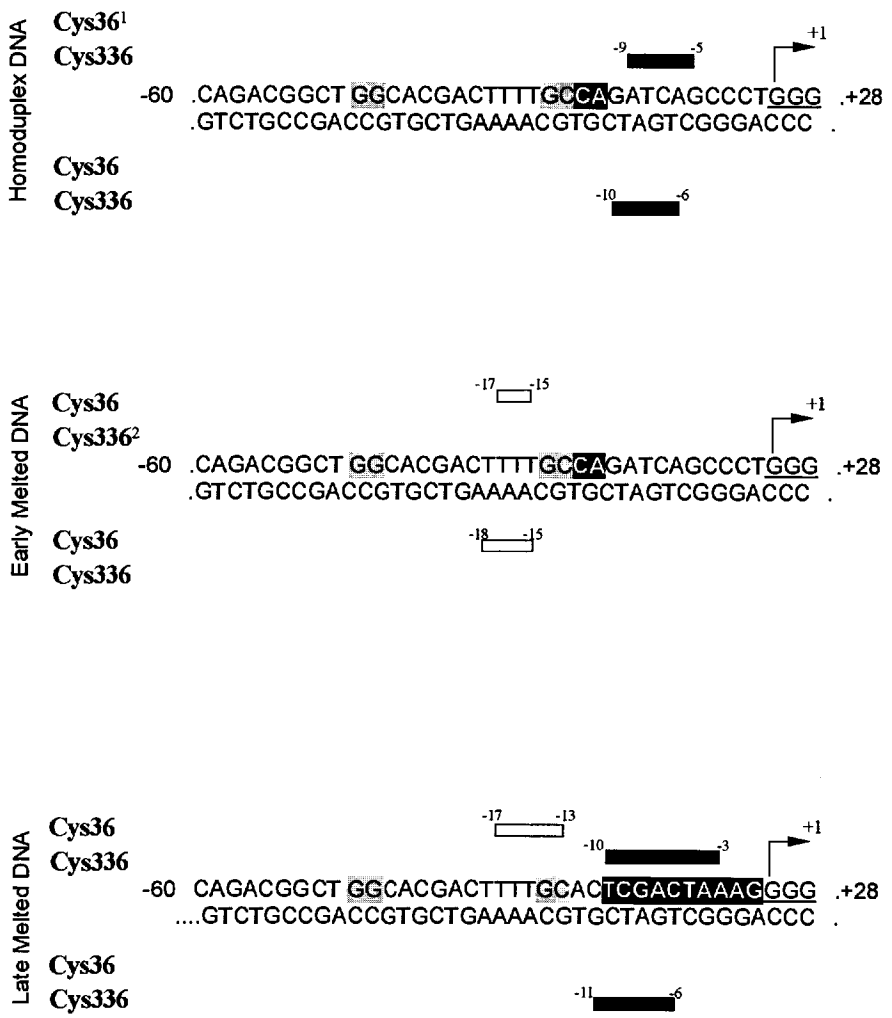
To explore the functional relationship of Region I and Region III residue R336, we conducted DNA binding activity assays using the wild-type σ^{54}, R336A and ΔIR336A mutant proteins (Figure 1(a)).

Homoduplex DNA binding

Using a native gel mobility shift assay, we examined binding to double-stranded promoter DNA to gauge contributions of Region I and R336. The results indicate that ΔIR336A has similar binding activity as the wild-type σ^{54}, and showed a three- to fourfold increased binding compared to the R336A mutant (Figure 7). Consistent with previous results using σ^{54} and $\Delta I\sigma^{54}$ (Cannon *et al.*, 1999; Casaz *et al.*, 1999), the lower DNA-binding activity of the R336A mutant in comparison to that of ΔIR336A argues for a role for Region I in masking some DNA-binding activity. It is clear that R336 is not a determinant of the extra binding activity masked by Region I, but is a determinant of another DNA-binding activity.

Binding to early melted DNA probe

In the next set of binding assays, we used heteroduplex *S. meliloti nifH* promoter DNA probes mismatched at positions −12 and −11 (next to the GC element) to simulate the early DNA melting found in closed complexes. σ^{54} binds the early melted DNA probe better than the homoduplex DNA probe (Figure 7). In contrast the $\Delta I\sigma^{54}$ bound the early melted DNA probe less well (Figure 7) and formed a complex with mobility slightly greater than that of the initial σ^{54}-DNA complex (Figure 8, compare lane 2 to lanes 6, 7 and 8; see also Gallegos & Buck, 2000; Cannon *et al.*, 2000). The R336A mutant was defective in binding to the early melted DNA probe: when added at up to 6 μM it did not detectably bind DNA (Figure 7 and Figure 8, lanes 9-11). Since R336A binding to the homoduplex probe is less impaired, there appears to be a specific contribution by R336 to bind the early melted DNA structure. The binding defects with R336A were greater than that associated with the removal of Region I (Figure 7 and in Figure 8, compare lanes 6-8 with lanes 9-11). Relative to R336A, ΔIR336A is able to bind well to both homoduplex and early melted DNA probes (Figure 7). The ΔIR336A showed an affinity for the early melted DNA probe approaching that of $\Delta I\sigma^{54}$ (Figure 7 and in Figure 8 compare lanes 6-8 with lanes 12-14) demonstrating that removal of Region I allows recovery of DNA binding, as seen with the homoduplex DNA. It seems that Region I masks a DNA-binding activity that is independent of R336. However, the ΔIR336A differed from $\Delta I\sigma^{54}$ in the sense that part of the DNA was bound in a complex with lower mobility (Figure 7, lanes 12-14, marked as ssσ). It seems that R336 contributes to the conformation of the complex with early melted DNA when Region I is removed. A low-mobility σ^{54}-DNA complex (called super-shifted DNA, ssσ) is usually seen with wild-type σ^{54} only in the presence of activator and hydro-



[1]Binding evident by gel mobility assays (Figure 2(d))
[2]Binding not evident by native gel mobility assays (Figure 2(d))

Figure 6 (*legend shown on page 692*)

lysable nucleotides (Figure 8, lane 3 and Cannon *et al.*, 2000). As anticipated from the lack of Region I and/or weak DNA binding, $\Delta I\sigma^{54}$, R336A and ΔIR336A did not respond to activator and nucleotides to form a new slow-running species (data not shown). Region I supplied *in trans* to a $\Delta I\sigma^{54}$-early melted DNA complex increased DNA binding and resulted in the formation of the low-mobility DNA complex (Cannon *et al.*, 2000; Gallegos & Buck, 2000) implying that Region I peptide directs σ^{54} binding to the early melted DNA probe. Presenting Region I *in trans* to the R336A and ΔIR336A did not detectably increase binding to the early melted DNA probe (data not shown). This is consistent with the dominant role of R336 seen in DNA-binding assays with Region I *in cis* (Chaney & Buck, 1999 and Figure 7).

Characterisation of the ΔIR336A supershifted complexes

Activator functions catalytically to cause σ^{54} bound to an early melted DNA probe to isomerise and melt DNA (Cannon *et al.*, 2000). Since the mutant ΔIR336A formed a lower-mobility complex with the early melted DNA (Figure 7, lanes 12-14)

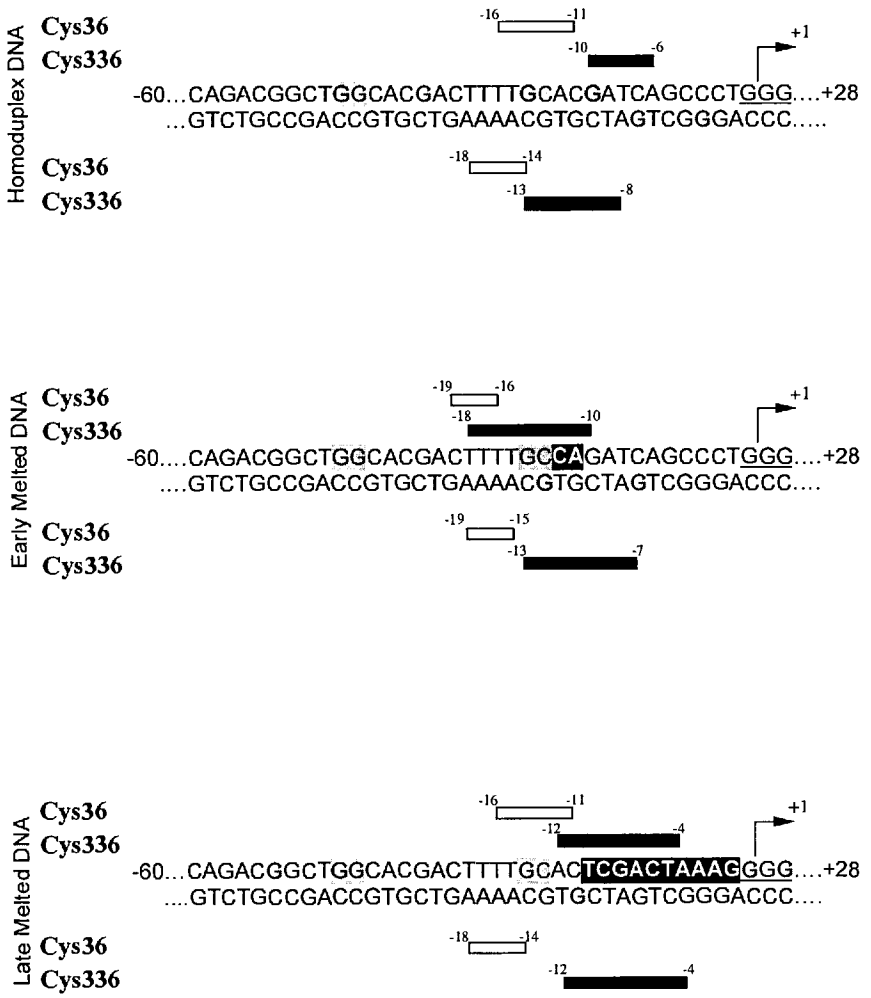


Figure 6 *(legend shown on page 692)*

independent of activation, we characterised the complex by DNAse I and $KMnO_4$ footprinting. This did not reveal an extended footprint (towards −1) nor any extra DNA melting as is seen for the wild-type σ^{54} in the isomerised DNA complex (data not shown; see Cannon *et al.*, 2000). It seems that ΔIR336A forms an altered complex with DNA, in contrast to σ^{54} and $\Delta I\sigma^{54}$, but one that is not distinguishable from $\Delta I\sigma^{54}$ by footprinting (data not shown).

Binding to late melted DNA probe

Next, we assayed the binding properties of the mutant σ^{54} to heteroduplex *S. meliloti nifH* DNA with mismatches from −10 to −1. Compared to homoduplex and early melted DNA binding, the R336A bound the late melted DNA probe best (Figure 7). However, the R336A protein showed a significant decrease in binding to late melted DNA probe with respect to wild-type σ^{54} (Figure 7). In contrast, similar affinities of wild-type σ^{54}, $\Delta I\sigma^{54}$ and ΔIR336A for the late melted DNA were evident. The reduced affinity of R336A mutant for late melted DNA structure is clearly compensated for by the removal of Region I. Overall, it seems defects in DNA binding of R336A are largely corrected by removal of Region I, irrespective of the DNA probe (Figure 7), to yield a binding pattern similar to that of $\Delta I\sigma^{54}$.

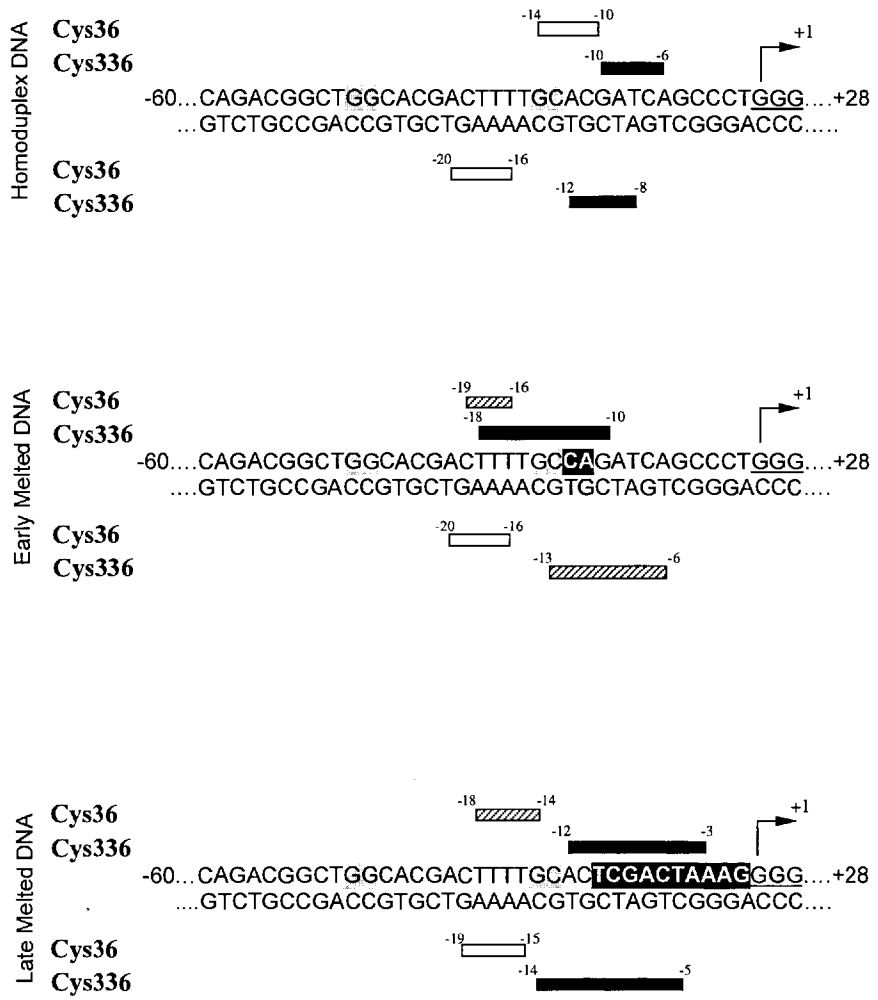


Figure 6. Summary of *S. meliloti nifH* cleavage results. Boxes above (non-template strand) and below (template strand) the DNA sequence indicate positions of cleavage by FeBABE-modifed σ^{54} proteins. Open boxes show cutting by Cys36-FeBABE and filled boxes by the Cys336-FeBABE protein. Shown are (a) σ^{54}-DNA complex, (b) closed complex and (c) open complex on homoduplex, early melted and late melted DNA structures, respectively. In (c) the hatched box indicates the "repositioning" of Cys36 from the non-template DNA strand and Cys336 from the template DNA strand under activating conditions (see the text).

Characterisation of R336A and ΔIR336A holoenzymes on early and late melted DNA structures

We determined the DNA binding activities of the mutant holoenzymes. Initially we showed that ΔIR336A bound *E. coli* core RNA polymerase at a 1:1 ratio like the R336A and $\Delta I\sigma^{54}$ proteins (Chaney & Buck, 1999; Cannon *et al.*, 1995; data not shown). Consistent with the DNA binding properties of R336A and ΔIR336A, ΔIR336A-holoenzyme bound the homoduplex DNA probe with higher affinity than the R336A holoenzyme (like the $\Delta I\sigma^{54}$-holoenzyme) (Chaney & Buck, 1999; data not shown). Therefore, it appears that removal of Region I increases DNA-binding activity that is not reliant on R336.

Holoenzyme interactions with early melted DNA

Holoenzymes formed with R336A and ΔIR336A bound the early melted DNA probe in a similar fashion as the $\Delta I\sigma^{54}$-holoenzyme and the σ^{54}-holoenzyme (Figure 9(a) lanes 1, 3 and Figure 9(b) lanes 1, 3). This contrasts the undetected binding of R336A with the early melted DNA probe (Figures 7 and 8), indicating that the interaction with core RNAP recovers R336A binding to the early melted DNA probe. However, as with the

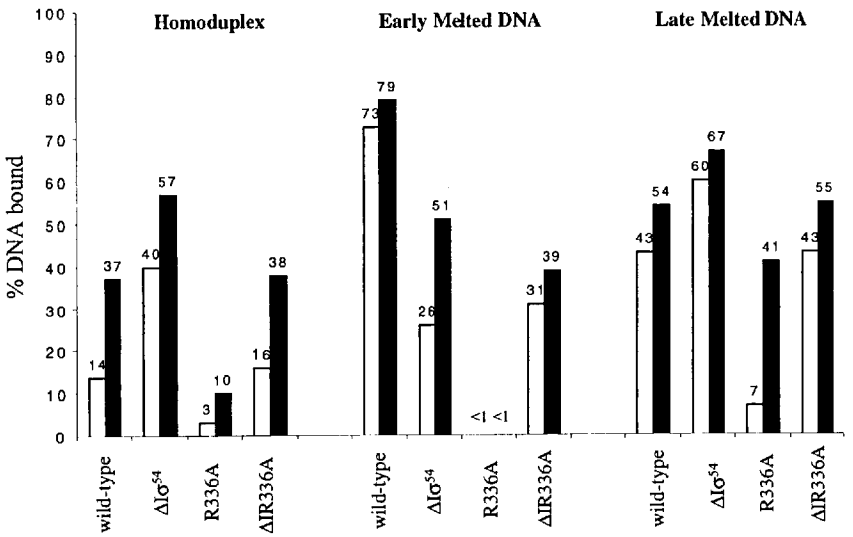


Figure 7. Binding of wild-type σ^{54}, $\Delta I\sigma^{54}$, R336A and ΔIR336A to homoduplex, early melted and late melted DNA probes at 1μM (open bars) and 2 μM (filled bars) protein concentrations, expressed as percentage DNA complexed with σ^{54}.

$\Delta I\sigma^{54}$-holoenzyme, the R336A and ΔIR336A holoenzymes appear to bind DNA in two different complexes, one with the same mobility as that found with the wild-type holoenzyme (marked with an arrow in Figure 9(a) and (b)), and another with further reduced mobility and which was very sensitive to the polyanion heparin. This low-mobility complex was only very weakly evident with wild-type σ^{54}-holoenzyme. Core RNAP alone did not produce the same defined slow-moving species (Oguiza *et al.*, 1999; data not shown). Unlike wild-type σ^{54} and the R336A mutant holoenzymes, $\Delta I\sigma^{54}$- and the ΔIR336A-holoenzymes were unable to form heparin-resistant complexes on the early melted DNA probe (Figure 9(a) and (b) lanes 5, respectively). Presenting Region I peptide *in trans* to the $\Delta I\sigma^{54}$-holoenzyme enables the $\Delta I\sigma^{54}$ holoenzyme to form heparin-stable complexes on the early melted DNA probe (Gallegos *et al.*, 1999 and Figure 9(a), compare lanes 5, 6, 9 and 10; Figure 9(c)). In marked contrast, Region I *in trans* did not diminish the heparin sensitivity of the ΔIR336A holoenzyme (Figure 9(b), compare lanes 5, 6, 9 and 10, and (c)). Thus, it appears that R336 is necessary for σ^{54}-holoenzyme to form heparin-stable complexes on the early melted DNA probe when Region I is supplied *in trans*, but not *in cis*. These data suggest an interaction between Region I and Region III within the holoenzyme, since R336 appears to influence the heparin sensitivity of complexes to the presence and absence of Region I. It seems that interactions made by Region I and R336 that influence properties of the holoenzyme are closely linked.

Holoenzyme interactions with late melted DNA probe

Previously, we showed that the R336A holoenzyme, like the $\Delta I\sigma^{54}$-holoenzyme, formed heparin-stable activator-independent complexes on the late melted DNA probe, whereas the wild-type holoenzyme did not (Chaney & Buck, 1999). The formation of heparin-resistant complexes on late melted DNA by σ^{54} holoenzyme is dependent on the presence of activators (Cannon *et al.*, 1999). Acquisition of heparin stability independent of activation on the late melted DNA probe is a property of deregulated σ^{54} holoenzymes (Cannon *et al.*, 1999, Casaz *et al.*, 1999, Chaney & Buck, 1999). Addition of Region I peptide *in trans* to the deregulated $\Delta I\sigma^{54}$ holoenzyme prior to the addition of the late melted DNA probe increases heparin sensitivity, causing it to behave like the unactivated wild-type holoenzyme (Cannon *et al.*, 1999). We therefore tested whether the same was true for the R336A and ΔIR336A-holoenzymes. It is remarkable that addition of Region I peptide *in trans* to the R336A-holoenzyme prior to binding the late melted DNA probe inhibited the formation of heparin-resistant complexes (Figure 10(a), compare lanes 4, 6, 8, 10 with 5, 7, 9). It seems that sites needed for the negative functions of Region I appear to be intact in the R336A-holoenzyme, thus allowing Region I to act *in trans* and drive the formation of heparin-sensitive complexes. These results imply that mutating R336 allows the holoenzyme to interact with the late melted DNA probe for stable complex formation by causing Region I *in cis* to adopt an altered conformation. However, a site of negative interaction appears to be intact within the R336A-holoenzyme, thus allowing Region I to act *in trans* and reduce stability. Again, the data suggest a functional interaction between σ^{54} Regions I and III within the holoenzyme.

The ΔIR336A-holoenzyme formed activator-independent, heparin-stable complexes on late melted DNA with a slightly reduced stability compared to that of the $\Delta I\sigma^{54}$-holoenzyme (Figure 10(b), lanes 5-8). However, in marked contrast to the $\Delta I\sigma^{54}$ (Figure 10(b), lanes 1-4) and the

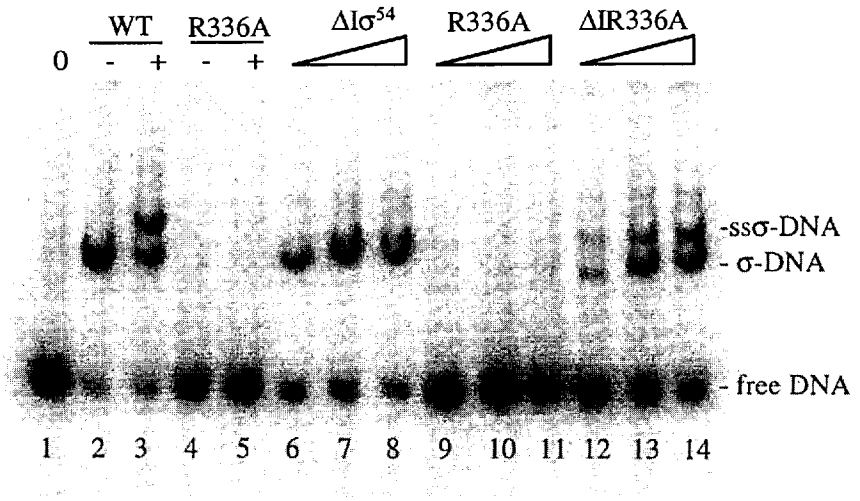


Figure 8. DNA-binding activity of σ^{54} proteins on *S. meliloti nifH* early melted promoter DNA probe in a gel mobility shift assay. Lane 1, DNA alone. Lanes 2 and 3, wild-type σ^{54} (1 μM) alone (−) or in the presence (+) of 4 μM PspFΔHTH and 4 mM dGTP. Activation results in the lower-mobility band (ssσ-DNA, lane 3) indicative of σ^{54}-isomerisation. ΔIσ^{54}, ΔIR336A and R336A were titrated against the early melted DNA at 2, 4 and 6 μM. The R336A mutant does not detectably bind the early melted DNA probe at up to 6 μM (lane 11) or at 4 μM in the presence of PspFΔHTH and dGTP (lane 5). The ΔIR336A forms an ssσ-DNA complex in the absence of activator and nucleotide (lanes 12 to 14).

R336A (Figure 10(a)) holoenzymes, presenting Region I peptide *in trans* to the ΔIR336A-holoenzyme prior to DNA binding increased rather than reduced its heparin stability (Figure 10(c) lanes 4, 6, 8 and 10, and (d)). Addition of Region I *in trans* to the ΔIσ^{54}-holoenzyme already bound to the late melted DNA probe has been shown to increase its heparin stability (Cannon *et al.*, 1999). Thus it appears that ΔIR336A forms a holoenzyme that is further (with respect to the ΔIσ^{54} and R336A-holoenzymes) deregulated for interaction with the late melted DNA probe. Region I supplied *in trans* is without order of addition effects and does not destabilise interactions of the ΔIR336A-holoenzyme with the late melted DNA probe.

In vitro activator-independent transcription activity of the mutant ΔIR336A

To extend the DNA binding assays, and explore the sensitivity of the R336A and ΔIR336A-holoenzymes to Region I further, we conducted activator-independent transcription assays. The ability of the ΔIR336A holoenzyme to support activator-independent (bypass) transcription *in vitro* from the supercoiled plasmid pMKC28 (Chaney & Buck, 1999), which contains the *S. meliloti nifH* promoter was examined using a single-round transcription assay. The ΔIR336A-holoenzyme was as active for activator-independent transcription as the ΔIσ^{54} and R336A holoenzymes (Figure 11(a), compare lanes 1, 5 and 9). Addition of Region I *in trans* to the reaction inhibits the activator independence of the ΔIσ^{54} and R336A-holoenzymes (Figure 11(a), lanes 2-4 and 6-8, respectively; Gallegos *et al.*, 1999). This is consistent with our previous observation that ΔIσ^{54} and R336A-holoenzymes became increasingly heparin sensitive on late melted DNA in the presence of Region I peptide supplied *in trans* (Gallegos & Buck, 1999; Cannon *et al.*, 1999; and Figure 10(a)). In marked contrast, addition of Region I *in trans* to the ΔIR336A-holoenzyme stimulated activator-independent transcription (Figure 11(a), lanes 10-12, and (b)). Recall that Region I peptide *in trans* increased the heparin stability of ΔIR336A holoenzyme on late melted DNA (Figure 10(c)). We suggest the enhancement of activator-independent transcription with the ΔIR336A-holoenzyme when Region I acts *in trans* is through stabilising the ΔIR336A-holoenzyme's interaction with a transiently melted state of the supercoiled DNA downstream of −11. It appears that the sites in the holoenzymes with which Region I interacts to inhibit activator-independent transcription (and reduce heparin stability on late melted DNA structures) are inaccessible or inactive in the ΔIR336A-holoenzyme. Holoenzymes with ΔIR336A appear to be further deregulated for transcription than the ΔIσ^{54} and R336A mutant proteins, since the negative functions of Region I supplied *in trans* cannot be demonstrated.

