HHS Public Access

Author manuscript

J Mol Biol. Author manuscript; available in PMC 2017 November 20.

Published in final edited form as:

J Mol Biol. 2016 November 20; 428(23): 4669–4685. doi:10.1016/j.jmb.2016.10.007.

Role of the σ^{54} Activator Interacting Domain in Bacterial Transcription Initiation

Alexander R. Siegel[a] and **David E. Wemmer**[a,b]

[a]Biophysics Graduate Group, University of California, Berkeley, CA 94720, USA

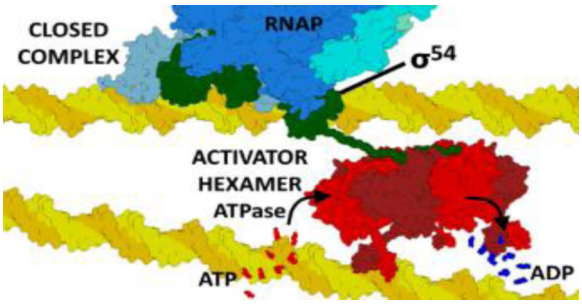
[b]Department of Chemistry, University of California, Berkeley, CA 94720, USA

Abstract

Bacterial sigma factors are subunits of RNA polymerase that direct the holoenzyme to specific sets of promoters in the genome, and are a central element of regulating transcription. Most polymerase holoenzymes open the promoter and initiate transcription rapidly after binding. However, polymerase containing members of the σ^{54} family must be acted on by a transcriptional activator before DNA opening and initiation occur. A key domain in these transcriptional activators forms a hexameric AAA+ ATPase that acts through conformational changes brought on by ATP hydrolysis. Contacts between the transcriptional activator and σ^{54} are primarily through an N-terminal σ^{54} activator interacting domain (AID). To better understand this mechanism of bacterial transcription initiation, we characterized the σ^{54} AID by NMR spectroscopy and other biophysical methods, and show that it is an intrinsically disordered domain in σ^{54} alone. We identified a minimal construct of the *A. aeolicus* σ^{54} AID that consists of two predicted helices and retains native-like binding affinity for the transcriptional activator NtrC1. Using the NtrC1 ATPase domain, bound with the non-hydrolyzable ATP analog ADP-beryllium fluoride, we studied the NtrC1-σ^{54} AID complex using NMR spectroscopy. We show that the σ^{54} AID becomes structured after associating with the core loops of the transcriptional activators in their ATP state and that the primary site of the interaction is the first predicted helix. Understanding this complex, formed as the first step toward initiation, will help unravel the mechanism of σ^{54} bacterial transcription initiation.

Graphical Abstract

Correspondence to David E. Wemmer: 508C Stanley Hall #3220, 1 Hearst Mining Circle, University of California, Berkeley, CA 94720-3220, dewemmer@berkeley.edu, Phone: (510) 666-2683.



Introduction

The five subunits of the bacterial “core” RNA polymerase, α, α, β, β’ and ω, are sufficient for transcribing mRNA once the promoter has been opened. However, in order to recognize promoter sequences, ‘melt’ the promoter DNA and initiate transcription, the core RNA polymerase requires an additional, modular subunit, the sigma factor [1]. The sigma factors bind to the core RNA polymerase, forming the RNA polymerase holoenzyme, and bind sequence specifically to DNA in the promoter region, with different sigma factors targeting different subsets of genes to accomplish differential transcriptional regulation [2]. For many sigma factors the regulation occurs by controlling the formation of the promoter-holoenzyme complex, either through anti-sigma proteins that compete with polymerase for a particular sigma factor [3], or through repressors that block the promoter [4]. Once the RNA polymerase holoenzyme-promoter complex forms, the sigma factors help open DNA and initiate transcription. After initiation the sigma factor can disassociate from the complex and the core RNA polymerase can continue to transcribe mRNA using the single stranded DNA template [5][6] (Figure 1).

Sigma factors fall into two broad families that share no sequence homology: the more common σ^{70} family and the rarer σ^{54} family [7][8]. All sigma factors serve the same purpose in directing RNA polymerase to specific promoters, but they differ in their mechanism of action and regulation. All sigma factors bind to core RNA polymerase to form a holoenzyme and all bind promoter regions slightly upstream from the transcription start site. σ^{70} RNA polymerase holoenzyme is capable of opening promoter DNA and initiating transcription immediately after binding the promoter [9]. However, σ^{54} polymerase requires an additional activation step, a conformational change that is driven by a transcriptional activator, before it can open the promoter [7]. The σ^{54}-RNAP holoenzyme recognizes conserved sequences −24 and −12 basepairs upstream of the transcription start site [10] where it binds and awaits activation by a transcriptional activator that assembles further upstream [11]. The transcriptional activators themselves must be triggered, often in response to an environmental stimulus [12], after which they act on the σ^{54}-RNAP holoenzyme, which then transcribes the DNA for the encoded protein initiating gene expression [13]. The additional activation requirement affords genes under control of σ^{54} an extra layer of control that both reduces background levels of transcription and gives a rapid cellular response when conditions are right. Consistent with this behavior, genes regulated by σ^{54} include those necessary for response to starvation and heat shock among others [14]. The detailed

mechanism by which these transcriptional activators reconfigure σ^{54} and the RNAP holoenzyme into a form capable of opening DNA is not known.

The σ^{54} transcriptional activators typically have three functional domains: (1) an N-terminal regulatory domain that receives a signal and promotes assembly of the active, hexameric form of the activator; (2) a central AAA+ ATPase domain that binds σ^{54} and hydrolyzes ATP; and (3) a C-terminal DNA binding domain that binds to enhancer sequences well upstream of the site of DNA melting [15]. Regulatory domains are quite diverse [16][17][18], responding to different kinds of signals including phosphorylation of a receiver domain [19], or ligand binding by a GAF domain [20]. Regulation can be positive, for example phosphorylation of the receiver domain promoting formation of the active hexamer ATPase, or negative where the domain inhibits formation of the hexamer until the signal is received. The central domain of negatively regulated activators may oligomerize into an active conformation when expressed without its regulatory and DNA binding domains, as is the case with the NtrC1 central domain construct ($NtrC1^{C}$) used in the experiments reported here [21].

Activated transcriptional activators assemble into hexameric rings with six ATP binding sites, each at the cleft between subunits [19]. A highly conserved loop at the top of the central pore, with the sequence GAFTGA, has been shown to be involved in the interaction between the activator and σ^{54} [22][23]. Crystal structures show that the GAFTGA loop extends upward on subunits bound to ATP (or a non-hydrolyzable ATP analog), but retracts inward for subunits bound to ADP [19][24][25].

σ^{54} has several functional domains, two of which had structures determined as individual domains [26][27] (Figure 2). The focus of the present work is the N-terminal ≈50 amino acids of σ^{54}, which are responsible for interacting with the assembled ATPase of the activator that we term the activator interacting domain (AID) and has also been called Region I [28]. This is followed by a variable length, low conservation linker, and then the core binding domain, which consists of two subdomains, a four-helix and three-helix bundle, that dock together [27]. The core binding domain is a primary region of interaction with core RNA polymerase subunits, making important contacts to core polymerase to form the holoenzyme. The next domain, which we term the −12 DNA binding domain, interacts with DNA in the −12 region of the promoter where DNA opening occurs [29]. While a consensus −12 DNA sequence has been identified, there is variability in this region of the promoter. DNA opening is initiated in the −12 region, and the part of σ^{54} in contact with it is likely responsible for helping to stabilize the opened ‘bubble’ that forms during transcription initiation [30][31][32]. The final, C-terminal segment of σ^{54} is a helix-turn-helix sequence-specific DNA binding domain [21]. This region binds to the strongly conserved −24 element of DNA called the RpoN box in σ^{54} driven promoters, making numerous sequence specific contacts [26]. This interaction also fixes σ^{54} in the correct position along the DNA such that the −12 binding domain can interact with the correct region of the promoter.

