A fork junction DNA–protein switch that controls promoter melting by the bacterial enhancer-dependent sigma factor

Yuli Guo, Lei Wang and Jay D.Gralla¹

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA

¹Corresponding author
e-mail: gralla@mbi.ucla.edu

Results of binding assays using DNA fork junction probes indicate that sigma 54 contains multiple determinants that regulate melting to allow RNA polymerase to remain in closed promoter complexes in order to respond to enhancers. Gel mobility shift studies indicate that the –12 promoter element and parts of sigma 54 act together to form a molecular switch that controls melting. The DNA sequences and the sigma 54 N-terminus help direct polymerase to the location within the –12 promoter element where melting will initiate. However, the fork junction that would lead to melting does not form, due to the action of an inhibitory DNA element. Such unregulated melting is inhibited further by the lack of availability of the single-strand binding elements, which are needed to spread opening from the junction to the transcription start site. Thus, in the absence of looping enhancer protein, proper regulation is maintained as the sigma 54 polymerase remains bound in an inactive state. These complex protein–DNA interactions allow the controls over protein recruitment and DNA melting to be separated, enhancing the diversity of accessible mechanisms of transcription regulation.

Keywords: DNA fork junction/enhancer/promoter opening/sigma 54/transcriptional initiation

Introduction

Core RNA polymerases in both bacterial and eukaryotic cells generally have a limited range of activities. They contain the catalytic RNA synthesis center and have a very limited ability to discriminate between different DNA sequences. The polymerases also contain determinants that mediate associations with the factors that bring them to promoters and allow them to transcribe in a regulated manner. In higher eukaryotes, dozens of polypeptides are typically involved in this process (Orphanides *et al.*, 1996; Hampsey, 1998). In simple bacteria, a single polypeptide, a sigma factor, is sometimes sufficient. These sigma factors typically contain multiple domains that collectively provide a range of activities (Gross *et al.*, 1992). The sigma domains always include determinants to bind the polymerase and to recognize the specific type of targeted promoter sequences. They can also include determinants that mediate transcriptional regulation (Busby and Ebright, 1994; Hochschild and Dove, 1998). Thus, sigma factors are essential intermediaries in the process of promoter

recognition and regulated transcription initiation. Bacterial cells contain numerous sigma factors, which function under different physiological conditions to direct transcription of different classes of promoters.

In *Escherichia coli* and many other bacteria, these sigma factors fall into two distinct classes with regard to both sequence similarity and transcription mechanism (Collado-Vides *et al.*, 1991; Gross *et al.*, 1992; Merrick, 1993). The minor class is represented uniquely by sigma 54, a protein that causes transcription to assume several features reminiscent of higher eukaryotic mechanisms; among these are a requirement to respond to remote enhancers and to hydrolyze ATP during initiation (Reitzer and Magasanik, 1986; Ninfa *et al.*, 1987; Weiss *et al.*, 1991; Austin and Dixon, 1992). The major class is the sigma 70 family of proteins. These sigma 70-type proteins bring RNA polymerase to targeted promoters, often with the assistance of activator proteins bound to sites directly adjacent to the sites recognized by domains within the sigma factor itself (Busby and Ebright, 1994; Gralla, 1996; Hochschild and Dove, 1998). The holoenzyme then opens the DNA to expose the start site and begins transcription without the need for ATP hydrolysis. Non-template (top) strand promoter sequences have a role in formation of the open complex (Marr and Roberts, 1997). We showed recently that this opening process relies on a previously unsuspected activity, one that binds the upstream fork junction that is created when DNA strands are to be separated (Guo and Gralla, 1998). This activity cooperates with two single-strand binding activities to spread melting downstream to complete promoter opening (Guo and Gralla, 1998).

The fork junction binding activity also appears in the sigma 54 machinery, although it has not yet been shown that this is related to DNA melting (Guo and Gralla, 1998). The issue is of central interest because sigma 54 transcription is typically regulated at the level of DNA melting (Ninfa *et al.*, 1987; Sasse-Dwight and Gralla, 1988; Morett and Buck, 1989; Popham *et al.*, 1989), i.e. typically, the holoenzyme does not need to be recruited to the promoter by activators but instead binds its target promoters in an inactive state. Activation occurs after signal transduction triggers an enhancer binding protein to become active (Ninfa and Magasanik, 1986; Keener and Kustu, 1988). This enhancer-bound activator loops to the inactive closed complex and causes an ATP-dependent melting event that leads to transcription (Su *et al.*, 1990).

These considerations indicate that control of melting is the key to understanding how sigma 54 uniquely directs the bacterial RNA polymerase to respond to enhancers. In this study we find that a unique multi-component DNA–protein interaction, not present at sigma 70 promoter assemblies, but still operating through fork junction binding, is of central importance to this control.

Results

Sigma 54 holoenzyme specifically recognizes junctions created within the –12 promoter consensus region and discriminates between the forks

Prior band-shift experiments showed that sigma 70 holoenzyme binds tightly to a specifically located fork junction structure within the consensus sequence of a target promoter (Guo and Gralla, 1998). Optimal binding is to a structure that mimics one fork of the open complex; it occurs when –12 is the terminal base pair, and the –11 nucleotide on the top (non-template) strand is present and unpaired. Several lines of evidence suggested that this binding is closely related to promoter melting. The preferred junction location is the same as occurs within the promoter when open complexes are created during transcription initiation. The junction is located within very highly conserved –10 region promoter consensus sequences. When these promoter sequences are experimentally varied, the affinity of holoenzyme for the fork junction variants correlates roughly with promoter opening rates for the same variants (Roberts and Roberts, 1996; Guo and Gralla, 1998). Holoenzyme binding to this fork junction confers substantial heparin resistance, which is a hallmark of open complex formation.