Discussion

Results overview

We have examined the inter-relationship of two sequences in σ^{54} that, when mutated, lead to overlapping new functionalities of the σ^{54}-holoenzyme and changed DNA binding activities of σ^{54}. The σ^{54} Region I and Region III point mutants used in

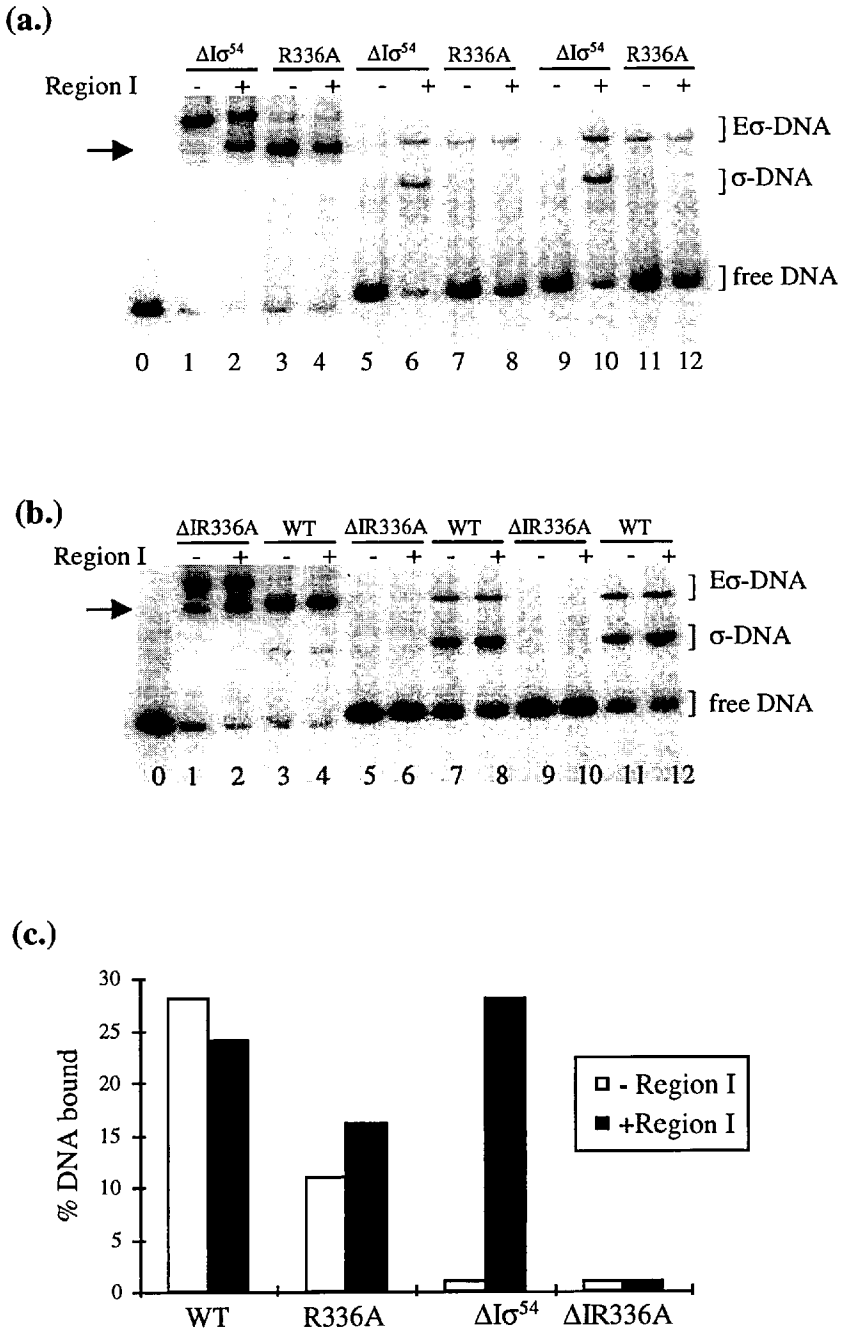


Figure 9. Holoenzyme stability on the *S. meliloti nifH* early melted DNA probe in a gel mobility shift assay and the effect of Region I *in trans*. Holoenzyme (100 nM) formed at a ratio of 1:6 (core RNAP to σ^{54}) was incubated with the early melted DNA probe (2.5 nM) and 1 mM GTP in the absence (−) or presence (+) of 2 μM Region I *in trans* at 30 °C. Region I was added to the reaction before DNA. After 15 minutes, 100 μg/ml heparin was added to the reactions, samples were taken directly before (time 0) and at five and ten minutes after the addition of heparin. Samples were immediately loaded onto a running native gel. The arrow indicates the position of the heparin-stable holoenzyme-DNA complexes. (a) and (b) Lanes 1 to 4, time 0; lanes 5 to 8, five minutes; and lanes 9 to 12, ten minutes after adding heparin, respectively. (c) Histogram showing the percentage of DNA bound by holoenzyme five minutes after the heparin challenge in the absence (open bars) and presence (filled bars) of Region I *in trans*.

this work depend upon activation for stable open complex formation, demonstrating that they productively interact with activator for open complex formation. Although the Region I and Region III sequences are separated by nearly 300 amino acid residues, they seem to occupy closely related positions with respect to the core RNAP (Wigneshweraraj *et al.*, 2000) and the GC promoter DNA region (this work, Figure 6). The GC promoter region is associated with determining regulata-

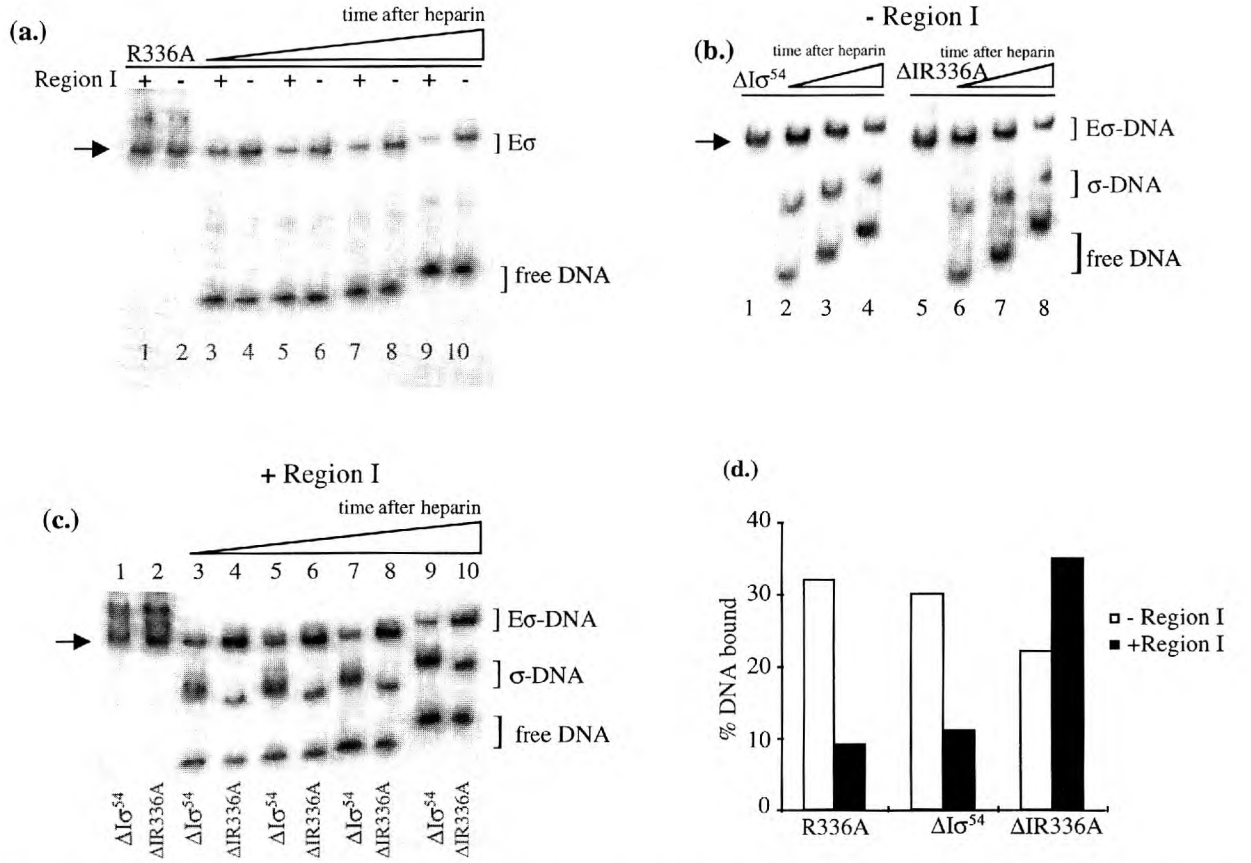


Figure 10. Holoenzyme stability on the *S. meliloti nifH* late melted DNA probe. (a) The R336A holoenzyme was incubated with the late melted DNA probe (2.5 nM) and 1 mM GTP in the presence (+) or absence (−) of 2 μM Region I added *in trans* prior to addition of DNA. Lanes 1 and 2, before heparin, remaining lanes, 2, 5, 10 and 20 minutes after heparin. (b) $\Delta I\sigma^{54}$ and $\Delta IR336A\sigma^{54}$-holoenzymes were incubated with the late melted DNA probe. Lanes 1 and 5 show complexes before the addition of heparin, remaining lanes at 2, ten and 20 minutes after the heparin challenge. (c) The same assay as in (b) except 2 μM Region I was added *in trans* prior to addition of DNA. Odd-numbered lanes contain $\Delta I\sigma^{54}$-holoenzyme, and even-numbered lanes contain $\Delta IR336A\sigma^{54}$-holoenzyme. Lanes 1 and 2 show complexes before heparin and the remaining lanes complexes at two, five, ten and 20 minutes after the heparin challenge. The arrow indicates the position of the heparin-stable holoenzyme-DNA complex. (d) Histogram showing the percentage of DNA bound by holoenzyme after a 20 minute heparin challenge in the absence (open bars) or presence (filled bars) of Region I *in trans*.

ble properties of the promoter, especially the dependence upon activator for open complex formation and DNA melting (Guo *et al.*, 1999, 2000; Cannon *et al.*, 2000; Wang, J.T. *et al.*, 1997; Wang, L. *et al.*, 1999). Hence, we suggest that parts of Region I and Region III form a centre in the holoenzyme closely associated with control of DNA melting. Changing relationships between σ^{54} and DNA were evident under activating conditions (Figures 4 and 5, summarised in Figure 6), suggesting that movements in the centre occur and are probably needed for open complex formation (Guo *et al.*, 2000). Some of these movements seem to involve parts of σ^{54}, in particular Region I sequences, that direct σ^{54} to melted DNA structures downstream of the promoter GC element (Cannon *et al.*, 2000; Gallegos & Buck, 2000; Guo *et al.*, 1999). The different DNA cutting patterns seen between FeBABE-modified σ^{54} proteins and their holoenzymes implies some re-arrangements in the σ^{54} DNA-interacting surfaces upon binding to core RNAP. FeBABE appears to be in a slightly different place in the σ^{54} *versus* the σ^{54}-holoenzyme. Strikingly, either the addition of core RNAP or the use of early melted DNA increases the non-template strand cutting by Cys36-FeBABE. This suggests that core RNAP and the distorted DNA structure created when holoenzyme binds homoduplex DNA can increase the interaction of Region I sequences with promoter DNA. This is consistent with differences seen between σ^{54} and σ^{54}-holoenzyme DNA −12 region footprints (Morris *et al.*, 1994).

As summarised in Figure 6, ~25 bp of promoter DNA between −18 and −5 are within cutting range (~12 Å from the cysteine sulphur atom, ±3-4 Å diffusion distance of the actual cutting species) of Region I residue 36 and Region III residue 336. Our cleavage data imply that Cys36-FeBABE cleavage sites are, in general, located upstream to

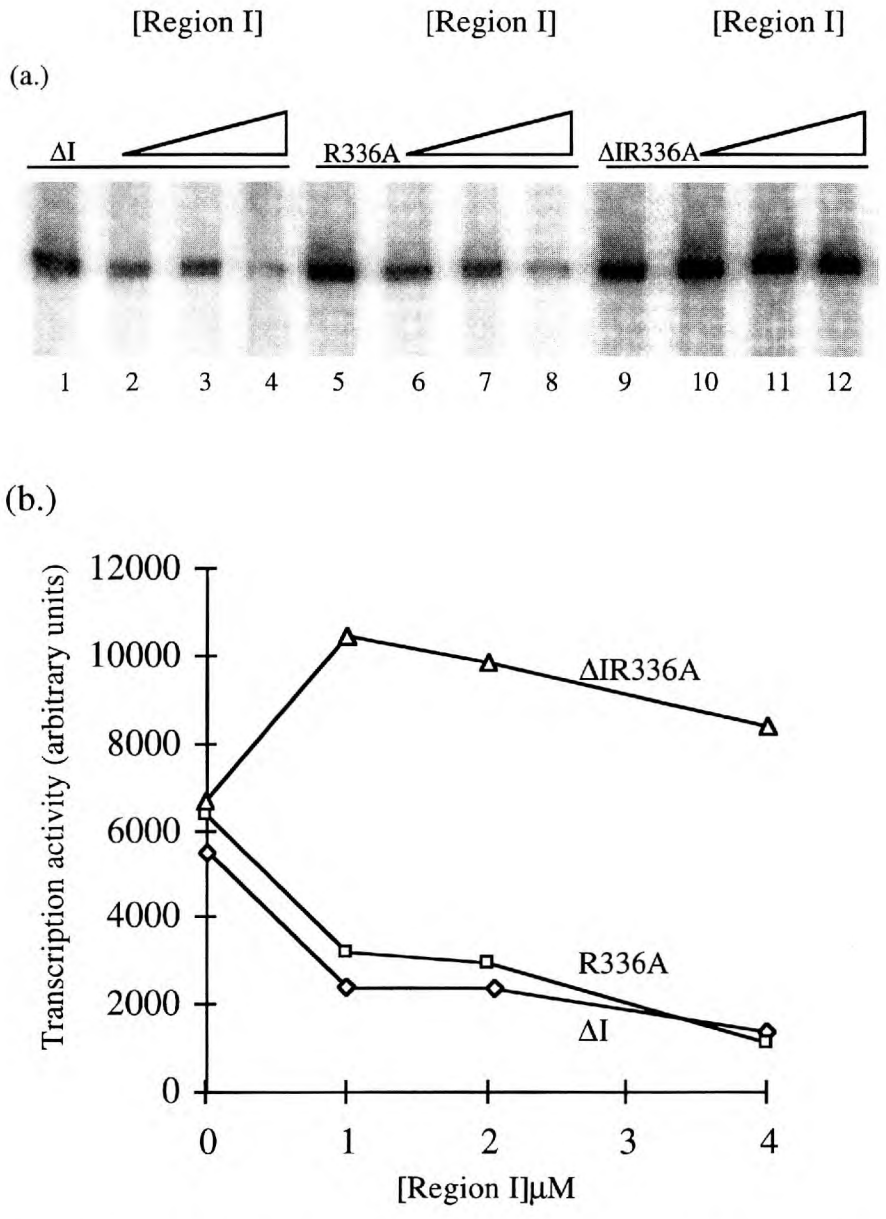


Figure 11. (a) Effect of Region I *in trans* on activator-independent transcription. Increasing amounts of Region I peptide (1, 2 and 4 μM) were added to the activator independent transcription assay prior to addition of DNA template. Holoenzymes (100 nM) in the absence (lanes 1, 5 and 9) or the presence of Region I *in trans* were incubated with template 20 nN DNA and 3 mM GTP for 15 minutes at 30 °C prior to addition of 100 μg/ml heparin and remaining NTPs plus [α-^{32}P]UTP. (b) Transcripts were quantified by phosphorimager and values plotted against concentration of Region I.

the consensus GC promoter element, whereas the Cys336-FeBABE cleavage sites are located downstream to the consensus GC element. This suggests parts of Region I and Region III could form or contribute to a "clamp" across the GC region to restrict or aid DNA melting, depending upon the functional state of the holoenzyme. It remains to be shown that probes attached to part of σ^{54} at sites that do not partially deregulate the holoenzyme lead to the same set of conclusions drawn using Cys36-FeBABE and Cys336-FeBABE. FeBABE conjugated at native cysteine 346, where the regulated properties of the holoenzyme are fully intact, results in −12 region DNA cutting (Wigneshweraraj *et al.*, 2000, 2001).

Region III R336 and relationship to Region I

The DNA proximity relationships suggested for R336 indicate that it interacts with sequences just downstream to the GC element, close to the early melted DNA. Consistent with this, R336A binds probes with early melted DNA less well than homoduplex DNA. The known roles of arginine

residues in DNA interactions suggests that R336 may well make a direct DNA contact. DNA melting within isomerised σ^{54}-DNA complexes is to −6 (Cannon *et al.*, 2000), suggesting that R336 could interact with DNA that exists in the double-stranded form, and is later melted out. R336 may make a direct interaction with the −12 to −6 promoter segment in both double and single-stranded states. This suggests that σ^{54} functions in DNA melting through a direct interaction with DNA that later melts.

Activation of the Cys336-FeBABE-holoenzyme complex on early melted DNA probe causes a reduction in the efficiency of template strand cutting, whereas activation leads to enhanced cutting of the non-template strand in the complex with late melted DNA. Cutting of non-template strand sequences by the Cys36-FeBABE seems always to be reduced under activating conditions on the early melted and late melted DNA probes. The FeBABE cutting data clearly provide evidence for movements in two σ^{54}-DNA relationships during transition from a closed complex to an open complex. Conformational changes in Region I upon activation has been observed (Casaz *et al.*, 1997). The α-helical secondary structure in which R336 is predicted to lie may make several contacts to promoter DNA during open complex formation, some closer to the GC element, possibly even overlapping those associated with Region I, and has been suggested to be involved in binding single-stranded DNA (Kelly *et al.*, 2000).

The properties of the double mutant lacking R336 and Region I reinforce a positive role for Region I in open complex formation and show that Region I masks DNA-binding activity. When supplied *in trans*, Region I clearly stimulates bypass transcription by the ΔIR336A-holoenzyme, an activity of Region I not evident in assays using $\Delta I\sigma^{54}$ or R336A-holoenzymes. It seems that on transiently melted DNA the absence of R336 makes accessible a complex that Region I can stabilise so as to increase initiation. The presence of R336 may normally disfavour such positive effects of Region I under non-activating conditions, since R336 may bind to and stabilise double-stranded DNA in the −12 to −6 area (Chaney *et al.*, 2000; Chaney & Buck, 1999).

Region I and Region III sequences form a regulatory centre

Region I appears to interact with the carboxyl terminal DNA-binding domain in Region III (Wigneshweraraj *et al.*, 2000; Casaz & Buck, 1999). The FeBABE cleavage data now shows that parts of Region I and Region III are close to the −12 promoter element, a DNA sequence known to influence the regulation of open complex formation (this work; Guo *et al.*, 1999; Wang & Gralla, 1998), and to core RNA polymerase subunits (Wigneshweraraj *et al.*, 2000). An advantage of this protein-DNA arrangements would be in co-ordinating concerted movements in σ^{54}, promoter DNA and core RNAP to orchestrate the formation of a stable open promoter complex upon activation. Region I may act as an organising centre that brings the key σ^{54}, core RNAP (Wigneshweraraj *et al.*, 2000) and DNA components together. Region I and Region III change their relationship to DNA in response to activation, suggesting that the organising centre includes a target for the activator NTPase activity (Guo *et al.*, 2000; Cannon *et al.*, 2000). In σ^{70}-mediated transcription initiation, where activation is often by recruitment of the RNAP, determinants within the holoenzyme needed for activation often localise at or upstream of the start-site distal −35 promoter element. In contrast, our work shows that for σ^{54}-holoenzyme, which is activated at the DNA melting step, activation determinants in the holoenzyme localise over the start-site proximal −12 promoter region from where DNA melting originates (Guo *et al.*, 1999, 2000; Cannon *et al.*, 2000; Morris *et al.*, 1994). DNA looping could easily deliver enhancer-bound activator to the centre formed by core RNAP, Region I and the R336 DNA crosslinking patch of σ^{54}. Access of activator to the centre from sites on DNA proximal to the holoenzyme would be sterically hindered by RNAP, suggesting a requirement for DNA looping in activation when activators are enhancer bound.

Materials and Methods

Proteins

The σ^{54} proteins used in this work are listed in Figure 1(a). Wild-type σ^{54} and the mutants $\Delta I\sigma^{54}$ and R336A were overexpressed and purified essentially as described by Gallegos & Buck (1999), Cannon *et al.* (1999) and Chaney & Buck (1999), respectively. The single-cysteine mutants E36C and R336C were made using Stratagene's Quickchange mutagenesis kit using pSRWCys(−) as the template (Wigneshweraraj *et al.*, 2000) The Cys(−), E36C and R336C σ^{54} mutant proteins were overproduced and purified as described by Wigneshweraraj *et al.* (2000). For the construction of ΔIR336A mutant, the *Msc*I-*Hind*III fragment (containing the R336A substitution) from pMKCRA336 (Chaney & Buck, 1999) was replacement cloned into p160HMK′ (Casaz *et al.*, 1999), which contains the *rpoN* gene missing the coding sequences for Region I, derived from pMB7.7160 (Cannon *et al.*, 1995). The resulting plasmid p160RA336 was sequenced to confirm the presence of the Region I-deleted *rpoN* gene harbouring the R336A substitution.

The ΔIR336A mutant was transformed into *E. coli* C41 (Miroux & Walker, 1996) strain and overexpressed essentially as described by Gallegos & Buck (1999) for the wild-type σ^{54}. For induction of ΔIR336A, the 2 × YT growth mediuma was supplemented with 2% (w/v) glucose, 0.2 mg/ml glutamine and 150 μg/ml carbenicillin (for selecting p160RA336). The cell pellets were lysed at 4 °C in TG_5ED (10 mM Tris-HCl (pH 8.0), 5% (v/v) glycerol, 0.1 mM EDTA, 1 mM DTT) containing 100 mM NaCl. The ΔIR336A protein was predominantly found in the soluble cell fraction, which was chromatographed on a HiTrap Q-Sepharose cartridge (Pharmacia). Peak

fractions containing the ΔIR336A were pooled and further purified on a HiTrap heparin cartridge (Pharmacia) as described by Cannon *et al.* (1995). Elution was achieved with a NaCl gradient in all cases. All working protein solutions were stored at −20 °C, stocks at −70 °C, in TG_5ED buffer (10 mM Tris-HCl (pH 8.0), 50% (v/v) glycerol, 0.1 mM EDTA, 50 mM NaCl and 1 mM DTT).

The activator protein PspF(HTH used in the *in vitro* transcription assays was purified and stored essentially as described by Casaz & Buck (1997) and Jovanovic *et al.* (1999). Purified *E. coli* core RNAP was purchased from Epicentre Technologies.

Homo- and heteroduplex promoter DNA formation

Either [^{32}P]end-labelled 88-mer double-stranded homoduplex or non-template strand mismatched between −12 to −11 (early melted DNA) and −10 to −1 (late melted DNA) heteroduplex DNA fragments (Figure 1(b)) consisting of the −60 to +28 *S. meliloti nifH* promoter was prepared as described by Cannon *et al.* (1999, 2000) and Gallegos & Buck (2000) and used as a template for gel mobility shift and FeBABE cleavage assays. Briefly, PAGE-purified 88-mer oligonucleotides were purchased (Genosys), and pairs of DNA strands were mixed with the unlabelled strand present at a two-fold molar excess (10 pmol in 20 μl) over [^{32}P]DNA, heated at 95 °C for three minutes in 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, and then rapidly chilled in iced water for five minutes to allow duplex formation.

FeBABE modification of E36C and R336C single-cysteine mutants

E36C and R336C were conjugated with the FeBABE reagent (Dojindo Chemicals, Japan) and assayed for the conjugation yield essentially as described (Wigneshweraraj *et al.*, 2000). All conjugated proteins were stored at −70 °C.

DNA Cleavage of the *S. meliloti nifH* promoter DNA

The σ^{54}-DNA (either homoduplex, early or late melted DNA) binding reactions (10 μl) were conducted in cleavage buffer (40 mM Hepes (pH 8.0), 10 mM $MgCl_2$, 5% (v/v) glycerol, 0.1 M KCl, and 0.1 mM EDTA) with 20 nM DNA and 2 μM σ^{54}, and incubated for ten minutes at 30 °C. For σ^{54}-isomerisation assays, 4 mM dGTP (where indicated, GTPγS) and 4 μM PspFΔHTH were added for a further ten minutes. Closed complexes were formed with 100 nM holoenzymes (ratio 1:2, core RNAP to Fe-BABE modified σ^{54}). For open complex formation, 4 mM dGTP (where indicated GTPγS) and 4 μM PspFΔHTH were added and incubated for a further ten minutes. Following DNA-binding or activation, cleavage of promoter DNA carrying FeBABE modified σ^{54} proteins or holoenzymes containing FeBABE modified σ^{54} proteins was initiated by the rapid sequential addition of 2 mM sodium ascorbate (pH 7.0) and 1 mM hydrogen peroxide. Reactions were allowed to proceed at 30 °C for ten minutes before quenching with 30 μl of stop buffer (0.1 M thiourea and 100 μg/ml sonicated salmon sperm DNA) and 80 μl of TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA. The stopped reactions were phenol/chloroform-extracted, precipitated with ethanol and electrophoresed on denaturing urea-10% polyacrylamide gels. The cleavage sites were determined by using end-labelled fragments of the *S. meliloti nifH* promoter DNA.

Native gel mobility shift assays

These were performed at 30 °C in STA buffer (25 mM Tris-acetate (pH 8.0), 8 mM magnesium acetate, 10 mM KCl, 1 mM DTT and 3.5% (w/v) PEG 8000). For σ binding assays, standard 10 μl binding reactions contained 2.5 nM DNA and 2 μM σ^{54}, unless otherwise indicated, and were incubated for ten minutes. For holoenzyme binding assays, 100 nM holoenzyme (ratio 1:6, core RNAP σ) was used. Where indicated, 2 μM Region I was added to the holoenzyme prior to addition of DNA. For open complex formation, 4 μM PspFΔHTH and 4 mM dGTP were added. Where indicated, 100 μg/ml heparin was added for a specified time prior to gel loading. Free DNA was separated from σ or holoenzyme-bound DNA on a native 4.5% (w/v) polyacrylamide gels run in 25 mM Tris/200 mM glycine buffer at room temperature. Quantitative data were from phosphorimager analyses of two or three independent experiments to enhance reliability of data.

In vitro activator-independent transcription assays

Supercoiled pMKC28 (Chaney & Buck, 1999) containing the *S. meliloti nifH* promoter was used as the template for the *in vitro* transcription assays. Activator-independent transcription was performed essentially as described by Chaney & Buck (1999) using 100 nM holoenzyme (ratio 1:6, core RNAP σ) and 20 nM supercoiled pMKC28 template DNA. Reactions (10 μl) containing promoter DNA, holoenzyme and 4 mM GTP (the *S. meliloti nifH* start sequence is GGG (+1 to +3) were incubated for ten minutes at 30 °C in STA buffer to allow activator-independent open complex formation and initiation. The remaining rNTPs (CTP and GTP at 0.1 mM and UTP at 0.05 mM), 1.5 μCi of [α-^{32}P]UTP and heparin (100 μg/ml) were added and incubated for a further ten minutes at 30 °C. Reactions were stopped with 4 μl of formamide loading buffer and, after heating for five minutes at 95 °C, 7 μl of each sample was loaded onto a denaturing 6% sequencing gel. The dried gel was analysed on a phosphorimager. Where indicated, 500 nM Region I was added together with the holoenzyme.

Acknowledgements

We thank Nobuyuki Fujita for help with early stages of this work, Wendy Cannon for oligonucleotides and members of the MB laboratory for comments on the manuscript. This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) project grant to M.B.; M.K.C. was supported by a BBSRC postgraduate studentship and S.R.W. from a postgraduate studentship from LEA of Karlsruhe, Germany.

References

Barrios, H., Valderrama, B. & Morett, E. (1999). Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucl. Acids Res.* **15**, 4305-4313.

Bown, J. A., Owens, J. T., Meares, C. F., Fujita, N., Ishihama, A., Busby, S. J. & Minchin, S. D. (1999). Organization of open complexes at *Escherichia coli* promoters. Location of promoter DNA sites close to

region 2.5 of the sigma7 subunit of RNA polymerase. *J. Biol Chem.* **274**, 2263-2270.

Bown, J. A., Kolb, A., Meares, C. F., Ishihama, A., Minchin, S. D. & Busby, S. J. (2000). Positioning of region 4 of the *Escherichia coli* RNA polymerase sigma(70) subunit by a transcription activator. *J. Bacteriol.* **182**, 2982-2984.

Buck, M., Gallegos, M.-T., Studholme, D. J., Guo, Y. & Gralla, J. D. (2000). The bacterial enhancer-dependent σ^{54} (σ^{N}) transcription factor. *J. Bacteriol.* **182**, 4129-4136.

Cannon, W., Claverie-Martin, F., Austin, S. & Buck, M. (1994). Identification of a DNA-contacting surface in the transcription factor sigma-54. *Mol. Microbiol.* **11**, 227-236.

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.

Cannon, W., Gallegos, M. T., Casaz, P. & Buck, M. (1999). Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization. *Genes Dev.* **13**, 357-370.

Cannon, W., Gallegos, M. T. & Buck, M. (2000). Isomerisation of a binary sigma-promoter DNA complex by enhancer binding transcription activators. *Nature Struct. Biol.* **7**, 594-601.

Casaz, P. & Buck, M. (1997). Probing the assembly of transcription initiation complexes through changes in sigma N protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145-12150.

Casaz, P. & Buck, M. (1999). Region I modifies DNA-binding domain conformation of sigma 54 withing the holoenzyme. *J. Mol. Biol.* **285**, 507-514.

Casaz, P., Gallegos, M. T. & Buck, M. (1999). Systematic analysis of σ^{54} N-terminal sequences identifies regions involved in positive and negative regulation of transcription. *J. Mol. Biol.* **292**, 229-239.

Chaney, M. K. & Buck, M. (1999). The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness. *Mol. Microbiol.* **33**, 1200-1209.

Chaney, M., Pitt, M. & Buck, M. (2000). Sequences within the DNA-crosslinking patch of sigma54 involved in promoter recognition, σ isomerisation and open complex formation. *J. Biol. Chem.* **275**, 22104-22113.

Gallegos, M. T. & Buck, M. (1999). Sequences in sigma N determining holoenzyme formation and properties. *J. Mol. Biol.* **288**, 539-553.

Gallegos, M. T. & Buck, M. (2000). Sequences in σ^{54} region I required for binding to early melted DNA and their involvement in sigma-DNA interactions. *J. Mol. Biol.* **297**, 849-859.

Gallegos, M. T., Cannon, W. & Buck, M. (1999). Functions of the sigma(54) region I *in trans* and implications for transcription activation. *J. Biol. Chem.* **274**, 25285-25290.

Guo, Y., Wang, L. & Gralla, J. D. (1999). A fork junction DNA-binding switch that controls promoter melting by the bacterial enhancer-dependent sigma factor. *EMBO J.* **18**, 3736-3745.

Guo, Y., Lew, C. M. & Gralla, J. D. (2000). Promoter opening by σ^{54} and σ^{70} RNA polymerases: σ factor-directed alterations in the mechanism and tightness of control. *Genes Dev.* **14**, 2242-2255.

Hsieh, M. & Gralla, J. D. (1994). Analysis of the N-terminal leucine heptad and hexad repeats of sigma 54. *J. Mol. Biol.* **239**, 15-24.

Ishihama, A. (1997). Adaptation of gene expression in stationary phase bacteria. *Curr. Opin. Genet. Dev.* **7**, 582-588.

