

The role of Region II in the RNA polymerase σ factor σ^N (σ^{54})

Emma Southern and Mike Merrick*

Department of Molecular Microbiology, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK

Received February 23, 2000; Revised April 10, 2000; Accepted May 12, 2000

ABSTRACT

Bacterial RNA polymerase holoenzymes containing the σ subunit σ^N (σ^{54}) can form a stable closed complex with promoter DNA but only undergo transition to an open complex and transcription initiation when acted on by an activator protein. Proteins of the σ^N family have a conserved N-terminal region of 50 amino acids (Region I) that is separated from a conserved C-terminal region of around 360 amino acids (Region III) by a much more variable sequence of between 30 and 110 residues (Region II). We have investigated the role of Region II in *Klebsiella pneumoniae* σ^N by studying the properties of deletions of all or part of the region both *in vivo* and *in vitro*. We found that whilst Region II is not essential, deletion of all or part of it can significantly impair σ^N activity. Deletions have effects on DNA binding by the isolated σ factor and on holoenzyme formation, but the most marked effects are on transition of the holoenzyme from the closed to the open complex in the presence of the activator protein.

INTRODUCTION

Bacterial RNA polymerases (RNAPs) are multi-subunit enzymes comprising a core enzyme ($\alpha_2\beta\beta'$) and an exchangeable σ subunit. RNAPs initiate transcription specifically when they are in the holoenzyme form ($\alpha_2\beta\beta'\sigma$) and the σ factor plays roles in both promoter recognition and the early stages of transcription. Most σ factors belong to the σ^{70} family of proteins (1), the exception to this being σ^N (or σ^{54}), which has a distinct primary amino acid sequence and constitutes a second minor σ family (2). Although both classes of σ factor bind the same core polymerase and perform similar functions, their mechanisms of action are apparently quite distinct. The σ^N holoenzyme binds promoters in a transcriptionally inactive complex (3) and isomerisation to the open complex requires an activator protein bound at an upstream activator site (UAS) 100 bp or more upstream of the transcription initiation site (4). Activation appears to require direct contact between the activator and the holoenzyme, which is achieved through DNA looping.

The primary amino acid sequence of σ^N proteins is highly conserved, suggesting that all members of this family probably have a similar tertiary structure. Alignment of the primary

sequences (of which more than 30 have now been determined) allows the protein to be divided into three distinct regions (2). Region I, which comprises the first 50 N-terminal residues, contains an excess of glutamine residues and characteristic regularly repeated leucines (5–7). A number of studies have shown that Region I is required for the response to activator proteins (8–13). Region III comprises the C-terminal 350 residues, which include the RNAP core binding determinants (14–16) and sequences that are important for promoter DNA binding (5,17–20).

Region II, which lies between the two conserved regions, is very variable both in length and in sequence, although it is characterised by a predominance of acidic residues. In some organisms these acidic residues have a marked periodicity, occurring every third residue so that they have been termed acidic trimer repeats (ATRs) (21). The length of Region II varies between 26 residues in *Rhodobacter capsulatus* and 110 residues in *Bradyrhizobium japonicum* RpoN₂, with a mean length of around 85 residues amongst those sequences presently determined. Much less is known about the role of Region II than about the conserved Regions I and III. *In vivo* studies of a deletion in the ATR region implicated Region II in promoter melting (21). Recent studies of an RNAP holoenzyme containing a σ^N in which both Regions I and II had been deleted concluded that Region II was not required for polymerase isomerisation and transcription initiation on pre-melted DNA templates. However, they suggested that Region II sequences may play a role in stabilising the holoenzyme and in suppressing non-specific DNA binding by the holoenzyme (22).

To investigate further the role of Region II sequences we have constructed a number of deletions within this part of *Klebsiella pneumoniae* σ^N and analysed the properties of each of these variants both *in vivo* and *in vitro*. Our results indicate that deletions within Region II can influence the binding of σ^N to core RNAP, have relatively little effect on DNA binding by the assembled holoenzyme and are most marked in their effects on open complex formation in the presence of the activator protein.

MATERIALS AND METHODS

Strains and plasmids

Escherichia coli strain JM109 was used as host in cloning experiments. *Escherichia coli* strain BL21-GOLD (DE3) (Stratagene, La Jolla, CA) was used as the host strain for protein expression and purification. *Klebsiella pneumoniae*

*To whom correspondence should be addressed. Tel: +44 1603 450749; Fax: +44 1603 450018; Email: mike.merrick@bbsrc.ac.uk

Table 1. Oligonucleotides used in this work

Oligonucleotide name/number	Sequence (5'→3')	Description
UNIV	GTA AAA CGA CGG CCA GT	PCR upstream primer to create probe <i>pnifH</i> – UAS
RSP	CAG GAA ACA GCT ATG AC	PCR downstream primer to create probe <i>pnifH</i> – UAS
36	TGG TTT CTA GAT CAT CGT GA	PCR downstream primer; adds <i>XbaI</i> site to 5'-end of region II coding sequence
55	CGC GCG CTG GTG AAG AAA	PCR upstream primer
146	CCC CAA GCT TGG GTC AAA CTA GTT GCT TGC	PCR downstream primer; adds <i>SpeI</i> site with <i>HindIII</i> site immediately downstream
152	CTA GCC CAT CAT CAT CAT CAT TGA	Histidine cassette, coding strand
153	A GCT TCA ATG ATG ATG ATG ATG ATG GG	Histidine cassette, non-coding strand
179	ATG ATT ACG AAT TCA GGA GG	PCR upstream primer to create pES11
180	CTG GCA TCT AGA GGC AG	PCR downstream primer to create pES11
181	CTG CCT CTA GAT GCC AG	PCR upstream primer to create pES16
182	GCG CGG ATC CAG CGA TTG	PCR downstream primer to create pES16
185	GCG AAT TCG CCC ATA CGA CAC TG	PCR upstream primer to create probe <i>pnifH</i> + UAS
<i>lacZ</i>	CAT TCG CCA TTC AGG CTG	PCR downstream primer to create probe <i>pnifH</i> + UAS

rpoN deletion strain UNF2927 (17) was used as the host strain for β -galactosidase assays.