We wished to learn if such junction binding was also important for sigma 54 promoter melting, which is highly regulated. For sigma 70 promoters, the preferred junction has base pair –12 intact and base pair –11 melted, and binding preferentially uses the non-template strand (Roberts and Roberts, 1996; Guo and Gralla, 1998). Sigma 54 polymerase uses a consensus element in this same location but the sequence is entirely different (Wang and Gralla, 1998). We initiated band-shift experiments using the tight binding *Rhizobium meliloti* nifH sigma 54 promoter (Buck and Cannon, 1992) to learn if this same junction location is preferred despite the difference in nucleotide sequence. As before, initially we used probes containing only one single strand of the fork, thereby avoiding introducing mutant nucleotides and creating artificial mismatches associated with mismatched heteroduplexes.

Initially, a series of probes were used that differ in the location of the double-strand–single-strand junction within the promoter (the locations are marked by dots in Figure 1C and D). Probes are created by annealing a labeled single strand to a series of progressively shortened complementary strands (Figure 1B shows the products of annealing). A low concentration of sigma 54 holoenzyme is used to bind these probes, thereby revealing differences in affinity. The results using this set of junction probes with bottom single-strand tails show that position –12 is strongly preferred as the terminal base pair. Approximately two-thirds of this probe is bound by holoenzyme (lane T-12 of Figure 1C, left panel). The binding requires sigma 54 (Figure 1C, right panel). Approximately two orders of magnitude more of this fork probe is bound by holoenzyme than when a double-strand probe of the same sequence is used (see quantitation).

Approximately half of the complexes with probe T-12 survive a heparin challenge (Figure 1C, central panel). This extent of resistance is associated with the type of

Fig. 1. EMSA using sigma 54 holoenzyme (7.5 nM) and DNA fork probes (1 nM) with differing junction positions. **(A)** The sequence of the parental double-stranded probe is shown. The numbering in all figures denotes the –12 nucleotide as the most downstream C:G, shown in bold. The four base pairs shown in bold are within the –12 and –24 consensus elements. The sequence is derived from the *R. meliloti* nifH promoter with a C:G pair deleted at –1. **(B)** An example of the annealing products used in this work [see (C) for designations]. The preparations are typically >90% double stranded. **(C)** EMSA using bottom strand fork probes. The identical 32 P-labeled (*) bottom strand was annealed to each top strand indicated. In all cases, the position of last 3' base on the top strand (T) is indicated above the appropriate lane. The top shows a schematic of the various probes with dots representing this terminal base. Each probe was subjected to EMSA with sigma 54 holoenzyme, with heparin challenge (+Hep) in the middle panel. In this and subsequent experiments, the unbound DNA probes were run off in order to enhance the resolution around the shifted bands. The appropriate shifted band, representing the complex of holoenzyme and DNA fork probe, is marked with an arrow. Controls: 'ssBot', bottom single-strand probe; 'Core', enzyme lacking sigma 54. As in the other figures, the numbers under each lane represent the percentage of probe bound; this was determined by comparing with unbound probes run in parallel lanes or gels and averaged from multiple experiments. **(D)** EMSA using top strand forks. The schematic and labels are as above except that the shorter strand in the duplex is always the bottom strand (B). Experiments were repeated at 0–4°C with no difference in interpretation.

interactions that can partially mimic formation of open complexes (for a discussion see Guo and Gralla, 1998). The preference for binding to a –12/–11 junction and the

Fig. 2. Preferred locations and nucleotides within the -12 consensus element. EMSA protocols and probes are as in Figure 1. (A) Effect of removing top single-strand sequences on the weak binding observed with -12 as the terminal base pair of a top strand fork. (B) Effect of nucleotide substitution for the wild-type adenosine at the top single-strand -11 position. (C) Effect of removing top single-strand sequences when -13 is the terminal base pair. (D) Effect of removing top strand sequences when -11 is the terminal base pair. (E) Effect of removing bottom strand sequences at the optimal -12 fork junction. (F) Effect of substitution for the bottom strand -11 nucleotide on the strong binding observed at an optimal fork junction. In all cases, detectable binding required the sigma component of the holoenzyme. Core enzyme alone does not give detectable bands (Figure 1) either with or without heparin (not shown).

substantial resistance of this interaction to heparin is as observed previously for sigma 70 promoters. Thus, despite the total absence of DNA sequence similarity at the target promoters, the sigma 70 and sigma 54 holoenzymes show a similar preference in locating and binding a -12/-11 fork junction.

Figure 1D shows the results of the same type of experiment except that the junction probes contained the top single strand of the fork. By far the greatest binding is to probes B-13 and B-11 within the -12 promoter element. Extraordinarily poor binding occurs with probe B-12 (terminal base pair -12) and this lack of binding is remarkable as the location is between the two strongest binding probes. The lack of binding is also remarkable because this same -12 junction location was the strongest binder when a probe with the bottom strand of the fork was used (T-12 in Figure 1C). Binding to the top fork probe B-12 is inhibited by approximately an order of magnitude compared with probes with adjacent junctions at -13 or -11. The unexpected inhibition of binding to probe B-12 persisted when several different preparations of probes were assayed. It was also confirmed using probes made with different sigma 54-dependent promoter sequences (not shown). We note that some bands are

doublets, with stronger binding yielding slightly retarded mobility, indicating that two conformers of bound holoenzyme might be present.

These results show that junction structures terminating at base pair -12 have unique properties. When the bottom fork is assayed, binding to the -12/-11 junction is preferred over all other junction positions. But when the top fork is assayed, binding to this same location is inhibited specifically (reduced 10-fold compared with adjacent locations). Such inhibition was not observed in previous experiments on the sigma 70 system (Guo and Gralla, 1998).