The mechanism by which the N-terminal activator interacting domain (AID) functions in the σ^{54} activation remains unclear. It has been shown that deletion of the first 50 residues of the N-terminal AID [33], or any mutation in the conserved GAFTGA loop of the transcriptional

activators [8], prevent binding of σ^{54} to the activator *in vitro* and activation is completely lost *in vivo*. However, most single amino acid substitutions and even deletions of small regions within the AID have little effect on activator binding and activity [34][35]. σ^{54} only binds the transcriptional activators with significant affinity when the activator subunits are primarily in the ATP state [22], which should be the default state in the cell. Activator binding is not detected when only ADP is present. Thus a primary role of the AID seems to be in forming the initial complex with the assembled activator ATPase domains.

Contact between σ^{54} and the activator alone is not enough to initiate transcription, ATP hydrolysis by the activator is required for the σ^{54}-RNAP holoenzyme bound to DNA to melt the DNA and transition from a closed to an open complex [8][36]. The extent of ATP hydrolysis required for opening has not been determined. Structural information about the initial encounter complex may shed light on how ATP hydrolysis by the activator ATPase domains bound to the AID is coupled to open complex formation through structural rearrangements of other domains of the σ^{54}-RNAP holoenzyme.

Attempts to study the structure and interactions of σ^{54} have been hindered by the poor solubility of σ^{54} constructs that include the AID. Crystallization of σ^{54} alone only occurred with the AID removed (S. Darst, personal communication). Recently full length σ^{54} was crystallized in the presence of RNA polymerase [37]. The structure is of modest resolution, and does not address how it interacts with the transcriptional activators. In the holoenzyme crystal structure, the AID appears to fold against RNAP as two helices near the −12 site of the DNA where the RNAP holoenzyme initiates DNA melting (Figure 3). It is unclear whether this is compatible with DNA bound polymerase, and how interaction with the activator would occur. Here we present evidence that the σ^{54} AID is an unstructured, intrinsically disordered domain in σ^{54} alone, with most of it becoming ordered when in complex with core polymerase. Studies in the presence of an activator ATPase domain show that a segment of the AID becomes immobilized in the initial activator complex, probably contacting the activator's GAFTGA loops.

Results

Disorder in the σ^{54} N-terminal Activator Interacting Domain

Our previous structural work was done using domains from the thermophile *Aquifex aeolicus* σ^{54} [27][38]. Several activators from this organism have also been characterized [19][20][39], which encouraged us to continue studies of it. However, a majority of genetic and biochemical work has used *Escherichia coli* (or closely related organisms), and so we have also made and studied similar constructs of the *E.c.* σ^{54}. We chose to express protein constructs that correspond to just the N-terminal activator interacting domain of σ^{54}, a central fragment of it, or including the full AID through neighboring domains. The longer constructs terminate with the four helix bundle of the core binding domain, or the full core binding domain, or correspond to full length σ^{54}. Amino acid sequence alignments (Figure 2) were used to design constructs covering the equivalent domains from the *A.a.* and *E.c.* versions of σ^{54}. Each protein was expressed in ^{15}N labeled growth medium, purified, and ^{1}H-^{15}N HSQC spectra were collected using an 800 MHz NMR spectrometer.

A.a. σ^{54}(1-135) includes the AID and the four-helix bundle of the core binding domain. HSQC spectra from this construct were compared with those from the previously studied *A.a.* σ^{54}(60-135), for which a structure was determined [27]. Nearly all of the amide peaks in the previously assigned σ^{54}(60-135) spectrum perfectly match the peaks in the σ^{54}(1-135) spectrum (Figure 4). Almost all of the additional peaks in the σ^{54}(1-135) spectrum have low dispersion of chemical shifts in the ^{1}H dimension. These peaks correspond to the first 60 residues of σ^{54} and the low shift dispersion is characteristic of an unfolded protein segment. The same comparison was done with the *E.c.* core binding domain construct σ^{54}(106-269) and the longer version containing the AID, linker, and CBD, *E.c.* σ^{54}(1-269). The results are the same, the σ^{54}(106-269) peaks are well-dispersed and nearly all of them overlay with a peak in the σ^{54}(1-269) spectrum. The extra peaks in the σ^{54}(1-269) construct, which must correspond to the first 105 residues of *E.c.* σ^{54}, are poorly dispersed and in the region of the spectrum that corresponds to unfolded residues. For both the *A.a.* and *E.c.* versions there are only very minor changes in the chemical shifts of CBD peaks in the presence of the AID and linker domain, which indicates that there is no significant contact between the CBD and the AID.

Other constructs of *E.c.* σ^{54} were also studied, including the full length protein (residues 1-477), the AID-linker-4 helix bundle σ^{54}(1-186), the AID alone σ^{54}(1-62), and a segment of the AID alone σ^{54}(11-48). In all of these the pattern of poorly dispersed peaks associated with the AID, with ^{1}H shifts between 8 and 9 ppm, occurs (Figure S1). The consistently low dispersion of AID peaks and the lack of change in the chemical shifts of well-folded residues in other σ^{54} domains in the presence of the AID show that the AID in free σ^{54} is unstructured, and does not interact significantly with other domains.

The Activator Interacting Domain interacts with core RNA Polymerase but only when part of full length σ^{54}

To determine whether the activator interacting domain makes contacts in the context of the full RNA polymerase (RNAP), we collected ^{1}H-^{15}N HSQC spectra of *E.c.* σ^{54}(1-269), which includes the full AID and core binding domain, in the presence of increasing concentrations of *E.c.* RNA polymerase. In the absence of RNAP, the chemical shifts of peaks from the folded CBD align well with those of the CBD alone, with the peaks showing low shift dispersion contributed by the AID (Figure 4). With increasing RNAP concentrations, the peaks from the folded CBD decrease, and vanish at a 1:1 ratio of RNAP:σ^{54} (Figure S2). This is expected because the 400 kDa polymerase tumbles slowly, giving rise to severe line broadening of immobilized residues. However, the low dispersion peaks associated with the AID remain relatively sharp even with excess RNAP present. These peaks remain poorly dispersed making the number of residues with remaining resonances difficult to determine, but the number roughly matches with the number from the AID alone σ^{54}(1-62).

To explore interactions between the AID and RNAP that require the presence of the remaining σ^{54} domains, as suggested by the crystal structure [37], we performed similar experiments with full length ^{15}N-labeled *E.c.* σ^{54} and RNA polymerase. With excess core RNAP present most of the peaks corresponding to the AID and linker are broadened (Figure

5). The ≈10 remaining peaks, many of which are broadened by the presence of the activator NtrC1^{C} and ADP-BeF_3^-, likely correspond to the N-terminal residues before the start of, and possibly including, the N-terminal helix of the AID. The presence of the CBD alone is not sufficient to promote AID binding to RNAP, consistent with the AID contacts with the −12 DNA binding domain in the structure of the holoenzyme (Figure 3) [37].