Ishihama, A. (2000). Molecular anatomy of RNA polymerase using protein-conjugated metal probes with nuclease and protease activities. *Chem. Commun.* **13**, 1091-1094.

Jovanovic, G., Rakonjac, J. & Model, P. (1999). *In vivo* and *in vitro* activities of the *Escherichia coli* sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFΔHTH. *J. Mol. Biol.* **285**, 469-483.

Kelly, M. T. & Hoover, T. R. (2000). The amino terminus of Salmonella enterica serovar Typhimurium sigma(54) is required for interactions with an enhancer-binding protein and binding to fork junction DNA. *J. Bacteriol.* **182**, 513-517.

Kelly, M. T., Ferguson, J. A. & Hoover, T. R. (2000). Transcription initiation-defective forms of σ^{54} that differ in ability to funtion with a heteroduplex DNA template. *J. Bacteriol.* **185**, 6503-6508.

Lee, J. H. & Hoover, T. R. (1995). Protein crosslinking studies suggest that Rhizobium meliloti C4- dicarboxylic acid transport protein D, a sigma 54-dependent transcriptional activator, interacts with sigma 54 and the beta subunit of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 9702-9706.

Merrick, M. & Chambers, S. (1992). The helix-turn-helix motif of sigma 54 is involved in recognition of the -13 promoter region. *J. Bacteriol.* **174**, 7221-7226.

Miroux, B. & Walker, J. E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289-298.

Morett, E. & Buck, M. (1989). *In vivo* studies on the interaction of RNA polymerase-sigma 54 with the Klebsiella pneumoniae and *Rhizobium meliloti nifH* promoters. The role of NifA in the formation of an open promoter complex. *J. Mol. Biol.* **210**, 65-77.

Morris, L., Cannon, W., Claverie-Martin, F., Austin, S. & Buck, M. (1994). DNA distortion and nucleation of local DNA unwinding within σ^{54} (σ^{N}) holoenzyme closed promoter complexes. *J. Biol. Chem.* **269**, 11563-11571.

Murakami, K., Owens, J. T., Belyaeva, T. A., Meares, C. F., Busby, S. J. & Ishihama, A. (1997a). Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors. *Proc. Natl Acad. Sci. USA*, **94**, 11274-11278.

Murakami, K., Kimura, M., Owens, J. T., Meares, C. F. & Ishihama, A. (1997b). The two alpha subunits of Escherichia coli RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proc. Natl Acad. Sci. USA*, **94**, 1709-1714.

Oguiza, J. A. & Buck, M. (1997). DNA-binding domain mutants of sigma-N (sigma N, sigma54) defective between closed and stable open promoter complex formation. *Mol. Microbiol.* **26**, 655-664.

Oguiza, J. A., Gallegos, M. T., Chaney, M. K., Cannon, W. & Buck, M. (1999). Involvement of the sigmaN DNA-binding domain in open complex formation. *Mol. Microbiol.* **33**, 873-885.

Owens, J. T., Chmura, A. J., Murakami, K., Fujita, N., Ishihama, A. & Meares, C. F. (1998). Mapping the promoter DNA sites proximal to conserved regions of sigma 70 in an *Escherichia coli* RNA polymerase-lacUV5 open promoter complex. *Biochemistry*, **37**, 7670-7675.

Popham, D. L., Keener, J. & Kustu, S. (1991). Purification of the alternative sigma factor, sigma 54, from Salmonella typhimurium and characterization of sigma 54-holoenzyme. *J. Biol. Chem.* **266**, 19510-19518.

Rippe, K., Guthold, M., von Hippel, P. H. & Bustamante, C. (1997). Transcriptional activation *via* DNA-looping: visualization of intermediates in the activation pathway of E. coli RNA polymerase sigma 54 holoenzyme by scanning force microscopy. *J. Mol. Biol.* **270**, 125-138.

Sasse-Dwight, S. & Gralla, J. D. (1988). Probing the *Escherichia coli* glnALG upstream activation mechanism *in vivo*. *Proc. Natl Acad. Sci. USA*, **85**, 8934-8938.

Sasse-Dwight, S. & Gralla, J. D. (1990). Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor sigma 54. *Cell*, **62**, 945-954.

Su, W., Porter, S., Kustu, S. & Echols, H. (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter. *Proc. Natl Acad. Sci. USA*, **87**, 5504-5508.

Syed, A. & Gralla, J. D. (1998). Identification of an N-terminal region of sigma 54 required for enhancer responsiveness. *J. Bacteriol.* **180**, 5619-5625.

Wang, J. T., Syed, A., Hsieh, M. & Gralla, J. D. (1995). Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH2-terminal leucine patch in sigma 54. *Science*, **270**, 992-994.

Wang, J. T., Syed, A. & Gralla, J. D. (1997). Multiple pathways to bypass the enhancer requirement of sigma 54 RNA polymerase: roles for DNA and protein determinants. *Proc. Natl Acad. Sci. USA*, **94**, 9538-9543.

Wang, L. & Gralla, J. D. (1998). Multiple in vivo roles for the −12-region elements of sigma 54 promoters. *J. Bacteriol.* **180**, 5626-5631.

Wang, L., Guo, Y. & Gralla, J. D. (1999). Regulation of sigma 54-dependent transcription by core promoter sequences: role of −12 region nucleotides. *J. Bacteriol.* **181**, 7558-7565.

Wedel, A. & Kustu, S. (1995). The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes Dev.* **9**, 2042-2052.

Wedel, A., Weiss, D. S., Popham, D., Droge, P. & Kustu, S. (1990). A bacterial enhancer functions to tether a transcriptional activator near a promoter. *Science*, **248**, 486-490.

Wigneshweraraj, S. R., Fujita, N., Ishihama, A. & Buck, M. (2000). Conservation of sigma-core RNA polymerase proximity relationships between the enhancer independent and enhancer dependent sigma classes. *EMBO J.* **19**, 3038-3048.

Wigneshweraraj, S. R., Ishihama, A. & Buck, M. (2001). *In vitro* roles of invariant helix-turn-helix motif residue R383 in σ^{54} (σ^{N}). *Nucl. Acids Res.* In the press.

Wong, C., Tintut, Y. & Gralla, J. D. (1994). The domain structure of sigma 54 as determined by analysis of a set of deletion mutants. *J. Mol. Biol.* **236**, 81-90.

Edited by J. Karn

(Received 14 August 2000; received in revised form 13 December 2000; accepted 13 December 2000)

Nucleic Acids Research, 2001, Vol. 29, No. 5 1163–1174

In vitro roles of invariant helix–turn–helix motif residue R383 in σ^{54} (σ^{N})

Siva R. Wigneshweraraj, Akira Ishihama[1] and Martin Buck*

Department of Biology, Imperial College of Science, Technology and Medicine, Sir Alexander Fleming Building, Imperial College Road, London SW7 2AZ, UK and [1]Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411, Japan

Received October 24, 2000; Revised and Accepted December 13, 2000

ABSTRACT

***In vitro* DNA-binding and transcription properties of σ^{54} proteins with the invariant Arg383 in the putative helix–turn–helix motif of the DNA-binding domain substituted by lysine or alanine are described. We show that R383 contributes to maintaining stable holoenzyme–promoter complexes in which limited DNA opening downstream of the –12 GC element has occurred. Unlike wild-type σ^{54}, holoenzymes assembled with the R383A or R383K mutants could not form activator-independent, heparin-stable complexes on heteroduplex *Sinorhizobium meliloti nifH* DNA mismatched next to the GC. Using longer sequences of heteroduplex DNA, heparin-stable complexes formed with the R383K and, to a lesser extent, R383A mutant holoenzymes, but only when the activator and a hydrolysable nucleotide was added and the DNA was opened to include the –1 site. Although R383 appears inessential for polymerase isomerisation, it makes a significant contribution to maintaining the holoenzyme in a stable complex when melting is initiating next to the GC element. Strikingly, Cys383-tethered FeBABE footprinting of promoter DNA strongly suggests that R383 is not proximal to promoter DNA in the closed complex. This indicates that R383 is not part of the regulatory centre in the σ^{54} holoenzyme, which includes the –12 promoter region elements. R383 contributes to several properties, including core RNA polymerase binding and to the *in vivo* stability of σ^{54}.**

INTRODUCTION

The promoter specificity of bacterial RNA polymerases (RNAP) is determined by the σ subunit present in the holoenzyme (Eσ). Two classes of σ factors, σ^{70} and σ^{54} (σ^{N}), have been identified. In marked contrast to the σ^{70} factor, σ^{54} associates with core RNAP to form a holoenzyme that binds to promoter DNA forming a closed complex that rarely spontaneously isomerises to the open complex. Conversion of the σ^{54} holoenzyme closed complex to a transcription-competent open complex is dependent upon γ–β bond hydrolysis of nucleoside triphosphates by activator proteins that bind DNA elements with enhancer-like properties. Activation is mediated by direct activator–closed complex interactions (1–6).

Promoter-specific DNA-binding activity of σ^{54} is central to formation of the Eσ^{54}–promoter complex. DNA binding by σ^{54} appears complex and the interaction between σ^{54} and DNA is modulated by core RNAP (7,8). The promoter sequence recognised by Eσ^{54} is generally characterised by the presence of GG and GC doublets 24 and 12 bp, respectively, upstream of the transcription initiation point (9). The specific DNA-binding determinants of σ^{54} are located in the C-terminal Region III (residues 329–477 in *Klebsiella pneumoniae*). Included are a putative helix–turn–helix (HTH) motif (residues 367–386) and a patch (residues 329–346) that UV cross-links to DNA, each located C-terminal to the core RNAP-binding domain (residues 120–215) (10–15).

The N-terminal Region I has important regulatory roles in Eσ^{54} function, including effects on DNA binding (8,16,17). Region I sequences also bind to core RNAP, an interaction suggested to control properties of the holoenzyme important for activator responsiveness, but dispensable for core RNAP binding *per se* (7,10,13,18,19). The solvent accessibility of sequences within the DNA-binding domain of σ^{54} is changed in the holoenzyme when Region I is deleted, suggesting that Region I contributes to physical properties of the holoenzyme, some of which involve sequences that are closely associated with the DNA-binding function of σ^{54} (7). Holoenzymes formed with mutant or deleted Region I σ^{54} function in activator-independent transcription, in which the promoter-bound Eσ^{54} isomerises and produces transcripts via an unstable open promoter complex (17,20,21–24). Mutant or deleted Region I σ^{54} proteins display changes in DNA-binding activity associated with recognition of the local DNA melting that occurs next to the consensus GC element upon closed complex formation (8,25,26). Proper recognition of this local DNA melting downstream to the GC is a hallmark for regulated transcription initiation by Eσ^{54} (8,10,15,26). The GC promoter region of σ^{54}-dependent promoters in known to be a key DNA element contributing to the network of interactions that keep the polymerase in the closed complex and limit DNA opening prior to activation (8,22,27). Region I, the σ^{54} UV cross-linking patch and the –12 promoter region form a centre in the

*To whom correspondence should be addressed. Tel: +44 20 7594 5442; Fax: +44 20 7594 5419; Email: m.buck@ic.ac.uk

holoenzyme that contains protein and DNA determinants for activator responsiveness and DNA melting (15,17,22,27,28).

Region III residues 367–386 of σ^{54} are proposed to form a HTH DNA-binding structure. R383 in the recognition helix is suggested to interact with bases in the –12 promoter element, in particular with the consensus G of the GC promoter doublet (14). Substitution of R383 with any other amino acid except lysine and, to a lesser extent, histidine was suggested to result in an inactive protein, implying that the nature of the charge on this residue is important for σ^{54} function (14). The suppression of –12 promoter-down mutations in the *K.pneumoniae glnAp2* promoter by R383K *in vivo* is considered as evidence for a role for R383 in recognition of the –12 promoter region. An extension of these conclusions was that the promoter interaction was direct, based largely on the idea that the suggested bi-helical structure would make specific contacts to promoter DNA and that the apparent suppression data might not be explained by indirect effects (14).

Here we have explored the functionality of purified σ^{54} proteins altered at position 383 to determine if R383 is part of the regulatory centre in the σ^{54} holoenzyme. Results indicate that R383 is not a part of the centre and that R383 may not establish a direct contact to DNA. Rather it seems that residue 346 is part of the centre and is close to the GC promoter region. However, it is clear that R383 contributes to DNA binding and discrimination between bases at the G of the GC. It is also required for σ^{54} stability *in vivo*. We show that R383 contributes to maintaining stable promoter complexes in which limited one base DNA opening downstream of the –12 GC element has occurred. Although R383 appears inessential for polymerase isomerisation, it appears to make a significant contribution to maintaining the holoenzyme in a stable complex when melting is initiating next to the GC element.

MATERIALS AND METHODS

Site-directed mutagenesis

Plasmids pSRW-R383K and pSRW-R383A expressing *K.pneumoniae* σ^{54} as an N-terminal His_6-tagged protein with alanine or lysine substitution, respectively, at residue R383 were created using the Quickchange mutagenesis kit (Stratagene) as previously described (18). Briefly, pET28::rpoN (pMTHσ^N) plasmid DNA (29) was used as template with a large molar excess of complementary mutagenic primers. Following mutagenesis PCR, DNA was transformed into *Escherichia coli* strain XL2B and mutant clones were identified by sequencing. The *Bam*HI–*Hin*dIII fragment carrying part of the C-terminal Region III of σ^{54} and harbouring the R383 mutations was cloned into pMT1/306 (29). A cysteine-free σ^{54} [pSRW-Cys(–)] was created by changing the native cysteines at positions 198 and 346 by site-directed mutagenesis using pMTHσ^N as template. pSRW-Cys(–) was used to introduce a cysteine at position R383 to generate pSRW-R383C (18).

Immunoblotting

Mutant plasmids pSRW-R383K and pSRW-R383K were transformed into *E.coli* strain TH1 (Δ*rpoN2518*, *end*A1, *thi*1, *hsdR*17, *supE*44, Δ*lacU*169), which has a deletion of chromosomal *rpoN*, and grown in Luria Broth (LB) to an OD_{600} of 1.0. Cells (1 ml) were collected by centrifugation and resuspended in 100 µl of sterile H_2O. Aliquots of 20 µl of concentrated cells were lysed with 20 µl of 2× SDS sample buffer, heated at 95°C and 10 µl used for loading. Proteins were separated on denaturing 7.5% SDS–PAGE mini-gels and blotted onto PVDF membranes (0.2 µm pore size for western blotting; Millipore). Anti-σ^{54} (30) and alkaline phosphatase-conjugated anti-rabbit IgG (Promega) antibodies were used for detection (20).

Protein expression and purification

The R383A and R383K mutant σ^{54} proteins were overexpressed in *E.coli* strain BL21 (pLysS). Freshly transformed *E.coli* BL21 (pLysS) cells (overnight growth) were used to inoculate (~100–200 c.f.u.) 1 l of 2× YT medium and grown at 37°C with 50 µg/ml kanamycin. The cultures were grown to an OD_{600} of between 0.5 and 0.7 and then induced with 1 mM IPTG at 25°C for 2 h. This temperature shift protocol increases the level of solubility of σ^{54} (29) and improves stability of the R383A mutant, which otherwise becomes severely proteolysed when overproduced at higher temperatures. The cells were collected by centrifugation and resuspended in cold 25 mM sodium phosphate (pH 7.0), 0.5 M NaCl, 5% (v/v) glycerol and 1 mM PMSF and lysed in a French press. The lysate was centrifuged at 20 000 *g* for 30 min and >50% of R383K and ~25% of R383A were found in the soluble fraction. The N-terminal His-tagged mutant proteins were partially purified by Ni affinity chromatography using FPLC and eluted with an imidazole gradient (29). Since R383K and, to a larger extent, R383A co-purified with a truncated form (implying proteolysis of R383K and R383A in the C-terminal domain), peak fractions from the Ni affinity column were dialysed into TGED buffer (10 mM Tris–HCl, pH 8.0, 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT and 5% v/v glycerol) overnight at 4°C for further purification. The truncated fragments in the two protein preparations were removed by heparin and finally Mono Q chromatography essentially as previously described (13). Elution was achieved with a NaCl gradient in both cases. Peak fractions from the Mono Q chromatography were pooled and dialysed against TGED buffer and stored at –70 (long-term storage) or –20°C (short-term storage). Cys383 protein was overexpressed and purified as previously described (18) using a Ni affinity column.

The activator proteins *E.coli* PspFΔHTH and *K.pneumoniae* NtrC were overexpressed and purified as His_6-tagged fusion proteins from pMJ15 (31) and pDW78 (provided by David Widdick and Ray Dixon). PspFΔHTH was stored at –70°C in TGED buffer with 50% (v/v) glycerol and NtrC in TGED buffer with 10% glycerol (50 mM Tris–HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole and 10% v/v glycerol). *Escherichia coli* core RNAP was purchased from Epicentre Technologies.

Assay for free sulfhydryl groups (CPM test)

A modified version of the method of Parvari *et al.* (32) was used. Briefly, β-mercaptoethanol standard solutions were prepared in MOPS buffer (10 mM MOPS pH 8.0, 0.1 mM EDTA and 50 mM NaCl) at concentrations of 100, 50, 20, 10, 5, 1, 0.5 and 0.1 µM. Cys383 was exchanged into MOPS buffer by dialysis at 4°C and protein concentration was determined by Bradford assay. A 15 µl aliquot of 0.4 mM 7-diethylamino-3-[4′-maleimidylphenyl)-4-methylcoumarin (CPM) in dimethylformamide was added to 15 µl of each standard and

Table 1. *Escherichia coli glnHp2* and *S.meliloti nifH* promoter sequences.

Promoter	Plasmid	Sequence			Reference
		–28	–13	+1	
E.coli glnHp2	pFC50	ACTGGCACGATTTTTTCATATATGTGAATGT			34
	pFC50-m12	ACTGGCACGATTTTTGCATATATGTGAATGT			34
	pFC50-m33	ACTGGCACGATTTTTCCATATATGTGAATGT			34
	pFC50-m11	ACTGGCACGATTTTTACATATATGTGAATGT			34
		–29	–14	+1	
S.meliloti nifH	pMKC28	GCTGGCACGACTTTTGCACGATCAGCCCTGGG			36
		–28	–13	+1	
K.pneumoniae glnAp2	–	GTTGGCACAGATTTCGCTTTATATTTTTTAA			

The *E.coli glnHp2* (shown from –28 to +3) and *S.meliloti nifH* (shown from –29 to +3) promoter sequences used for the *in vitro* transcription assays. The –13GC (for the *glnHp2* promoters) and –14GC (for the *nifH* promoter) promoter elements are underlined and the consensus G is shown in bold. The *K.pneumoniae glnAp2* promoter sequence (from –28 to +3) is also given for comparison.

each protein sample. After 1 h incubation at 37°C the reaction was stopped by adding 3 ml of 1% (v/v) Triton X-100. The intensity of fluorescence emission was measured on 1 ml samples, using a Perkin-Elmer 2000 fluorescence spectrophotometer. The excitation wavelength was 390 nm and the emission wavelength was 473 nm.

Core RNAP binding assays

These were performed essentially as previously described as 10 µl reactions in Tris–NaCl buffer (40 mM Tris–HCl pH 8.0, 10% v/v glycerol, 0.1 mM EDTA, 1 mM DTT and 100 mM NaCl) (29). Briefly, *E.coli* core RNAP (250 nM) and different amounts of mutant σ^{54} proteins were mixed together and incubated at 30°C for 10 min, followed by addition of glycerol–bromophenol blue loading dye. Aliquots of 10 µl of the samples were loaded onto Bio-Rad native 4.5% polyacrylamide Mini-Protean II gels and run at 50 V for 2 h at room temperature in Tris–glycine buffer (25 mM Tris and 200 mM glycine). Complexes were visualised by Coomassie blue staining of the gels.

Gel mobility shift assays

^{32}P-end-labelled, fully complementary 88 bp homoduplex or heteroduplex fragments mismatched at positions –12, –12 to –11, –12 to –6, –12 to –1, –5 to –10 and –10 to –1 (heteroduplexes 1–6, respectively, consisting of the –60 to +28 *S.meliloti nifH* promoter sequence; Table 2) were formed as described (15) and used as probes. *Escherichia coli glnHp2* promoter fragments were obtained by PCR using pFC50 (33) and pFC50-m12 as templates with primers FC5 and FC6 (34). The promoter fragments were gel purified and end-labelled with ^{32}P. A typical σ^{54} or Eσ^{54} (formed with σ^{54} at a 2-fold molar excess over core RNAP) binding assay contained 16 nM DNA and σ^{54} or Eσ^{54} (concentrations as indicated in figures or corresponding legends) in STA buffer (25 mM Tris–acetate pH 8.0, 8 mM magnesium acetate, 10 mM KCl, 1 mM DTT and 3.5% w/v PEG 6000) and incubated for 10 min at 30°C. For activation, 4 µM PspFΔHTH activator protein and 4 mM dGTP were used. Briefly, core RNAP, σ^{54} and DNA were pre-incubated at 30°C for 10 min and then nucleotide and activator were added for 10 min and, if required, heparin (final concentration 100 µg/ml) for a further 5 min prior to gel loading. Samples were then loaded onto native 4.5% polyacrylamide gels and run at 60 V for 80 min (for the *E.coli glnHp2* promoter fragments 60 V for 150 min) at room temperature in Tris–glycine buffer. DNA–protein complexes were detected and quantified by phosphorimager analysis.

In vitro transcription assays

The template for transcription assays was either the supercoiled plasmid pMKC28 carrying the *S.meliloti nifH* promoter in pTE103 (35,36) or pFC50 containing the *E.coli glnHp2* promoter and its mutant derivatives (33) harbouring –13 GC element mutations (Table 1): –13T→G (pFC50-m12), –13T→C (pFC50-m33) and –13T→A (pFC50-m11) (the nucleotide numbering system used here is based on *E.coli glnHp2* and differs from that used for *S.meliloti nifH* due to minor variations in the location of the transcription start site; Table 1). The transcription assays were performed in STA buffer as previously outlined (35), except that 30 nM Eσ^{54} (30 nM core RNAP:120 nM σ^{54}) and 10 nM DNA was used. For activation, 4 µM PspFΔHTH or 100 nM NtrC were added with 4 mM ATP (plus 10 mM carbamyl phosphate, used for NtrC phosphorylation). The reactions were incubated for 20 min to allow open complexes to form. The remaining rNTPs (100 nM), 3 µCi [α-^{32}P]UTP and heparin (100 µg/ml) were added and incubated for a further 20 min at 30°C. The reactions were stopped with 4 µl of formamide loading buffer and 7 µl aliquots were loaded on 6% denaturing sequencing gels. The dried gel was analysed on a phosphorimager.

DNA cleavage of the *S.meliloti nifH* promoter DNA

DNA cleavage was conducted essentially as previously described (28). Briefly, closed complexes were formed with 100 nM holoenzyme (ratio 1:2, core RNAP to FeBABE-modified σ^{54}) and incubated at 30°C for 10 min. Cleavage was initiated by rapid sequential addition of 2 mM sodium ascorbate (pH 7.0) and 1 mM hydrogen peroxide. Reactions were allowed to proceed at 30°C for 10 min before quenching with 30 µl of stop buffer (0.1 M thiourea and 100 µg/ml sonicated salmon sperm DNA) and 80 µl of TE buffer (10 mM Tris–HCl pH 8.0, and 1 mM EDTA pH 8.0). The stopped reactions were

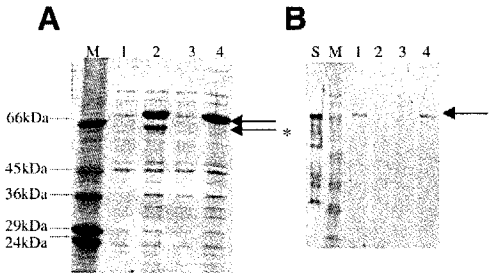


Figure 1. Overexpression and *in vivo* stability of R383K and R383A mutant σ^{54} proteins. (**A**) Uninduced whole cell extracts from *E.coli* BL21 pLysS carrying pSRW-R383A (lane 1) and pSRW-R383K (lane 3). The arrow indicates expression of R383A (lane 2) and R383K (lane 4) in whole cell extracts after 2 h induction. The arrow with asterisk indicates the proteolysed R383A protein (lane 2). The marker (lane M) is BroadRange SDS-7L (Sigma). (**B**) Whole cell extracts from *E.coli* strain TH1 carrying pMTHσ (lane 1), pET28b$^+$ (lane 2), pSRW-R383A (lane 3) or pSRW-R383K (lane 4) were probed with anti-σ^{54} antibodies. The arrow indicates σ^{54}. The prestained marker M (lane M) is from BioRad (broad range) and lane S contains a partially purified sample of σ^{54}.

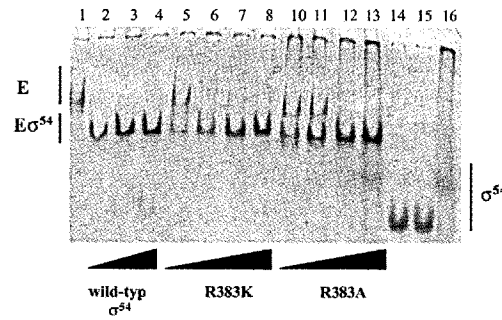


Figure 2. Binding of σ^{54} to *E.coli* core RNAP. Native gel holoenzyme assembly assays were used to detect complexes forming between core RNAP and R383K and R383A, respectively. The formation of holoenzyme ($E\sigma^{54}$) was detected as the presence of a faster migrating species when compared with core (**E**, lane 1) alone. Titrations of core RNAP with σ^{54} were carried out using 250 nM core RNAP and increasing concentrations of σ^{54} at ratios of 1:1 (lanes 2, 5 and 10), 1:2 (lanes 3, 6 and 11), 1:4 (lanes 4, 7 and 12) and 1:8 (lanes 8 and 13). Wild-type σ^{54} shifted nearly all the core into the holoenzyme form at a 1:1 molar ratio of core to σ^{54} (lane 2); in contrast, R383K shifted all the core to the holoenzyme form at a ratio of 1:2 (lane 6) and R383A at 1:4 (lane 12). Free σ^{54} (2.5 μM) proteins are also shown: lane 14, wild-type; lane 15, R383K; lane 16, R383A.

phenol/chloroform extracted, precipitated with ethanol and electrophoresed on 10% denaturing urea–polyacrylamide gels. The cleavage sites were determined using end-labelled fragments of the *S.meliloti nifH* promoter DNA.

RESULTS

Expression and stability of the R383K and R383A σ^{54} mutant proteins

We constructed the σ^{54} mutants R383K and R383A to explore their activities *in vitro*. Denaturing gel analysis of whole cell extracts from induced *E.coli* BL21 (pLysS) cultures revealed proteolysis of the overexpressed R383A protein (Fig. 1A, lane 2). In contrast, the R383K protein, harbouring the more conservative substitution, appeared to be more stable and migrated mainly as a single band, as did the wild-type protein during denaturing gel electrophoresis (Fig. 1A, lane 4). Overexpression of the R383A protein in different *E.coli* backgrounds and altering overexpression conditions (time and temperature) did not improve the stability of the R383A protein (data not shown). Since the truncated R383A co-purified with full-length R383A protein during Ni affinity chromatography, R383A appears to be proteolysed in its C-terminal domain. These observations indicate that R383 is structurally important, not readily predicted from the suggestion that R383 is solvent exposed (14). We therefore constructed the R383C mutant protein in a cysteine-free σ^{54} background to measure solvent accessibility (18). The CPM reactivity of R383C in its native state showed that R383C is indeed solvent accessible. Furthermore, the R383C protein was more stable upon overexpression and more active in transcription *in vitro* than were R383A and R383K (18; data not shown). We therefore infer that R383 has a structural role related to the bulk of the side chain and that 383 is a surface accessible residue.

Previous *in vivo* studies led to the conclusion that R383A is unable to initiate transcription from the *E.coli glnAp2* promoter (14). The instability of R383A we observed (Fig. 1A) suggested that the apparant inactivity of the R383A protein could be due to proteolytic cleavage *in vivo* rather than solely a functional defect caused by the mutation. We therefore conducted *in vivo* promoter activation assays (β-galactosidase promoter fusion assays) and western blots with the R383A protein. Leaky expression of *rpoN* in pET28b+ allows use of the overexpression plasmid in these assays (20). Consistent with the previous *in vivo* results (14), the R383A mutant did not support activation *in vivo* (data not shown). Analysis of whole cell extracts containing pSRW-R383A prepared from *E.coli* TH1 cells under activating conditions with anti-σ^{54} polyclonal antibodies failed to detect full-length R383A σ^{54} protein (Fig. 1B, lane 3); wild-type and R383K proteins were detected (Fig. 1B, lanes 1 and 4). It appears that the activity of R383A may not be easily judged by *in vivo* activity assays. We therefore conducted a series of *in vitro* assays to explore the activity of the R383A and R383K mutants.