The high copy number plasmid pES6 was created to facilitate overexpression and purification of wild-type *K.pneumoniae* σ^N . PCR using as template pWVC93025, in which the ribosome-binding site of *rpoN* was modified to give a consensus sequence (23), and primers 55 and 146 generated a product which was cloned into pRJ21 (19) at the *SacII* and *HindIII* sites (all oligonucleotides used in this study are listed in Table 1). This introduced a novel *SpeI* site at the very 3'-end of *rpoN*, and a hexahistidine cassette, created by annealing oligonucleotides 152 and 153, was ligated into this *SpeI* site. The modified 3'-end of *rpoN* was then substituted in pWVC93025, using the *SacII* and *HindIII* sites, to give pES6.

Three derivatives of pES6 were created which encode different deletions of Region II (Fig. 1). pES8, which encodes a deletion of residues 56–106, was derived by introduction of novel *XbaI* and *NheI* restriction sites at 165 and 312 bp, respectively, from the *rpoN* ATG (forming pES10) and subsequent deletion of the intervening DNA by restriction and ligation. pES11 encodes a deletion from residues 86–106 and was generated by cloning a 280 bp fragment with *EcoRI* and *XbaI* ends (generated by PCR using pES6 and primers 179 and 180) between the *EcoRI* and *NheI* sites in pES10. pES16 encodes a deletion from residues 56–83 and was created in two stages. First, a 530 bp fragment (generated by PCR using pES6 and primers 181 and 182) was ligated into the *XbaI* and *BamHI* sites of pES10, yielding pES13. Secondly, to add the modified RBS sequence, a 440 bp fragment (generated by PCR using pES6 and primers 179 and 36) was cloned into pES13 at the *EcoRI* and *XbaI* sites. Low copy number derivatives of each of the overexpression plasmids were created by cloning the respective *EcoRI*–*HindIII* fragments into pHSG575 (24), thereby creating pES15 (wild-type), pES17 (Δ 56–83), pES12 (Δ 86–106) and pES9 (Δ 56–106). Sequences of all constructs containing PCR-generated DNA were verified.

Purification of σ^N proteins

The σ^N variants were overexpressed from the appropriate plasmid in 1–4 l of JM109 cell culture by addition of 1 mM IPTG at a cell density ($A_{600\text{nm}}$) of ~0.6. Induced growth was continued at 37°C for 3 h, then cells were harvested by centrifugation and stored at –20°C overnight. The pellet was resuspended in 15 ml of ice-cold lysis buffer [300 mM NaCl, 50 mM Tris–HCl pH 8.0, 1 mM PEFABLOC (Boehringer-Mannheim, Lewes, UK), 5% glycerol, 130 $\mu\text{g ml}^{-1}$ lysozyme] and passed twice through a French press at 690 bar. The lysate was centrifuged at 12 000 r.p.m. for 45 min at 4°C and the supernatant was then decanted and stored at –20°C. DNase I was added to the lysate when required, to reduce viscosity.

The proteins were purified using a Pharmacia Biotech FPLC apparatus. Cell lysates were run on a 1 ml Hi-Trap chelating column (Pharmacia) primed with NiCl₂ and eluted with a 0–500 mM gradient of imidazole. Pooled fractions were dialysed in TGED (10 mM Tris–HCl pH 8.0, 0.1 mM EDTA, 0.1 mM DTT, 50 mM NaCl) containing 5% glycerol, then loaded onto a 1 ml Hi-Trap heparin column (Pharmacia). Proteins were eluted with a 0–1.0 M gradient of NaCl, dialysed into TGED containing 50% glycerol and stored in liquid nitrogen. All samples were analysed on 10% SDS–polyacrylamide gels.

Purification of *Azotobacter vinelandii* NifA

Azotobacter vinelandii native NifA protein was purified from 3 l of BL21-GOLD (DE3) (pLysS) (pDB737) culture grown at 37°C in Luria broth to an OD of 0.6 and then induced by addition of 0.1 mM IPTG. Protein purification was then carried out as described previously (25) and final fractions were stored in 50% glycerol, in liquid nitrogen.