To test whether the AID-RNAP interaction could occur in trans, NMR spectra of a partial AID construct, *E.c.* σ^{54}(11-48), were collected in the presence of polymerase and a σ^{54} N-terminal deletion that removed the AID. Alone, peaks from ^{15}N σ^{54}(11-48) show the expected low dispersion in the hydrogen dimension of the ^{1}H-^{15}N HSQC spectrum (Figure S3). The peaks neither shift nor broaden in the presence of excess σ^{54}(106-477), showing there is no interaction between the AID and the folded domains of σ^{54} in the absence of core RNA polymerase. When excess RNA polymerase is added, there are only very minor perturbations to the chemical shifts and slight line-broadening. The slight broadening in this spectrum affects all peaks, and can be attributed to the increased viscosity of the samples containing high concentrations of RNA polymerase rather than any interaction between AID and the high molecular weight RNA polymerase holoenzyme. The ordering of the AID relative to the holoenzyme is greatly diminished when it is present as a separate molecule, indicating a weak interaction that depends on the AID being tethered near the holoenzyme by the linker region for interaction. This is also consistent with the relatively poor packing interface between the AID and −12 binding domain in the crystal structure of the holoenzyme [36].

A Segment of the Activator Interacting Domain Drives Complex Formation with the activator ATPase domain NtrC1^{C} in its ATP state

To study the interaction between the activator interacting domain and the transcriptional activators, we analyzed ^{1}H-^{15}N HSQC spectra of ^{15}N-labeled *A.a.* σ^{54} constructs containing the AID in the presence of unlabeled central ATPase domain from the transcriptional activator NtrC1. The activator was maintained in its ATP state by the addition of the non-hydrolyzable ATP analog ADP-BeF_3^-, which causes NtrC1^{C} to assemble into heptamers with uniform ATP-like sites [22] that mimic the binding behavior of full length hexamers. Alternately, we used E239A NtrC1^{C}, which forms an analogous, uniform ATP state heptamer. This variant binds to, but does not hydrolyze, ATP and its biochemical properties and structure resemble WT NtrC1^{C} with ADP-BeF_3^- [24]. Titrating either heptameric NtrC1^{C} into a solution of the ^{15}N labeled AID caused changes in the chemical shifts and linewidths of some AID peaks. Broadening arises from slowed tumbling of the small AID when bound to the higher molecular weight (≈210 kDa) NtrC1^{C} heptamer. With *A.a.* AID(1-62), many peaks in the σ^{54} construct remain sharp and unshifted even at a ratio of 1 AID per heptamer ring. This indicates that many residues in the AID do not contact NtrC1^{C}, and hence remain flexible (giving sharp resonances) even in the complex. To identify the key residues required to form the σ^{54}-activator encounter complex, we made constructs reducing the number of residues from the AID, while verifying that these AID constructs still bound to the activator ATPase oligomers.

Minimal AID construct

We considered available data to try to identify what region of the σ^{54} activator interacting domain might be responsible for specific binding to activators. Although mutagenesis failed to identify individual residues critical for binding, changes that have a moderate effect occur mostly in the range of residues 20–40. Furthermore, this region contains two predicted helices [40][41]. We therefore prepared a construct of *A.a.* σ^{54} that begins with the first predicted α helix and ends after the second one, comprising residues 16-41, σ^{54}(16-41). As shown below, this construct still binds well to the activator. The ^{1}H-^{15}N HSQC peaks for this construct cluster in the same region of the spectrum as the longer versions, consistent with a lack of folded structure as in the full AID (Figure S4).

For the peptide alone the ^{1}H-^{15}N HSQC peaks from all 25 of the residues are sharp. We assigned amide peaks in σ^{54}(16-41) (Figure 6) by using sequential connectivities in 3D ^{15}N-resolved [^{1}H-^{1}H]-NOESY spectra together with residue type identifications from intra-residue peaks. Neither patterns of NOEs or chemical shifts indicate any regions of substantial secondary structure for the peptide alone in spite of the two predicted helices.

Characterization of the σ^{54}-NtrC1^{C} Binding

We cloned a version of the reduced AID segment 16-41, adding an N-terminal tryptophan followed by a cysteine before the start of the σ^{54} sequence at residue 16. The cysteine was derivatized with the fluorescent marker Alexa Fluor 488 by forming a linkage between the cysteine and a maleimide on the dye. Binding was assayed by adding increasing concentrations of NtrC1^{C} (with ADP or ATP analog ADP-BeF_3^- at 500 μM) to 0.16 μM Alexa488-σ^{54}(16-41). Each sample was run on a native gel, and visualized using a Typhoon at 488nm. In the gel, when ADP-BeF_3^- is present, the NtrC1^{C} oligomer band fluoresces at 488 nm showing that the Alexa488 labeled σ^{54}(16-41) bound to NtrC1^{C} as the complex traveled through the gel (Figure S5). Thus a complex between NtrC1^{C} and σ^{54} forms, even with only 25 residues from the predicted helices in the activator interacting domain. However, when only ADP is present, the NtrC1^{C} oligomer has little emission at 488nm, indicating binding in the ADP state is substantially reduced relative to the ATP state, as expected from previous work [22][24].

To assess the binding affinity of our reduced AID construct, we measured the changes in the fluorescence anisotropy of Alexa488 labeled σ^{54}(16-41) upon adding NtrC1^{C} in different nucleotide states. Binding of the small, fluorescently labeled σ^{54}(16-41) peptide to the high molecular weight NtrC1^{C} greatly lengthens the rotational time of the peptide, resulting in a change in anisotropy in the dye fluorescence. Titrations were done by adding increasing concentrations of ADP-BeF_3^--NtrC1^{C}, ADP-NtrC1^{C} or just NtrC1^{C} to a fixed concentration of peptide, and following the anisotropy change. In the presence of ADP-BeF_3^-, σ^{54} AID(16-41) binds with a K_d of 1 μM (Figure 7A). In the absence of any nucleotide, or in the presence of ADP, we observed no change in the fluorescence anisotropy, and therefore no binding, even with a significant excess of NtrC1^{C} (Figure 7B). This confirms that the minimal AID construct σ^{54}(16-41) is sufficient to form the activator complex.

Fluorescence anisotropy experiments were also performed with full length σ^{54} carrying a K2C mutation and with an Alexa Fluor 488 dye attached at residue 2. The anisotropy changes during titrations with ADP-BeF_3^--$NtrC1^C$ indicate a similar Kd of 2 μM (Figure S6). The equivalence of the binding constants for σ^{54}(16-41) and full length σ^{54} to the 'ATP' form of $NtrC1^C$ shows that residues between 16 and 41 are responsible for binding to the activator.

NMR characterization of the AID-$NtrC1^C$ complex

Given the moderate binding affinity of the peptide for $NtrC1^C$ and the high molecular weight of the complex, we thought it might be possible to use exchange transferred NOEs within the peptide to characterize the bound state. However, titrating in sub-stoichiometric $NtrC1^C$ with ADP-BeF_3^- produced no new NOE peaks even at long mixing times, indicating slow exchange between the free AID and the AID-$NtrC1^C$ encounter complex. We also examined HSQC spectra collected with varying ratios of peptide and activator. Addition of $NtrC1^C$ alone had no effect on the peptide spectrum (consistent with fluorescence data), but the addition of ADP-BeF_3^- together with $NtrC1^C$ (not shown) or ATP with $NtrC1^C$(E239A) (Figure 8), reduced the intensity of many peaks. Most peaks from the ^{15}N-labeled σ^{54}(16-41) are so broad that they cannot be observed when excess ATP and $NtrC1^C$(E239A) are present, but amides for residues E38, E39, V40, and L41, remain sharp even in the complex and amides for K34, L35, I36 and H37 appear but are very weak. The peaks with the least broadening are from the C-terminal part of the σ^{54}(16-41) AID construct, indicating that the N-terminus and the first predicted helix are likely the main contributors to binding affinity.