Interaction of R383K and R383A with the *E.coli* core RNAP

Native gel holoenzyme assembly assays were used to detect complexes forming between core RNAP and σ^{54} based on the different mobilities of core versus holoenzyme. Results showed that R383K has a slightly reduced affinity for core RNAP (Fig. 2). In contrast, R383A had a significantly reduced affinity and, compared to wild-type σ^{54}, forms a holoenzyme with an increased mobility on native gels. The R383A protein did not produce a characteristic σ^{54} band but was diffuse and slower running (Fig. 2, lane 16), in contrast to the R383K and wild-type proteins (Fig. 2, lanes 14 and 15). Previously we showed, using Cys383-tethered FeBABE footprinting methods, that R383 is not proximal to the core subunits β and β′ (18). We conclude that changing the invariant R383 to A results in a conformational change that may not be localised and which results in significant changes in core RNAP binding and in formation of holoenzyme with a different conformation. These observations further support a structural role for R383.

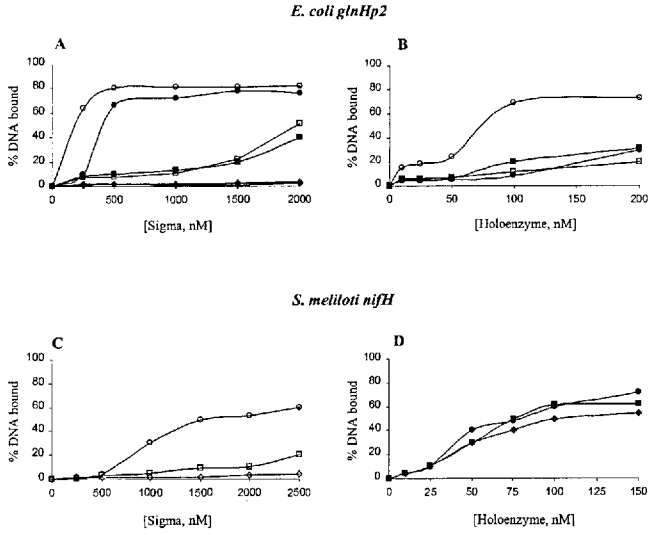


Figure 3. Interactions of R383K and R383A with *E.coli glnHp2* and *S.meliloti nifH* promoter fragments. Gel mobility shift assays were used to detect the binding activities of the σ mutants and their holoenzymes to *E.coli glnHp2* promoter fragments. Binding of (**A**) wild-type σ^{54} (closed circles), R383K (closed squares) and R383A (closed diamonds) and (**B**) their holoenzymes to Hp2-13T (closed symbols) and Hp2-13T→G (open symbols). Binding of (**C**) wild-type σ^{54} (open circles), R383K (open squares) and R383A (open diamonds) and (**D**) their holoenzymes [as in (**C**), but with closed symbols] to the *S.meliloti nifH* promoter fragment.

DNA-binding activities of the R383K and R383A mutant σ^{54} proteins and their holoenzymes

The R383K mutant was suggested to show an altered DNA-binding preference for the promoter GC element (14). Using a gel mobility shift assay we compared the DNA-binding activities of the R383A, R383K and wild-type σ^{54} proteins and their holoenzymes for *E.coli glnHp2* (termed Hp2-13T) and a mutant derivative with a T→G substitution at position –13 (termed Hp2-13T→G) (Fig. 3A and B, respectively). The *glnHp2* promoter is very close in sequence to the *K.pneumoniae glnAp2* promoter used in previous *in vivo* work with R383 mutants (14; Table 1). Previous *in vivo* and *in vitro* studies have shown that the *glnHp2* promoter with a G at –13 is a better substrate for σ^{54} holoenzyme function than one with a T at –13 (33). Binding of σ^{54} confirmed the promoter with a G at position –13 (Hp2-13T→G) as the preferred template, to which σ^{54} had 7-fold increased binding (at 250 nM) compared to the Hp2-13T promoter. The R383K mutant bound both promoter sequences similarly, but had a 7- to 8-fold reduced overall binding (at 1 μM) to the Hp2-13T and Hp2-13T→G templates compared to wild-type σ^{54} (Fig. 3A). This observation, together with the inability of R383A to detectably bind either promoter probe even at higher protein concentrations (1.5 and 2 μM; Fig. 3A) establishes that R383 is important for DNA binding by σ^{54}.

By comparing the binding activities of the wild-type, R383A and R383K holoenzymes for the two promoter probes we determined that the wild-type holoenzyme (at 100 nM) had a 4- to 5-fold higher binding to the Hp2-13T→G than the Hp2-13T promoter sequence (Fig. 3B). The biphasic nature of the graph in Figure 3B (wild-type holoenzyme binding to Hp2-13T→G) probably reflects the complex binding mode of σ^{54} holoenzyme to DNA. Like R383K σ^{54}, the R383K holoenzyme showed a similar binding preference to both promoter probes, being unable to distinguish between them (Fig. 3B). Clearly, σ^{54}–core RNAP interactions are significant in determining promoter binding (compare Fig. 3A and B) and it seems that σ^{54} dominates the promoter binding preference. Any binding preferences of the R383A holoenzyme for Hp2-13T and Hp2-13T→G could not be determined due to low binding of the R383A holoenzyme to both probes (data not shown). Overall, our data show that R383K and its holoenzyme bind the Hp2-13T→G and Hp2-13T templates equally. In contrast, wild-type σ^{54} and its holoenzyme bind to the Hp2-13T→G probe better than to the Hp2-13T probe. R383 is significant for DNA binding by σ^{54} and is needed for preferential binding to –13G rather than the –13T probe.

To further explore the DNA-binding properties of the mutant σ^{54} proteins and their holoenzymes we compared the binding activities of R383K and R383A to *S.meliloti nifH* promoter DNA, a higher affinity binding site used for many σ^{54} activity measurements (see below). As shown (Fig. 3C), R383K bound less *S.meliloti nifH* probe (3-fold reduced) compared to wild-type σ^{54}. In contrast, the R383A mutant appeared defective for DNA binding. Next, holoenzyme binding to the *S.meliloti nifH* promoter was assayed using saturating ratios of σ to core RNAP. The wild-type holoenzyme shifted 70% of the *S.meliloti nifH* promoter probe DNA at 150 nM, whereas mutant holoenzymes shifted 60 (R383K) and 45% (R383A) of the probe (Fig. 3D). This observation contrasts with the behaviour of the R383A holoenzyme on Hp2-13T→G and suggests that sequences in the *S.meliloti nifH* promoter rescue promoter binding by R383A holoenzyme.

In vitro transcription activity of the R383K and R383A holoenzymes

To begin to examine the consequences of altered DNA binding by R383K and R383A upon later steps in activation, we next examined the ability of the R383K and R383A holoenzymes to support transcription *in vitro* from supercoiled plasmid pFC50, which contains the wild-type *E.coli glnHp2* promoter or GC promoter region mutant derivatives of this promoter (Table 1). Changing –13T to a C (pFC50-m33) or A (pFC50-m11) results in a strong promoter-down phenotype or a largely inactive mutant promoter, respectively (33).

Initially, the response of wild-type and mutant holoenzymes to saturating concentrations of the *E.coli* activator protein PspFΔHTH, which functions in solution, was tested. The ability of the holoenzymes to promote transcription at *glnHp2*-13T, *glnHp2*-13T→G and *glnHp2*-13T→C was expressed as a percentage of wild-type holoenzyme activity at the *glnHp2*-13T→G promoter (Fig. 4A). Assays were conducted with sub-saturating amounts of holoenzyme to allow quantitative detection of holoenzyme activities (see Materials and Methods). Experiments were performed at least six times to enhance reliability. The standard error range for the data shown in Figure 4A and B was ±4%. The results clearly show that the R383A holoenzyme is active and able to support transcription *in vitro* (40–50% of wild-type activity on the *glnHp2*-13T→G promoter; Fig. 4A). Transcription by R383A is apparently greater than promoter DNA binding by the holoenzyme (Fig. 3B). Formation of stable open complexes

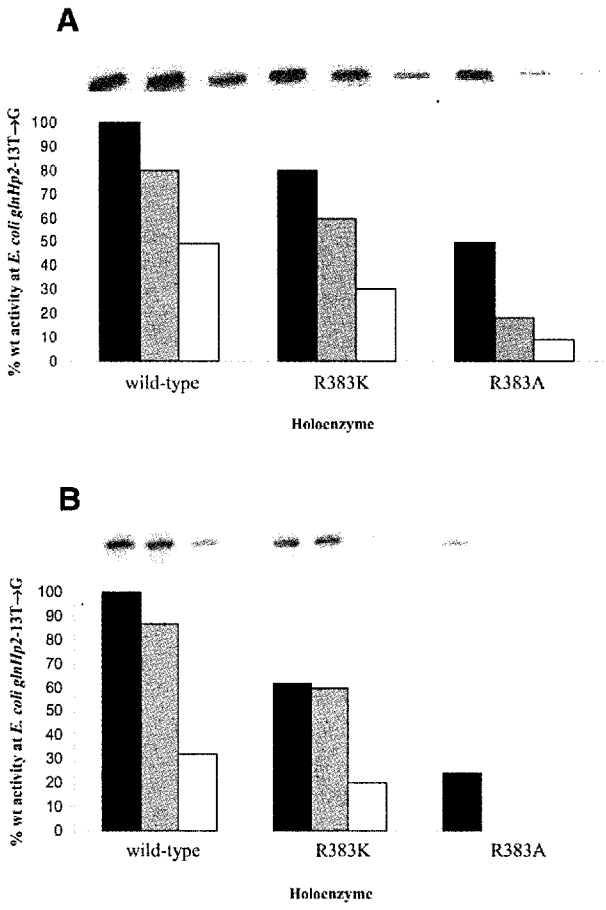


Figure 4. *In vitro* activator-dependent transcription assays on *E.coli glnHp2* wild-type and mutant promoter sequences. The sequences of the promoters used for transcription are as shown in Table 1. (**A**) PspFΔHTH-activated transcription and (**B**) NtrC-activated transcription on *glnHp2*-13T→G (pFC50-m12) (black bars), *glnHp2*-13T (pFC50) (grey bars) and *glnHp2*-13T→C (pFC50-m33) (white bars). The transcription activities are expressed as a percentage of wild-type activity at the *glnHp2*-13T→G (pFC50-m12) promoter.

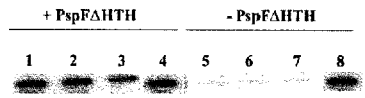


Figure 5. *In vitro* activator-independent transcription assays on the *E.coli glnHp2*-13T→G (pFC50-m12) promoter. Activator-dependent (lanes 1–3) and activator-independent 'bypass' transcription (lanes 5–7) for wild-type σ^{54}, R383K and R383A, respectively, are shown. The R336A mutant σ^{54} (35) was used as the positive control in both cases (lane 4 for activator-dependent and lane 8 for activator-independent reactions).

that do not rapidly decay to heparin-sensitive closed complexes could explain this.

The R383K holoenzyme was 20 ± 4% less efficient in transcription than the wild-type holoenzyme at all the promoters tested (Fig. 4A). This result differs from the *in vivo* assays on *K.pneumoniae glnAp2* (Table 1), which showed that the R383K holoenzyme transcribed better from –13T or –13T→C promoters than did wild-type σ^{54} (14). To explore the potential for suppression *in vitro* we varied the assay conditions. We failed to see any significant R383K-specific suppression at the promoter-down mutant (*glnHp2*-13T and *glnHp2*-13T→C) sequences in the presence of nucleotides that facilitate formation of initiated complexes prior to heparin challenge, at higher temperatures (37 instead of 30°C) used to stimulate transient DNA melting, at different (10–8000 nM) PspFΔHTH concentrations or with varying incubation times (5, 10, 20 and 30 min) before and after addition of heparin (data not shown).

The possibility that the suppression of promoter-down *K.pneumoniae glnAp2* mutants by R383K seen *in vivo* could be either activator-specific or require an enhancer-bound activator was considered. We used NtrC instead of PspFΔHTH and examined *in vitro* transcription activity of the wild-type and mutant holoenzymes at the *glnHp2* promoters. The results showed that when activated by NtrC, the R383K holoenzyme transcribed from the *glnHp2*-13T and *glnHp2*-13T→G promoters at levels consistent with our observation using *glnHp2*-13T→G and PspFΔHTH (compare Fig. 4A and B). In contrast to the PspFΔHTH results, a much lower activity of R383A holoenzyme at the *glnHp2*-13T→G or no detectable activity at the *glnHp2*-13T and *glnHp2*-13T→C promoters, even at higher holoenzyme and NtrC concentrations (data not shown), was evident. This could be linked to an altered holoenzyme conformation (Fig. 2) and reduced DNA binding by the R383A holoenzyme (Fig. 3B).

In conclusion, our *in vitro* transcription results do not show the suppression of promoter-down phenotypes of the *E.coli glnHp2* promoter by R383K holoenzyme reported for *in vivo K.pneumoniae glnAp2* promoter assays (14). The *in vitro* activities of the R383K and R383A holoenzymes argue that R383, at least *in vitro*, is not absolutely required for productive transcription initiation by Eσ^{54}. It seems that R383 is not needed to allow preferential initiation of transcription in which the –13 base is G rather than T (Fig. 4).

Activator-independent transcription activity of the R383K and R383A holoenzymes

Maintaining the transcriptionally silent state of Eσ^{54} in closed complexes depends upon the interaction of σ^{54} with locally distorted promoter DNA downstream of the consensus –12 GC promoter element (10,26,27). σ^{54} proteins defective in recognition of the –12 GC promoter element proximal DNA distortion are capable of increased activator-independent transcription *in vitro*, so called bypass transcription (8,20,25). We used the *in vitro* bypass assay to see whether holoenzymes formed with R383K and R383A were active in unregulated transcription from the *glnHp2*-13T→G promoter. We used R336A mutant σ^{54} as a positive control for the bypass transcription assay (35) and PspFΔHTH for activator-dependent transcript formation. As shown, no bypass transcription was observed with R383K or R383A (Fig. 5). Additional assays from *glnHp2*-13 variants (33) or pMKC28 carrying the *S.meliloti nifH* promoter (33) failed to give unregulated transcription with the R383 mutants (data not shown). The failure to detect bypass transcription with R383K and R383A suggests tight binding of these mutant σ^{54} proteins to the early melted DNA formed in closed complexes, as seen in heteroduplex DNA-binding assays (25,27). Bypass transcription correlates with strong defects in the binding of σ^{54} to early melted DNA, a defect that is only weakly evident with the R383 mutants (see below). Thus we infer that the R383A and R383K mutants appear functionally

Table 2. The *S.meliloti nifH* and *E.coli glnHp2* DNA fragments used for the gel mobility shift assays

Heteroduplex		Sequence
1	-12	-60....GCT**GG**CACGACTTTT**GC**CCGATCAGCCCTGGG....+28
2	-12 to -11	-60....GCT**GG**CACGACTTTT**GC**CAGATCAGCCCTGGG....+28
3	-12 to -6	-60....GCT**GG**CACGACTTTT**GC**CATCGACGCCCTGGG....+28
4	-12 to -1	-60....GCT**GG**CACGACTTTT**GC**CATCGACTAAAGGGG....+28
5	-10 to -1	-60....GCT**GG**CACGACTTTT**GC**ACTCGACTAAAGGGG....+28
6	-5 to -1	-60....GCT**GG**CACGACTTTT**GC**ACGATCATAAAGGGG....+28
7	wild-type	-60....GCT**GG**CACGACTTTT**GC**ACGATCAGCCCTGGG....+28
glnHp2-13T→G(-12)		-60....ACT**GG**CACGATTTTT**GC**CTATATGTGAATGT....+28

The *S.meliloti nifH* and *E.coli glnHp2*-13T→G (–12) heteroduplex promoter DNA fragments (from –60 to +28) used for the gel mobility shift assays. Shown are sequences from –26 to +3, where the σ^{54} consensus promoter sequences are in bold and mutant sequences introduced in the top strand to create the mismatch are boxed in black (16,27).

intact in the generation and maintenance of locally melted –12 proximal promoter structures associated with the closed complex (8,25). Further, chemical footprinting with copper *o*-phenanthroline of closed complexes formed with R383 mutant holoenzymes revealed a local distortion of promoter DNA 3′ adjacent to the GC element, as seen with the wild-type holoenzyme but not in bypass mutants (26,35; data not shown). Overall, R383 is neither directly nor indirectly associated with inhibition of unregulated bypass transcription *in vitro*. By inference, R383 does not closely interact with the elements of the –12 promoter region that are involved in maintaining the stable conformation of the closed complex and in limiting DNA opening prior to activation.

Interaction of R383K and R383A mutant proteins and their holoenzymes with heteroduplex *S.meliloti nifH* promoter DNA probes

In the course of the *in vitro* transcription experiments we observed that the R383K and R383A holoenzymes were essentially inactive for transcription at the *S.meliloti nifH* promoter even though promoter binding was sufficiently efficient to expect transcripts (Fig. 3D and data not shown, respectively). The transcriptional inactivity of the mutant proteins at the *nifH* promoter but not at the *glnHp2*-13T→G promoter with essentially the same –12 region sequences (Table 1) prompted us to further explore the properties of the R383K and R383A mutant proteins. *Sinorhizobium meliloti nifH* (from –11 to –1) is rich in G and C residues, whereas *E.coli glnHp2*, like the *K.pneumoniae glnAp2* promoter used in the *in vivo* assays, is AT-rich in this region, which is melted in open complexes (Table 1). This led us to consider that the R383K and R383A mutant holoenzymes might be defective in some aspect of DNA melting or single-stranded DNA binding at the *S.meliloti nifH* promoter. We used heteroduplex DNA that mimics the DNA at different stages of open complex formation to test this idea (Table 2). In marked contrast to the failure to transcribe from the *S.meliloti nifH* promoter (data not shown), both the R383K and R383A mutant holoenzymes gave heparin-stable, activator- and nucleotide hydrolysis-dependent complexes on promoters with a region of heteroduplex from –10 to –1 (Table 2, heteroduplex 5) (Fig. 6A). On this DNA structure the mismatched region includes the non-conserved sequence from –10 to –1 that interacts with σ^{54} within the closed complex (13) or with σ^{54} holoenzyme in the open promoter complex (37–39). The ability of the wild-type, R383K and R383A holoenzymes to form activator- and nucleotide hydrolysis-dependent, heparin-stable complexes when bound to promoter DNA where the sequence from –10 to –1 is heteroduplex (Table 2, heteroduplex 5) argues that the R383K and R383A holoenzymes are not *per se* defective in polymerase isomerisation at the *nifH* promoter. Also, pre-opening from –10 to –1 appears to allow a range of activities with R383K and R383A similar to that seen with the *glnHp2* promoters in transcription assays. As expected from the *in vitro* activator-dependent transcription results, the wild-type holoenzyme, like the R383K and R383A holoenzymes, does not form heparin-stable, activator- and nucleotide hydrolysis-independent complexes on heteroduplex 5 (16; data not shown).

Heteroduplex with early melted sequences. Next we examined whether the R383A and R383K mutants were defective in interacting with DNA structures representing the early stages of DNA melting. We used heteroduplex promoter DNA fragments unpaired at –12 (Table 2, heteroduplex 1) and at –12/–11 (Table 2, heteroduplex 2) to mimic the structure believed to be involved in initial DNA opening (25,27). These heteroduplexes allow the wild-type holoenzyme to form complexes that survive a heparin challenge independently of activator and nucleotide hydrolysis (25). As shown (Fig. 6B), we were unable to form heparin-stable complexes with the R383K and R383A holoenzymes on either of the heteroduplexes, even under activating conditions. This defect correlates with the

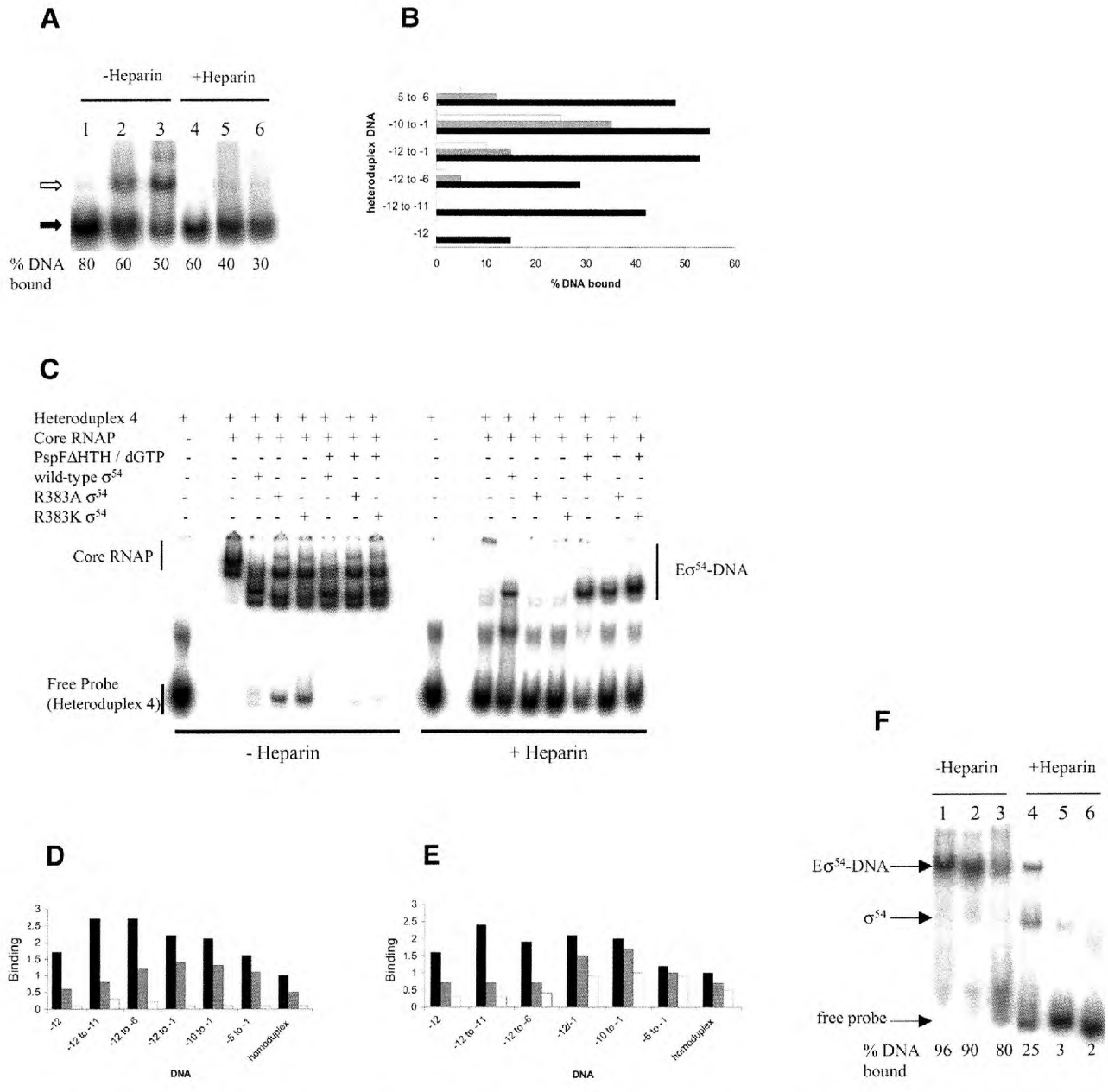


Figure 6. Binding of R383K and R383A to homo- and heteroduplex promoter DNA probes. (**A**) Activator- and nucleotide-dependent, heparin-resistant open complex formation on the *S.meliloti nifH* heteroduplex –10 to –1 (heteroduplex 5, Table 2) by wild-type σ^{54} (lane 1), R383K (lane 2) and R383A (lane 3) holoenzymes (100 nM). Black arrows indicate the position of the Eσ^{54}–DNA complex; white arrows the position of core RNAP complex. Reactions without (lanes 1–3) and with (lanes 4–6) heparin (100 µg/ml) challenge for 5 min prior to loading are shown. (**B**) Activator-dependent, heparin-resistant Eσ^{54}–DNA complex formation on heteroduplexes (heteroduplexes 1–6, Table 2). Percent DNA shifted with the wild-type σ^{54} (black bars), R383K (grey bars) and R383A (white bars) holoenzymes are shown. (**C**) Stability of wild-type and mutant Eσ^{54}–DNA complexes on *S.meliloti nifH* heteroduplex –12 to –1 (heteroduplex 4) under non-activating and activating conditions in the presence and absence of heparin (100 µg/ml), respectively (lanes as indicated on figure). The slower running band in the lanes containing Eσ^{54} is a heparin-unstable complex of Eσ^{54} with DNA (15,35,40). (**D**) DNA-binding activities of R383K (grey bars) and R383A (white bars) to homo- and heteroduplex DNA (see Table 2) expressed as a fraction of wild-type σ^{54} (black bars) binding. (**E**) As (D) but for holoenzymes. (**F**) Activator- and nucleotide hydrolysis-independent, heparin-stable Eσ^{54}–DNA complex formation on *E.coli glnHp2*-13T→G(–12) (see Table 2). Reactions without (lanes 1–3) and with (lanes 4–6) heparin challenge are shown. Lanes 1 and 4, wild-type σ^{54}; lanes 2 and 5, R383K; lanes 3 and 6, R383A.

inability of the R383 mutants to transcribe from the *nifH* promoter. However, when using heteroduplex promoter DNA where the sequences from –12 to –6 (Table 2, heteroduplex 3) and –12 to –1 (Table 2, heteroduplex 4) were unpaired, the R383K and R383A holoenzymes survived the heparin challenge, but only with activator and nucleotide hydrolysis (compare Fig. 6B and C). In contrast, the wild-type holoenzyme formed heparin-stable complexes on heteroduplex 4 in the absence of activator and nucleotide, as predicted from the results with heteroduplexes 1 and 2, opened at –12 (Table 2, heteroduplex 1) and from –12 to –11 (Table 2, heteroduplex 2) (16,25,27,40). Although non-native structures near –12 have

the property of allowing the wild-type holoenzyme to form heparin-stable complexes independently of activation, the R383K and R383A holoenzymes formed heparin-stable and activator- and nucleotide hydrolysis-dependent complexes with promoter probes containing –12 proximal melts (Table 2, heteroduplexes 1 and 2) only when these heteroduplexes included further regions of heteroduplex proximal to the start site (Table 2, heteroduplexes 3 and 4). These results show that R383 specifies an interaction in the closed complex associated with stable complex formation by the holoenzyme when the –12 proximal sequences are melted. Other interactions required to acquire heparin-stable complex formation involving late melted sequences appear intact in the R383K and R383A mutants.

DNA-binding activities on heteroduplexes. Next, we measured the DNA-binding activities of the mutant proteins (Fig. 6D) and their holoenzymes (Fig. 6E) on *S.meliloti nifH* heteroduplex DNA with –12 promoter element proximal melts (Table 2, heteroduplexes 1 and 2), start site proximal melts (Table 2, heteroduplexes 3 and 4) and on heteroduplexes with mismatches between –10 and –1 (Table 2, heteroduplex 5) and –5 to –1 (Table 2, heteroduplex 6). Results are shown relative to binding of wild-type σ^{54} and its holoenzyme to *S.meliloti nifH* homoduplex DNA. It is evident that R383K and, especially, its holoenzyme bound more (2- to 3-fold) of the heteroduplex DNA probes which contain start site proximal melts, whilst wild-type σ^{54} and its holoenzyme prefer heteroduplex DNA with –12 proximal melts. Since certain Region I mutants of σ^{54} have defects in binding to early melted DNA structures (8,25,27), the DNA-binding properties of R383 mutants (Fig. 6) may reflect indirect effects upon the function of Region I (28).

The results (Fig. 6) clearly imply a role for R383 in interactions within the closed complexes in which limited DNA opening next to the –12 element has occurred. To further explore this idea we used the *E.coli glnHp2*-13T→G promoter mismatched at position –11 (i.e. *glnHp2*-13T→G equivalent to heteroduplex 1; Table 2). R383K had 80% of wild-type transcription activity on this promoter *in vitro*. As shown (Fig. 6F), whilst the wild-type holoenzyme was able to form activator- and nucleotide hydrolysis-independent, heparin-stable complexes on the *glnHp2* heteroduplex, the R383K holoenzyme did not. Activation enabled the R383K and, to a lesser extent, R383A holoenzymes to form some heparin-stable complexes (<10% DNA shifted) on this heteroduplex (data not shown). The ease of opening of the *glnHp2* AT-rich sequence (from –11 to –1) may explain why the activator allows acquisition of heparin stability, and also why a longer segment of heteroduplex is needed at *nifH*. Therefore, binding assays with two different σ^{54}-dependent promoters are consistent with the view that R383 has a role in interactions within the initial closed complexes in which limited DNA opening next to the –12 element has occurred. Unless DNA opening occurs easily, the defects associated with R383 dominate and few open complexes form.