β -Galactosidase assays

β -Galactosidase assays were performed as previously described (17). For assays using pUJ100 (26) as the reporter, cells were grown aerobically overnight in LB medium supplemented with

Table 2. Results of β -galactosidase assays to determine dependence on local activator concentration

σ^N variant	β -galactosidase activity (Miller units)					
	pxyIS + UAS		pnifL + UAS		pnifL – UAS	
None	350 \pm 26	(4%)	210 \pm 20	(5%)	90 \pm 30	(14%)
Wild-type (pES15)	8690 \pm 2920	(100%)	4320 \pm 930	(100%)	640 \pm 160	(100%)
Δ 56–83 (pES17)	10 990 \pm 1480	(126%)	3230 \pm 870	(75%)	260 \pm 70	(41%)
Δ 86–106 (pES12)	2700 \pm 980	(31%)	1370 \pm 220	(32%)	110 \pm 30	(18%)
Δ 56–106 (pES9)	3880 \pm 680	(45%)	2100 \pm 260	(49%)	240 \pm 70	(37%)

Activities are the means of at least three experiments with four replicates per experiment. All assays were conducted in the *K.pneumoniae* *rpoN* deletion strain UNF2792 carrying the appropriate *lacZ* promoter fusion and, where indicated, a second plasmid carrying *rpoN*. Percentage activity is calculated relative to that of a C-terminally hexahistidine tagged wild-type σ^N on pES15, which has a comparable *in vivo* transcriptional activity to non-His tagged σ^N (data not shown).

2.5 mM 3-methylbenzyl alcohol and the appropriate antibiotics. All assays were repeated at least three times with four replicates per assay.

Core binding assay

Binding of the σ^N variants to *E.coli* core RNA polymerase was analysed by native gel electrophoresis as described (19).

Gel mobility shift assays

DNA binding assays were carried out essentially as described (19). Complexes were visualised by electrophoresis on a 3% native polyacrylamide gel (1:0.05 acrylamide:bis-acrylamide), followed by imaging on a Fujix Phosphorimager or exposure to Biomax film overnight. Open complex assays were carried out as described (27) with some modifications. Complexes were visualised by electrophoresis on a 3% native polyacrylamide gel (80:1 acrylamide:bis-acrylamide) containing 2.5% glycerol, followed by imaging as before. The 260 bp *Rhizobium meliloti* *nifH* promoter – UAS fragment was that described (19) and a 470 bp fragment carrying the *R.meliloti* *nifH* promoter + UAS was derived by PCR from pMB210 (28) using primers 185 and *lacZ*. Reaction mixes used an initiating nucleotide (4 mM GTP) to drive open complex formation.

RESULTS

Previous work that implicated Region II of σ^N in open complex formation was carried out using *in vivo* dimethyl sulphate footprinting to assess promoter binding and *in vivo* KMnO₄ footprinting to assess promoter melting (21). We have carried out both *in vivo* and *in vitro* assays to critically assess the role of Region II residues in each of the steps leading up to formation of the open complex. The activities of three σ^N variants, lacking Region II residues 56–106, 56–83 and 86–106 (Fig. 1), were first analysed *in vivo* and the purified proteins were then used for *in vitro* assays of holoenzyme formation, promoter binding, closed complex formation and formation of open complexes in the presence of activator protein.

In vivo analysis of the effects of deletions in Region II

We compared the *in vivo* transcriptional activities of the three σ^N deletion proteins using two different σ^N -dependent



Figure 1. Amino acid sequence of the wild-type (residues 1–120) and deletion variants of Region II. Region II residues are shown in bold and residues flanking the deletions that were altered as a consequence of introducing restriction sites are shown in lower case. The relative positions of the restriction sites used in the constructions are indicated. The *EcoRI* site in pES6 is 19 bp upstream of the ATG of *rpoN* and the *BamHI* site located 770 bp downstream of the ATG is that of the wild-type *rpoN* sequence.

promoters fused to *lacZ*. These promoters were pxyIS on pUJ100, where the activator protein is XylR which is also encoded by the plasmid (26), and *pnifL*, for which the activator is the chromosomally encoded NtrC protein, either with (pRD554) or without a UAS (pRD535) (29) (Table 2). The relative activities of the σ^N variants compared to the wild-type were similar at both complete promoters (*pnifL* + UAS and *pxyIS* + UAS), suggesting that the activities reflect properties of the variant σ subunits rather than properties of particular promoter sequences.

Deletion of the majority of Region II (Δ 56–106) had a marked effect on *in vivo* transcription activation, reducing it to less than 50% of the wild-type value. A similar but even more marked effect was seen when the C-terminal part of Region II was deleted (Δ 86–106). However, removal of the N-terminal part of Region II (Δ 56–83) had much less effect and was only

markedly different from wild-type σ^N when assayed on a promoter lacking a UAS (Table 2). In the absence of a UAS transcription is effected by the activator acting either from solution or by binding to non-specific sites in the promoter rather than from a specific upstream DNA-bound site and, consequently, transcription initiation is much more sensitive to activator concentration (30). As expected, the activity of *pnifL* lacking its UAS was significantly reduced compared to *pnifL* + UAS. In the absence of a UAS the relative activities of each of the σ^N variants was further reduced when compared to the relevant wild-type activity, suggesting that all of the deletions, but especially $\Delta 56-83$ and $\Delta 86-106$, are sensitive to reductions in the effective activator concentration.

Deletion of Region II affects holoenzyme formation

A prerequisite for the formation of the closed complex at promoter DNA is association of the σ factor with core RNA polymerase to form a holoenzyme. The holoenzymes formed between the σ^N variants and *E.coli* core polymerase were analysed on native gradient gels. It should be noted that on these gels the core enzyme does not run as a discrete band, but as a reproducible pattern of bands, and the holoenzyme shows a greater electrophoretic mobility than all of the lower molecular weight core enzyme bands. These anomalies may be explained by differences in conformation and charge of the proteins, and have been observed in previous studies (19,20). All *in vitro* work used the purified, His-tagged forms of the σ^N variants, herein designated σ^N (wild-type), ΔII ($\Delta 56-106$), $\Delta 56-83$ and $\Delta 86-106$.