Figure 2A and B shows that this inhibition is due to the presence of the unpaired top strand nucleotide at position -11. Removal of the top strand sequences downstream of -7 or -11 does not restore binding (Figure 2A, lanes T-7 and T-11). However, removal of a single unpaired top strand nucleotide at position -11 does increase binding (cf. T-12 and T-11; see quantitation in Figure 2). The identity of this top -11 nucleotide is also important as shown by assay of substitutions at this position. Purines lead to a 20-fold reduction in binding and pyrimidines lower binding by 2- to 3-fold (Figure 2B, including quantitation of the weakest binding).

Fig. 3. Sigma 54 itself contains the determinants for use of the fork junction. Unless indicated specifically in the figure, EMSA used isolated sigma (370 nM) and DNA fork probes (1 nM). The sigma–probe complexes gave two conformer bands, with mobilities as described previously (Guo and Gralla, 1998). The sigma–DNA complexes are indicated as ‘σ’ and the holoenzyme–DNA complexes are indicated with ‘H’. (A) Bottom single-strand fork structures. (B) Top single-strand fork structures. (C) Contribution of the bottom strand –11 nucleotide using the optimal fork junction.

In contrast, the identity of the –11 nucleotide on the bottom strand is not important for the high affinity binding that occurs when this strand is single stranded. Figure 2F shows that –11 bottom strand substitutions do not alter this strong binding. However, as expected, a single stranded nucleotide of some type must be present, as its removal leads to a great loss of binding (lanes ‘No-11N’ in Figure 2F). Thus, when in single stranded form, the partner nucleotides in base pair –11 have very different effects on RNA polymerase binding.

The inhibition by the top unpaired nucleotide is specific with regard to the –12/–11 junction location as well as to identity. Assay of a junction probe terminated one base pair downstream does not show such an inhibitory effect; at a –11/–10 junction removal of the unpaired nucleotide (at –10) is without effect (Figure 2D, compare T-11 and T-10). When the same comparison is made at a –13/–12 junction (Figure 2C; T-13 and T-12) no information is available because binding to these probes is weak. This is because affinity for these probes requires the single-strand sequences downstream of –7; without the –7 to +1 segment there is no detectable signal (lane T-7 of Figure 2C; see below for further discussion).

We assessed the relative binding to the various fork structures by quantitative comparisons. When comparing various top strand forks, the junction that terminates at base pair –11 is bound slightly better than that terminating at –13, which in turn binds an order of magnitude more probe than the inhibited junction at –12 (Figure 1D, central and left panels). It takes a separate experiment to compare these top strand forks to bottom strand forks because the sets of probes were prepared separately and labeled in a different manner. Figure 2D shows that binding to the optimal bottom strand fork probe (T-12, which has the same structure as the optimal B-11 in Figure 2E) is ~7-fold greater than binding to the optimal top strand fork (e.g. T+1). Overall, the binding hierarchy is as follows: the bottom strand fork junction terminated at base pair

–12 is best; a top strand fork with the junction at –11 binds 5- to 10-fold less well; a top strand probe terminating at –12 is inhibited by another order of magnitude.

This analysis demonstrates that both the best and worst binding locations in the consensus region terminate at base pair –12. The bottom fork is bound nearly two orders of magnitude better than the top fork at this same junction location. Binding here confers substantial resistance to heparin (~50% in Figure 2). Thus, the holoenzyme has the ability to discriminate between the two forks, strongly preferring the bottom. In addition, the exposure of base pair –11 would be expected to induce remarkable opposing properties; it would expose the best binding structure, a bottom strand fork, while also exposing an inhibitory component, a top strand fork. The potential regulatory consequences of this opposition will be discussed below.

Sigma 54 itself contains the determinants for use of the fork junctions

Next we determined that the specificity for many of these properties is contained within the sigma 54 component of the holoenzyme. Preliminary experiments showed that sigma 54 alone can bind the optimal probe efficiently but an ~50-fold higher molar concentration is required to obtain similar extents of binding compared with holoenzyme (Figure 3A). Using these higher concentrations, the data show that the properties just described for holoenzyme are preserved using sigma 54 alone. The tightest binding is still to the same –12/–11 bottom strand fork (Figure 3A, lane T-12). Nearly 50% of this probe can be bound (Figure 3A) and this is completely dependent on an unpaired bottom strand nucleotide (Figure 3C, as shown in a previous related experiment by Guo and Gralla, 1998). This binding is only modestly reduced by heparin (data not shown), a reagent long known to bind so tightly to core RNA polymerase as to inactivate it, indicating that sigma 54 binds more tightly to the fork junction probe than to heparin. The shifted band appears as a doublet, as

Fig. 4. The effects of the N-terminal region deletion on fork junction binding preference. The experiments are the same as in Figure 1 except that the N-terminal deletion mutant of sigma 54 is used. The upper panels use holoenzyme and the lower panels show the weaker binding of mutant sigma alone. For holoenzyme assay, protein and DNA concentrations are the same as in Figure 1; for the mutant sigma assay, the sigma concentration is 460 nM and the probe is 10 nM. In the lower panel the total bound probe was in the order of 3% using optimal probes. For this mutant sigma, the binding was increased substantially in the absence of heparin (~10–20% on the optimal probes).

seen in several previous experiments (where the bands display as a broad smear; Tintut *et al.*, 1994; Wong *et al.*, 1994). The relative amount of each depends on whether the sigma is wild-type or mutant and varies somewhat between preparations; we surmise that they represent two different bound conformers of sigma.