Proximity of residue 383 to promoter DNA

To examine the physical proximity of R383 to promoter DNA we constructed a σ^{54} with FeBABE located at 383. A single cysteine substitution at 383 was made, the naturally occurring cysteines of *K.pneumoniae* σ^{54} at 198 and 346 having been replaced by alanine to allow 383-specific conjugation of FeBABE. DNA cleavage by tethered FeBABE is achieved through the generation of hydroxyl radicals coordinated to the Fe^{2+}, which attack the deoxyribose–sugar backbone of nucleic acids within a radius of 12 Å from the FeBABE attachment site (reviewed in 41). Using the *S.meliloti nifH* homoduplex DNA probe the DNA-binding activity of the R383C mutant was 90% that of the wild-type and Cys-free σ^{54} activity (data not shown). Upon conjugation with FeBABE (76% efficiency) the DNA-binding activity remained largely unchanged. This suggests that even a relatively bulky substituent at 383 is tolerated for DNA binding, consistent with R383 being dispensable for DNA binding (this paper) and non-essential for transcription (this paper; 18). σ^{54} containing Cys383-tethered FeBABE failed to produce detectable cutting of several different *S.meliloti nifH* promoter templates (homoduplex and heteroduplexes 2 and 5), both in the presence and absence of core RNAP and under activating conditions that allow open complex formation (see below and data not shown). We considered the possibility that σ^{54} containing Cys383-tethered FeBABE and its holoenzyme might have dissociated from the promoter DNA under DNA cleavage conditions. However, binding assays conducted with the Cys383-tethered FeBABE protein under DNA cleavage conditions showed that 91% of σ^{54} containing Cys383-tethered FeBABE and 48% of holoenzyme containing Cys383-tethered FeBABE remained bound to DNA (Fig. 7A). Repeated experiments failed to show DNA cutting by σ^{54} containing Cys383-tethered FeBABE and its holoenzyme. As one positive control σ^{54} with FeBABE conjugated to Cys346, at the edge of the DNA cross-linking patch of σ^{54}, produced cutting proximal to the GC element on *S.meliloti nifH* homoduplex promoter DNA (Fig. 7B). The putative HTH motif in σ^{54} is C-terminal to a patch of amino acids that UV cross-links to promoter DNA. In this patch FeBABE conjugated to residue 336 cut DNA downstream of the conserved GC promoter element (28). As shown in Figure 7B, holoenzyme containing Cys346-tethered FeBABE strongly cut homoduplex promoter DNA between positions –14 and –7, mostly downstream of the GC element, but this cutting was upstream of that seen with the Cys336-tethered FeBABE derivative (28). The C-terminal to N-terminal orientation of the cross-linking patch is therefore 5′→3′ with respect to the template strand of the promoter DNA.

DISCUSSION

Specific recognition of promoter DNA by σ factors has an essential role in locating the RNAP at the correct site for initiation. However, the function of residues in σ^{54} that contact DNA are likely to be more complex than just facilitating promoter location. Emerging functions associated with σ^{54}–DNA binding include recognition of the DNA fork junction created when the DNA starts to melt and keeping the polymerase silent for transcription (8,10,15,35). Protein footprints suggest that the DNA-binding domain of σ^{54} is part of the interface with core RNAP and properties of mutants indicate a role in generating polymerase isomerisation and facilitating promoter opening (7,27,35). Our results address the functions of invariant residue R383 of σ^{54}, previously implicated in interactions with the –12 promoter element (14).

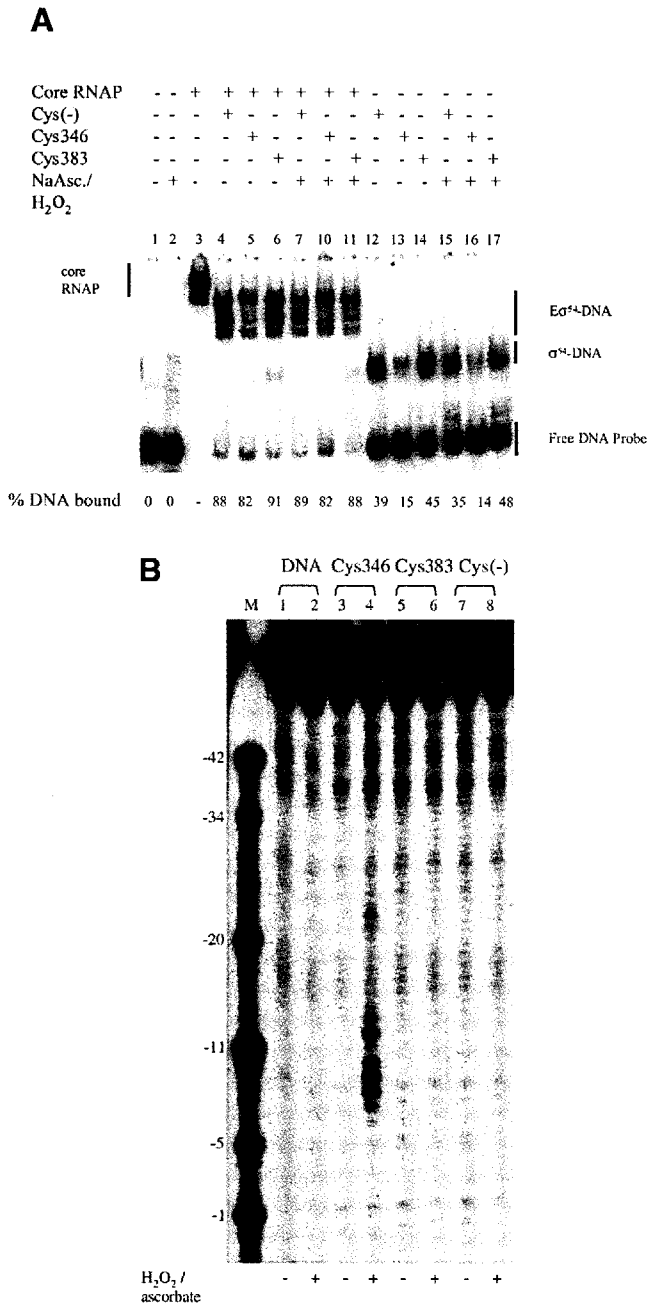


Figure 7. *Sinorhizobium meliloti nifH* promoter DNA template strand cleavage by RNAP holoenzyme containing FeBABE-modified σ^{54} proteins. (A) Cys383-FeBABE σ^{54} and Eσ^{54} binding to the *S.meliloti nifH* homoduplex probe under DNA cleavage conditions (lanes as marked on figure). (B) Homoduplex cleavage. Reactions to which hydrogen peroxide and ascorbate were added to initiate DNA cleavage are marked +; control reactions to which no ascorbate and hydrogen peroxide were added are marked –. Lane M contains a mixture of end-labelled *S.meliloti nifH* promoter DNA fragments as molecular weight markers.

Protein stability

R383 is clearly required for protein stability *in vivo*. The instability of R383A may be associated with an unfavourable change in structure directly increasing its proteolytic sensitivity. Alternatively, the reduced DNA-binding activity of R383A may indirectly increase proteolysis by changing the intracellular location of σ^{54}. The suggestion from *in vivo* promoter activation assays that R383 is essential for σ^{54} function may be incorrect, as purified R383A does support transcription from the *E.coli glnHp2* promoter *in vitro*.

Core RNAP binding

That part of σ^{54} strongly footprinted by the core RNAP is centred on residue 397 and in the absence of Region I ($\Delta I\sigma^{54}$) much of the 325–440 sequence is protected by core RNAP (7). R383A had reduced core RNAP binding and formed a holoenzyme with altered mobility on native gels, suggesting that R383 contributes to core RNAP binding. Possibly, the core interface of σ^{54} is altered in the R383A mutant, potentially in a manner that involves Region I sequences (7,18,19). Interestingly, Cys383-tethered FeBABE footprinting of core RNAP failed to show any proximity of residue 383 (within at least 12 Å) to the core subunits β and β′ (18). Thus we suggest that the defects in core RNAP binding by the R383 mutants are indirect. The R383K mutant was less disrupted for core binding, consistent with the conservative nature of the mutation.

DNA binding

Gel mobility shift assays showed that R383A was very defective in DNA binding, R383K less so. R383K did not preferentially bind *E.coli glnHp2*-13T promoter DNA. R383K is reported to transcribe more efficiently than wild-type σ^{54} from a similar promoter sequence (mutant *K.pneumoniae glnAp2*) (14). Clearly, the increased transcription reported may not simply correlate with increased DNA binding of σ^{54}. R383K might therefore influence other steps to allow increased transcription *in vivo*. The R383K holoenzyme did not footprint the *glnAp2*-13G→T promoter *in vivo*, but wild-type σ^{54} did, consistent with the view (developed below) that promoter occupancy may not be dominant for the increased transcription observed (14). It was striking that the defects in *in vitro* transcription at *E.coli glnHp2* with the R383K and R383A mutants were less than the defects in *in vitro* σ and holoenzyme DNA binding. We interpret this to mean that promoter occupancy is not reduced to a point that severely limits transcription *in vitro*. It is plausible that tight binding of the holoenzyme to the promoter increases a transition barrier for open complex formation. The R383K and R383A mutants may reduce this barrier, compensating for reduced promoter occupancy. This favourable effect might contribute to the elevated activities observed with R383K *in vivo* and the good level of transcription detected *in vitro*. It may also partly compensate for the defect in forming stable complexes with early melted DNA (discussed below).

Interactions with heteroduplex DNA

Results from transcription assays with heteroduplexes suggested that slow opening of the DNA at the *S.meliloti nifH* promoter might mean that the closed complex or an activator-dependent isomerised holoenzyme formed with R383K dissociates prior to full strand opening. In contrast, more frequent opening of the *E.coli glnHp2* promoter or stable opening as in heteroduplexes may explain why these templates support stable open complex formation with R383K. Even so, these complexes decay more rapidly than those formed with wild-type σ^{54} (data not shown), suggesting that R383 contributes to DNA binding within the open complex. However, R383 is not

essential for transcription, at least *in vitro*. Our ability to recover activator-dependent stable complex formation using pre-melted *S.meliloti nifH* DNA templates with heteroduplex –10 to –1 sequences suggests that the mutant holoenzymes are limited at some promoters in steps leading to full DNA melting. Compared to DNA templates with start site proximal melts, DNA templates with melted sequences proximal to the –12 promoter element were poor DNA-binding sites for the R383K and R383A mutants. The failure of the R383A and R383K mutants to bind well or make stable complexes on the heteroduplex promoter fragments with –12 proximal melts suggests that the R383 mutant proteins are directly or indirectly defective for interaction with such structures. With the mutant proteins the activator did not allow the use of heteroduplex promoters with –12 proximal melts for efficient stable complex formation on either the *S.meliloti nifH* or *E.coli glnHp2* promoters, but in the context of the *S.meliloti nifH* promoter the activator allowed formation of stable complexes on promoter sequences which were further opened to include the –1 residue. Overall, the results suggest that DNA melting from –10 to –1 stabilises the promoter complexes that form with the R383K and R383A mutant holoenzymes and that the initial melting at –12 is unfavourable for stable $E\sigma^{54}$–DNA binding when R383 is altered to A or K. Rapid melting at the AT-rich *E.coli glnHp2* may allow stable complexes to form, but slow melting at the GC-rich *S.meliloti nifH* may result in reduced stable complex formation.

Bypass transcription

Interaction of σ^{54} with the –12 GC promoter element appears important in maintaining the transcriptionally silent state of the holoenzyme. Mutations that change the sequences adjacent to the GC or substitution of certain amino acids in σ^{54} that interact with it allow transcription independent of activator (8,20,22,24,42). The σ^{54} DNA-binding domain mutant R336A gives strong bypass transcription (35). R383K and R383A did not, despite supporting levels of activated transcription and binding to –10 to –1 heteroduplex DNA (heteroduplex 5), which suggested that bypass activity might be readily detected. If R383 interacts with the –12 GC promoter sequence, it would appear not to be an interaction contributing to maintaining the silent state of the holoenzyme prior to activation, in contrast to the properties of the R336A and Region I mutants (20,24,35).

DNA proximities

Residue 383 was suggested to be within a HTH motif, expected to establish direct contacts with DNA and thought to interact with the –12 region of the promoter (14). Although our data show that substitutions at 383 influence DNA binding, other data suggest that a simple direct interaction of R383 with DNA may not occur. Cleavage of the promoter DNA by Cys383-tethered FeBABE was not evident. Closed complex promoter DNA cleavage by the FeBABE derivative of σ^{54} in the UV cross-linking patch (Cys336-tethered FeBABE) is centred around position –9 (±1) (28), while DNA cutting by the Cys346-tethered FeBABE derivative is centred further upstream at position –11 (±1) (Fig. 7). Given that the UV cross-linking patch is α-helical in structure, the FeBABE cleavage data suggest that the UV cross-linking patch (N-terminal to the HTH motif) aligns with or is inclined towards the promoter DNA template strand in closed complexes. The lack of any discernible cleavage by the Cys383-tethered FeBABE derivative suggests that either residue 383 is not involved in a direct DNA contact or that some structural consequences of making substitutions at 383 do not allow detection using the FeBABE methodology. However, R383C was active for transcription *in vitro*, more so than R383A (18).

Summary

Overall, our data are consistent with a requirement for R383 to distinguish between T or G at –13 and overall reduced DNA binding when it is substituted by K or A (14). It is possible that some of the overall loss in DNA-binding activity has a basis in an altered protein structure rather than simple loss of a DNA-interacting side chain, a view supported by the *in vivo* instability of R383A. Although instability is unexpected on substitution of a surface exposed residue in an α-helix by alanine, there are suggestions from secondary structure predictions that R383 may exist within a non-helical structure (www.http://jura.ebi.ac.uk:8888). The data presented in Figures 1–5 suggest that the interpretation placed on the *in vivo* activation data may need reconsideration and suggest that R383 has a previously unexpected role in the stability of σ^{54}.

In the absence of additional structural or proximity data, any suggestion that the HTH motif lies within a fold that localises sufficiently near the –12 region to contact DNA is speculative. A colinear arrangement, N-terminal to C-terminal (beginning at the –24 promoter element and ending at the start site proximal sequences), of the UV cross-linking patch, the HTH motif and the RpoN box is possible, but unproven. Further, the clear involvement of the σ^{54} Region I sequences in promoter binding has shown that determinants outside the DNA-binding domain make critical contributions to the DNA binding function of σ^{54} (8,20,25,28).

Specialisation of function across the σ^{54} DNA-binding domain is evident. Residues F402, F403, F354 and F355 appear to be associated with interactions needed for efficient polymerase isomerisation (40,43), R383 with forming stable initially melted DNA complexes and R336 with maintaining the inhibited silent state of the polymerase (10,35). We note functional similarities between the putative α-helix, in which C346 and R336 in σ^{54} lie, and helix 14 of *E.coli* σ^{70}. Both helices interact with promoter sequences that include recognition sequences (15,44). They also contain determinants for binding locally single-stranded DNA structures from which melting originates and spreads (8,26,28,45). These and other considerations lead to the view that a series of linked interactions that involve σ^{54}–DNA interaction and σ^{54}–core interfaces are required to change in order to allow the polymerase to progress to the open complex. It seems that the putative HTH motif of σ^{54} contributes as a structural element rather than as a major direct DNA-contacting surface. Nevertheless, several conformational changes in σ^{54} are probably necessary for open complex formation. Transient contacts between σ^{54} and core RNAP or between σ^{54} and promoter DNA may have escaped our analysis of Cys383-tethered FeBABE proximities to DNA. Current data suggest that the centre formed by Region I, the UV cross-linking patch of σ^{54} and the –12 promoter region does not include the HTH motif as an element in proximity (15,18,28).

ACKNOWLEDGEMENTS

We thank Nobuyuki Fujita for help with early stages of this work, Wendy Cannon for oligonucleotides and members of the MB laboratory for comments on the manuscript. This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) project grant to M.B. and by a post-graduate studentship from the LEA, Karlsruhe, Germany, to S.R.W.

REFERENCES

1. Kelly,M.T. and Hoover,T.R. (2000) The amino terminus of *Salmonella enterica* σ^{54} defective in transcription initiation but not promoter binding activity. *J. Bacteriol.*, **182**, 513–517.
2. Lee,J.H. and Hoover,T.R., (1995) Protein cross-linking studies suggest that *Rhizobium meliloti* C-4-dicarboxylic acid transport protein-D, a σ^{54}-dependent transcriptional activator, interacts with σ^{54}-subunit and the β-subunit of the RNA-polymerase. *Proc. Natl Acad. Sci. USA*, **92**, 9702–9706.
3. Rippe,K., Guthold,M., von Hippel,P.H. and Bustamante,C. (1997) Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of *E. coli* RNA polymerase and σ^{54} holoenzyme by scanning force microscopy. *J. Mol. Biol.*, **270**, 125–138.
4. Weiss,D.J., Batut,J., Klose,K.E., Keener,J. and Kustu,S. (1991) The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell*, **67**, 155–167.
5. Su,W., Porter,S., Kustu,S. and Echols,H. (1990) DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial *glnA* promoter. *Proc. Natl Acad. Sci. USA*, **87**, 5504–5508.
6. Wedel,A. and Kustu,S. (1995) The bacterial enhancer-binding protein NtrC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes Dev.*, **9**, 2042–2052.
7. Casaz,P. and Buck,M. (1999) Region I modifies DNA-binding domain conformation of σ^{N} within the holoenzyme. *J. Mol. Biol.*, **285**, 507–514.
8. Guo,Y., Wang,L. and Gralla,J.D. (1999) A fork junction DNA–protein switch that controls promoter melting by the enhancer-dependent sigma factor. *EMBO J.*, **18**, 3736–3745.
9. Barrios,H., Valderrama,B. and Morett,E. (1999) Compilation and analysis of σ^{54}-dependent promoter sequences. *Nucleic Acids Res.*, **15**, 4305–4313.
10. Gallegos,M.T., Cannon,W. and Buck,M. (1999) Functions of the σ^{54} Region I *in trans* and implications for transcription activation. *J. Biol. Chem.*, **274**, 25285–25290.
11. Guo,Y. and Gralla,J.D. (1997) DNA binding determinants of σ^{54} as deduced from libraries of mutations. *J. Bacteriol.*, **179**, 1239–1245.
12. Taylor,M., Butler,R., Chambers,S., Casimiro,M., Badii,F. and Merrick,M. (1996) The RpoN-box motif of the RNA polymerase sigma factor σ^{N} plays a role in promoter recognition. *Mol. Microbiol.*, **22**, 1045–1054.
13. Cannon,W., Missailidis,S., Smith,C., Cottier,A., Austin,S., Moore,M. and Buck,M. (1995) Core RNA polymerase and promoter DNA interactions of purified domains of σ^{N}: bipartite functions. *J. Mol. Biol.*, **248**, 781–803.
14. Merrick,M. and Chambers,S. (1992) The helix–turn–helix motif of σ^{54} is involved in recognition of the −13 promoter region. *J. Bacteriol.*, **174**, 7221–7226.
15. Chaney,M., Pitt,M. and Buck,M. (2000) Sequences within the DNA-crosslinking patch of σ^{54} involved in promoter recognition, sigma isomerisation and open complex formation. *J. Biol. Chem.*, **275**, 22104–22113.
16. Cannon,W., Gallegos,M.T., Casaz,P. and Buck,M. (1999) Amino-terminal sequences of σ^{54} (σ^{N}) inhibit RNA polymerase isomerisation. *Genes Dev.*, **13**, 357–370.
17. Syed,A. and Gralla,J.D. (1998) Identification of an N-terminal region of σ^{54} required for enhancer-responsiveness. *J. Bacteriol.*, **180**, 5619–5625.
18. Wigneshweraraj,S.R., Fujita,N., Ishihama,A. and Buck,M. (2000) Conservation of sigma-core RNA polymerase proximity relationships between the enhancer independent and enhancer dependent sigma classes. *EMBO J.*, **19**, 3038–3048.
19. Casaz,P. and Buck,M. (1997) Probing the assembly of transcription initiation complexes through changes in σ^{N} protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.
20. Casaz,P., Gallegos,M.T. and Buck,M. (1999) Systematic analysis of σ^{54} N-terminal sequences identifies regions involved in positive and negative regulation of transcription. *J. Mol. Biol.*, **292**, 229–239.
21. Cannon,W., Chaney,M. and Buck,M. (1999) Characterisation of holoenzyme lacking σ^{N} regions I and II. *Nucleic Acids Res.*, **27**, 2478–2486.
22. Wang,L. and Gralla,J.D. (1998) Multiple *in vivo* roles for the −12-region elements of σ^{54} promoters. *J. Bacteriol.*, **180**, 5626–5631.
23. Wang,J.T. and Gralla,J.D. (1996) The transcription initiation pathway of σ^{54} mutants that bypass the enhancer protein requirement. Implications for the mechanism of activation. *J. Biol. Chem.*, **271**, 32707–32713.
24. Wang,J.T., Syed,A., Hsieh,M. and Gralla,J.D. (1995) Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH_2-terminal leucine patch in σ^{54}. *Science*, **270**, 992–994.
25. Gallegos,M.T. and Buck,M. (2000) Sequences in σ^{54} Region I required for binding to early melted DNA and their involvement in sigma-DNA isomerisation. *J. Mol. Biol.*, **297**, 849–859.
26. Morris,L., Cannon,W., Clever-Martin,F., Austin,S. and Buck,M. (1994) DNA distortion and nucleation of local DNA unwinding within σ^{54} (σ^{N}) holoenzyme closed promoter complexes. *J. Biol. Chem.*, **269**, 11563–11571.
27. Cannon,W., Gallegos,M.T. and Buck,M. (2000) Isomerisation of a binary sigma-promoter DNA complex by enhancer binding transcription activators. *Nature Struct. Biol.*, **7**, 594–601.
28. Wigneshweraraj,S.R., Chaney,M.K., Ishihama,A. and Buck,M. (2001) Regulatory sequences in σ^{54} localise near the start of DNA melting. *J. Mol. Biol.*, in press.
29. Gallegos,M.T. and Buck,M. (1999) Sequences in σ^{54} determining holoenzyme formation and properties. *J. Mol. Biol.*, **288**, 539–553.
30. Jishage,M. and Ishihama,A. (1996) Regulation of RNA polymerase sigma subunit levels in *Escherichia coli*: intracellular levels of four species of sigma subunits under various growth conditions. *J. Bacteriol.*, **178**, 5447–5451.
31. Jovanovic,G., Rakonjac,J. and Model,P. (1999) *In vivo* and *in vitro* activities of *Escherichia coli* σ^{54} transcription activator, PspF and its DNA-binding mutant PspDeltaHTH. *J. Mol. Biol.*, **285**, 469–483.
32. Pavari,R., Pecht,I. and Soreq,H. (1983) A microfluorometric assay for cholinesterases, suitable for multiple kinetic determinations of picomoles of released thiocholine. *Anal. Biochem.*, **133**, 450–456.
33. Claverie-Martin,F. and Magasanik,B. (1992) Positive and negative effects of DNA bending on activation of transcription from a distant site. *J. Mol. Biol.*, **227**, 996–1008.
34. Claverie-Martin,F. and Magasanik,B. (1991) Role of integration host factor in the regulation of the glnHp2 promoter of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **88**, 1631–1635.
35. Chaney,M.K. and Buck,M. (1999) The σ^{54} DNA-binding domain includes a determinant of enhancer responsiveness. *Mol. Microbiol.*, **33**, 1200–1209.
36. Elliot,S. and Geiduschek,E.P. (1984) Defining a bacteriophage T4 late promoter: absence of a −35 region. *Cell*, **36**, 211–219.
37. Popham,D.L., Szeto,D., Keener,J. and Kustu,S. (1989) Function of a bacterial activator protein that binds to transcriptional enhancers. *Science*, **243**, 629–635.
38. Morett,E. and Buck,M. (1989) *In vivo* studies on the interaction of RNA polymerase σ^{54} with the *Klebsiella pneumoniae* and *Rhizobium meliloti* nifH promoters. *J. Mol. Biol.*, **210**, 65–77.
39. Sasse-Dwight,S. and Gralla,J.D. (1988) Probing the *E. coli* gln ALG upstream activation mechanism *in vivo*. *Proc. Natl Acad. Sci. USA*, **85**, 8934–8938.
40. Oguiza,J.A., Gallegos,M.T., Chaney,M.K., Cannon,W. and Buck,M. (1999) Involvement of σ^{N} DNA-binding domain in open complex formation. *Mol. Microbiol.*, **33**, 873–885.
41. Ishihama,A. (2000) Molecular anatomy of RNA polymerase using protein-conjugated metal probes with nuclease and protease activities. *Chem. Commun.*, **13**, 1091–1094.
42. Wang,L., Guo,Y. and Gralla,J.D. (1999) Regulation of σ^{54}-dependent transcription by core promoter sequences: role of −12 region nucleotides. *J. Bacteriol.*, **181**, 7558–7565.
43. Oguiza,J.A. and Buck,M. (1997) DNA binding domain mutants of sigma-N (σ^{N}, σ^{54}) defective between closed and stable open promoter complex formation. *Mol. Microbiol.*, **26**, 655–64.
44. Gross,C.A., Chan,C., Drombroski,A., Gruber,T., Sharp,M., Tupy,J. and Young,B. (1998) Mechanism of transcription. *Cold Spring Harbor Symp. Quant. Biol.*, **63**, 141–155.
45. Fenton,M.S., Lee,S.J. and Gralla,J.D. (2000) *E. coli* promoter opening and −10 recognition: mutational analysis of sigma 70. *EMBO J.*, **19**, 1130–1137.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 6, Issue of February 11, pp. 4210–4214, 2000
Printed in U.S.A.

Low Resolution Structure of the σ54 Transcription Factor Revealed by X-ray Solution Scattering*

(Received for publication, October 26, 1999)

Dmitri I. Svergun‡§¶, Marc Malfois‡, Michel H. J. Koch‡, Siva R. Wigneshweraraj‖, and Martin Buck‖

From the ‡European Molecular Biology Laboratory, Hamburg Outstation, EMBL c/o DESY, Notkestraße 85, D-22603 Hamburg, Germany, the §Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia, and the ‖Department of Biology, Imperial College of Science, Technology and Medicine, London SW7 2BB, United Kingdom

The σ54 RNA polymerase holoenzyme functions in enhancer-dependent transcription. The structural organization of the σ54 subunit of bacterial RNA polymerase in solution is analyzed by synchrotron x-ray scattering. Scattering patterns are collected from the full-length protein and from a large fragment able to bind the core RNA polymerase, and their low resolution shapes are restored using two *ab initio* shape determination techniques. The σ54 subunit is a highly elongated particle, and the core binding fragment can be unambiguously positioned inside the full-length protein. The boomerang-like shape of the core binding fragment is similar to that of the atomic model of a fragment of the *Escherichia coli* σ70 protein, indicating that, although the σ54 and σ70 factors are unrelated by primary sequence, they may share some structural similarity. Potential DNA binding surfaces of σ54 are also predicted by comparison with the σ54 core binding fragment.

Transcriptional regulation of gene expression is a major area where information flow from DNA is controlled, often by activation of RNA polymerase function, and is central to regulating the initiation rates (1, 2). The expression of diverse gene sets in bacteria requires the specialized σ54-containing RNA polymerase (2), likened in its properties to the eukaryotic RNA polymerase II (3). Transcription activation involves a poorly defined interaction of activator proteins, bound at remote sites, with the σ54 holoenzyme bound at a promoter as a transcriptionally inactive closed complex (4, 5). Our objectives are to determine the contributions of the σ54 protein to RNA polymerase function and activation through structure-function studies. The σ54 protein has a modular domain organization (4–8), and individual domains can be prepared as active partial σ54 sequences, potentially allowing the assignment of structural features seen in the complete σ54 molecule (9).

σ factors contribute to several functions of the holoenzyme, and although discrete activities are known to reside within individual domains, the domains appear to interact with each other for the full function of the σ54. In the present paper, the structural organization of the 54-kDa monomeric σ54 protein from *Klebsiella pneumoniae* bacteria is analyzed using x-ray solution scattering (SAXS),[1] the method yielding low resolution structural information at nearly physiological conditions. Two methods of *ab initio* shape restoration (10–12) are used to establish the shapes of the full-length protein and of its 30-kDa fragment (previously obtained as a stable product of σ54 proteolysis). The latter shape can be unambiguously positioned within the former, providing insight about the structural organization of the σ54 molecule. Further, the shape of the 30-kDa fragment exhibits similarity with that of the crystal structure of a fragment of the *Escherichia coli* σ70 protein that also binds the core RNA polymerase. A prediction of potential σ54 DNA binding surfaces is also made.

EXPERIMENTAL PROCEDURES

Sample Preparation—K. pneumoniae σ54 protein (amino acids 1–477) and the 30-kDa core binding fragment (amino acids 70–324) with an 11-amino acid amino acid extension (MARIRARGSSR) at its N terminus were overproduced as soluble proteins in *E. coli* and purified as described previously (13). They were concentrated to 15–20 mg/ml and buffer-exchanged into 10 mM Tris-HCl, 5% glycerol, pH 8.0, at 4 °C by centrifugal ultrafiltration. Concentrated samples were stored at −70 °C and thawed on ice prior to the scattering experiments. Protein concentrations of σ54 and the 30-kDa fragment were determined side by side using the Bio-Rad dye assay and bovine serum albumin as standard.