To try to directly characterize the high molecular weight activator-σ^{54} AID complex we made ^{2}H/^{15}N σ^{54}(16-41) with specific ^{1}H-^{13}C methyl labels by expressing it in ^2H_2O/$^{15}NH_4Cl$ in the presence of labeled metabolic precursor α-ketoisovalerate to label the protein only at the δ-methyls of leucine and γ-methyls of valine [42]. The labeling approach produces a deuterated σ^{54}(16-41) peptide with nine leucines and one valine each labeled with ^{1}H and ^{13}C at the δ_1 or δ_2 methyls of leucine and γ_1 or γ_2 methyls of valine. We then ran ^{1}H-^{13}C methyl-TROSY HMQC experiments [43] to detect signals from the labeled residues in the complex.

For the peptide alone the leucine δ methyl peaks were poorly dispersed and all located close to the values predicted for random coil chemical shifts (Figure 9), as might be expected given the ^{1}H-^{15}N HSQC spectrum. The single valine is easily identified from its γ_1 and γ_2 peaks that appear with equal intensities and with nearly equal ^{1}H chemical shift but different ^{13}C chemical shifts. The nine sets of leucine peaks overlap significantly and cannot clearly be distinguished, though the nine δ_1 peaks are distinct from the nine δ_2 peaks, and are consistent with chemical shifts expected for a leucine in a random coil. Though individual leucine δ_1 and δ_2 peaks cannot be identified, integration of the region of overlapping peaks suggests that they do correspond to 9 times as many methyls as a single valine peak, which is expected from the ratio of nine leucines and one valine in the σ^{54}(16-41) sequence.

Addition of NtrC1^{C} alone had no effect on the chemical shifts of the leucines and valines in the labeled σ^{54}(16-41) minimal AID peptide, again indicating no or very weak interaction between the two in the absence of ATP (Figure 9). In the presence of the ATP analog ADP-BeF_3^- and ^{2}H-NtrC1^{C}, the 2D ^{1}H-^{13}C methyl-TROSY HMQC spectrum changes, reflected also in the 1D projections. Many peaks in the leucine δ_1 and δ_2 regions are broadened, but three remain relatively sharp along with the γ_1 and γ_2 peaks from the single valine. The integration of the leucine δ_1 and δ_2 regions is consistent with the broadening (loss of signal) of six leucines. The three relatively non-overlapping leucine peaks show only small chemical shift changes. These observations are consistent with the ^{1}H-^{15}N HSQC data indicating that NtrC1^{C} binding is localized to the N-terminus of the AID peptide, which has 6 leucines, and that the C-terminal tail of the AID, which has 3 leucines and 1 valine, remains free in the high molecular weight complex. Surprisingly, the six N-terminal leucines bound to the ≈210 kDa NtrC1^{C} complex could not be detected even using TROSY experiments designed to minimize the broadening of high molecular weight species. This suggests that in spite of the relatively tight binding there may be conformational exchange in the complex such that contact residues are broadened.

Discussion

Previous structural, biochemical, and genetic studies of σ^{54} have identified functional regions in the protein that play different roles in transcription: interaction with activators; binding to core polymerase; DNA interactions in the −12 region of the promoter where opening occurs; and sequence specific recognition of the −24 region of the promoter. The latter three functions are carried out by folded domains of σ^{54}; structures of the core binding domain [27] and −24 recognition domain [26] have been solved as independent domains. DNA opening is primarily controlled by the −12 DNA binding domain of σ^{54} in the RNAP holoenzyme [29]. The activation event that brings about these conformational changes is initiated through the first 50 N-terminal residues of σ^{54}, the activator interacting domain (AID), and a transcriptional activator. The AID has relatively low sequence conservation, and the linker region between the AID and core binding domain is quite variable in length. Neither genetic studies nor a recent crystal structure of the polymerase-σ^{54} complex [37] has provided insight into how the AID functions. Here we further characterize the nature of the AID and its interaction with core polymerase and the activator ATPase domain. These studies of the σ^{54} AID are a step toward understanding the changes in the σ^{54}-polymerase holoenzyme that lead to transcription initiation.

The Activator Interacting Domain is Intrinsically Disordered in σ^{54} Alone

The ^{1}H-^{15}N HSQC spectra from all *A.a.* and *E.c.* σ^{54} constructs that include the N-terminal domain, ranging in size from full length down to a minimal construct of the AID, display many relatively sharp resonances with low chemical shift dispersion, particularly in the proton dimension. The low shift dispersion is a hallmark of an unfolded domain, the lack of secondary structure leading to a similar environment for the backbone amides [44]. Rapid local motion of the backbone of unfolded domains also reduces the linewidths from those expected for a folded domain of the same molecular weight [45]. Comparing spectra of σ^{54} constructs containing the AID with those lacking it shows that the residues in the AID and

linker domain of σ^{54} account for the vast majority of the poorly dispersed, unstructured residues. The peaks from the AID have the same chemical shifts when part of constructs with the σ^{54} CBD or the full length σ^{54} present. Upon addition of RNA polymerase to ^{15}N labeled σ^{54} the interaction with polymerase subunits leads to slower tumbling broadening almost all of the resonances beyond detection. This includes most resonances from the AID, indicating its interaction with core polymerase and/or other σ^{54} domains, consistent with the crystal structure reported. However, the AID does not bind the ΔN-σ^{54} holoenzyme when present *in trans* and therefore may only have a low affinity for its binding site. The crystal structure of holo-σ^{54} polymerase places the AID in a region close to the –12 region binding domain, which would appear to conflict with DNA interaction. If the first helix of the AID is displaced by DNA binding then a substantial part of the AID maybe again become disordered, and free to interact with an activator. It has been proposed that binding of the AID in the cleft of polymerase could lead to the AID blocking the DNA from accessing the active site of the RNA polymerase inhibiting transcription initiation [37]. However, deletion of this domain does not lead to constitutive activation, so it appears that more than displacement is needed for activation.

^{1}H-^{15}N HSQC spectra of ^{15}N-labeled *A.a.* σ^{54} constructs were also studied in the presence of the ATPase domain from *A.a.* activator NtrC1^{C}. In this case, the AID peaks are strongly broadened due to immobilization in the presence of NtrC1^{C} in its ATP state (mimicked by ADP-BeF_3^-), but they are unaffected by apo NtrC1^{C} or NtrC1^{C} in the ADP state where the activator-σ^{54} interaction is weak. The observed broadening indicates that the AID becomes immobilized in the complex with the oligomeric, high molecular weight transcriptional activator oligomer, but only when the activators are in their ATP state with the GAFTGA loops raised above the ring around the central pore [19][22].