Comparisons of wild-type His-tagged and non-His-tagged σ^N showed that the presence of the His-tag had no effect on the core binding ability of the σ factor. In both cases the holoenzyme complex appeared when σ^N was present at half the molar concentration of core enzyme and the core enzyme was completely titrated out when σ^N was present in 2-fold molar excess (data not shown).

Both the $\Delta 56-83$ and $\Delta 86-106$ variants were able to bind core enzyme in a manner comparable to wild-type, though $\Delta 86-106$ formed complexes at a slightly lower molar ratio σ :core, perhaps suggesting a greater affinity for the core enzyme (Fig. 2a-c). However assays with ΔII revealed a different pattern of holoenzyme formation. The discrete holoenzyme band was not present, even when ΔII was present in 5-fold molar excess over core enzyme, but instead a smear was consistently and reproducibly seen, suggestive of an unstable, dissociating holoenzyme complex (Fig. 2d). The absence of the holoenzyme is not due to insufficient σ factor, as the excess unbound ΔII can be clearly seen at the bottom of the gel.

A number of naturally occurring σ^N proteins have very short Regions II, the smallest being in *R.capsulatus* where Region II is a similar length (26 residues) to that in the *K.pneumoniae* $\Delta 56-106$ mutant (21 residues). We examined the formation of a holoenzyme comprising *R.capsulatus* σ^N and *E.coli* core enzyme and found that this resulted in a similar unstable complex as judged by its behaviour on native gradient gels (data not shown).

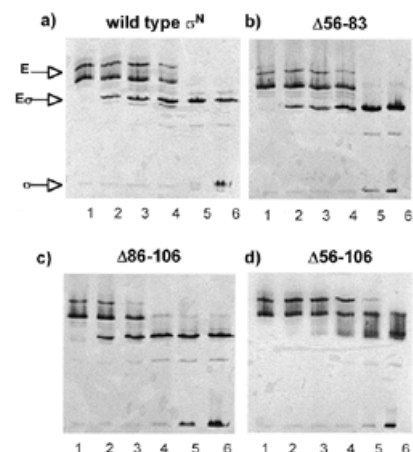


Figure 2. Holoenzyme formation by variants of σ^N . (a) Wild-type σ^N ; (b) $\Delta 56-83$; (c) $\Delta 86-106$; (d) $\Delta 56-106$. Complexes were visualised on Coomassie stained 4–15% native gels. E, core polymerase; E σ , holoenzyme; σ , free σ factor. For all gels the concentration of core enzyme was 2.9 μ M and the ratios of core polymerase (E) to σ factor (σ) were as follows: lane 1, E alone; lane 2, 5E:1 σ ; lane 3, 2E:1 σ ; lane 4, 1E:1 σ ; lane 5, 1E:2 σ ; lane 6, 1E:5 σ .

DNA binding by isolated σ^N is affected by deletion of residues 56–83 or 56–106, but not by deletion of residues 86–106

Formation of a heparin-sensitive closed complex involves binding of the holoenzyme containing σ^N to a $-24, -12$ type promoter. However, σ^N has also been shown to bind to promoters with consensus $-24, -12$ sequences in the absence of core enzyme (31). We therefore assayed the ability of each σ^N variant to bind alone and as holoenzyme to a radiolabelled *R.meliloti nifH* promoter fragment derived from pMB210 (28) and visualised the complexes by autoradiography.

Figure 3 shows binding to the *nifH* promoter by σ^N , ΔII , $\Delta 56-83$ and $\Delta 86-106$. These shift assays are unusual in that free DNA completely disappears with the lowest concentration of protein rather than being titrated out gradually. Also, instead of a discrete, one-step shift from one gel mobility to another, the addition of increasing concentrations of the protein causes the shift to occur in a number of steps until a maximum is reached. Addition of greater concentrations of protein after this point does not cause any further shift in the position of the complex. These phenomena have been observed in previous studies with σ^N (19,32).

Binding of the promoter DNA by σ^N -His was complete at a concentration of 2.0 μ M protein, in agreement with previous data (19,32), and a similar result was obtained with $\Delta 86-106$. However, promoter binding by ΔII and $\Delta 56-83$ was significantly impaired and an incomplete shift was observed even at the higher concentrations of protein (4 μ M) used in each of these assays. We conclude that residues 86–106 of σ^N are not essential for the specific DNA-binding function of the σ factor, but deletion of residues 56–83 (in either $\Delta 56-83$ or ΔII) does affect the ability of σ^N to bind to the *nifH* promoter and results in an unstable σ -DNA complex.

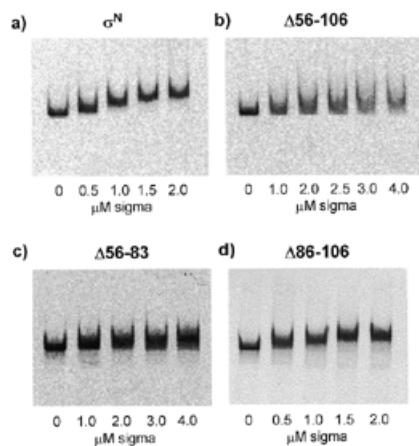


Figure 3. Binding of each isolated σ^N variant to the *R. meliloti nifH* promoter. (a) Wild-type σ^N ; (b) $\Delta 56-106$; (c) $\Delta 56-83$; (d) $\Delta 86-106$. A constant amount (150 nM) of radiolabelled promoter probe was used in each 10 μ l reaction with various concentrations of σ factor, indicated at the bottom of each lane. The mobility of free DNA can be seen in the first lane of each gel, where no σ factor was added to the assay.