The interaction of sigma 54 alone with various top strand forks is much weaker, as predicted based on the properties of the holoenzyme containing sigma 54; 10 times less sigma is bound even at high concentration (Figure 3B). As with holoenzyme, the top strand fork terminated by base pair –12 is selectively inhibited in binding (Figure 2B, lane B-12). Overall, the data show that it is the sigma 54 component of the holoenzyme that has the information to act both positively and negatively at the –12/–11 junction.

Based on previous experiments, the N-terminus of sigma 54 is an obvious candidate to contribute to these interactions. Various mutations in the N-terminus lead to altered control of melting (J.T.Wang *et al.*, 1995, 1997; Syed and Gralla, 1997; Cannon *et al.*, 1999). N-terminal deletion allows promoter DNA to be melted transiently by the mutant holoenzyme without activator, leading to some unregulated transcription. If the properties of sigma 54 at the –12/–11 fork junction are directly related to regulation of melting then N-terminal deletions should lead to altered use of this fork junction.

This connection is supported by the data of Figure 4,

using the sigma 54 N-terminal deletion mutant ΔN (J.T.Wang *et al.*, 1997) and the related mutant Sal58 (Sasse-Dwight and Gralla, 1990; data not shown). When ΔN holoenzyme is used to bind different bottom strand fork probes the usage is drastically altered from the wild-type pattern. The tightest binding probe has a junction located near –9/–8 (Figure 4A, lane T-9) rather than at –12/–11 as seen with the wild type. Indeed, the –12/–11 junction now binds relatively weakly (lane T-12) and virtually no binding here survives the heparin challenge (Figure 4A, central panel). Thus, the loss of control over melting observed previously with this mutant is mirrored by a loss of selective binding to the –12/–11 junction.

Drastic alterations are also seen when the top strand forks are assayed with the mutant holoenzyme. Preferred binding is again shifted towards downstream locations (Figure 4B). However, the residual binding to the –12/–11 junction has a revealing property. For both holoenzyme (Figure 4B, top) and sigma alone (Figure 4B, bottom), the inhibition associated with the use of this top strand fork junction appears to be retained (minimum binding to probes B-12). This apparent retention of inhibition suggests that protein domains responsible for the top –11 nucleotide inhibition include critical determinants outside the N-terminus of sigma 54.

Overall, the data demonstrate that the N-terminus is required for recognition of the –12/–11 fork junction. It is apparently not required for the inhibition of binding that can occur there along the top strand of the fork. When the N-terminus is deleted recognition occurs preferentially at a downstream location. This alternative location is likely to be a cryptic binding determinant as it represents a weak local maximum in the pattern of usage of wild-type holoenzyme (Figure 1C, left panel). When the mutant enzyme binds to this cryptic determinant it forms a heparin-sensitive complex (Figure 4, showing >80% dissociation by heparin), mirroring the heparin-sensitive unregulated transcription by this same mutant holoenzyme (J.T.Wang *et al.*, 1997).

The N-terminus of sigma 54 and the –12 DNA element cooperate to determine how fork junctions are used

The data have indicated that sigma 54 can use its N-terminus to direct holoenzyme to the incipient –12/–11 junction where melting originates during transcription initiation. We proposed previously that there may be regulatory cooperation between the N-terminus and the conserved –12 region promoter sequences. Part of the basis for this proposal was that unregulated transient melting can be caused by mutations within either the N-terminus of sigma 54 or within the –12 sequences of the promoter (J.T.Wang *et al.*, 1995, 1997; Syed and Gralla, 1997; Wang and Gralla, 1998). If this proposal is true then one might see the same altered junction recognition patterns whether it is the sigma 54 that is mutated or the promoter DNA. That is, sequence changes in the DNA that bypass regulation may lead to the use of downstream determinants as was just shown with bypass sequence alterations in sigma 54 protein.

We used the D3 promoter to test this proposal. D3 is a double promoter substitution mutation at positions –12 and –14; the conserved –14T and –12C are changed to

Fig. 5. The effects of an alternative -12 region promoter D3 sequence (Wang and Gralla, 1998) on the fork junction probe (5 nM) binding preference for wild-type sigma 54 holoenzyme (15 nM). (A) Bottom single strand fork structures. (B) Top single strand fork structures.

$-14A$ and $-12T$. D3 was identified as a sequence change that leads to low-level transcription without activator (Wang and Gralla, 1998), a phenotype similar to that caused by certain N-terminal sigma 54 mutants. We wished to learn if wild-type holoenzyme interacts with the D3 promoter to give altered fork junction preference patterns.

The data represented in Figure 5A show that the use of wild-type holoenzyme on the mutant D3 promoter causes the binding preference to shift as predicted. Just as with the sigma mutation the promoter DNA mutation causes a switch from the normal to the cryptic junction; binding is maximal at the $-9/-8$ junction (T-9) and weak at the $-12/-11$ junction (T-12). The same switch in preference is seen when probes with top strand tails are used (Figure 5B). We conclude that mutations in either the promoter or the protein can lead to common changes in fork junction usage and that these mimic the common changes in regulation of melting.

The result of Figure 5B shows an additional feature of fork junction recognition using the D3 promoter sequence. Note that there is residual recognition of the $-12/-11$ fork junction (T-12), although it has switched to being secondary to the downstream location. However, this $-12/-11$ recognition is not inhibited when the top -11 nucleotide is exposed (probe B-12). Recall that in the context of the wild-type promoter sequence, usage of this junction is strongly inhibited by unpairing of the -11 nucleotide on the top strand. Thus, the D3 promoter eliminates the inhibitory effect even when holoenzyme is associated with the normal fork junction location. The inhibitory effect was also reduced by nucleotide substitutions at -11 (Figure 2B) and D3 is substituted at -12 and -14 , so inhibition depends critically on the sequence context of an extended segment of the -12 region consensus element as well as the integrity of the protein.