Scattering Experiments and Data Treatment—The synchrotron radiation x-ray scattering data were collected using standard procedures on the X33 camera (14–16) of the EMBL on the storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY) and multiwire proportional chambers with delay line readout (17). Samples at concentrations between 2 and 15 mg/ml were measured at a wavelength $\lambda = 0.15$ nm for sample-detector distances of 2.9 and 1.4 m covering the momentum transfer ranges $0.20 < s < 2.0$ nm^{-1} and $0.35 < s < 5.0$ nm^{-1}, respectively ($s = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle). The data were normalized to the intensity of the incident beam, corrected for the detector response; the scattering of the buffer was subtracted; and the difference curves were scaled for concentration using the program SAPOKO.[2] The curves recorded at a sample-detector distance of 2.9 m were extrapolated to zero concentration and merged with the data obtained at 1.4 m.

The maximum dimensions D_{max} of the σ54 and the 30-kDa core were estimated from the experimental data using the orthogonal expansion program *ORTOGNOM* (19). The distance distribution functions $p(r)$ and the radii of gyration R_g were evaluated by the indirect Fourier transform program *GNOM* (20–21). The molecular masses of the solutes were estimated by comparison of the extrapolated forward scattering I(0) with that of a reference solution of bovine serum albumin.

Prior to the shape analysis, undesirable contributions from the scattering due to the internal particle structure at higher scattering angles that become significant above approximately $s = 2.0$ nm^{-1} were re-

* This work was supported by European Commission Grant BIO4-CT97-2143 (to D. I. S. and M. B.) and by Grant B05984 from the Biotechnology and Biological Sciences Research Council (to M. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: EMBL c/o DESY, Notkestraße 85, D-22603 Hamburg, Germany. Tel.: 49 40 89902 125; Fax: 49 40 89902 149; E-mail: Svergun@EMBL-Hamburg.de.

[1] The abbreviations used are: SAXS, small angle x-ray scattering; DAM, dummy atoms model.

[2] D. I. Svergun and M. H. J. Koch, unpublished results.

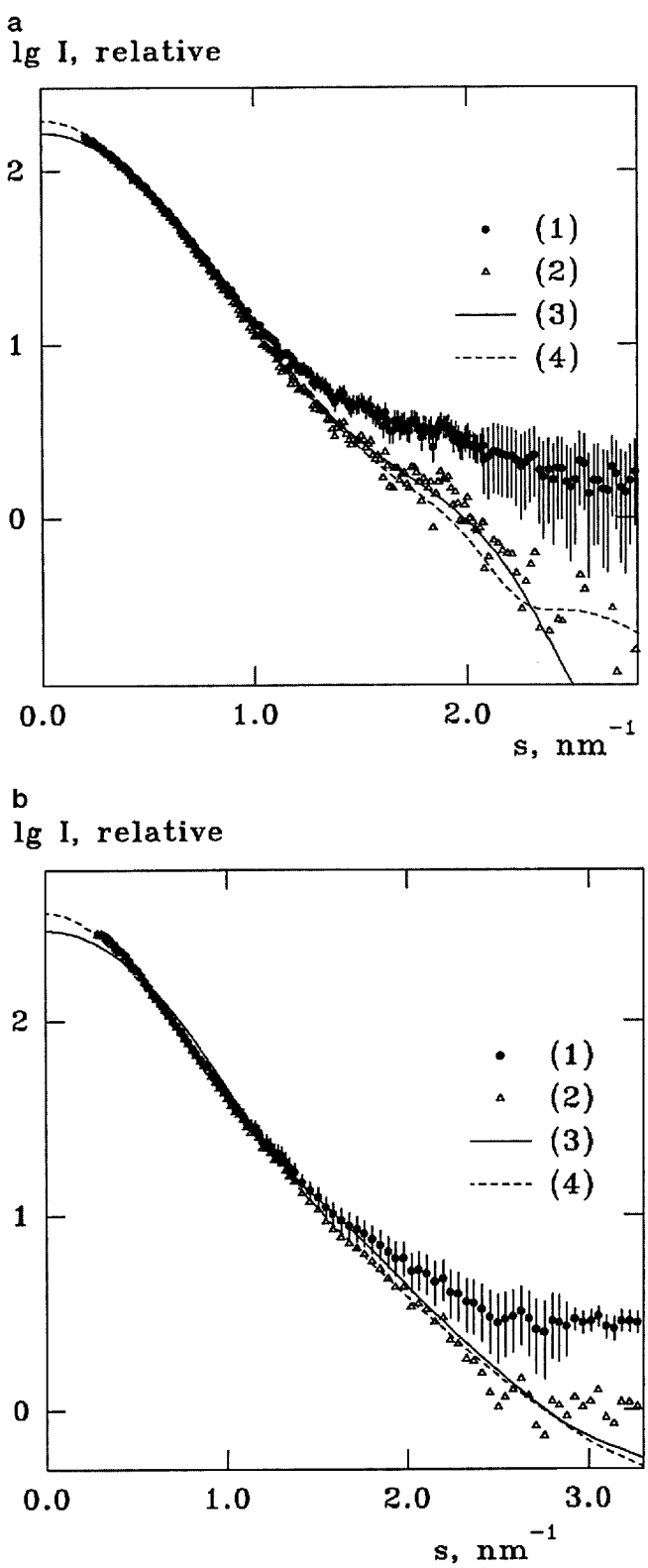


FIG. 1. **Experimental solution scattering curves of σ54 (*a*) and the 30-kDa fragment (*b*) and calculated scattering from the models.** *1*, composite experimental curves obtained by merging the data for the two sample-detector settings; *2*, shape scattering curves after modifying for the scattering from internal structure; *3*, scattering from the envelope models; *4*, scattering from the DAMs.

moved by subtracting a constant from the experimental data. This procedure ensures that the intensity decays as s^{-4} following Porod's (22) law for homogeneous particles and yields an approximation of the "shape scattering" curve (*i.e.* scattering due to the excluded volume of the homogeneous particle with a constant density). The shape scattering curves in the ranges up to s_{max} = 2.8 nm^{-1} (σ54) and 3.3 nm^{-1} (30-kDa core) were used to compute the excluded volumes V of the hydrated particles (22) and for the *ab initio* shape determination. The outer portions of the curves dominated by the scattering from the internal structure were discarded in the further analysis.

Shape Determination—The low resolution particle shapes were restored from the experimental data using two *ab initio* procedures. In the method of Svergun *et al.* (10, 11), the shape is represented by an angular envelope function $r = F(\omega)$, where (r,ω) are spherical coordinates. The envelope is parameterized as follows (23),

$$F(\omega) = \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{lm} Y_{lm}(\omega) \qquad \text{(Eq. 1)}$$

where $Y_{lm}(\omega)$ are spherical harmonics and the multipole coefficients f_{lm} are complex numbers. The maximum order of the spherical harmonics L is selected to keep the number of free parameters $M = (L + 1)^2 - 6$ close to the number of Shannon channels $N_s = D_{max}s_{max}/\pi$ in the experimental data (10, 24–25). The scattering intensity of the envelope is as follows,

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} |A_{lm}(s)|^2 \qquad \text{(Eq. 2)}$$

where the partial amplitudes $A_{lm}(s)$ are calculated from the coefficients f_{lm} using the recurrence relation of Svergun and Stuhrmann (26). These coefficients are determined by the program *SASHA* (10, 11) by minimizing the discrepancy χ between the calculated and the experimental curves as follows,

$$\chi = \sqrt{\frac{1}{N-1} \sum_{j=1}^{N} \left[\frac{I(s_j) - I_{exp}(s_j)}{\sigma(s_j)} \right]^2} \qquad \text{(Eq. 3)}$$

where N is the number of the experimental points and $I_{exp}(s_k)$ and $\sigma(s_k)$ are the experimental intensity and its S.D. in the kth point, respectively. The spatial resolution of the envelope function representation is approximately $\delta r = \sqrt{5}\pi R_g/(\sqrt{3}\,(L + 1))$.

Another *ab initio* procedure employs a particular implementation of a general method (12). A sphere of diameter D_{max} is filled by densely packed small spheres (dummy atoms) of radius $r_0 \ll D_{max}$. The structure of this dummy atoms model (DAM) is defined by a configuration X, assigning an index to each atom corresponding to solvent (0) or solute particle (1). The scattering intensity from the DAM is computed using Equation 2 with the following partial amplitudes,

$$A_{lm}(s) = i^l \sqrt{2/\pi} \sum_j j_l(sr_j) Y^*_{lm}(\omega_j) \qquad \text{(Eq. 4)}$$

where the sum runs over the dummy atoms with $X_j = 1$ (particle atoms); r_j,ω_j are their polar coordinates; and $j_l(x)$ denotes the spherical Bessel function. The number of dummy atoms in the search model $M_{DAM} \approx (D_{max}/r_0)^3 \gg 1$ significantly exceeds the number of Shannon channels. To reduce the effective number of free parameters in the model, the method searches for a configuration X minimizing $f(X) = \chi^2 + \alpha P(X)$, where the looseness penalty $P(X)$ with a positive weight $\alpha > 0$ ensures that the DAM has low resolution with respect to the packing radius r_0. The weight of the penalty is selected to have significant penalty contribution at the end of the minimization. The latter is performed starting from a random approximation using the simulated annealing method (27); details of the procedure are described elsewhere (12). The spatial resolution provided by the DAM is defined by the range of the momentum transfer as $\delta r = 2\pi/s_{max}$.

Since the shape determination methods yield models with an arbitrary orientation and handedness, they were appropriately rotated for comparison with each other and also with the atomic model of a σ70 subunit fragment from *E. coli* RNA. The latter was taken from the Protein Data Bank (28), entry 1sig (29). The model of the fragment of the B-DNA was generated with the program Turbo-FRODO (30), and the three-dimensional models were displayed using the program ASSA (31).

RESULTS

The experimental scattering curves from the σ54 and its 30-kDa fragment are presented in Fig. 1, and the structural parameters computed from these curves are listed in Table I.

TABLE I
Experimental and model parameters of σ54 and of its 30-kDa fragment

Parameter/Sample	σ54			30-kDa fragment		
	Exp.[a]	Env.[b]	DAM[c]	Exp.	Env.	DAM
Mass (kDa)	58 ± 6			32 ± 4		
D_{max} (nm)	13.0 ± 0.5	10.4	12.5	11.0 ± 0.5	8.3	11.0
R_g (nm)	3.43 ± 0.06	3.04	3.39	3.05 ± 0.05	2.51	3.08
V (nm^3)	155 ± 10	144	152	85 ± 6	72.2	86.7
Free parameters	11.6[d]	19	2243	11.6[d]	19	295
Discrepancy χ		1.79	1.18		1.49	0.88
Resolution (nm)		2.8	2.2		2.4	1.9

[a] Exp, calculated from the experimental data.
[b] Env, parameters of the envelope models.
[c] DAM, parameters of the dummy atoms models.
[d] Number of Shannon channels is given.

The molecular masses of the solutes indicate that both σ54 and the 30-kDa fragment are monomeric in solution, in agreement with the results of analytical ultracentrifugation.[3] The experimental values of D_{max} and R_g suggest that the two macromolecules, and especially σ54, are rather anisometric. The volumes of the hydrated σ54 and the 30-kDa fragment are twice the dry volumes expected from the molecular masses (about 75 and 42 nm^3, respectively). This reveals a very high hydration (0.4 g of H_2O/g of protein) characteristic for extended macromolecules with high specific surface accessible to the solvent. The profiles of the distance distribution functions $p(r)$ in Fig. 2 are typical for elongated particles (32). It follows that σ54 is a very elongated protein and that the 30-kDa fragment, although having a 2-nm smaller maximum diameter, is still rather elongated.

The portions of the scattering curves used for *ab initio* shape determination contained N_s = 11.6 Shannon channels both for σ54 and for the 30-kDa fragment, allowing the use of harmonics up to L = 4 (19 independent parameters) in the shape determination using the envelope functions (10). The restored envelopes are presented in Fig. 3 (*left column*), and the fits to the experimental data are displayed in Fig. 1, *a* and *b*. The envelope of the 30-kDa fragment can be unambiguously positioned inside that of σ54. Moreover, a boomerang-like shape of the core correlates well with the gross structure of the core RNA polymerase-binding fragment of the *E. coli* σ70 protein (the atomic model of the latter is superimposed with the low resolution models in Fig. 3). Nevertheless, the shape reconstruction using envelope functions has its limitations for elongated particles like σ54. To keep the model at low resolution and the number of model parameters reasonably small, higher spherical harmonics are omitted in the description of the envelope (Equation 1). This ensures uniqueness of the model restoration (10) but leads to a termination effect especially noticeable for anisometric particles for which the higher harmonics play a significant role. As a result, the radius of gyration and the maximum size of the restored envelopes are lower than those determined experimentally (Table I), and the calculated curves display systematic deviations from the experimental data (Fig. 1).

A much less restricted description of the particle shape is provided by the alternative DAM technique. The search volume for the σ54 contained M_{DAM} = 2243 dummy atoms with a packing radius r_a = 0.45 nm within a sphere with the diameter D_{max} = 13 nm. Of those, 295 dummy atoms were attributed to the particle in the final model of σ54, and they further defined the search volume for the 30-kDa fragment. The obtained DAMs presented in Fig. 3 (*right column*), yield better fits to the experimental data in Fig. 1 than those provided by the envelope models and also the structural parameters that neatly agree with the experimental values (Table I). To verify the uniqueness of the shape restoration using DAM, several independent restorations were performed using different conditions (by varying the packing radius r_a and by starting from random initial configurations). These restorations yielded reproducible results similar to that in Fig. 3 (*right column*). In particular, the shape of the 30-kDa fragment restored *ab initio* within the spherical search volume with D_{max} = 11 nm was very similar to that presented in Fig. 3.

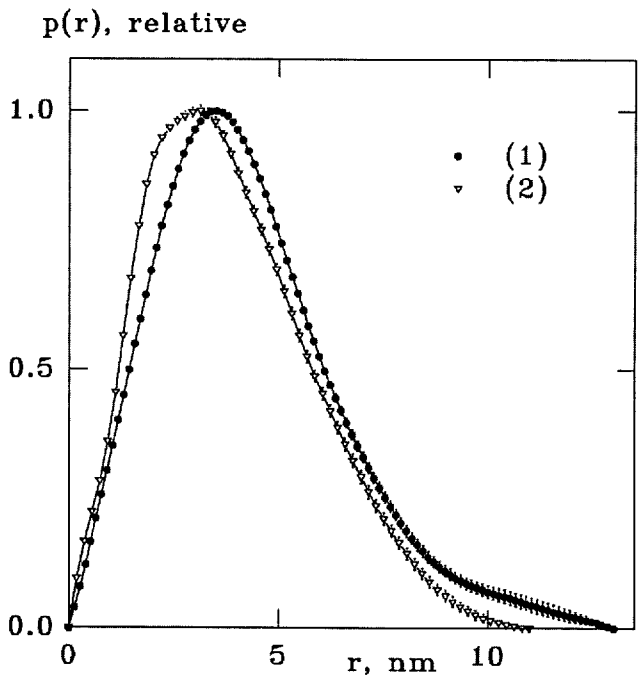


FIG. 2. **Distance distribution functions of σ54 (*1*) and the 30-kDa fragment (*2*).**

DISCUSSION

An *ab initio* shape restoration from solution scattering data is not unique unless the search model is kept at low resolution (10). The two *ab initio* procedures employed here use different restrictions. In the envelope restoration, higher spherical harmonics are omitted in the description of the envelope function so that the number of the adjustable parameters is comparable with that of the Shannon channels in the scattering data. In the DAM retrieval, the number of parameters is much larger ($M_{DAM} \approx 10^3$), and the looseness penalty ensures that the final model is compact and interconnected. Given the differences in parameterization and in the numbers of model parameters, similarities in the low resolution models independently restored by the two methods are remarkable. The differences between the two restorations in Fig. 3 are largely attributed to the termination effects in the envelope method, which are

[3] D. Scott and J. Hoggett, personal communication.

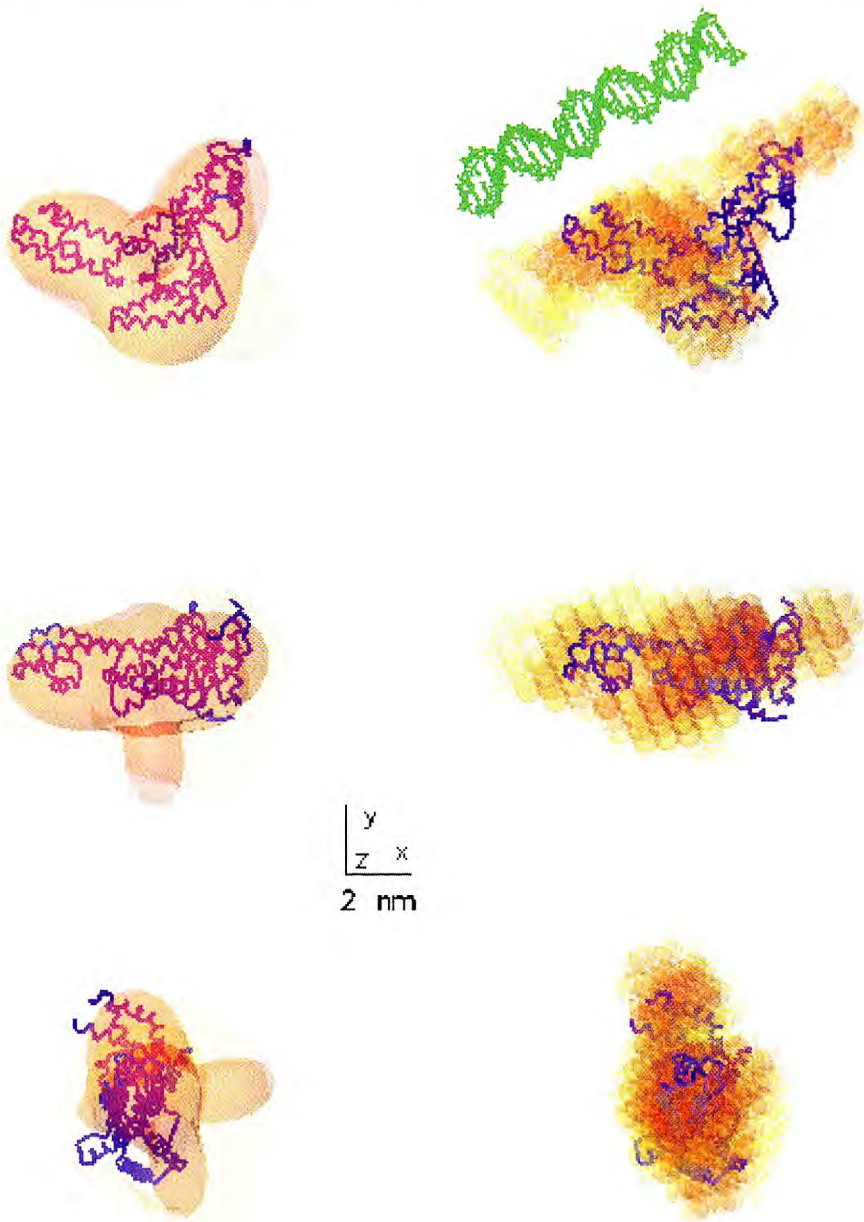


FIG. 3. **Low resolution models of σ54 (*yellow*) and 30-kDa fragment (*red*).** *Left*, envelope models (semitransparent envelopes); *right*, DAMs (semitransparent spheres of radius r_a = 0.45 nm). The *middle* and *bottom rows* are rotated counterclockwise by 90° around the *x* and *y* axes, respectively. The models are superimposed with the crystal structure of the σ70 subunit fragment from *E. coli* RNA (*blue* C-α trace). A 28-base pair-long B-DNA fragment comprising the σ54 binding site (*green*) is shown *above* potential DNA binding surfaces (*top orientation* in the *right panel*).

significant for elongated particles like σ54. It should be stressed that although the shapes obtained with the DAM technique are still low resolution models, their better nominal resolution (Table I) and more detailed appearance compared with the envelope models are fully justified by better fits to the experimental data.

DNA footprinting experiments have shown that σ54 interacts with promoter DNA from −33 to −5. Major points of contact are within four base pairs (TGCA) centered at −12.5 and seven (CTGGCAC) centered at −24, within the consensus sequence of σ54 DNA-binding sites. The distance between these contact centers suggests that unless σ binds to a non-B-DNA conformation, interactions between σ and DNA will involve contacts made by structures within more than one of the well developed lobes revealed by the SAXS analysis. Superposition of the low resolution dummy atoms models with a 28-base pair B-DNA fragment in Fig. 3 (*top right orientation*) summarizes a relationship between σ54, the 30-kDa fragment, and promoter DNA that can account for the known extent of σ-promoter DNA interactions. The 30-kDa fragment analyzed by SAXS lacks the σ DNA-binding domain and region I sequences that also contribute to DNA recognition. It therefore seems probable that the lobe missing from the 30-kDa fragment (Fig. 3) but present in σ54 includes some of the structures that direct promoter binding. The upper surface of σ54 depicted in Fig. 3 (*top right orientation*) further distinguishes the full-length protein from the 30-kDa fragment and may contribute additional DNA contacting structures. Together, the lobe and the upper surface can provide DNA-contacting surfaces of approximately the necessary extent to fully interact with promoter DNA. The direction of the DNA sequence displayed in Fig. 3 is arbitrary, since no experimental data correlating which structure recognizes the −12 or −24 sequences are available.

The low resolution solution structure for σ54 determined in this work provides the first initial model for a protein that functions to convert *E. coli* RNA polymerase into an enhancer-dependent enzyme (3). Clearly developed structures are present that probably represent either a discrete individual domain or alternately domains that are directly interacting within the

tertiary structure. The amino-terminal domain required for transcription silencing and activator responsiveness and the C-terminal DNA binding domain are thought to interact (9). Comparison of the structure of the 30-kDa fragment that lacks both of these domains with the full-length protein structure is consistent with such an interaction, since one lobe is clearly absent (Fig. 3). Possibly, both the amino-terminal domain and C-terminal domain contribute to this lobe and do directly interact with each other. Another striking observation is that the fragment of σ70 for which the atomic structure has been determined has a similar shape and size compared with the 30-kDa fragment of σ54 (Fig. 3). Both fragments bind to the same core RNA polymerase. Assuming the shape of the fragments is related to core binding, it seems that σ70 and σ54 are probably located in similar places on the core. The sequence differences that exist between σ54 and σ70 are, however, sufficient to account, at least in part, for the different properties of the two holoenzymes. Some interactions with core may nevertheless be conserved between the two σ types.

CONCLUSION

The low resolution solution structure of σ54 shows a well developed anisometric shape. Similarity to a part of σ70 that binds to the core RNA polymerase and interacts with the −10 promoter DNA is evident. Now that crystallography and cryo-electron microscopy are beginning to reveal the structures of multisubuit RNA polymerases (18), one can also begin to evaluate the functional significance of this similarity. These structural approaches together with the use of tethered iron chelate methods for examining proximity relationships within transcription complexes should help determine the organization of RNA polymerase holoenzymes.

REFERENCES

1. Helmann, J. D., and de Haseth, P. L. (1999) *Biochemistry* **38,** 5959–5967
2. Merrick, M. J. (1993) *Mol. Microbiol.* **10,** 903–909
3. Wang, J. T., Syed, A., Hsieh, M., and Gralla, J. D. (1995) *Science* **270,** 992–994
4. Cannon, W., Missailidis, S., Smith, C., Cottier, A., Austin, S., Moore, M., and Buck M. (1995) *J. Mol. Biol.* **248,** 781–803
5. Wong, C., Tintut, Y., and Gralla, J. D. (1994) *J. Mol. Biol.* **236,** 81–90
6. Cannon, W. V., Chaney, M. K., Wang X-Y. and Buck, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 5006–5011
7. Buck, M., and Cannon, W. (1992) *Nature* **358,** 422–424
8. Cannon, W., Claverie-Martin, F., Austin, S., and Buck, M. (1994) *Mol. Microbiol.* **11,** 227–236
9. Casaz, P., and Buck, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12145–12150
10. Svergun, D. I., Volkov, V. V., Kozin, M. B., and Stuhrmann, H. B. (1996) *Acta Crystallogr. Sec. A* **52,** 419–426
11. Svergun, D. I., Volkov, V. V., Kozin, M. B., Stuhrmann, H. B., Barberato, C., and Koch, M. H. J. (1997) *J. Appl. Crystallogr.* **30,** 798–802
12. Svergun, D. I. (1999) *Biophys. J.* **76,** 2879–2886
13. Missailidis, S., Cannon, W. V., Drake, A., Wang, X. Y., and Buck, M. (1997) *Mol. Microbiol.* **24,** 653–664
14. Koch, M. H. J., and Bordas, J. (1983) *Nucl. Instrum. Methods* **208,** 461–469
15. Boulin, C., Kempf, R., Koch, M. H. J., and McLaughlin, S. M. (1986) *Nucl. Instrum. Methods Sec. A* **249,** 399–407
16. Boulin, C. J., Kempf, R., Gabriel, A., and Koch M. H. J. (1988) *Nucl. Instrum. Methods Sec. A* **269,** 312–320
17. Gabriel, A., and Dauvergne, F. (1982) *Nucl. Instrum. Methods* **201,** 223–224
18. Mooney, R. A., and Landick, R. (1999) *Cell* 687–690
19. Svergun, D. I. (1993) *J. Appl. Crystallogr.* **26,** 258–267
20. Svergun, D. I., Semenyuk, A. V., and Feigin, L. A. (1988) *Acta Crystallogr. Sec. A* **44,** 244–250
21. Svergun, D. I. (1992) *J. Appl. Crystallogr.* **25,** 495–503
22. Porod, G. (1982) in *Small-angle X-ray Scattering* (Glatter, O., and Kratky, O., eds) pp. 17–51, Academic Press, Inc., London
23. Stuhrmann, H. B. (1970) *Z. Phys. Chem.* **72,** 177–198
24. Shannon, C. E., and Weaver, W. (1949) *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL
25. Moore, P. B. (1980) *J. Appl. Crystallogr.* **13,** 168–175
26. Svergun, D. I., and Stuhrmann, H. B. (1991) *Acta Crystallogr. Sec. A* **47,** 736–744
27. Kirkpatrick, S., Gelatt, C. D., Jr., and Vecci, M. P. (1983) *Science* **220,** 671–680
28. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542
29. Malhotra, A., Severinova, E., and Darst, S. A. (1996) *Cell* **87,** 127–136
30. Roussel, A., and Cambilleau, C. (1989) *Silicon Graphics Geometry Partners Directory*, pp. 77–79, Silicon Graphics, Mountain View, CA
31. Kozin, M. B., Volkov, V. V., and Svergun, D. I. (1997) *J. Appl. Crystallogr.* **30,** 811–815
32. Feigin, L. A., and Svergun, D. I. (1987) *Structure Analysis by Small-angle X-ray and Neutron Scattering*, pp. 83–87, Plenum Press, New York

References

Aiba H, Nakasai F, Mizushima S & Mizuno T (1989) "Evidence for the physiological importance of the phosphotransfer between the two regulatory components, EnvZ and OmpR, in osmoregulation in Escherichia coli". J Biol Chem, vol. 264, pp. 14090-14094.

Aki T, Choy HE, Adhya S (1996) "Histone-like protein HU as a specific transcriptional regulator: co-factor role in repression of gal transcription by GAL repressor". Genes Cells, vol. 1, pp.179-188.

Allison L A (2000) "The role of sigma factors in plastid transcription". Biochemie, vol. 82, pp. 537-548.

Anand G S, Goudreau P N & Stock A M (1998) "Activation of methylesterase CheB: evidence of a dual role for the regulatory domain". Biochemistry, vol. 37, pp.14038-14047.

Ansari A, Chael M & O'Halloran T (1992) "Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg-MerR". Nature, vol. 355, pp.87-89.

Antson A A, Otridge J, Brzozowski A M, Dodson E J, Dodson G G, Wilson K S, Smith T M, Yang M, Kurecki T & Gollnick P "The structure of trp RNA-binding attenuation protein." Nature, vol. 374, pp. 693-700.

Arthur M T, Anthony L C & Burgess R R (2000) "Mutational Analysis of $\beta'_{260-309}$, a σ^{70} Binding Site Located on Escherichia coli Core RNA Polymerase". J Biol Chem, vol. 275, pp. 23113-23119.

Arthur T M & Burgess R R (1998) "Localization of a σ^{70} binding site on the N-terminus of the Escherichia coli RNA polymerase β' subunit". J Biol Chem 273, pp. 31381-31387.

Angerer A, Enz S, Ochs M & Braun V (1995) "Transcriptional regulation of ferric citrate transport in Escherichia coli. FecI belongs to subfamily of σ^{70}-type factors that respond to extracytoplasmic stimuli." Mol Microbiol, vol. 18, pp. 163-174.

Austin S & Dixon R (1992) "The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent". EMBO J vol. 11, pp. 2219-2228.

Barrios H, Valderrama B & Morett E (1999) "Compilation and analysis of sigma(54)-dependent promoter sequences". Nucl Acids Res 15, pp. 4305-4313.