Defects in DNA binding by the σ^N variants are overcome in the holoenzyme

Although σ^N can bind to certain promoters in the absence of core polymerase, the holoenzyme has a significantly increased promoter affinity (32). The holoenzyme is also ultimately the catalytic form of the enzyme, the σ factor alone having no ability to promote open complex formation (31). The results of assays in which holoenzymes incorporating each of the σ^N variants were incubated with *nifH* promoter DNA to form closed complexes are shown in Figure 4.

As with binding of the σ factors alone to the DNA, the shift to closed complex does not occur as one discrete step, but in a continuum. Wild-type holoenzyme at 50 nM induces a detectable mobility shift in the DNA and a complete shift occurs with 100 nM holoenzyme. All three variants showed a marked shift at 100 nM holoenzyme, although the Δ II holoenzyme apparently had a somewhat lower affinity than the wild-type and the $\Delta 86-106$ holoenzyme showed a higher affinity than the wild-type, with a detectable shift at 25 nM enzyme.

We conclude that the reduced ability of the Δ II and $\Delta 56-83$ σ factors to bind alone to the *nifH* promoter is rescued when the σ factors are bound to core polymerase in the holoenzyme. This suggests that binding of the σ factor to core enzyme constrains the conformation such that any defects in binding are effectively overcome.

Deletion of residues 86–106 or 56–106 causes a defect in activator response and open complex formation

Once the closed complex has formed on the promoter DNA, an activator protein is required to catalyse open complex formation. The *in vitro* formation of an open complex can be assayed using a promoter sequence with or without a UAS such that the activator is either tethered upstream of the holoenzyme or exerts its effect from solution. This allows assessment of the

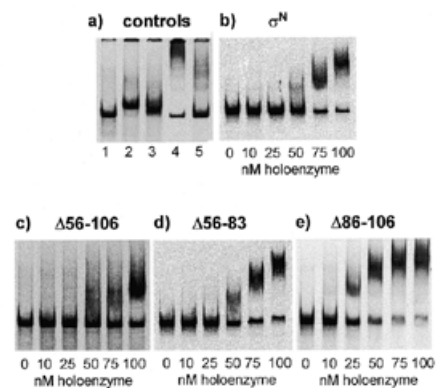


Figure 4. Holoenzyme binding to the *R. meliloti nifH* promoter. (a) Controls; (b) wild-type σ^N ; (c) $\Delta 56-106$; (d) $\Delta 56-83$; (e) $\Delta 86-106$. Holoenzymes were formed by preincubation of *E. coli* core polymerase and the relevant σ factor in a ratio of 1:10. An aliquot of 150 nM radiolabelled promoter probe was used in each reaction. Each reaction contained 60 ng of single-stranded salmon sperm DNA to titrate out non-specific binding by core polymerase. Controls are shown in the first gel. Lane 1, DNA probe, no enzymes; lane 2, 2.0 μ M σ^N bound to DNA; lane 3, 3.0 μ M Δ II bound to DNA; lane 4, DNA probe plus 145 nM core enzyme (the band at the very top of the gel is indicative of core polymerase bound non-specifically to the DNA); lane 5, DNA probe plus 145 nM core enzyme and 60 ng salmon sperm DNA. The concentration of holoenzyme present in assays of holoenzyme binding to DNA is indicated at the bottom of each lane.

sensitivity of the closed complex to the local activator concentration.

Previous studies have shown that a concentration of ~400 nM NifA is required to activate the formation of open complexes from solution at the *R. meliloti nifH* promoter with no UAS (33). With *R. meliloti nifH* as the template the use of GTP in our open complex assays allows the formation of a trinucleotide and hence the observed complex is formally an activator-dependent initiated complex, however, for simplicity in the following discussion we will refer to this complex as an open complex. Using wild-type σ^N , 200 nM NifA was sufficient for open complex formation, with 400 nM NifA giving a reproducible and distinct open complex band on the gel (Fig. 5a). In contrast, when a single UAS was included on the promoter fragment a reproducibly distinct open complex band was produced with 75–100 nM NifA (Fig. 5a). Hence, with wild-type σ^N at the *R. meliloti nifH* promoter 4- to 5-fold less NifA is required for open complex formation *in vitro* in the presence of the UAS. We then carried out assays at *pnifH* with and without a UAS using each of the σ^N deletion proteins to examine open complex formation and the response to NifA concentration.