Taken together, the data indicate that both the N-terminus of sigma 54 and the -12 promoter sequence are intimately involved in determining whether the $-12/-11$ fork junction is recognized and whether inhibition

Fig. 6. The contributions of single-stranded sequences to binding at the alternative junctions. The concentrations used are as in Figure 4. Heparin was not included. E refers to the presence of core polymerase. Quantitation of the data shown in (C) shows that the upper band is always 3–5% of the probe and the amount of lower band binding is indicated.

will be operative there. A defect in either partner can lead to the use of a normally cryptic downstream location that does not contain an inhibitory component. It seems plausible that this property can account for the previous observations of low-level unregulated transcription and melting under these circumstances.

Single-strand binding components

In previous studies, two single-strand binding activities were found to be associated with sigma 70 holoenzyme interacting with fork probes. These activities are believed to be functionally important but are lower affinity interactions compared with the dominant effect of binding to fork junctions. We explored whether related activities are also associated with sigma 54 holoenzyme (Cannon *et al.*, 1999). The data of Figure 2 suggested that one such activity exists, as removal of the top strand segment between -6 and $+1$ lowered affinity in the context of a weakly bound $-13/-12$ junction probe (Figure 2C, lane T+1 versus T-7).

This top single-strand interaction is easily detected in the context of binding to the weak cryptic fork junctions. Figure 6 (left panel) shows a strong reduction in binding when this -6 to $+1$ segment is removed (lanes T-1 and T-7). This activity cannot be detected in all contexts; when the fork junction terminates at base pairs -12 and -11 it makes a negligible contribution to stability (Figure 2A and D). Thus, it appears to be a context-dependent stabilizing interaction with top single-strand DNA just upstream from the transcription start site. There

are no indications in any of these experiments that the activity extends to sequences upstream from position -6 (Figures 2A, C, D and 6). In contrast, Figure 6A (top right) shows that shortening the bottom single strand leads to a gradual, progressive reduction in binding. This is in contrast to the strong selective reduction in binding caused by shortening of the top strand (Figure 6A, left). We infer that bound holoenzyme contains single-strand binding potential with the most specific interaction between +1 and -6 on the top strand.

However, the data also shows that the single-strand binding can be greatly strengthened by removal of the N-terminus of sigma 54. Quantitation of the data shows that N-terminal deletions lead to binding of 5–10 times more probe for holoenzyme (Figure 6A and B) and a substantial increase is also noted for sigma alone (Figure 6C; data not shown). This is true for both the top and bottom strand fork probes. The sites of single-strand interactions are preserved. On the top strand, the +1 to -6 segment is critical, whereas the removal of the bottom strand leads to a strong but progressive reduction in binding (Figure 6B and shown in Figure 6C for ΔN sigma 54). We conclude that top and bottom single-strand binding activities exist in holoenzyme but that they are largely masked when the N-terminus of sigma 54 is intact.

Discussion

Sigma 54 binds the same core RNA polymerase as does sigma 70, but each confers different properties on the holoenzymes formed. The association with sigma 54 causes the holoenzyme to alter its regulatory properties and to assume an enhancer responsive mechanism. Because such a fundamental shift in regulatory response may be relevant to understanding the wide diversity of prokaryotic and eukaryotic mechanisms a number of groups have investigated how sigma 54 directs this alternative behavior of the polymerase. It is known that sigma 54 holoenzyme is regulated primarily at the level of post-recruitment promoter melting (Ninfa *et al.*, 1987; Sasse-Dwight and Gralla, 1988; Morett and Buck, 1989; Popham *et al.*, 1989). A recent study showed that melting by sigma 70 holoenzyme relies on its ability to recognize the upstream fork junction that is created when melting initiates (Guo and Gralla, 1998). In this paper we find that a key property of sigma 54 is its ability to modify this interaction with the fork junction in a way that should make the promoter susceptible to regulation at the melting step. This interaction can be thought of as a kind of molecular switch that controls the on/off state of the melting reaction.

A sigma 54-specific mechanism for keeping melting under control

Both sigma 54 and sigma 70 promoters have two consensus elements, although the recognition sequences are entirely different (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 (Spassky *et al.*, 1985; Morris *et al.*, 1994; Meier *et al.*, 1995; Chen and Helmann, 1997). The new data show that both types of holoenzymes bind optimally to a fork junction at a common -12/-11 location (Guo and Gralla, 1998). This

location coincides with the location of the natural upstream fork junction within transcriptionally active open complexes (Kainz and Roberts, 1992; Morris *et al.*, 1994; Chen and Helmann, 1997). Thus, the sigma factors have in common an ability to direct selective recognition of the fork junctions that must be created when melting originates during transcription initiation.

Nonetheless, the mode of binding to this common location and structure is quite different for sigma 54 and sigma 70. There are several major differences that probably relate to how melting is controlled. First, the two holoenzymes prefer different forks. The data show that sigma 54 has a strong preference for the bottom strand of the fork and uses this to help acquire the heparin-resistance characteristic of activated open complexes. In contrast, sigma 70 has a top strand fork preference (Guo and Gralla, 1998).

A second difference lies in how the unpaired nucleotides at the -11 position are used. The unpaired non-template nucleotide at this position is important in both cases (Figure 2B; Roberts and Roberts, 1996; Guo and Gralla, 1998). However, it plays opposite roles in the two systems. In the case of sigma 70, the consensus nucleotide directs tight binding along the top strand fork. In the case of sigma 54, the unpaired consensus nucleotide inhibits tight binding. This inhibition would be expected to come into play when the -11 base pair becomes unpaired during initiation of promoter melting. A tendency to open this base pair to create the optimal bottom strand fork junction would be antagonized by the creation of an inhibitory unpaired top strand nucleotide. This inhibition of creation of the fork junction during melting would only happen in the case of sigma 54.