Bell A, Bains M, Hancock R E (1991) "Pseudomonas aeruginosa outer membrane protein OprH: expression from the cloned gene and function in EDTA and gentamicin resistance." J Bacteriol, vol. 173, pp. 6657-6666.

Belyaeva T A, Bown J A, Fujita N, Ishihama A & Busby S J (1996) "Location of the C-terminal domain of the RNA polymerase alpha subunit in different open complexes at the Escherichia coli galactose operon regulatory region. Nucl Acids Res, vol. 24, pp. 2242-2251.

Bertoni G, Fujita N, Ishihama A & de Lorenzo V (1998) "Active recruitment of σ^{54}-RNA polymerase to the Pu promoter of Pseudomonas putida: role of IHF and αCTD". EMBO J, vol. 17, pp. 5120-5128.

Bhende P M & Egan S M (2000) "Genetic evidence that transcription actvation by RhaS involves specific amino acid contacts with sigma 70". J Bacteriol, vol. 182, pp. 4959-4969.

Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, Kahn D & Samama J P (1999)

"Conformational changes induced by phosphorylation of the FixJ receiver domain". Structure Fold Des., vol. 7, pp. 1505-1515.

Blat Y & Eisenbach M (1994) "Phosphorylation-dependent binding of the chemotaxis signal molecule CheY to its phosphatase, CheZ". Biochemistry, vol. 33, pp., 902-906.

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

Blattner F R, Plunkett G III, Bloch C A, Perna N T, Burland V et al. (1997) "The complete genome sequence of Escherichia coli K-12." Science, vol. 277, pp. 1453-1462.

Bourret R B, Borkovich K A & Simon M I (1991) "Signal transduction pathways involving protein phosphorylation in prokaryotes". Annual Rev Biochem, vol. 60, pp. 401-441.

Bown J A, Kolb A, Meares C F, Ishihama A, Minchin S D & Busby S J (2000) "Positioning of region 4 of the Escherichia coli RNA polymerase sigma(70) subunit by a transcription activator". J Bacteriol, 182, 2982-2984.

Bown J A, Owens J T, Meares C F, Fujita N, Ishihama A, Busby S J & Minchin S D (1999) "Organization of open complexes at Escherichia coli promoters. Location of promoter DNA sites close to region 2.5 of the sigma70 subunit of RNA polymerase". J Biol Chem, 274, 2263-2270.

Brodolin K, Mustaev A, Severinov K & Nikiforov V (2000) "Identification of RNA Polymerase β Subunit Segment Contacting the Melted Region of the lacUV5 Promoter". J Biol Chem, 275, 3661-3666

Brown J L, North S & Bussey H (1993) "SKN7, a yeast multicopy suppressor of a mutation affecting cell wall beta-glucan assembly, encodes a product with domains homologous to prokaryotic two-component regulators and to heat shock transcription factors" J Bacteriol, vol. 175, pp. 6908-15.

Brown K L & Hughes K T (1995) "The role of anti-sigma factors in gene regulation". Mol Microbiol, vol. 16, pp. 397-404.

Bu Z, Koide S & Engelman D M (1998) "A solution SAXS study of Borrelia burgdorferi OspA, a protein containing a single-layer beta-sheet". Protein Sci, 7, pp. 2681-2683.

Buck M, Gallegos M T, Studholme D J, Guo Y & Gralla J D (2000) "The Bacterial Enhancer-Dependent σ^{54} (σ^{N}) Transcription Factor". J Bacteriol, vol. 182, pp. 4129-4136.

Buck M & Cannon W V (1992) "Specific binding of the transcription factor σ^{54} to promoter DNA". Nature, vol. 358, pp. 422-424.

Buck M & Cannon W V (1989) "Mutations in the RNA polymerase recognition sequence of the Klebsiella pneumoniae nifH promoter permitting transcriptional activation in the absence of NifA binding to upstream activator sequences." Nucl Acids Res, vol. 17, pp. 2597-2612.

Buck M, Khan H, Dixon R (1985) "Site-directed mutagenesis of the Klebsiella pneumoniae nifL and nifH promoters and in vivo analysis of promoter activity". Nucl Acids Res, vol. 13, pp. 7621-7638.

Busby S, Lloyd G, Belyaeva T, Rhodius V & Savery N (1998) "Bacterial gene regulatory proteins: organisation and mechanism of action". Mol Microbiol, vol. 103, pp. 125-139.

Busby S & Ebright RH (1997) "Transcription activation at class II CAP-dependent promoters". Mol Microbiol, vol. 23, pp. 853-859.

Busby S & Ebright R (1994) "Promoter structure, promoter recognition, and transcription activation in procaryotes". Cell, vol. 79, pp.889-896.

Burgess R R, Arthur T M & Pietz B C (1998) "Interaction of Escherichia coli sigma 70 with core RNA polymerase". Cold Spring Harb Symp Quant Biol, vol. 63, pp. 277-287.

Callaci S & Heyduk T (1998) "Conformation and DNA binding properties of the single-stranded DNA binding region of σ^{70} subunit from Escherichia coli RNA polymerase are modulated by an interaction with the core enzyme." Biochemistry, vol. 37, 3312-3320.

Callaci S, Heyduk E & Heyduk T (1998) "Conformational Changes of Escherichia coli RNA polymerase σ^{70} Factor Induced by Binding of the Core Enzyme." J Biol Chem, vol. 273, pp. 32995-33001.

Cannon W, Gallegos M T & Buck M (2000) "Isomerisation of a binary sigma-promoter DNA complex by enhancer binding transcription activators". Nat Struct Biol, vol. 7, pp. 594-601.

Cannon W, Gallegos M T, Casaz P & Buck M. (1999) "Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization". Genes Dev, vol. 13, 357-370.

Cannon W V, Chaney M K, Wang X-Y & M Buck (1997) "Two domains within σ^{N} (σ^{54}) cooperate for DNA binding", Proc Natl Acad Sci USA, vol. 94, pp. 5006-5011.

Cannon W V, Claverie-Martin F, Austin S & M Buck (1994) "Identification of a DNA-contacting surface in the transcription factor sigma 54." Mol Microbiol, vol. 11, pp. 227-236.

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 sigma N: bipartite functions" J Mol Biol, vol. 248, pp. 781-803.

Cannon W, Claverie-Martin F, Austin S & Buck M (1993) "Core RNA polymerase assists binding of the transcription factor σ^{54} to promoter DNA". Mol Microbiol, vol. 8, pp. 287-298.

Cannon W & Buck M (1989) "Mutations in the RNA polymerase recognition sequence of the Klebsiella pneumoniae nifH promoter permitting transcriptional activation in the absence of NifA binding to upstream activator sequences". Nucl Acids Res, vol. 17, pp. 2597-2612.

Carmona M & Magasanik B (1996) "Activation of Transcription at σ^{54}-dependent Promoters on Linear Templates Requires Intrinsic or Induced Bending of the DNA." J Mol Biol, vol, 261, pp. 348-365.

Carmona M & de Lorenzo V (1999) "Involvement of the FtsH (HflB) Protease in the activity of σ^{54} promoter." Mol Microbiol, vol. 31, pp. 261-270.

Casaz P & Buck M (1999) "Region I modifies DNA-binding domain conformation of Sigma 54 withing the holoenzyme". J Mol Biol, 285, 507-514.

Casaz P, Gallegos M T & Buck M (1999) "Systematic analysis of σ^{54} N-terminal sequences identifies regions involved in positive and negative regulation of transcription". J Mol Biol, 292, 229-239.

Casaz P & Buck M (1997) "Probing the assembly of transcription initiation complexes through changes in sigma N protease sensitivity". Proc Natl Acad Sci USA, 94, 12145-12150.

Chaney M, Pitt M & Buck M (2000) "Sequences within the DNA-crosslinking patch of sigma54 involved in promoter recognition, sigma isomerisation and open complex formation". J Biol Chem, vol. 275, 22104-22113.

Chaney M K & Buck M (1999) "The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness". Mol Microbiol, 33, 1200-1209.

Chang C, Kwok S F, Bleecker A B & Meyerowitz E M (1993) "Arabidopsis ethylene-response gene ETR1: similarity of product to two-component regulators". Science, vol. 262, pp. 539-544.

Cheetham GM & Steitz TA (2000) "Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases". Curr Opin Struct Biol, vol. 10, pp. 117-123.

Cheetham G M T, Jeruzalmi D & Steiz T A (1999) "Structural basis for initiation of transcription from an RNA polymerase-promoter complex". Nature, vol. 399, pp. 80-82.

Chen Y, Ebright Y W & Ebright R H (1994) "Identification of the target of a transcription activator by protein-protein photocrosslinking". Science, vol. 265, pp. 90-92.

Chen Y F & Helmann J D (1997) "DNA-melting at the Bacillus subtilis flagellin promoter nucleates near –10 and expands unidirectionally." J Mol Biol, vol. 267, pp. 47-59.

Claverie-Martin F & Magasanik B (1992) "Positive and Negative Effects of DNA Bending on Activation of Transcription from a Distant Site". J Mol Biol, vol. 227, pp. 996-1008.

Claverie-Martin F & Magasanik B (1991) "Role of integration host factor in the regulation of the glnHp2 promoter of Escherichia coli". Proc Natl Acad Sci U S A, vol. 88, pp. 1631-1635.

Colland F, Fujita N, Kotlarz D, Bown J A, Meares C F, Ishihama A, Kolb A (1999) "Positioning of sigma(S), the stationary phase sigma factor, in Escherichia coli RNA polymerase-promoter open complexes." EMBO J, vol.18, pp. 4049-4059.

Colland F, Orsini G, Brody EN, Buc H & Kolb A (1998) "The bacteriophage T4 AsiA protein: a molecular switch for sigma 70-dependent promoters". Mol Microbiol, vol. 27, pp. 819-829.

Coppard J R & Merrick M J (1991) "Cassette mutagenesis implicates a helix-turn-helix motif in promoter recognition by the novel RNA polymerase sigma factor sigma 54". Mol Microbiol, vol. 5, pp. 1309-1317.

Coulombe B & Burton Z F (1999) "DNA Bending and Wrapping around RNA Polymerase: a "Revolutionary" Model Describing Transcriptional Mechanisms". Microbiol Biol Rev, vol. 63, pp. 457-478.

Cramer P, Bushnell D A, Fu J, Gnatt A L, Maier-Davis B, Thompson N E, Burgess R R, Edwards A M, David P R & Kornberg R D (2000) "Architecture of RNA polymerase II and implications for the transcription mechanism. Science, vol. 288, pp. 640-649.

Datwyler S A & Meares C F (2000) "Protein-protein interactions mapped by artificial proteases: where σ bind to RNA polymerase" Trend Biol Sci, vol. 25, pp. 408-414.

Debarbouille M, Martin-Verstraete I, Kunst F & Rapoport G. (1991) "The Bacillus subtilis sigL gene encodes an equivalent of sigma 54 from gram-negative bacteria". Proc Natl Acad Sci USA, 88, 9092-9096.

Dove S L, Huang F W & Hochschild A (2000) "Mechanism for a transcriptional activator that works at the isomerisation step". Proc Natl Acad Sci, vol. 97, pp. 13215-13220.

Dove S L, Jound J K & Hochschild A (1997) "Activation of prokaryotic transcription through arbitrary protein-protein contacts." Nature, vol. 386, pp. 627-630.

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

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

Drummond M, Whitty P & Wootton J (1986) "Sequence and domain relationships of ntrC and nifA from Klebsiella pneumoniae: homologies to other regulatory proteins". EMBO J, vol. 5, pp. 441-447.

Dworkin J, Jovanovic G, Model P (2000) "The PspA protein of Escherichia coli is a negative regulator of sigma(54)-dependent transcription". J Bacteriol, vol. 182, pp. 311-319.

Dworkin J, Jovanovic G & Model P (1997) "Role of Upstream Activation Sequences and Integration Host Factor in Transcriptional Activation by the Constitutively Active Prokaryotic Enhancer-binding Protein PspF." J Mol Biol, vol. 273, pp. 377-388.

Ebright R H (2000) "RNA Polymerase: Structural Similarities Between Bacterial RNA Polymerase and Eukaryotic RNA Polymerase II." J Mol Biol, vol. 304, pp. 687-698.

Ebright R H (1993) "Transcription activation at Class I CAP-dependent promoters". Mol Microbiol, vol 8, 797-802.

Estrem S T, Ross W, Gaal T, Chen Z W S, Niu W, Ebright R H et al. (1999) "Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxyl-terminal domain of the RNA polymerase α subunit". Genes Devel, vol. 13, pp. 2134-2147.

Fabret C, Feher V A & Hoch J A (1999) "Two-component signal transduction in Bacillus subtilis: how one organism sees its world". J Bacteriol vol. 181, pp. 1975-1983.

Flashner Y, Weiss D S, Keener J & Kustu S (1995) "Constitutive forms of the enhancer-binding protein NtrC: evidence that essential oligomerization determinants lie in the central activation domain." J Mol Biol, vol. 249, , pp. 700-713.

Flanagan J M, Wall J S, Capel M S, Schneider D K, Shanklin J (1992) "Scanning transmission electron microscopy and small-angle scattering provide evidence that native Escherichia coli ClpP is a tetradecamer with an axial pore." Biochemistry, vol. 34, pp. 10910-10917.

Fenton M S, Lee S J & Gralla J D (2000) "Escherichia coli promoter opening and -10 recognition: mutational analysis of sigma70". EMBO J, vol. 19, pp. 1130-1137.

Fiedler U & Weiss V (1995) "A common switch in activation of the response regulators NtrC and PhoB: phosphorylation induces dimerization of the receiver modules". EMBO J, vol. 14, pp., 3696-3705.

Finn R D, Orlova E V, Gowen B, Buck M & van Heel M (2000) "Escherichia coli RNA polymerase core and holoenzyme structures", EMBO J, vol. 19, 6833-6844.

Fujita N, Nomura T, & Ishihama A. (1987) "Promoter selectivity of Escherichia coli RNA polymerase. Purification and properties of holoenzyme containing the heat-shock sigma subunit". J Biol Chem, 262, 1855-1859.

Gallegos M T & Buck M (2000) "Sequences in σ^{54} Region I required for binding to early melted DNA and their involvement in sigma-DNA interactions". J Mol Biol, 297, 849-859.

Gallegos M T & Buck M, (1999) "Sequences in sigmaN determining holoenzyme formation and properties". J Mol Biol, 288, 539-553.

Gallegos M T, Cannon W & Buck M (1999) "Functions of the sigma(54) region I in trans and implications for transcription activation". J Biol Chem. 274, 25285-25290.

Garcia E, Bancroft S, Rhee S G & Kutsu S (1977) "The product of a newly identified gene, glnF, is required for the synthesis of glutamine synthetase in Salmonella" Proc Natl Acad Sci USA, vol. 74, pp. 1662-1666.

Gentry D R & Burgess R R (1989) "rpoZ, encoding the omega subunit of Escherichia coli RNA polymerase, in the same operon as spoT". J Bacteriol, vol. 171, pp. 1271-1277.

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

Gralla J (2000) "Signaling through Sigma", Nat Struct Biol, vol. 7, pp. 530-532.

Grande RA, Valderrama B & Morett E (1999) "Suppression analysis of positive control mutants of NifA reveals two overlapping promoters for Klebsiella pneumoniae rpoN." J Mol Biol, vol. 294, pp. 291-298.

Gray M W & Lang B F (1998) "Transcription in choloroplasts and mitochondria: a tale of two polymerases". Trends Microbiol, vol. 6, pp. 1-3.

Greiner D P, Hughes K A, Gunasekera A H & Meares C F (1996) "Binding of the sigma 70 protein to the core subunits of Escherichia coli RNA polymerase, studied by iron-EDTA protein footprinting" . Proc Natl Acad Sci USA, 93, 71-5.

Greiner D P, Miyake R, Moran J K, Jones A D, Negishi T, Ishihama & Meares C F (1997) "Synthesis of the Protein Cutting Reagent Iron (S)-1-(p-Bromoacetamidobenzyl) ethylenediaminetetraacetate and Conjugation to Cysteine Side Chains." Bioconjugate Chemistry, vol. 8, pp.44-48.

Gross C A, Chan C, Dombroski A, Gruber T, Sharp M, Tupy J & Young B (1998) The Functional and Regulatory Roles of Sigma Factors in Transcription. In Cold Spring Harbor Symposia on Quantitative Biology, Volume LXIII - Mechanisms of Transcription. Cold Spring Harbor Press, Cold Spring Harbor, New York.

Gross C A, Lonetto M & Losick R (1992) "Sigma Factors" In Transcriptional Regulation, Yamamoto, K R & McKnight, S L (eds.). Cold Spring Harbour, NY, Cold Spring Harbour Laboratory Press, pp. 129-176.

Gourse R L, Ross W & Gaal T (2000) "UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition". Mol Microbiol, vol. 37, pp. 687-695.

Guo Y, Lew C M & Gralla J D (2000) "Promoter opening by sigma(54) and sigma(70) RNA polymerases: sigma factor-directed alterations in the mechanism and tightness of control". Genes Dev, vol. 14, pp. 2242-2255.

Guo Y, Wang L & Gralla J D (1999) "A fork junction DNA-binding switch that controls promoter melting by the bacterial enhancer-dependent sigma factor". EMBO J, 18, 3736-3745.

Guo Y & Gralla J D (1997) "DNA binding determinants of sigma 54 as deduced from libraries of mutations." J Bacteriol, vol. 179, pp. 1239-1245.

Guo Y & Gralla J D (1998) "Promoter opening via a DNA fork junction binding activity." Proc Natl Acad Sci, vol. 95, pp. 11655-11660.

Hedtke B, Börner T & Weihe A (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis", Science, vol. 277, pp. 809-811.

Helmann J D & deHaseth P I (1999) "Protein-Nucleic Acid Interactions during Open Complex Formation Investigated by Systematic Alteration of the Protein and DNA Binding Partners", Biochemistry, vol. 28, pp. 5659-5967.

Helmann J D (1991) "Alternative sigma factors and the regulation of flagellar gene expression". Mol Microbiol, vol. 5, pp. 2875-2882.

Helmann J D & Chamberlin M J (1988) "Structure and function of bacterial sigma factors." Annual Rev Biochem, vol. 57, pp. 839-872.

Hernandez V J & Cashel M (1995) "Changes in conserved region 3 of Escherichia coli sigma 70 mediate ppGpp-dependent functions in vivo". J Mol Biol, vol. 252, pp. 536-549.

Heyduk T, Heyduk E, Severinov K, Tang H & 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, vol. 93, pp. 10162-10166.

Heyduk T, Lee J C, Ebright Y W, Blatter E E, Zhou Y, Ebright R H (1993) "CAP interacts with RNA polymerase in solution in the absence of promoter DNA". Nature, vol. 364, pp. 548-549.

Hochschild A & Dove S L (1998) "Protein-protein contacts that activate and repress prokaryotic transcription." Cell, vol. 92, pp. 597-600.

Hoover T R, Santero E, Porter S & Kustu S (1990) "The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons." Cell, vol. 63, pp. 11-22.

Hopper S, Vasquez B, Merz A, Clary S, Wilbur J S & So M (2000) "Effects of the immunoglobulin A1 protease on Neisseria gonorrhoeae trafficking across polarized T84 epithelial monolayers." Infect Immun, vol. 68, pp. 906-911.

Hsieh M & Gralla J D (1994) "Analysis of the N-terminal leucine heptad and hexad repeats of sigma 54". J Mol Biol, vol. 239, pp. 15-24.

Hsieh M, Tintut Y & Gralla J. D. (1994) "Functional roles for the glutamines within the glutamine-rich region of the transcription factor sigma 54". J Biol Chem, 269, 373-378.

Huang X, Lopez de Saro F J & Helmann J D (1997) "Sigma factor mutations affecting the sequence selective interaction of RNA polymerase with –10 region single-stranded DNA." Nucl Acids Res, vol. 25, pp. 2603-2609.

Inouye S, Nakazawa T & Nakazawa A (1987) "Expression of the regulatory gene xylS on the TOL plasmid is positively controlled by the xylR gene product". Proc Natl Acad Sci USA, vol. 84, pp. 5182-5186.

Igarashi K, Hanamura A, Makino K, Aiba H, Aiba H, Mizuno T, Nakata A & Ishihama A (1991) "Functional map of the alpha subunit of Escherichia coli RNA polymerase: two modes of transcription activation by positive factors." Proc Natl Acad Sci U S A, vol. 88, pp. 8958-8962.

Igarashi K, Fujita N & Ishihama A (1989) "Promoter selectivity of Escherichia coli RNA polymerase: omega factor is responsible for ppGpp sensitivity". Nucl Acids Res, vol. 17, pp. 8755-8765.

Igloi G L & Kössel H (1992) "Transcriptional apparatus of chloroplasts". Crit Rev Plant Sci, vol. 10, pp. 525-558.

Ishihama A (2000a) "Molecular anatomy of RNA polymerase using protein-conjugated metal probes with nuclease and protease activities". Chem Comm, vol. 13, pp. 1091-1094.

Ishihama A (2000b) "Functional modulation of Escherichia coli RNA polymerase". Annu Rev Microbiol, vol. 54, pp. 466-518.

Ishihama A (1999) "Modulation of the nucleod, the transcription apparatus, and the translation machinery in bacteria for stationary phase survival." Genes to Cells, 1999, vol. 4, pp. 135-143.

Ishihama A (1997) "Adaptation of gene expression in stationary phase bacteria". Curr Opin Genet, vol. 7, 582-588

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

Ishihama A (1981) "Subunit assembly of Escherichia coli RNA polymerase". Adv Biophys, vol. 14, pp. 1-35.

Ishihama A, Taketo M, Saitoh T & Fukuda R (1976) "Control of formation of RNA polymerase in Escherichia coli". In RNA polymerase. ed. M Chamberlin, R Losick, pp. 475-502. Cold Spring Harbor, NY: Cold Spring Harbour Lab. Press.

Ishimoto K & Lory S (1989) "Formation of pilin in Pseudomonas aeruginosa requires the RpoN subunit of RNA polymerase". Proc Natl Acad Sci USA, vol. 86, pp. 1954-1957.

Iuchi S & Lin E C (1987) "Molybdenum effector of fumarate reductase repression and nitrate reductase induction in Escherichia coli". J Bacteriol, vol. 169, pp. 3720-3725.

Iyoda S & Kutsukake K (1995) "Molecular dissection of the flagellum-specific anti-sigma factor, FlgM, of Salmonella typhimurium". Mol Gen Genet, vol. 249, pp. 417-424.

Jacobson G R, Schaffer M H, Stark G R & Vanaman T. C. (1973) Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues. J Biol Chem., 248, 6583-6591.

Jin D J & Gross C A (1989) "Characterization of the pleiotropic phenotypes of rifampin-resistant rpoB mutants of Escherichia coli." J Bacteriol, vol. 171, pp. 5229-5231.

Jishage M & Ishihama A (1999) "Transcriptional organisation and in vivo role of the Escherichia coli rsd gene, encoding the regulator of RNA polymerase sigma D". J Bacteriol, vol. 181, pp. 3768-3776.

Jishage M & Ishihama A (1998) "A stationary phase protein in Escherichia coli with binding activity to the major σ subunit of RNA polymerase." Proc Natl Acad Sci USA, vol. 95, pp. 4953-4958.

Jishage M & Ishihama A (1997) "Variation in RNA polymerase sigma subunit composition within different stocks of Escherichia coli W3110". J Bacteriol, 179, 959-963.

Jishage M, Iwata A, Ueda S & Ishihama A (1996) "Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of four species of sigma subunit under various growth conditions". J Bacteriol, vol. 178, pp. 5447-5451.

Jishage M & Ishihama A (1995) "Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of σ^{70} and σ^{28}". J Bacteriol, vol. 177, pp. 6832-6835.

Jovanovic G, Rakonjac J & Model P (1999) "In vivo and in vitro activities of the Escherichia coli sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFΔHTH". J Mol Biol 285, 469-83.

Jovanovic G, Weiner L & Model P (1996) "Identification, nucleotide sequence, and characterization of PspF, the transcriptional activator of the Escherichia coli stress-induced psp operon". J Bacteriol, 178, 1936-1945.

Juang Y L & Helmann J D (1995) "Pathway of promoter melting by Bacillus subtilis RNA polymerase at a stable RNA promoter: effects of temperature, delta protein and sigma factor mutations". Biochemistry, vol. 34, pp. 8465-8473.

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, vol. 235, pp. 1470-1488.

Kahn D & Ditta G (1991) "Modular structure of FixJ: homology of the transcriptional activator domain with the -35 binding domain of sigma factors". Mol Microbiol, vol. 5, pp. 987-997.

Kashlev M, Nudler E, Severinov K, Borukhov S, Komissarova N & Goldfarb A (1996) "Histidine-tagged RNA polymerase of Escherichia coli and transcription in solid phase". Methods Enzymol., vol. 274, pp. 326-334.

Katayama A, Fujita N & Ishihama A (2000) "Mapping of subunit-subunit contact surfaces on the β' subunit of Escherichia coli RNA polymerase". J. Biol. Chem., 275, 3583-3592.

Keener J & Kustu S (1988) "Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC". Proc. Natl. Acad. Sci. USA, vol. 85, pp. 4976-4980.

Kelly M T, Ferguson J A & Hoover T R (2000) "Transcription initiation-defective forms of sigma(54) that differ in ability to function with a heteroduplex DNA template". J. Bacteriol., vol. 182, pp. 6503-6508.

Kelly M T & Hoover T R (2000) "The amino terminus of Salmonella enterica serovar Typhimurium sigma(54) is required for interactions with an enhancer-binding protein and binding to fork junction DNA". J. Bacteriol., 182, 513-517.

Klose K E, Weiss D S & Kustu S (1993) "Glutamate at the site of phosphorylation of nitrogen-regulatory protein NTRC mimics aspartyl-phosphate and activates the protein". J. Mol. Biol., vol. 232, pp. 67-78.

Kohler T, Alvarez J F & Harayama S (1994) "Regulation of the rpoN, ORF102 and ORF154 genes in Pseudomonasputida." FEMS Microbiol Lett, vol. 115, pp. 177-184.

Kolb A, Igarashi K, Ishihama A, Lavigne M, Buckle M & Buc H (1993) "E. coli RNA polymerase, deleted in the C-terminal part of its alpha-subunit, interacts differently with the cAMP-CRP complex at the lacP1 and at the galP1 promoter." Nucleic Acids Res, vol. 21, pp. 319-326.

Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A & Darst S A (2000) "A structural model of transcription elongation." Science, vol. 289, pp. 619-625.

Kroos L, Zhang B, Ichikawa H & Yu Y T (1999) "Control of sigma activity during Bacillus subtilis sporulation". Mol. Microbiol., vol. 31, pp. 1285-1294.

Kuldell N & Hochschild A (1994) "Amino acid substitutions in the –35 recognition motif of sigma 70 that result in defects in phage lambda repressor-stimulated transcription." J. Bacteriol., , vol. 176, pp. 2991-2998.

Kulik E A, Kalinin I D & Sevastianov V I (1991) "The heterogeneity of protein/surface interactions and structural alterations of adsorbed albumin and immunoglobulin G." Artif Organs, vol. 15, pp. 386-391.

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., vol. 235, pp. 403-413.

Kusano S., Ding Q., Fujita N & Ishihama A (1996) "Promoter selectivity of Escherichia coli RNA polymerase E sigma 70 and E sigma 38 holoenzymes". J Biol Chem. 271, 1998-2004.

Kustu S, Santero E, Keener J, Popham D & Weiss D (1989) "Expression of σ^{54} (ntrA)-dependent genes is probably united by a common mechanism". Microbiol. Rev., vol. 53, pp. 367-376.

Lau P C, Wang Y, Patel A, Labbe D, Bergeron H, Brousseau R, Konishi Y & Rawlings M (1997) "A bacterial basic region leucine zipper histidine kinase regulating toluene degradation". Proc. Natl. Acad. Sci. USA, vol. 94, pp.1453-1458.

Lee J H & Hoover T R (1995) "Protein crosslinking studies suggest that Rhizobium meliloti C4-dicarboxylic acid transport protein D, a sigma 54-dependent transcriptional activator, interacts with sigma 54 and the beta subunit of RNA polymerase". Proc. Natl. Acad. Sci. U.S.A., vol. 92, pp. 9702-9706.

Lee J H, Scholl D, Nixon B T & Hoover T R (1993) "Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of Rhizobium meliloti DCTD, a sigma-54 dependent transcriptional activator". J. Biol. Chem, vol. 269, pp. 20401-20409.

Lemmon MA, Bu Z, Ladbury JE, Zhou M, Pinchasi D, Lax I, Engelman DM, Schlessinger J (1997) "Two EGF molecules contribute additively to stabilization of the EGFR dimer." EMBO J, vol. 16, pp. 281-294.

Lerbs-Mache S (1993) "The 110kDa polypeptide of spinach plastid DNA-dependent RNA polymerase: single subunit enzyme or catalytic core of multimeric enzyme complex ?". Proc. Natl. Acad. Sci. USA, vol. 90, pp. 5509-5513.

Li M, Moyle H & Susskind M M (1994) "Target of the transcriptional activation function of phage lambda cI protein." Science, vol. 263, pp. 75-77.