An intrinsically unfolded domain becoming immobilized and partly structured upon encountering a specific binding partner is common behavior for such domains [46]. One common role for intrinsically unfolded domains is to facilitate assembly of functional complexes [47], for example to initiate transcription in eukaryotes. In that case, transcription factor activation domains serve to form tethers with other proteins required to assemble a polymerase complex competent for initiation [48][49]. For bacterial σ^{54}-polymerase, it is expected that the polymerase holoenzyme is normally bound at the promoter waiting for the action of the activator, which may already be bound to the enhancer region in its unactivated dimeric state. Upon receiving its signal, the activators assemble to the active hexamer [50], which should have primarily ATP bound since ATP is much more abundant than ADP in the cell. At this point, flexibility of the AID of σ^{54} would facilitate making the initial ‘encounter complex’ with the ATPase domain of the assembled activator, completing assembly of all the components needed for initiation of transcription upon subsequent ATP hydrolysis. At present it is not clear whether there are contacts between the activator ATPase domain and other parts of the polymerase holoenzyme, how much ATP hydrolysis is required for promoter opening, or the extent to which the complex changes structure when ATP hydrolysis occurs. Studies of the activator ATPase domains have shown that the GAFTGA loops that contact the AID in the initial complex are raised above the ring in their ATP bound state [22] but are down when ADP is bound [19]. A recent structure of the NtrC1^{C} ATPase crystallized in the presence of ADP-BeF_3^- gave a mixed nucleotide state for the

hexameric ring, with a ‘lock-washer’ spiral structure [25] reminiscent of the Rho-ATPase [51]. For Rho a functional model has been developed in which sequential hydrolysis drives loops down pushing RNA through the central pore of the ring through multiple rounds of ATP hydrolysis. Although the organization of the ATPase domains is different, there is some structural similarity between these systems.

A minimal Activator Interacting Domain is sufficient for formation of the encounter complex

Mutational studies of the N-terminal activator interacting domain of σ^{54} show that no single residue is critical for activator binding and function, rather changes in numerous residues over a segment of about 25 amino acids result in modest decreases in activity [52][53][35] [54]. Secondary structure predictions using PSIPRED indicate that there may be two conserved helices in σ^{54}, predicted consistently across multiple species [40]. Our experiments with σ^{54} AID constructs show that just having residues from these two predicted helices is sufficient for activator binding in the presence of an ATP analog, but the NMR data do not suggest significant helix formation in the peptide alone. Just as with full length σ^{54}, the formation of the encounter complex between the minimal σ^{54} AID construct and the transcriptional activators requires activator subunits in the ATP state in which GAFTGA loops are extended above the ATPase ring. The dependence of peptide binding on ATP, as well as the similar binding affinities of the peptide and the full length σ^{54} to the transcriptional activators, confirms that we are observing native-like interactions of this segment of the AID rather than a non-specific binding event.

Examining the sequences of the *E.c.* and *A.a.* AIDs shows that if they fold into the predicted α-helices these would be amphipathic. Regular repeats of hydrophobic residues (primarily leucine) and hydrophilic residues (primarily glutamine) cluster on opposite sides of the predicted α-helices (Figure 10), which would be connected by a short linker. This suggests that formation of the encounter complex between σ^{54} and the GAFTGA loops of the transcriptional activators could be driven by hydrophobic interactions between the leucine-rich side of the AID and the hydrophobic GAFTGA loops that form a surface around the pore in the ATPase ring.

σ^{54} AID to activator binding affinity and slow exchange

Full length σ^{54} and σ^{54}(16-41) bind $NtrC1^{C}$ in the ATP state with approximately the same dissociation constant of ≈ 2 μM, while neither binds to the activator with measurable affinity in the ADP or apo state. This indicates that residues between 16 and 41 are not only sufficient for binding, but that they account for essentially the full binding affinity of σ^{54}. Thus it is likely that this segment of σ^{54} contains the large majority, if not the entirety, of the $NtrC1^{C}$ binding surface during the formation of the initial encounter complex.

When the ^{15}N-labeled σ^{54} AID is present in excess of the ATP-state $NtrC1^{C}$ rings, there are no discernible chemical shift changes in the unbound AID peaks. In these samples, σ^{54} AID is present in two states: the bound encounter complex and the unbound free form. The lack of chemical shift changes or broadening during the titrations shows that the AID is in slow exchange on the chemical shift timescale. This is somewhat surprising given the moderate

binding affinity of the peptide, if binding were even near the diffusion limit then the predicted dissociation rate would be sufficient to give intermediate to fast exchange behavior. To be consistent with the observed dissociation constant and the slow exchange behavior both the rate of binding and the rate of dissociation must be quite low. In the presence of excess $NtrC1^{C}$-ADP-BeF_3^- (above one ATPase ring per peptide) all but a few C-terminal peaks of the AID broaden and disappear, indicating complete binding of the AID.

Characterization of the σ^{54}-$NtrC1^{C}$ Encounter Complex

The minimal AID construct σ^{54}(16-41) yielded NMR spectra with little overlap of amide residues and uniformly sharp peaks, which enabled doing peak assignments using conventional sequential methods [56]. In the NOESY spectra we observed no crosspeaks between i and i+3 residues as would be expected in an α-helical structure, in which the α-proton of a residue is in proximity of the amide of the residue 3 amino acids later in the sequence. This indicates that the free AID peptide is not significantly populating folded α-helices by itself in solution, but this does not rule out the formation of well-folded α-helix in the encounter complex with the ATPase.

In the presence of $NtrC1^{C}$ in the ATP state, most of the σ^{54}(16-41) peaks were broadened upon binding to the high molecular weight $NtrC1^{C}$ oligomers. However, a few peaks remained sharp enough for detection and underwent slight chemical shift changes. The relatively sharp resonances indicate that these residues remain quite flexible even when the AID is in the encounter complex. Our assignments (Figure 6) show that these residues arise from E38, E39, V40, and L41 at the end of the second predicted helix. Further very weak (broadened) resonances are consistent with chemical shifts of K34, L35, I36 and H37. To try to directly observe the bound state, methyl-TROSY experiments were done on uniformly deuterated σ^{54}(16-41) labeled with a ^{13}C-$^{1}H_3$ methyl at the δ carbon on the leucine and at the γ carbon of the valine side chains. In the free peptide these peaks exhibited low chemical shift dispersion as expected, consistent with other evidence that the AID is unstructured. Integration of the peaks shows that all nine leucines and one valine were labeled as expected. When AID binds to the transcriptional activators there are six leucine peaks that broaden and disappear from the methyl-TROSY spectra, while three leucine peaks and one valine remain sharp. There are six leucines in the first predicted helix and three leucines and one valine in the second predicted helix. One of the remaining leucines has a slightly different shift (unaffected by complex formation) and is very likely the C-terminal residue. The other two leucines that are observed, together with the only valine, which neighbors the C-terminal leucine, remain almost as sharp as in the free peptide. The combination of ^{15}N and ^{13}C observations show a gradient in mobility of the bound peptide residues, with flexibility starting around residue 33. The backbone appears to be affected more than the methyl-containing sidechains, but this might reflect sidechain dynamics.

In methyl-TROSY experiments, the methyl ^{1}H-^{13}C correlations are not broadened as much by the slow tumbling times of the high molecular weight activator-σ^{54} AID complex as those from other ^{1}H-^{13}C or amide ^{1}H-^{15}N pairs. Nevertheless, we still failed to detect even the methyl peaks corresponding to leucines in the N-terminal half of the AID in the ATPase encounter complex. This obviously prevents further analysis of the induced structure in the

AID. It must also mean that the AID in the encounter complex is not bound to the transcriptional activator in a single, well-defined configuration, but instead is bound in multiple different conformations experiencing different chemical environments. The lack of a single, well-defined induced structure in the AID encounter complex is consistent with an activation mechanism that involves changes in the AID-ATPase complex rather than serving a role simply as a defined docking site. Given that many AAA+ ATPases thread substrates through a central pore to act on them [57], a dynamic activation mechanism in which the AID is threaded through the central pore of the transcriptional activator by successive rounds of ATP hydrolysis and motion of the GAFTGA loops seems consistent with the properties of this initial encounter complex.