Another difference relates to how melting spreads downstream (Figure 7). In the case of sigma 70 (Guo and Gralla, 1998), single-strand binding activities exist over the entire downstream region over which melting must spread to reach the start site. In the sigma 54 case, however, there are barriers to spreading the melting. First, the single strand binding activities are masked (Figure 6 and associated discussion; Cannon *et al.*, 1999). Thus, it might require activation to unmask them, helping to establish the melting control. In addition, there is no obvious protein determinant that binds the single-stranded DNA region just adjacent to the fork junction, as exists in the case of sigma 70. Thus, there appear to be two barriers to the spread of melting, both specific for the sigma 54 system. These barriers would cooperate with the -11 inhibitory nucleotide to keep melting in check within sigma 54 holoenzyme closed complexes. The activator would then presumably be required to help initiate melting at -11 and to spread it to downstream positions, thereby exposing the start site and allowing transcription.

Control uses a multi-component switch involving -12 region DNA elements and multiple parts of sigma 54

As discussed above, sigma 54 itself contributes to at least four activities within the holoenzyme that are related to opening the DNA (Figure 7): a -12/-11 bottom fork junction recognition component, two downstream single-strand binding components, and a component that allows

Fig. 7. Hypothetical model for the role of the protein–DNA switch in sigma 54-dependent transcriptional control. Multiple DNA protein interactions are thought to direct specific recognition within the -12 promoter element. Prior to activation, a stable fork junction cannot be created there, due to the inhibition associated with unpairing the top -11 nucleotide. In addition, the single-strand DNA binding activities shown are masked. Thus, the melting switch remains off. We speculate that activators and ATP hydrolysis throw the switch by overcoming this inhibition to create the fork junction and then unmasking single-strand binding to allow the spread of melting. Certain mutations within the N-terminus of sigma or the -12 region of the DNA can destroy the switch; this leads to a partial loss of regulation via engagement of the single-strand binding regions in downstream locations. Sigma 54 contributes to the specificity of all four activities: fork junction recognition, -11 inhibition and the two single-strand binding activities.

inhibition via the top strand -11 nucleotide. Core polymerase enhances the strength of these activities (for examples see Figure 3). The data show that more than one region of sigma 54 is required (see Results). The data also show that the DNA sequence of the -12 region strongly influences how these activities may be used (see Results). Thus, the complex interactions involved in promoting the melted state involve multiple components of both protein and DNA. We consider these interactions to constitute a kind of molecular switch. Together they direct sigma 54 holoenzyme to the $-12/-11$ position where melting will originate. However, in the absence of activator, initiation of melting is unfavorable due to the -11 nucleotide inhibition and the barriers to the spread of melting. We speculate that the activator ‘throws the switch’, in a manner yet to be determined, and overcomes these collective impediments to melting.

The N-terminus of sigma 54 appears to play a central role in this regulatory switch. The N-terminus contains motifs including a leucine patch (Wang *et al.*, 1995; Syed and Gralla, 1997) which are required for regulated melting *in vivo* and *in vitro* (Sasse-Dwight and Gralla, 1990; Wong *et al.*, 1994; J.T.Wang *et al.*, 1997). The N-terminus also contains sequences proposed to mediate the effect of activator (Syed and Gralla, 1998). It has been implicated in conformational changes that allow heteroduplex templates to be bound in heparin-resistant complexes (Cannon *et al.*, 1999) and in exposure of single-strand binding determinants (Figure 6, top panels). Thus, the N-terminus is required to mediate activation by DNA melting.

One role of the N-terminus appears to be to help holoenzyme to select the $-12/-11$ fork junction location. When the N-terminus is deleted the $-12/-11$ fork junction is not recognized (Figure 4). Instead, heparin-sensitive complexes form in downstream locations near -8 . The data show that this downstream location is not associated with an inhibitory component and is not separated from single-strand binding components (Figure 6, and model in Figure 7). This downstream binding would then be free to promote unregulated melting; indeed, unregulated heparin-sensitive transcription is seen with this same mutant (J.T.Wang *et al.*, 1997). Moreover, these effects of the sigma mutation are mimicked by DNA mutations in the -12 promoter element, which also re-directs binding

and leads to unregulated transcription (Figure 5; Wang and Gralla, 1998). Thus, the N-terminus plays a critical role in this regulatory switch, in part by restricting recognition to the correct, regulated fork junction.

A variety of data from the literature are consistent with formation of this regulatory protein–DNA switch at $-12/-11$. When sigma 54 holoenzyme binds DNA prior to activated melting, the DNA is distorted precisely at the optimal $-12/-11$ fork junction identified here (Morris *et al.*, 1994). If the integrity of the N-terminus is impaired this distortion disappears and is replaced by a weaker distortion in downstream sequences (Morris *et al.*, 1994). This is mimicked perfectly by the changed pattern of junction binding using wild-type and mutant enzymes; mutation causes a disappearance of the $-12/-11$ preference and an appearance of a downstream preference. This mutation also causes some unregulated melting and transcription (J.T.Wang *et al.*, 1995, 1997; Syed and Gralla, 1997; Wang and Gralla, 1998), as do -12 region changes which also give rise to identical changes in patterns (see Results).