Neither Δ II nor $\Delta 86-106$ showed any open complex formation at *pnifH* – UAS or *pnifH* + UAS using 100–1200 nM NifA (data not shown). In contrast, open complex formation by $\Delta 56-83$ holoenzyme at *pnifH* – UAS was just detectable at 600 nM NifA (data not shown) and at *pnifH* + UAS open complexes were detectable at 200 nM NifA (Fig. 5b). The $\Delta 56-83$ holoenzyme also produced a distinct band of lower mobility that runs higher on the gel than the open complex (Fig. 5b). The identity of this band, which is also seen very

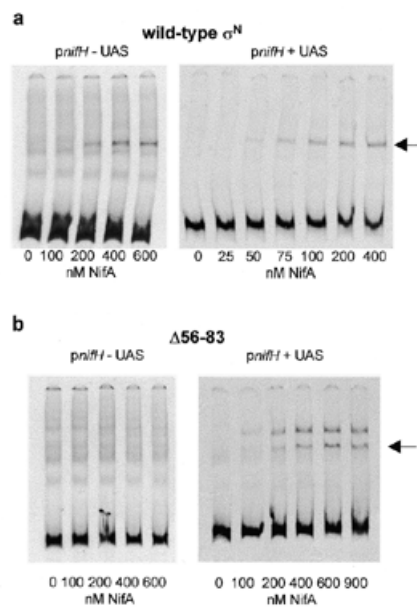


Figure 5. Open complex formation at the *R.meliloti nifH* promoter with and without a UAS (*nifH* ± UAS). (a) Assays of wild-type σ^N activity; (b) assays of $\Delta 56-83$. Assays contained 150 nM *nifH* – UAS or *nifH* + UAS. Holoenzymes were formed by preincubation of 125 nM *E.coli* core polymerase with ~320–330 nM σ factor (i.e. a 1:3 ratio) in a 10 μ l reaction volume, which also contained 4 mM GTP and 55 ng salmon sperm DNA. *Azotobacter vinelandii* NifA protein was added at various concentrations (indicated in the figure) with 2.5 μ M BSA and 50 nM IHF (for assays at *nifH* + UAS only) to a final volume of 15 μ l. Complexes were challenged with 2 μ g of heparin added with the loading dye. Control assays were carried out to verify the nature of each band on the gel (data not shown). The arrow indicates the position of the open complex.

faintly with the wild-type enzyme, is unknown, although it could represent an intermediate with a discrete conformation. This band was not observed with the Δ II holoenzyme but was clearly present with the $\Delta 86-106$ holoenzyme, even though the latter did not show a band equivalent to the open complex (data not shown).

We conclude that under these experimental conditions deletion of residues 56–106 and 86–106 results in a σ^N holoenzyme that is unable to respond correctly to activator protein and/or melt the DNA to form open complexes even when the activator is bound to the UAS. However, holoenzyme containing $\Delta 56-83$ σ^N is able to form a significant number of open complexes when the UAS is present on the promoter probe, at concentrations of NifA around 4-fold higher than required by wild-type holoenzyme. Therefore, deletion of residues 56–83 impairs open complex formation, but does not abolish the activity.

DISCUSSION

Whilst current models of transcription initiation by σ^N RNA polymerase have identified distinct functions for the conserved Regions I and III, the role of Region II (defined here for *K.pneumoniae* as residues 49–118 inclusive) has received relatively little attention. Initial *in vivo* studies on Region II

used a $\Delta 51-77$ deletion strain and a strain in which residues 40–81 were duplicated and concluded that Region II played a role in promoter melting (21). A more recent study compared the *in vitro* activities of a Region I deletion (residues 1–57) with a Region I+II deletion (residues 1–107) (22). In the latter case Region II was proposed to be dispensable for polymerase isomerisation on pre-melted DNA templates but was implicated in holoenzyme formation and the DNA-binding properties of that holoenzyme.

Our *in vivo* analysis showed that Region II is certainly not essential for transcription initiation, a result that is consistent with the virtual absence of Region II in σ^N from some organisms such as *R.capsulatus* or *Bacillus subtilis* (34,35). Nevertheless, whilst deletion of the N-terminal part of Region II ($\Delta 56-83$) had relatively little effect on transcription, deletion of the C-terminal part ($\Delta 86-106$) or of most of the region ($\Delta 56-106$) caused a very notable reduction in transcription initiation. We therefore dissected the various properties of σ^N *in vitro* to investigate the function(s) affected.

In constructing the deletions described in this paper we have introduced three alterations to residues within Region II; E55D, Q105L and D106A, of which all three only occur in the $\Delta 56-106$ variant (Fig. 1). As already mentioned, Region II is an extremely variable part of σ^N and indeed comparison of the very closely related σ^N proteins from *K.pneumoniae*, *E.coli* and *Salmonella typhimurium* shows that 15 of the 70 residues in the region vary between these organisms, with 11 of these being non-conservative changes. Of the changes we made E55D is a conservative change, residue 105 is isoleucine in both *E.coli* and *S.typhimurium* so that Q105L is very similar to the *E.coli* sequence and D106A results in the loss of just one charged residue. We therefore consider that it is relatively unlikely that any of the changes we have introduced will have major effects. Of potentially greater significance are the possible effects of the deletions on the conformation of Region II and the potential consequences for interactions between Regions I and III. Other studies have suggested that interactions between Region I and Region III are involved in the formation of a stable open complex (36,37) and hence the phenotypes of deletions in Region II may reflect the absence of specific residues in the Region, e.g. the change in overall charge, and/or changes in interactions between other regions of the protein.