Future Directions

Studies of the interaction of the intrinsically disordered activator interacting domain bound with $NtrC1^{C}$ provide some hints to the way transcriptional activators drive σ^{54} transcription initiation. Further structural studies may be possible, if conditions for forming a better-defined complex can be identified, either using NMR techniques or crystallography. A high resolution structure of the complex could provide a more complete picture of the specific interactions between σ^{54} and the activators and may give some further evidence about the activation mechanism. However, even a more complete understanding of the structure of the static encounter complex may not be enough to identify the precise mechanism of activation. Experiments that probe the dynamics of the AAA+ ATPase transcriptional activators and σ^{54} as they go through multiple ATP hydrolysis cycles, as well as *in vivo* experiments to test various changes to σ^{54} that disrupt activation, could be valuable techniques to study the dynamics of the mechanism of σ^{54} transcription initiation.

Materials and Methods

Protein expression and purification

Aquifex aeolicus and *Escherichia coli* σ^{54} constructs were cloned from full length σ^{54} plasmids and placed into pET28a expression vectors. *E.c.* σ^{54}(1-477), *E.c.* σ^{54}(1-269), *E.c.* σ^{54}(1-186), *E.c.* σ^{54}(1-62), and *A.a.* σ^{54}(1-135) all contained a C-terminal His_6 tag. *E. coli* Rosetta cells with the plasmid were grown at 37°C in 1L of isotope-labeled M9 minimal media to an optical density at 600 nm of 0.6 and induced with 0.5 mM isopropyl thiogalactopyranoside then harvested after 8–12 hours. Extracts for *A.a.* constructs were heated at 75°C for 20 minutes, and both *E.c.* and *A.a.* constructs were purified on a NiNTA column. Some constructs went to inclusion bodies and were purified on a NiNTA column under denaturing conditions in 8M urea, then refolded on the column by changing to buffer without urea.

E.c. σ^{54}(11-48), *A.a.* σ^{54}(10-47), *A.a.* σ^{54}(16-41), and *A.a.* σ^{54}(WC-16-41) were placed into a pET His6 MBP TEV LIC cloning vector (1M) obtained from the UC Berkeley MacroLab. His-MBP-AID was expressed by growing *E. coli* Rosetta cells at 37°C in 1L of isotope-labeled M9 minimal media to an optical density of 0.6 and induced with 0.5 mM isopropyl thiogalactopyranoside then harvested after 8–12 hours. Leucine, valine labeled peptide was obtained by growth on 1L $^{2}H_2O$ M9 media with deuterated glucose and the addition of 100

mg of the Leu-Val precursor α-ketoisovalerate added 1 hour before induction. This produces u-[^{2}H,^{12}C], δ1,δ2-[^{1}H,^{13}C] resulting in NMR active nuclei on the δ carbon of the leucine side chains and on the γ carbon of the valine side chains. The expressed protein was found in inclusion bodies after sonicating and pelleting the lysates. Proteins in the pelleted inclusion bodies were unfolded in Denaturing Wash Buffer (8 M urea, 20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4), residual insoluble protein was pelleted and then soluble supernatant was passed through a 0.2 μm filter before loading onto a NiNTA column. Samples were washed and refolded on the column by passing through 5 column volumes of Ni Wash Buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4), then eluted with Ni Elution buffer (20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4). His-MBP-AID was dialyzed into 20 mM Tris, 1 mM DTT, 0.1 mM EDTA, pH 7.4 buffer overnight, transferred to a falcon tube with the addition of 20% glycerol. This protein was cut with TEV protease and passed through 30k and 10k MWCO Amicon Ultra centrifugal filters, after which the majority of cleaved AID peptide (3.5–4.5 kDa) remained in the flow through. Peptide was reconcentrated into NMR buffer (20 mM Tris, 200 mM KCl, pH 7.0) in a 3 kDa MWCO Amicon Ultra centrifugal filter.

NtrC1^{C}(121–387) was cloned into pET28a vectors with an N-terminal, TEV protease cleavable His tag. NtrC1^{C}(E239A) was prepared using QuikChange site directed mutagenesis on the NtrC1^{C} plasmid. Protein was expressed in either Luria-Bertani media or $^{2}H_2O$ M9 minimal media, grown to an optical density at 600 nm of 0.6, induced with 0.5 mM isopropyl thiogalactopyranoside. Pelleted cells were sonicated and heated at 75°C for 20 minutes, which precipitated most of the proteins other than the thermophilic *A.a.* NtrC1^{C}. Protein was purified on an Ni-NTA column and dialyzed overnight back into Ni Wash Buffer with TEV protease. Cut NtrC1^{C} was passed through an NiNTA column to separate it from TEV and the flow through was collected and dialyzed into NMR Buffer (20 mM Tris, 200 mM KCl, 5% glycerol pH 7.0).

The RNAP(ααββ'ω) expression plasmid was a gift from the Buck Lab [58]. RNAP protein complex was expressed in cells grown on Luria-Bertani media at 37°C to an optical density of 0.6 at 600 nm. Cells were cold shocked on ice for 30 minutes before induction with 0.5 mM isopropyl thiogalactopyranoside and then grown for 6–8 hours at 25°C. The RNAP β subunit contains a C-terminal His tag, and all subunits of the complex were eluted together on an NiNTA column, then dialyzed into low salt buffer and further purified using a heparin column eluted with an NaCl gradient.

Preparation of the ATP analog ADP-BeF_3^-

ADP-BeF_3^- analog was prepared with a 1:1:8:1 ratio of ADP:$BeCl_2$:NaF:$MgCl_2$ by thawing 0.1 M ADP to room temperature and mixing it with NMR buffer to a final concentration of 20 mM. $BeCl_2$ was added to a final concentration of 20 mM causing a precipitate to be observed. The precipitate disappeared when NaF was added to a final concentration of 160 mM. Finally, slight excess of $MgCl_2$ was added to achieve a final concentration of 25 mM and precipitate was observed again. The solution was passed through a 0.2 μm filter to remove remaining precipitates.

NMR spectroscopy of σ^{54} and RNA Polymerase

Samples were obtained by mixing excess RNAP with ^{15}N-labeled σ^{54} constructs at low concentration and concentrated to 25 μM as a complex with a 10k MWCO Amicon Ultra centrifugal filter. NMR data were collected at 298 K on Bruker Avance 800 MHz or 600 MHz spectrometers. Chemical shifts were referenced to that of water in order to properly align them across different experiments and chemical shift changes were observed with a ^{1}H-^{15}N HSQC. Data were processed with NMRPipe [59] and chemical shift analysis was undertaken with CARA [60] or MestReNova [61].

NMR spectroscopy of σ^{54}(AID) and NtrC1^{C}

Samples were prepared by mixing ^{15}N-labeled, and later deuterated with ^{1}H-^{13}C labeled Leu δ and Val γ, σ^{54} AID constructs with either unlabeled or ^{2}H NtrC1^{C} and premixed ADP-BeF_3^-, ADP alone, or with the ATP-hydrolysis deficient mutant NtrC1^{C}(E239A) and ATP. ADP-BeF_3^- was used over NtrC1^{C}(E239A) and ATP for long NMR experiments to avoid spectral changes due to low background levels of ATP hydrolysis during the acquisition. NMR data were collected at 298 K on Bruker Avance 800 MHz or 600MHz spectrometers. Chemical shifts were referenced to that of water in order to properly align them across different experiments. Chemical shift changes of all σ^{54} AID amides were observed using ^{1}H-^{15}N HSQC experiments and chemical shift changes of Leu and Val side chains using ^{1}H-^{13}C methyl-TROSY experiments [43]. Amide assignments were carried out using 3D ^{15}N-NOESY-HSQC experiments to both identify the amino acid type of amides by their side chain shifts and connect amides to their neighbor H_α resonances with the sequential assignment approach [56]. Data were processed with NMRPipe [59] and assignment analysis used the programs CARA [60] or MestReNova [61].