These considerations suggest that the N-terminus may act as an organizing center that brings different protein: DNA components together near position -12 to create a molecular switch that controls melting. The N-terminus of sigma 54 appears to interact with other regions of sigma (Cannon *et al.*, 1995; Casaz and Buck, 1999) probably including the major C-terminal DNA binding region (Wong *et al.*, 1994; Taylor *et al.*, 1996; Guo and Gralla, 1997). The model requires conformational changes in the -12 region of the DNA (Sasse-Dwight and Gralla, 1988; X.Wang *et al.*, 1997) and in the protein as part of organizing a fully regulated open complex. The N-terminus is known to change its conformation upon activation (Casaz and Buck, 1997). This overall view is consistent with prior studies that showed that melting created artificially on heteroduplex templates is not sufficient for activation (Wedel and Kustu, 1995); such templates would still need to throw the switch to become fully active. As discussed above, when either the protein or the DNA components (Wang and Gralla, 1998) of the switch are destroyed, other unregulated promoter locations may be recognized. This downstream binding would bypass the

inhibitory features associated with use of the -12/-11 switch, leading to unregulated transcription.

The architecture of sigma 54 enhances diversity in transcription regulation

The addition of the sigma 54 mechanism to the repertoire of bacterial controls increases the diversity of steps at which regulation can occur (Bowman and Kranz, 1998). Although bacteria contain numerous sigma factors, all but sigma 54 belong to the same sequence family (the sigma 70 family) and no fundamental mechanistic differences have been detected within members of that family (Busby and Ebright, 1994; Hochschild and Dove, 1998). Sigma 54 adds an additional type of activation to the repertoire of available mechanisms. This is activation via directed melting within pre-assembled stable closed complexes, providing a stable target for looping enhancer proteins (Ninfa *et al.*, 1987; Sasse-Dwight and Gralla, 1988, 1990; Morett and Buck, 1989; Popham *et al.*, 1989). Sigma 54 mechanisms can also include sigma 70-type features, including activation from near sites (Collado-Vides *et al.*, 1991), stabilization of unstable closed complexes (Bertoni *et al.*, 1998), and co-activation by DNA bending proteins and by DNA bending (Carmona *et al.*, 1997; Bertoni *et al.*, 1998; Dworkin *et al.*, 1998). Thus, the architecture of sigma 54 can be seen as an evolutionary design that increases the diversity of potential transcription regulatory mechanisms.

The current data argue that a key aspect of sigma 54 function is a multi-component DNA-protein switch that allows melting to be turned on or off. This switch involves the formation and use of the DNA fork junction that is created when melting begins. The spread of melting after fork junction creation is also different for sigma 54 and sigma 70 systems. This design can be seen a step along the path to the even greater diversity possible in mammalian systems, where enhancers, co-activators, DNA-bending and ATP-dependent melting are even more common than in sigma 54 transcription. Understanding how the domain structure of sigma 54 enhances this diversity should further understanding of how individual polypeptides in mammalian transcription complexes cooperate to allow diversity of regulation.

Materials and methods

Proteins and DNAs

Escherichia coli RNA polymerase core enzyme is a commercial product of Epicenter Technologies. *Escherichia coli* sigma 54 was purified as reported (J.T.Wang *et al.*, 1997). The ΔN mutant sigma 54, with residues 2–40 removed, has been described (J.T.Wang *et al.*, 1997). DNA probes were prepared as follows: the strand whose length was constant was labeled with [γ - 32 P]ATP and mixed with complementary strands of different length. The 40 μ l mixture, containing 4.0 pmol 32 P-end-labeled DNA and 6.0 pmol complementary strand in 20 mM Tris-HCl pH 7.5/80 mM NaCl, was boiled briefly and cooled slowly to room temperature. The resulting annealed probes were then diluted in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) containing 80 mM NaCl to the desired concentration. Proper annealing was monitored by electrophoresis as shown in Figure 1. RNA polymerase holoenzyme containing sigma 54 was constructed by incubating purified sigma 54 and core with a molar ratio of 2.5:1 on ice for 30 min.

EMSA

Mobility assays were performed as follows: holoenzyme or sigma 54 was added to a 10 μ l reaction mixture containing 1 $\times$ buffer T (50 mM

HEPES-HCl pH 7.9, 100 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM DTT, 0.05 μ g/ml BSA, 2.8% PEG-8000, 6.0 ng/ μ l dIdC) with annealed DNA probe. The amounts of enzymes and probes are indicated in the figure legends. After a 10 min incubation at 37°C, 1 μ l of 1 mg/ml heparin was added into the reaction mixture when needed. For sigma 54 alone, heparin was added before the probe as described previously (Guo and Gralla, 1998). The sample was then incubated at 37°C for 5 min longer and run on ice using 5% polyacrylamide gel electrophoresis (PAGE) with 1 $\times$ TBE buffer. Following electrophoresis, the radioactive bands were visualized and quantified by phosphorimager analysis.

Acknowledgement

This research was supported by a grant from the National Institutes of Health (GM35754).