The partial deletions had no discernable effect on holoenzyme formation but the complete deletion did not form a stable complex when analysed on native gels, suggesting that Region II either interacts directly with the core or affects the conformation of σ^N so as to alter its ability to form a stable holoenzyme. Studies with σ^N in which both Regions I and II had been deleted also concluded that Region II played a role in stabilising the holoenzyme (22), and comparative studies of the protease sensitivity of σ^N in its free and holoenzyme forms came to similar conclusions (36). DNA binding studies with the isolated σ^N variants showed that both the complete deletion and $\Delta 56-83$ were impaired in binding to the *R.meliloti nifH* promoter. Here again this could reflect some impairment of the conformation in the deleted proteins. However, these defects were essentially overcome when we analysed holoenzyme binding to the same promoter fragment for each of the σ^N deletions. It would therefore appear that any conformational alterations are largely suppressed once the σ^N holoenzyme is

bound in a closed complex. These observations are consistent with studies using a Region I+II deletion of σ^N that also suggested a role for Region II in stabilising the holoenzyme and identified an influence of Region II on DNA binding (22).

Despite the fact that the deleted variants of σ^N formed closed complexes at the *R.meliloti nifH* promoter they were all very significantly impaired in open complex formation *in vitro*. However, it should be noted that our assays do not distinguish between the formation of fewer open complexes than seen with the wild-type or of a similar number of less stable complexes. We also cannot exclude the possibility that, as the assay used actually allows formation of an initiated complex, the mutations could be affecting initiation rather than the transition from the closed to the open complex. Only the $\Delta 56-83$ σ^N holoenzyme showed some activity, which was improved when the promoter fragment contained a NifA UAS. In this sense the *in vitro* data mirror the *in vivo* data but it is very apparent that the effects of the mutations are far less pronounced *in vivo*. This may reflect the fact that our *in vitro* conditions do not sufficiently resemble the physiological conditions *in vivo*. Our data also agree in a qualitative sense with the *in vivo* open complex analyses of Wong and Gralla, who concluded that the ATR region is required for a complete melting transition at physiological temperatures (21). However, RNA polymerase isomerisation and initiation of transcription *in vitro* on pre-melted DNA templates has been reported not to require Region II sequences (22). When considered with our data and those of Wong and Gralla (21), it would appear that Region II plays a role in formation of the open complex prior to the holoenzyme adopting an isomerised state. It should be noted that the studies on pre-melted DNA utilised σ^N variants lacking Region I and the effects of Region II reported in this work may well be mediated through Region I-dependent conformational changes.

When all known σ^N amino acid sequences (in excess of 30) are compared, two features of Region II are particularly apparent, namely the predominance of acidic residues and the variation in length. Whilst our data for σ^N confirm that open complex formation is impaired in the absence of Region II, they do not demonstrate a role for the acidic residues *per se*. It is notable that models involving σ acidic residues and open complex formation have been invoked on a number of occasions, though to date no clear or unifying picture explaining the role of these acidic residues has emerged.

The negative charge conferred on Region II of σ^N by the large number of acidic residues suggested a model involving electrostatic interactions between this region and promoter DNA to explain its apparent role in open complex formation (5). The presence of a non-conserved highly acidic patch is not a unique feature of σ^N , as a comparable region is found between regions 1.2 and 2.1 in the major vegetative σ factor of the proteobacteria, typified by σ^{70} of *E.coli*. In this acidic loop (residues 188–209 in *E.coli* σ^{70}) 18 of the 22 amino acids are negatively charged and a role for electrostatic interactions with promoter DNA has again been considered. Most of the loop is disordered in the crystal structure but it could potentially occupy a cleft adjacent to the -10 promoter recognition helix (region 2.4), leading to the suggestion that the loop could sterically inhibit promoter- σ interaction and repel the negatively charged DNA (38). However, studies of DNA binding by σ^{70}

derivatives in which fluorescent probes were introduced on specific tryptophan residues near the -10 helix provided no evidence for a highly negatively charged loop in the vicinity of region 2.4 (39).

Finally a number of conserved acidic amino acids within region 1.1 of *E.coli* σ^{70} (in this case between residues 18 and 75) have also been implicated in open complex formation, but subsequent mutational analysis showed that several of the acidic residues were tolerant of substitution (40,41).

With regard to the variable length of σ^N Region II, there are now two cases, namely *B.japonicum* (42) and *Rhizobium etli* (43), where two copies of *rpoN* are present in the genome. In both cases the two σ^N proteins differ by some 40–50 amino acids in the primary sequence and alignments indicate that this variation is essentially all located in the N-terminal end of Region II. This is consistent with our observation that deletions in this part of the protein are the least deleterious. In both *B.japonicum* and *R.etli* the smaller of the two σ^N homologues is preferentially expressed at low oxygen tensions, consistent with specific induction to accommodate the high levels of σ^N -dependent expression of nitrogen fixation genes during symbiosis. It may be that the difference in Region II in the two σ^N proteins facilitates differential interaction with core RNA polymerase under particular physiological conditions.

ACKNOWLEDGEMENTS

We thank Bob Kranz for σ^N from *R.capsulatus*, Martin Buck for plasmid pWVC93025, Gary Sawers, Richard Little and Jason Barrett for technical advice, and Sara Austin and Ray Dixon for valuable comments on the manuscript. This work was supported by a grant in aid to the John Innes Centre from the BBSRC and a BBSRC studentship to E.S.