Lilley D M, Chen D & Bowater R P (1996) "DNA supercoiling and transcription: topological coupling of promoters." Q Rev Biophys. vol. 29, pp. 203-225.

Lois A F, Ditta G S & Helinski D R (1993)"The oxygen sensor FixL of Rhizobium meliloti is a membrane protein containing four possible transmembrane segments". J. Bacteriol.,vol. 175 pp.1103-1109.

Lonetto M, Rhodius V, Lamberg K, Kiley P, Busby S & Gross C (1998) "Identification of a contact site for different transcription activators in Region 4 of the Escherichia coli RNA polymerase σ^{70} subunit". J. Mol. Biol., vol. 284, pp. 1353-1365.

Lonetto M, Gribskov M & Gross C A (1992) "The σ^{70} family: sequence conservation and evolutionary relationships." J. Bacteriol., vol. 174, pp. 3834-3849.

Losick R P, Youngman P & Piggot P J (1986) "Genetics of endospore formation in Bacillis subtilis" Annual Rev. Gen., vol. 20, pp. 625-670.

Lukat G S, Stock A M & Stock J B (1990) "Divalent metal ion binding to the CheY protein and its significance to phosphotransfer in bacterial chemotaxis". Biochemistry, vol. 29, pp. 5436-5442.

Lukat G S, McCleary W R, Stock A M & Stock J B (1992) "Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors". Proc. Natl. Acad. Sci. USA., vol. 89, pp., 718-722.

Luo J, Sharif K. A, Jin R, Fujita N, Ishihama A & Krakow JS (1996). "Molecular anatomy of the beta' subunit of the E. coli RNA polymerase: identification of regions involved in polymerase assembly". Genes to Cells, 1, 819-27.

Mandel-Gutfreund Y, Schueler O & Margalit H (1995) "Comprehensive Analysis of Hydrogen Bonds in Regulatory Protein DNA-Complexes" J Mol Biol, vol. 253, pp. 370-382.

MacFarlane S A & Merrick M (1985) "The nucleotide sequence of the nitrogen regulation gene ntrB and the glnA-ntrBC intergenic region of Klebsiella pneumoniae". Nucl. Acids Res. vol. 13, pp. 7591-606.

Maeda H, Fujita N & Ishihama A (2000) "Competition among seven Escherichia coli σ subunits: relative binding affinities to the core RNA polymerase". Nucl. Acids. Res., vol. 28, pp. 3497-3503.

Maeda H, Jishage M, Nomura T, Fujita N & Ishihama A (2000) "Promoter selectivity of the RNA polymerase holoenzyme containing the extracytoplasmic function (ECF) sigma subunit E or FecI". J. Bacteriol., vol. 182, pp. 1181-1184.

Maeda T, Wurgler-Murphy S M & Saito H (1994) "A two-component system that regulates an osmosensing MAP kinase cascade in yeast". Nature, vol. 369, pp. 242-245.

Malhotra A, Severinova E & Darst S A (1996) "Crystal structure of a σ^{70} subunit fragment from E.coli RNA polymerase". Cell, vol. 87, pp. 127-136.

Marr M T & Roberts J W (1997) "Promoter recognition as measured by binding of polymerase to nontemplate strand." Science, vol. 276, pp. 1258-1260.

McClure W R (1985) "Mechanism and control of transcription initiation in prokaryotes." Annual Rev. Biochem., vol. 54, pp. 171-204.

McMahan S A & Burgess R R (1994) "Use of aryl azide cross-linkers to investigate protein-protein interactions: an optimization of important conditions as applied to Escherichia coli RNA polymerase and localization of a sigma 70-alpha cross-link to the C-terminal region of alpha." Biochemistry, vol. 33, pp. 12092-12099.

Merrick M J (1993) "In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)". Mol Microbiol. 10, 903-909.

Merrick M & Chambers S (1992) "The helix-turn-helix motif of sigma 54 is involved in recognition of the -13 promoter region". J Bacteriol., 174, 7221-7226.

Merrick M, Gibbins J & Toukdarian A (1987) "The nucleotide sequence of the sigma factor gene ntrA (rpoN) of Azotobacter vinelandii: analysis of conserved sequences in NtrA proteins." Mol Gen Genet, vol. 210, pp. 323-330.

Merrick M & Gibbins J R (1985) "The nucleotide sequences of the nitrogen-regulation gene ntrA of Klebsiella pneumoniae and comparison with conserved features in bacterial RNA polymerase sigma factors" Nucl Acids Res, vol. 13, pp. 7607-7620.

Meng W, Savery N J, Busby S J W & Thomas M S (2000) "The Escherichia coli RNA polymerase α-subunit linker: length requirements for transcription activation at CRP-dependent promoters". EMBO J., vol. 19, pp. 1555-1566.

Meyer T F, Billyard E, Haas R, Storzbach S & So M (1984) "Pilus genes of Neisseria gonorrhoea: chromosomal organisation and DNA sequence". Proc. Natl. Acad. Sci. USA, vol. 81, pp. 6110-6114.

Michiels J, Verreth C & Vanderleyden J (1998) "Sequence analysis of the Rhizobium etli ribose kinase gene rbsK and its phylogenetic position." DNA Seq, vol. 9, pp. 317-321.

Missailidis S, Cannon W V, Drake A, Wang X Y & Buck M (1997) "Analysis of the architecture of the transcription factor sigma N (sigma 54) and its domains by circular dichroism." Mol Microbiol, vol. 24, pp. 653-664.

Missiakas D, Mayer M P, Lemaire M, Georgopoulos C & Raina S (1997) "Modulation of the Escherichia coli σ^E (RpoE) heat-shock transcription factor activity by the RseA, RseB and RseC proteins." Mol. Microbiol., vol. 24, pp. 355-371.

Miyake R, Murakami K, Owens J T, Greiner D P, Ozoline O N, Ishihama A & Meares C F (1998) "Dimeric association of E.coli RNA polymerase α subunits, studied by cleavage of single-cysteine α subunits conjugated to Iron-(S)-1-[p-(Bromoacetadimo)benzyl]ethylenediaminetetraacetate". Biochemistry, vol. 37, pp. 1344-1349.

Mizuno T (1998) "His-Asp phosphotransfer signal transduction". J. Biochem., vol. 123, pp. 555-563.

Mizuno T (1997) "Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of Escherichia coli". DNA Res., vol. 4, pp. 161-168.

Mooney R A & Landick R (1999) "RNA polymerase unveiled". Cell, vol. 98, pp. 687-690.

Morett E & Segovia L (1993) "The σ^{54} bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains". J. Bacteriol., vol. 175, pp. 6067-6074.

Morett E & Buck M (1989) "In vivo studies on the interaction of RNA polymerase-sigma 54 with the Klebsiella pneumoniae and Rhizobium meliloti nifH promoters. The role of NifA in the formation of an open promoter complex." J Mol Biol, vol. 210, pp. 65-77.

Morris L, Cannon W, Claverie-Martin F, Austin S & Buck M (1994) "DNA Distortion and nucleation of local DNA unwinding within sigma-54 (sigma N) holoenzyme closed promoter complexes." J. Biol. Chem., vol. 269, pp. 11563-11571.

Muller-Hill B (1996) "The lactose operon". Walter de Gruyter, Berlin.

Mullin A, Minnich S, Chen L-S & Newton A (1987) "A set of positively regulated flagellar gene promoters in Caulobacter crescentus with sequence homology to the nif gene promoters of Klebsiella pneumoniae". J. Mol. Biol., vol. 195, pp. 939-943.

Mustaev A, Kozlov M, Markovtsov V, Zaychikov E, Denissova L & Goldfarb A (1997) "Modular organization of the catalytic center of RNA polymerase". Proc. Natl. Acad. Sci. USA, 94, 6641-5.

Mukherjee K & Chatterji D (1997) "Studies on the omega subunit of Escherichia coli RNA polymerase – its role in the recovery of denatured enzyme activity". Eur J. Biochem., vol. 247, pp. 884-889.

Murakami K, Owens J T, Belyaeva T A, Meares C F, Busby S J & Ishihama A. (1997a) "Positioning of two alpha subunit carboxy-terminal domains of RNA polymerase at promoters by two transcription factors". Proc. Natl. Acad. Sci. U S A., 94, 11274-11278.

Murakami K, Kimura M, Owens J T, Meares C F & Ishihama A (1997b) "The two alpha subunits of Escherichia coli RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element". Proc. Natl. Acad. Sci. U S A., 94, 1709-1714.

Nedea E C, Markov D, Naryshkina T & Severinov K (1999) "Localisation of E. coli rpoC mutations that affect RNA polymerase assembly and activity at high temperature". J. Bacteriol., vol. 181, pp. 2663-2665.

Nelson K E, Clayton R A, Gill S R, Gwimm M L, Dodson R J et al. (1999) "Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of Thermotoga maritima". Nature, vol. 399 pp. 323-329.

Niu W, Zhou Y, Dong Q, Ebright Y W & Ebright R H (1994) "Characterization of the activating region of Escherichia coli catabolite gene activator protein (CAP). I. Saturation and alanine-scanning mutagenesis." J Mol Biol, vol. 243, pp. 595-602.

Oguiza J A, Gallegos M T, Chaney M K, Cannon W V & Buck M (1999) "Involvement of the sigmaN DNA-binding domain in open complex formation." Mol Microbiol, vol. 33, pp. 873-885.

Oguiza J A & Buck M (1997) "DNA-binding domain mutants of sigma-N (σ^N, σ^{54}) defective between closed and stable open promoter complex formation". Mol Microbiol, vol. 26, pp. 655-664.

Okamoto K, Hara S, Bhasin R & Freundlich M (1988) "Evidence in vivo for autogenous control of the cyclic AMP receptor protein gene (crp) in Escherichia coli by divergent RNA." J Bacteriol, vol. 170, pp. 5076-5099.

O'Neill E, Ng L C, Sze C C & Shingler V (1998) "Aromatic ligand binding and intramolecular signalling of the phenol-responsive sigma54-dependent regulator DmpR. Mol Microbiol vol. 28, pp. 131-141.

Opalka N, Mooney R A, Richter C, Severinov K, Landick R & Darst S A (2000) "Direct localisation of a β-subunit domain on the three-dimensional structure of Escherichia coli RNA polymerase". Proc. Natl. Acad. Sci., vol. 97, pp. 617-622.

Orsini G, Ouhmmouch M, Le Caer J P & Brody E N (1993) "The asiA gene of bacteriophage T4 codes for the anti-σ^{70} protein." J. Bacteriol., vol. 175, pp. 85-93.

Owens J T, Miyake R, Murakami K, Chmura A J, Fujita N, Ishihama A & Meares C F (1998) "Mapping the σ^{70} subunit contact sites on Escherichia coli RNA polymerase with a σ^{70}-conjugated chemical protease." Proc. Natl. Acad. Sci. USA, vol. 95, pp. 6021-6026.

Ozoline O N, Uteshev T A & Kamzolova S G (1991) "Specific modification of Escherichia coli RNA polymerase with monomercury derivative of fluorescein acetate." Biochim Biophys Acta, vol. 1076, pp. 387-394.

Pallen M (1999) "RpoN-dependent transcription of rpoH?" Mol Microbiol, vol. 31pp. 393.

Panaghie G, Aiyar S E, Bobb K L, Hayward R S & de Haseth P (2000) "Aromatic amino acids in region 2.3 of Escherichia coli sigma 70 participate collectively in the formation of an RNA polymerase-promoter open complex". J. Mol. Biol., vol. 299, pp. 1217-1230.

Parvari R, Pecht I & Soreq H (1983) A Microfluorometric Assay for Cholinesterases, Suitable for Multiple Kinetic Determinations of Picomoles of Released Thiocholine. Anal. Biochem., 133, 450-456.

Penas D, Connolly L & Gross C A (1997) "The σ^E-mediated response to extracytoplasmic stress in Escherichiae coli is transduced by RseA and RseB, two negative regulators of σ^E." Mol. Microbiol., vol. 24, pp. 373-385.

Perez-Martin J & de Lorenzo V (1997) "Clues and consequences of DNA bending in transcription." Annu Rev Microbiol, vol. 51, pp. 593-628.

Perez-Martin J & De Lorenzo V (1995) "The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor." Proc Natl Acad Sci U S A, vol. 92, pp. 9392-9396.

Pitt M, Gallegos M T & Buck M (2000) "Single amino acid substitution mutations of Klebsiella pneumoniae σ^{54}". Nucl. Acids. Res., vol. 28, pp. 4419-4427.

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, vol. 83, pp. 365-373.

Popham D, Keener J & Kustu S (1991) "Purification of the alternative σ factor, σ^{54}, from S.typhimurium and characterisation of σ^{54} holoenzyme". J. Biol. Chem., vol. 266, pp. 19518-19519.

Popham D L, Szeto D, Keener J & Kustu S (1989) "Function of a bacterial activator protein that binds to transcriptional enhancers". Science. 243, 629-635.

Phillips S & Stockley P (1996) "Structure and Function of the Escherichia coli met repressor: similarities and contrasts with the trp repressor". Phil Trans Roy Soc Lond B vol. 351, pp.527-535.

Powell B S, Court D L, Inada T, Nakamura Y, Michotey V, Cui X, Reizer A, Saier M H Jr & Reizer J (1995) "Novel proteins of the phosphotransferase system encoded within the rpoN operon of Escherichia coli. Enzyme IIANtr affects growth on organic nitrogen and the conditional lethality of an erats mutant." J Biol Chem, vol. 270, pp. 4822-4839.

Ptashne M & Gann A (1997) "Transcriptional activation by recruitment". Nature, vol. 386, pp. 569-577.

Qiagen (1997) "The QIAexpressionist - A handbook for high-level expression and purification of 6xHis-tagged proteins". July 1997.

Rana T M & Meares C F (1991) "Transfer of oxygen from an artificial protease to peptide carbon during proteolysis". Proc. Acad. Natl. Sci. USA, 88, 10578-10582.

Raina S, Missiakis D & Georgopoulos C (1995) "The rpoE gene encoding the σ^E (σ^{24}) heat shock sigma factor of Escherichia coli". EMBO J, vol. 14, pp. 1043-1055.

Reddy P, Hoskins J & McKenney K (1995) "Mapping domains in proteins: dissection and expression of Escherichia coli adenylyl cyclase." Anal Biochem, vol. 231, pp. 282-286.

Rippe K, Guthold M, von-Hippel P H & Bustamente C (1997) "Transcriptional activation via DNA-looping: visualisation of intermediates in the activation pathway of E.coli RNA polymerase sigma 54 holoenzyme by scanning force microscopy". J. Mol. Biol., vol. 270, pp. 125-138.

Rojo F (1999) "Repression of Transcription Initiation in Bacteria". J. Bacteriol., vol. 181, pp. 2987-2991.

Ronson C W, Nixon B T, Albright L M & Ausubel F M (1987) "Rhizobium meliloti ntrA (rpoN) gene is required for diverse metabolic functions". J. Bacteriol., vol. 169, pp. 2424-2430.

Ross W, Aiyar S, Salmon J & Gourse R L (1998) "Escherichia coli Promoters with UP Elements of Different Strengths: Modular Switch of Bacterial Promoters.". J. Bacteriol., vol. 180, pp. 5375-5383.

Roy S, Garges S, Adhya S (1998) "Activation and Repression of Transcription by Differential Contact: Two Sides of a Coin". J. Biol. Chem., vol. 273, pp. 14059-14062.

Sasse-Dwight S & Gralla J D (1990) "Role of eukaryotic-type domains found in the prokaryotic enhancer receptor factor σ^{54}". Cell, vol. 62, pp. 945-950.

Sasse-Dwight S & Gralla J D (1988) "Probing the Escherichia coli glnALG upstream activation mechanism in vivo". Proc. Natl. Acad. Sci. U.S.A., vol. 85, pp. 8934-8938.

Scott D J, Ferguson A L, Gallegos M T, Pitt M, Buck M & Hoggett J G (2000) "Interaction of sigma factor sigmaN with Escherichia coli RNA polymerase core enzyme." Biochem J, vol. 352, pp. 539-547.

Sauer R T, Jordan S R & Pabo C O (1990) "λ repressor: a model system for understanding protein DNA interaction and protein stability." Advanced Protein Chemistry, vol. 40, pp. 1-61 (1990).

Schutz H & Reinert K E (1991) "DNA-bending on ligand binding; change of DNA persistence length." J Biomol Struct Dyn, vol. 9, pp. 315-329.

Severinova E, Severinov K, Fenyo D, Marr M, Brody E N, Roberts J W, Chait T C & Darst S A (1996) "Domain organisation of the Escherichia coli RNA polymerase σ^{70} subunit". J. Mol. Biol., vol. 263, pp. 637-647.

Severinov K (2000) "RNA polymerase structure-function: insights into points of transcriptional regulation". Curr. Op. Microbiol., vol. 3, pp. 118-125.

Severinov K & Darst S A (1997) "A mutant RNA polymerase that forms unusual open promoter complexes." Proc Natl Acad Sci U S A, vol. 94, pp. 13481-13486.

Severinov K, Fenyo D, Severinova E, Mustaev A, Chait B T, Goldfarb A & Darst S A (1994) "The σ subunit conserved region 3 is part of "5'-face" of active center of Escherichia coli RNA polymerase". J. Biol. Chem., vol. 268, pp. 14820-14830.

Severinov K, Kashlev M, Severinova E, Bass I, McWilliams K, Klutter E, Nikiforov V, Snyder L & Goldfarb A (1994) "A non-essential domain of E. coli RNA polymerase required for the action of the termination factor Alc". J. Biol. Chem., vol. 269, pp. 14254-14259.

Severinov K, Mustaev A, Kashlev M, Borukhov S, Nikiforov V & Goldfarb A (1992) "Dissection of the β subunit in the E. coli RNA polymerase into domains by proteolytic cleavage". J. Biol. Chem., vol. 267, pp. 12813-12819.

Sharp M M, Chan C L, Lu C Z, Marr M T, Nechaev S, Merritt E W, Severinov K, Roberts J W & Gross C A (1999) "The interface of σ with core RNA polymerase is extensive, conserved and functionally specialised". Genes & Devel., vol. 13, pp. 3015-3026.

Shieh J C, Wilkinson M G, Buck V, Morgan B A, Makino K & Millar J B (1997) "The Mcs4 response regulator coordinately controls the stress-activated Wak1-Wis1-Sty1 MAP kinase pathway and fission yeast cell cycle". Genes Dev., vol. 11, pp. 1008-1022.

Shingler V (1996) "Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism." Mol Microbiol, vol. 19, pp. 409-416.

Shingler V & Pavel H R (1995) "Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds." Mol Microbiol, vol. 17, pp. 505-513.

Shilton B, Svergun D I, Volkov V V, Koch M H, Cusack S & Economou A (1998) "Escherichia coli SecA shape and dimensions." FEBS Lett, vol. 436, pp. 277-282.

Singh M, Berger B, Kim P S, Berger J M & Cochran A G (1998) "Computational learning reveals coiled coil-like motifs in histidine kinase linker domains". Proc. Natl. Acad. Sci. USA., vol. 95 pp. 2738-2743.

Sogaard-Andersen L, Pedersen H, Holst B & Valentin-Hansen P (1991) "A novel function of the cAMP-CRP complex in Escherichia coli: cAMP-CRP functions as an adaptor for the CytR repressor in the deo operon". Mol. Microbiol., vol. 5, pp.969-975.

Southern E & Merrick M (2000) "The role of Region III in the RNA polymerase σ factor σ^N (σ^{54})". Nucl. Acids. Res., vol. 28, pp. 2563-2570.

Stock J (1990) "A membrane receptor kinase that regulates development in Bacillus subtilis." Bioessays, vol. 12, pp. 387-388.

Stock J B, Ninfa A J & Stock A M (1989) "Protein phosphorylation and regulation of adaptive responses in bacteria". Microbiol. Rev. vol. 53, pp. 450-490.

Studholme D J, Buck M & Nixon T (2000) "Identification of potential σ^N-dependent promoters in bacterial genomes". Microbiology, vol. 146, pp. 3021-3023.

Studholme D J & Buck M(2000a) "Novel roles of σ^N in small genomes". Microbiology, vol. 146, pp. 3-5.

Studholme D J & Buck M (2000b) "The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences" FEMS Microbiol. Let., vol. 186, pp. 1-9.

Studholme D J, Finn R D, Chaney M K & Buck M (1999) "The C-terminal 12 amino acids of sigma(N) are required for structure and function." Arch Biochem Biophys, vol. 371, pp. 234-240.

Su W, Porter S, Kustu S & Echols H (1990) "DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter". Proc Natl Acad Sci U S A., 87, 5504-5508.

Sullivan S M & Maddock J R (2000) "Bacterial sporulation: pole-to-pole oscillation". Curr. Biol., vol. 10, pp. 156-161.

Summers A (1992) "Untwist and shout- a heavy metal – responsive transcription regulator". J. Bacteriol., vol.174, pp. 3097-3101.

Svergun D I, Malfois M, Koch, M H J, Wigneshweraraj S R & Buck M. (2000) "Low resolution structure of the σ^{54} transcription factor revealed by X-ray solution scattering". J. Biol. Chem., 275, 4210-4214.

Syed A & Gralla J D (1998) "Identification of an N-terminal region of sigma 54 required for enhancer responsiveness." J Bacteriol, vol. 180, pp. 5619-5625.

Syed A & Gralla J D (1997) "Isolation and Properties of enhancer-bypass mutants of sigma 54". Mol. Microbiol., vol. 23, pp. 987-995.

Syed A & Gralla (1998) "Identification of an N-Terminal Region of Sigma 54 Required for Enhancer Responsiveness." J. Bacteriol., vol. 180, no. 21, pp. 5619-5625.

Taylor M, Butler R, Chambers S, Casimiro M, Badii F & Merrick M (1996) The RpoN-box motif of the RNA polymerase sigma factor sigma N plays a role in promoter recognition. Mol Microbiol. 22, 1045-1054.

Tintut Y, Wang J T & Gralla J D (1995) "Abortive cycling and the release of polymerase for elongation at the sigma 54-dependent glnAp2 promoter". J Biol Chem. 270, 24392-24398.

Tintut Y, Wong C, Jiang Y, Hsieh M & Gralla J D (1994) "RNA polymerase binding using a strongly acidic hydrophobic-repeat region of σ^{54}." Proc. Natl. Acad. Sci. USA, vol. 91, pp. 2120-2124.

Tracy R L & Stern D B (1995) "Mitochondrial transcription initiation: promoter structures and RNA polymerases." Curr Genet, vol. 28, pp. 205-216.

Traviglia S L, Datwyler S A, Meares C F (1999) Mapping protein-protein interactions with a library of tethered cutting reagents: the binding site of sigma 70 on Escherichia coli RNA polymerase". Biochemistry, vol. 38, pp. 4259-4265.

Tuzikov F V, Zinoviev V V, Vavilin V I & Malygin E G (1996) "Application of the small-angle X-ray scattering technique for the study of two-step equilibrium enzyme-substrate interactions." Biopolymers, vol. 38, pp. 131-139.

van Aalten D M, Erlanson D A, Verdine G L & Joshua-Tor L (1999) "A structural snapshot of base-pair opening in DNA". Proc Natl Acad Sci U S A., 96, 11809-11814.

Wang L & Gralla J (2000) "Roles for the C-terminal region of Sigma 54 in transcriptional silencing and DNA-binding". J. Biol. Chem. [epub ahead of print]

Wang L & Gralla J D (1998) "Multiple In Vivo Roles for the –12 Region Elements of Sigma 54 Promoters." J. Bacteriol., vol. 180, pp. 5626-5631.

Wang J T, Syed A, Gralla J D (1997) "Multiple pathways to bypass the enhancer requirement of sigma 54 RNA polymerase: roles for DNA and protein determinants". Proc. Natl. Acad. Sci. USA., vol. 94, pp. 9538-9543.

Wang J T & Gralla J D (1996) "The transcription initiation pathway of sigma 54 mutants that bypass the enhancer protein requirement: implications for the mechanism of activation." J. Biol. Chem., vol. 271, pp. 32707-32713.

Wang J T, Syed A, Hsieh M & Gralla J D (1995) "Converting Escherichia coli RNA polymerase into an enhancer-responsive enzyme: role of an NH2-terminal leucine patch in sigma 54". Science, vol. 270, pp. 992-994.

Wang C, Tintut Y & Gralla J D (1994) "The domain structure of sigma 54 as determined by analysis of a set of deletion mutants." J. Mol. Biol., vol. 236, pp. 81-90.

Wang J C & Lynch A S (1993) "Transcription and DNA supercoiling." Curr Opin Genet Dev, vol. 3, pp. 764-768.

Webster C, Gaston K & Busby S (1988) 2Transcription from the Escherichia coli melR promoter is dependent on the cyclic AMP receptor protein." Gene, vol. 68, pp. 297-305.

Wedel A & Kustu S (1995) "The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation". Genes Dev. 9., 2042-2052.

Weihe A & Börner T (1999) "Transcription and the architecture of promoters in chloroplasts". Trends Plant Sci., vol. 4, pp. 169-170.

Weiner L, Brissette J L & Model P (1991) "Stress-induced expression of the Escherichia coli phage shock protein operon is dependent on σ^{54} and modulated by positive and negative feedback mechanisms" Genes Dev., vol. 5, pp. 1912-1923.

Weiss D S, Batut J, Klose K E, Keener J & Kustu S (1991) "The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription". Cell, 67, 155-167.

Weiss V, Claverie-Martin F & Magasanik B (1992) "Phosphorylation of nitrogen regulator I of Escherichia coli induces strong cooperative binding to DNA essential for activation of transcription". Proc. Natl. Acad. Sci. USA., vol. 89, pp. 5088-5092.

Wigneshweraraj S R, Fujita N, Ishihama A & Buck M (2000). "Conservation of sigma-core RNA polymerase proximity relationships between the enhancer independent and enhancer dependent sigma classes". EMBO J., vol. 19, pp. 3038-3048.

Wigneshweraraj S R, Ishihama A & Buck M (2001) "In vitro roles of invariant helix-turn-helix motif residue R383 in $\sigma^{54}(\sigma^{N})$". Nucl. Acids. Res., in press.

Wigneshweraraj S R, Ishihama A & Buck M (2001) "Regulatory sequences in sigma 54 localise near the start of DNA melting". J. Mol. Biol., in press.

Wösten M M S M (1998) "Eubacterial sigma-factors". FEMS Microbiol. Rev., vol. 22, pp. 127-150.

Wong C, Tintut Y & Gralla J D (1994) "The domain structure of sigma 54 as determined by analysis of a set of deletion mutants". J. Mol. Biol., vol. 236, pp. 81-90.

Wong C & Gralla J D (1992) "A role for the acid trimer repeat region of transcription factor σ^{54} in setting the rate and temperature dependence of promoter melting in vivo". J. Biol. Chem., vol. 267, pp. 24762-24765.

Wyman C, Rombel I, North A K, Bustamente C & Kustu S (1997) "Unusual Oligomerization required for activity of NtrC, a bacterial enhancer-binding protein". Science, vol. 275, pp. 1658-1661.

Yanofsky C, Horn V & Nakamura Y (1996) "Loss of overproduction of polypeptide release factor 3 influences expression of the tryptophanase operon of Escherichia coli." J Bacteriol, vol. 178, pp. 3755-3762.

Yura T, Nagai H & Mori H (1993) "Regulation of the heat-shock response in bacteria." Annual Rev. Microbiol., vol. 47, pp. 321-350.

Zhou Y N & Gross C A (1992) "How a mutation in a gene encoding sigma-70 suppresses the defective heat shock response caused by the gene encoding sigma-32". J. Bacteriol, vol. 174, pp. 7128-7137.

Zakharova N, Bass A, Arsenieva E, Nikiforov V & Severinov K (1998) "Mutations in and monoclonal antibody binding to evolutionarily hypervariable region of E. coli RNA polymerase β' subunit inhibit transcript cleavage and transcript elongation". J. Biol. Chem., vol. 273, pp. 24912-24920.

Zaychikov E, Martin E, Denissova L, Kozlov M, Markovtsov V, Kashlev M, Heumann H, Nikiforov V., Goldfarb A & Mustaev A (1996) "Mapping of catalytic residues in the RNA polymerase active center". Science, 273, 107-109.

Zhang G, Campbell E A, Minakhin L, Richter C, Severinov K & Darst S A (1999) "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution". Cell, 98, 811-24.

Zhang G & Darst S A (1998) "Structure of the Escherichia coli RNA polymerase alpha subunit amino-terminal domain." Science, vol. 281, pp. 262-266.

Zhou Y N & Gottesman S (1998) "Regulation of proteolysis of stationary phase sigma factor RpoS." J. Bacteriol., vol. 180, pp. 1154-1158.

Zhou W & Doetsch P W (1993a) "Effects of abasic sites and DNA single-strand breaks on prokaryotic RNA polymerases." Proc Natl Acad Sci U S A, vol. 90, pp. 6601-6605.

Zhou Y, Busby S & Ebright R H (1993) "Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers." Cell, vol. 73, pp. 375-379.