References

- Austin, S. and Dixon, R. (1992) The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *EMBO J.*, **11**, 2219–2228.
- Bertoni, G., Fujita, N. and de Lorenzo, V. (1998) Active recruitment of σ^{54} -RNA polymerase to the Pu promoter of *Pseudomonas putida*: role of IHF and α CTD. *EMBO J.*, **17**, 5120–5128.
- Bowman, W.C. and Kranz, R.G. (1998) A bacterial ATP-dependent, enhancer binding protein that activates the housekeeping RNA polymerase. *Genes Dev.*, **12**, 1884–1893.
- Buck, M. and Cannon, W. (1992) Specific binding of the transcription factor σ^{54} to promoter DNA. *Nature*, **358**, 422–424.
- Busby, S. and Ebright, R.H. (1994) Promoter structure, promoter recognition and transcription activation in prokaryotes. *Cell*, **79**, 743–746.
- 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., Gallegos, M., Casaz, P. and Buck, M. (1999) Amino-terminal sequences of σ^N (σ^{54}) inhibit RNA polymerase isomerization. *Genes Dev.*, **13**, 357–370.
- Carmona, M., Clavierie-Martin, F. and Magasanik, B. (1997) DNA bending and the initiation of transcription at sigma54-dependent bacterial promoters. *Proc. Natl Acad. Sci. USA*, **94**, 9568–9572.
- Casaz, P. and Buck, M. (1997) Probing the assembly of transcription initiation complexes through changes in sigmaN protease sensitivity. *Proc. Natl Acad. Sci. USA*, **94**, 12145–12150.
- Casaz, P. and Buck, M. (1999) Region I modifies DNA-binding domain conformation of sigma 54 within the holoenzyme. *J. Mol. Biol.*, **285**, 507–514.
- Chen, Y.F. and Helmann, J.D. (1997) DNA-melting at the *Bacillus subtilis* flagellin promoter nucleates near -10 and expands unidirectionally. *J. Mol. Biol.*, **267**, 47–59.
- Collado-Vides, J., Magasanik, B. and Gralla, J.D. (1991) Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.*, **55**, 371–394.
- Dworkin, J., Ninfa, A.J. and Model, P. (1998) A protein-induced DNA bend increases the specificity of a prokaryotic enhancer-binding protein. *Genes Dev.*, **12**, 894–900.
- Gralla, J.D. (1996) Activation and repression of *E.coli* promoters. *Curr. Opin. Genet. Dev.*, **6**, 526–530.
- Gross, C.A., Lonetto, M. and Losick, R. (1992) Bacterial sigma factors. In McKnight, S.L. and Yamamoto, K.R. (eds) *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 129–176.
- Guo, Y. and Gralla, J.D. (1997) DNA-binding determinants of sigma 54 as deduced from libraries of mutations. *J. Bacteriol.*, **179**, 1239–1245.
- Guo, Y. and Gralla, J.D. (1998) Promoter opening via a DNA fork junction binding activity. *Proc. Natl Acad. Sci. USA*, **95**, 11655–11660.
- Hampsey, M. (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol. Mol. Biol. Rev.*, **62**, 465–503.
- Hochschild, A. and Dove, S.L. (1998) Protein-protein contacts that activate and repress prokaryotic transcription. *Cell*, **92**, 597–600.
- Kainz, M. and Roberts, J. (1992) Structure of transcription elongation complexes *in vivo*. *Science*, **255**, 838–841.
- Keener, J. and 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*, **85**, 4976–4980.
- Marr, M.T. and Roberts, J.W. (1997) Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, **276**, 1258–1260.
- Meier, T., Schickor, P., Wedel, A., Cellai, L. and Heumann, H. (1995) *In vitro* transcription close to the melting point of DNA: analysis of *Thermotoga maritima* RNA polymerase–promoter complexes at 75°C using chemical probes. *Nucleic Acids Res.*, **23**, 988–994.
- Merrick, M.J. (1993) In a class of its own—the RNA polymerase sigma factor σ^{54} (sigma N). *Mol. Microbiol.*, **10**, 903–909.
- Morett, E. and 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. and Buck, M. (1994) DNA distortion and nucleation of local DNA unwinding within sigma-54 (sigma N) holoenzyme closed promoter complexes. *J. Biol. Chem.*, **269**, 11563–11571.
- Ninfa, A.J. and Magasanik, B. (1986) Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **83**, 5909–5913.
- Ninfa, A.J., Reitzer, L.J. and Magasanik, B. (1987) Initiation of transcription at the bacterial glnAp2 promoter by purified *E. coli* components is facilitated by enhancers. *Cell*, **50**, 1039–1046.
- Orphanides, G., Lagrange, T. and Reinberg, D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.*, **10**, 2657–2683.
- 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.
- Reitzer, L.J. and Magasanik, B. (1986) Transcription of glnA in *E. coli* is stimulated by activator bound to sites far from the promoter. *Cell*, **45**, 785–792.
- Roberts, C.W. and Roberts, J.W. (1996) Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. *Cell*, **86**, 495–501.
- Sasse-Dwight, S. and 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. and Gralla, J.D. (1990) Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor sigma 54. *Cell*, **62**, 945–954.
- Spassky, A., Kirkegaard, K. and Buc, H. (1985) Changes in the DNA structure of the lac UV5 promoter during formation of an open complex with *Escherichia coli* RNA polymerase. *Biochemistry*, **24**, 2723–2731.
- 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.
- Syed, A. and Gralla, J.D. (1997) Isolation and properties of enhancer-bypass mutants of sigma 54. *Mol. Microbiol.*, **23**, 987–995.
- Syed, A. and Gralla, J.D. (1998) Identification of an N-terminal region of sigma 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 sigma factor sigma N plays a role in promoter recognition. *Mol. Microbiol.*, **22**, 1045–1054.
- Tintut, Y., Wong, C., Hsieh, M. and Gralla, J.D. (1994) RNA polymerase binding using a strongly acidic hydrophobic-repeat region of sigma 54. *Proc. Natl Acad. Sci. USA*, **91**, 2120–2124.
- 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 NH₂-terminal leucine patch in sigma 54. *Science*, **270**, 992–994.
- Wang, J.T., Syed, A. and 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. and Gralla, J.D. (1998) Multiple *in vivo* roles for the –12-region elements of sigma 54 promoters. *J. Bacteriol.*, **180**, 5626–5631.
- Wang, X., Kolb, A., Cannon, W. and Buck, M. (1997) Nucleoprotein complex formation by the enhancer binding protein nifA. *Nucleic Acids Res.*, **25**, 3478–3485.
- 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.
- 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 sigma 54 as determined by analysis of a set of deletion mutants. *J. Mol. Biol.*, **236**, 81–90.

Received March 29, 1999; revised and accepted May 18, 1999