REFERENCES

- Lonetto, M., Gribskov, M. and Gross, C.A. (1992) *J. Bacteriol.*, **174**, 3843–3849.
- Merrick, M.J. (1993) *Mol. Microbiol.*, **10**, 903–909.
- Popham, D.L., Szeto, D., Keener, J. and Kustu, S. (1989) *Science*, **243**, 629–635.
- Buck, M., Miller, S., Drummond, M. and Dixon, R. (1986) *Nature*, **320**, 374–378.
- Sasse-Dwight, S. and Gralla, J.D. (1990) *Cell*, **62**, 945–954.
- Hsieh, M. and Gralla, J.D. (1994) *J. Mol. Biol.*, **239**, 15–24.
- Hsieh, M., Tintut, Y., Jiang, Y. and Gralla, J.D. (1994) *J. Biol. Chem.*, **269**, 373–378.
- Wong, C., Tintut, Y. and Gralla, J.D. (1994) *J. Mol. Biol.*, **236**, 81–90.
- Wang, J.T., Syed, A., Hsieh, M. and Gralla, J.D. (1995) *Science*, **270**, 992–994.
- Syed, A. and Gralla, J. (1998) *J. Bacteriol.*, **180**, 5619–5625.
- Gallegos, M.T., Cannon, W.V. and Buck, M. (1999) *J. Biol. Chem.*, **274**, 25285–25290.
- Cannon, W., Gallegos, M.T., Casaz, P. and Buck, M. (1999) *Genes Dev.*, **13**, 357–370.
- Casaz, P., Gallegos, M.T. and Buck, M. (1999) *J. Mol. Biol.*, **292**, 229–239.
- Tintut, Y., Wong, C., Jiang, Y., Hsieh, M. and Gralla, J.D. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 2120–2124.
- Tintut, Y. and Gralla, J.D. (1995) *J. Bacteriol.*, **177**, 5818–5825.
- Gallegos, M.T. and Buck, M. (1999) *J. Mol. Biol.*, **288**, 539–553.
- Coppard, J.R. and Merrick, M.J. (1991) *Mol. Microbiol.*, **5**, 1309–1317.
- Merrick, M. and Chambers, S. (1992) *J. Bacteriol.*, **174**, 7221–7226.
- Taylor, M., Butler, R., Chambers, S., Casimiro, M., Badii, F. and Merrick, M. (1996) *Mol. Microbiol.*, **22**, 1045–1054.
- Cannon, W., Missailidis, S., Smith, C., Cottier, A., Austin, S., Moore, M. and Buck, M. (1995) *J. Mol. Biol.*, **248**, 781–803.
- Wong, C. and Gralla, J.D. (1992) *J. Biol. Chem.*, **267**, 24762–24768.
- Cannon, W., Chaney, M. and Buck, M. (1999) *Nucleic Acids Res.*, **27**, 2478–2486.

23. Cannon,W., Austin,S., Moore,M. and Buck,M. (1995) *Nucleic Acids Res.*, **23**, 351–356.
24. Takeshita,S., Sato,M., Toba,M., Masahashi,W. and Hashimoto-Goto,T. (1987) *Gene*, **61**, 63–74.
25. Austin,S., Buck,M., Cannon,W., Eydmann,T. and Dixon,R. (1994) *J. Bacteriol.*, **176**, 3460–3465.
26. Du,Y., Holtel,A., Reizer,J. and Saier,M.H.,Jr (1996) *Res. Microbiol.*, **147**, 129–132.
27. Hill,S., Austin,S., Eydmann,T., Jones,T. and Dixon,R. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 2143–2148.
28. Better,M., Ditta,G. and Helinski,D.R. (1985) *EMBO J.*, **4**, 2419–2424.
29. Drummond,M., Clements,J., Merrick,M. and Dixon,R. (1983) *Nature*, **301**, 302–307.
30. Austin,S., Henderson,N. and Dixon,R. (1987) *Mol. Microbiol.*, **1**, 92–100.
31. Buck,M. and Cannon,W. (1992) *Nature*, **358**, 422–424.
32. Cannon,W., Claverie-Martin,F., Austin,S. and Buck,M. (1993) *Mol. Microbiol.*, **8**, 287–298.
33. Wang,X.Y., Kolb,A., Cannon,W. and Buck,M. (1997) *Nucleic Acids Res.*, **25**, 3478–3485.
34. Alias,A., Cejudo,F.J., Chabert,J., Willison,J.C. and Vignais,P.M. (1989) *Nucleic Acids Res.*, **17**, 5377.
35. Debarbouille,M., Martin-Verstraete,I., Kunst,F. and Rapoport,G. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9092–9096.
36. Casaz,P. and Buck,M. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.
37. Chaney,M. and Buck,M. (1999) *Mol. Microbiol.*, **33**, 1200–1209.
38. Malhotra,A., Severinova,E. and Darst,S.A. (1996) *Cell*, **87**, 127–136.
39. Callaci,S. and Heyduk,T. (1998) *Biochemistry*, **37**, 3312–3320.
40. Wilson,C. and Dombroski,A.J. (1997) *J. Mol. Biol.*, **267**, 60–74.
41. Bowers,C.W., McCracken,A. and Dombroski,A.J. (2000) *J. Bacteriol.*, **182**, 221–224.
42. Kullik,I., Fritsche,S., Knobel,H., Sanjuan,J., Hennecke,H. and Fischer,H.-M. (1991) *J. Bacteriol.*, **173**, 1125–1138.
43. Michiels,J., Moris,M., Dombrecht,B., Verreth,C. and Vanderleyden,J. (1998) *J. Bacteriol.*, **180**, 3620–3628.