Fluorescence Anisotropy and Native Gels of σ^{54} and NtrC1^{C}

A tryptophan and a cysteine residue were introduced in front of the start site of the σ^{54}(16-41) construct by QuikChange site directed mutagenesis. Expression and purification techniques were the same as σ^{54}(16-41). Purified peptide was incubated with excess Alexa Fluor 488 maleimide which attached the dye to the single cysteine residue. Reactions were kept in the dark and excess dye was removed by reconcentrating the peptide in a 3k MWCO Amicon Ultra centrifugal column. 0.16 μM concentration of Alexa488-σ^{54}(16-41) (or 5 μM of Alexa488-full length σ^{54} (K2C)) was mixed with varying concentrations of NtrC1^{C} and either ADP or the ATP analog ADP-BeF_3^-. Fluorescence anisotropy was measured on a DTX880 (Beckman Coulter) plate reader. Samples were also run on a native PAGE (4–15%) gel in running buffer that included ADP-BeF_3^- or ADP and visualized with a Typhoon (GE Life Sciences) to look for dye fluorescence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH grant GM 62163 (to DEW), and instrumentation grants NSF BBS 87-20134 for the 600 MHz NMR spectrometer, and NSF (BBS 01-19304) and NIH (RR15756) for the 800 MHz NMR

spectrometer and the Central California 900 MHz Facility (supported by NIH-GM68933). We also thank Seh-hyeon Jin and Zhijuan Gao for help with the *E. coli* σ^{54} core binding domain; Cristhian Canari for help with the studies of the σ^{54}-RNAP binding; Kwang Seo Kim for help with the fluorescence anisotropy measurements of full length σ^{54}; and Jeff Pelton for his help collecting NMR spectra.

References

1. Burgess RR, Travers AA, Dunn JJ, Bautz EKF. Factor Stimulating Transcription by RNA Polymerase. Nature. 1969; 221:43–46. [PubMed: 4882047]
2. Losick R, Pero J. Cascades of Sigma Factors. Cell. 1981; 25:582–584. [PubMed: 6793235]
3. Campbell EA, Westblade LF, Darst SA. Regulation of bacterial RNA polymerase σ factor activity: a structural perspective. Curr. Opin. Microbiol. 2008; 11:121–127. [PubMed: 18375176]
4. Colland F, Orsini G, Brody EN, Buc H, Kolb A. The bacteriophage T4 AsiA protein: A molecular switch for sigma 70-dependent promoters. Mol. Microbiol. 1998; 27:819–829. [PubMed: 9515707]
5. Liu C, Martin CT. Promoter clearance by T7 RNA polymerase. Initial bubble collapse and transcript dissociation monitored by base analog fluorescence. J. Biol. Chem. 2002; 277:2725–2731. [PubMed: 11694519]
6. Tahirov TH, Temiakov D, Anikin M, Patlan V, McAllister WT, Vassylyev DG, Yokoyama S. Structure of a T7 RNA polymerase elongation complex at 2.9Å resolution. Nature. 2002; 420:43–50. [PubMed: 12422209]
7. Buck M, Gallegos M-T, Studholme DJ, Guo Y, Gralla JD. MINIREVIEW The Bacterial Enhancer-Dependent σ54 (σN) Transcription Factor. J. Bacteriol. 2000; 182:4129–4136. [PubMed: 10894718]
8. Chaney MK, Grande R, Wigneshweraraj SR, Cannon W, Casaz P, Gallegos M-T, Schumacher J, Jones S, Elderkin S, Dago AE, Morett E, Buck M. Binding of transcriptional activators to sigma 54 in the presence of the transition state analog ADP-aluminum fluoride: insights into activator mechanochemical action. Genes Dev. 2001; 15:2282–2294. [PubMed: 11544185]
9. Campbell EA, Muzzin O, Chlenov M, Sun JL, Olson CA, Weinman O, Trester-Zedlitz ML, Darst SA. Structure of the bacterial RNA polymerase promoter specificity σ subunit. Mol. Cell. 2002; 9:527–539. [PubMed: 11931761]
10. Barrios H, Valderrama B, Morett E. Compilation and analysis of σ-dependent promoter sequences. Nucleic Acids Res. 1999; 27:4305–4313. [PubMed: 10536136]
11. Sasse-Dwight S, Gralla JD. Probing the Escherichia coli glnALG upstream activation mechanism in vivo. Proc. Natl Acad. Sci. USA. 1988; 85:8934–8938. [PubMed: 2904147]
12. Keener J, Kustu S. 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. 1988; 85:4976–4980. [PubMed: 2839825]
13. Kustu S, Santero E, Keener J, Popham D, Weiss D. Expression of σ54 (ntrA)-dependent genes is probably united by a common mechanism. Microbiol. Rev. 1989; 53:367–376. [PubMed: 2677638]
14. Kazmierczak MJ, Wiedmann M, Boor KJ. Alternative sigma factors and their roles in bacterial virulence. Microbiol. Mol. Biol. Rev. 2005; 69:527–543. [PubMed: 16339734]
15. Rappas M, Bose D, Zhang X. Bacterial enhancer-binding proteins: unlocking σ54-dependent gene transcription. Curr. Opin. Struct. Biol. 2007; 17:110–116. [PubMed: 17157497]
16. Nixon BT, Ronson CW, Ausubel FM. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. Proc. Natl. Acad. Sci. USA. 1986; 83:7850–7854. [PubMed: 3020561]
17. Ronson CW, Nixon BT, Ausubel FM. Conserved Domains in Bacterial Regulatory Proteins That Respond to Environmental Stimuli. Cell. 1988; 49:579–581.
18. Studholme DJ, Dixon R. Domain architectures of σ54-dependent transcriptional activators. J. Bacteriol. 2003; 185:1757–1767. [PubMed: 12618438]
19. Lee S-Y, De La Torre A, Yan D, Kustu S, Nixon BT, Wemmer DE. Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. Genes Dev. 2003; 17:2552–2563. [PubMed: 14561776]

20. Batchelor JD, Doucleff M, Lee C-J, Matsubara K, De Carlo S, Heideker J, Lamers MH, Pelton JG, Wemmer DE. Structure and Regulatory Mechanism of Aquifex aeolicus NtrC4: Variability and Evolution in Bacterial Transcriptional Regulation. J. Mol. Biol. 2008; 384:1058–1075. [PubMed: 18955063]
21. Batchelor JD, Sterling HJ, Hong E, Williams ER, Wemmer DE. Receiver Domains Control the Active-State Stoichiometry of Aquifex aeolicus σ54 Activator NtrC4, as Revealed by Electrospray Ionization Mass Spectrometry. J. Mol. Biol. 2009; 393:634–643. [PubMed: 19699748]
22. Chen B, Doucleff M, Wemmer DE, De Carlo S, Huang HH, Nogales E, Hoover TR, Kondrashkina E, Guo L, Nixon BT. ATP ground- and transition states of bacterial enhancer binding AAA plus ATPases support complex formation with their target protein, σ54. Structure. 2007; 15:429–440. [PubMed: 17437715]
23. Rappas M, Schumacher J, Niwa H, Buck M, Zhang X. Structural basis of the nucleotide driven conformational changes in the AAA+ domain of transcription activator PspF. J. Mol. Biol. 2006; 357:481–492. [PubMed: 16430918]
24. Chen B, Sysoeva TA, Chowdhury S, L Guo, De Carlo S, Hanson JA, Yang H, Nixon BT. Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase. Structure. 2010; 18:1420–1430. [PubMed: 21070941]
25. Sysoeva TA, Chowdhury S, Guo L, Nixon BT, Tracy Nixon B, Nixon BT. Nucleotide-induced asymmetry within ATPase activator ring drives σ54-RNAP interaction and ATP hydrolysis. Genes Dev. 2013; 27:2500–2511. [PubMed: 24240239]
26. Doucleff M, Pelton JG, Lee PS, Nixon BT, Wemmer DE. Structural basis of DNA recognition by the alternative sigma-factor, σ54. J. Mol. Biol. 2007; 369:1070–1078. [PubMed: 17481658]
27. Hong E, Doucleff M, Wemmer DE. Structure of the RNA polymerase core-binding domain of σ54 reveals a likely conformational fracture point. J. Mol. Biol. 2009; 390:70–82. [PubMed: 19426742]
28. Gallegos M-T, Cannon W, Buck M. Functions of the σ54 region I in Trans and implications for transcription activation. J. Biol. Chem. 1999; 274:25285–25290. [PubMed: 10464252]
29. Wong C, Gralla JD, Tintut Y. The Domain Structure of Sigma 54 as Determined by Analysis of a Set of Deletion Mutants. J. Mol. Biol. 1994; 236:81–90. [PubMed: 8107127]
30. Oguiza JA, Gallegos M-T, Chaney MK, Cannon W, Buck M. Involvement of the σN DNA-binding domain in open complex formation. Mol. Microbiol. 1999; 33:873–885. [PubMed: 10447895]
31. Buck M, Bose D, Burrows PC, Cannon W, Joly N, Pape T, Rappas M, Schumacher J, Wigneshweraraj SR, Zhang X. A second paradigm for gene activation in bacteria. Biochem. Soc. Trans. 2006; 34:1067–1071. [PubMed: 17073752]
32. Bose D, Pape T, Burrows PC, Rappas M, Wigneshweraraj SR, Buck M, Zhang X. Organization of an Activator-Bound RNA Polymerase Holoenzyme. Mol. Cell. 2008; 32:337–346. [PubMed: 18995832]
33. Wang JT, Syed A, Gralla JD. Multiple pathways to bypass the enhancer requirement of sigma 54 RNA polymerase: roles for DNA and protein determinants. Proc. Natl. Acad. Sci. USA. 1997; 94:9538–9543. [PubMed: 9275158]
34. Pitt M, Gallegos M-T, Buck M. Single amino acid substitution mutants of Klebsiella pneumoniae sigma(54) defective in transcription. Nucleic Acids Res. 2000; 28:4419–4427. [PubMed: 11071928]
35. Gallegos M-T, Buck M. Sequences in σ54 region I required for binding to early melted DNA and their involvement in sigma-DNA isomerisation. J. Mol. Biol. 2000; 297:849–859. [PubMed: 10736222]
36. Wedel AB, Kustu S. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. Genes Dev. 1995; 9:2042–2052. [PubMed: 7649482]
37. Yang Y, Darbari VC, Zhang N, Lu D, Glyde R, Wang Y, Winkelman JT, L Gourse R, Murakami KS, Buck M, Zhang X. Structures of the RNA polymerase-σ54 reveal new and conserved regulatory strategies. Science (80-.). 2015; 349:882–886.

38. Doucleff M, Malak LT, Pelton JG, Wemmer DE. The C-terminal RpoN domain of σ54 forms an unpredicted helix-turn-helix motif similar to domains of σ70. J. Biol. Chem. 2005; 280:41530–41536. [PubMed: 16210314]
39. Batchelor JD, Lee PS, Wang AC, Doucleff M, Wemmer DE. Structural Mechanism of GAF-Regulated σ54 Activators from Aquifex aeolicus. J. Mol. Biol. 2013; 425:156–170. [PubMed: 23123379]
40. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 2013; 41:349–357.
41. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. 1999; 292:195–202. [PubMed: 10493868]
42. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE. A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated 15N-, 13C-, 2H-labeled proteins. J. Biomol. NMR. 1999; 13:369–374. [PubMed: 10383198]
43. Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE. Cross-correlated relaxation enhanced 1H-13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. J. Am. Chem. Soc. 2003; 125:10420–10428. [PubMed: 12926967]
44. Wishart DS, Sykes BD, Richards FM. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. J. Mol. Biol. 1991; 222:311–333. [PubMed: 1960729]
45. Dyson HJ, Wright PE. Unfolded Proteins and Protein Folding Studied by NMR. Chem. Rev. 2004; 104:3607–3622. [PubMed: 15303830]
46. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic Disorder and Protein Function. Proteins. 2002; 41:6573–6582.
47. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. Nat. Publ. Gr. 2015; 16:18–29.
48. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic Disorder in Transcription Factors. 2006:6873–6888.
49. Tantos A, Han K-H, Tompa P. Intrinsic disorder in cell signaling and gene transcription. Mol. Cell. Endocrinol. 2012; 348:457–465. [PubMed: 21782886]
50. Porter SC, North AK, Wedel AB, Kustu S. Oligomerization of NTRC at the glnA enhancer is required for transcriptional activation. Genes Dev. 1993; 7:2258–2273. [PubMed: 7901122]
51. Skordalakes E, Berger JM. Structural Insights into RNA-Dependent Ring Closure and ATPase Activation by the Rho Termination Factor. Cell. 2006; 127:553–564. [PubMed: 17081977]
52. Xiao Y, Wigneshweraraj SR, Weinzierl R, Wang Y-P, Buck M. Construction and functional analyses of a comprehensive σ54 site-directed mutant library using alanine-cysteine mutagenesis. Nucleic Acids Res. 2009; 37:4482–4497. [PubMed: 19474350]
53. Syed A, Gralla JD. Identification of an N-terminal region of Sigma 54 required for enhancer responsiveness. J. Bacteriol. 1998; 180:5619–5625. [PubMed: 9791110]
54. Hsieh M, Tintut Y, Gralla JD. Functional roles for the glutamines within the glutamine-rich region of the transcription factor σ54. J. Biol. Chem. 1994; 269:373–378. [PubMed: 8276822]
55. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 1982; 157:105–132. [PubMed: 7108955]
56. Wüthrich, K. NMR of Proteins and Nucleic Acids. New York: 1986.
57. Tucker PA, Sallai L. The AAA+ superfamily - a myriad of motions. Curr. Opin. Struct. Biol. 2007; 17:641–652. [PubMed: 18023171]
58. Wigneshweraraj SR, Nechaev S, Bordes P, Jones S, Cannon W, Severinov K, Buck M. Enhancer-Dependent Transcription by Bacterial RNA Polymerase: The β Subunit Downstream Lobe Is Used by σ54 during Open Promoter Complex Formation. Methods Enzymol. 2003; 370:646–657. [PubMed: 14712681]
59. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR. 1995; 6:277–293. [PubMed: 8520220]
60. Keller, R. The computer aided resonance assignment tutorial. Cantina Verlag; 2004.

61. Cobas JC, Sardina FJ. Nuclear magnetic resonance data processing. MestRe-C: A Software package for desktop computers. Concepts Magn. Reson. 2003; 19A:80–96.

Highlights

- In free σ^{54}, the AID (residues 1–50) is intrinsically disordered.
- The AID becomes ordered upon core polymerase binding.
- The AID alone binds transcriptional activators in their ATP state.
- The AID binds the activator with native-like affinity.
- σ^{54} residues 16–25 are the major contact region to the ATPase.

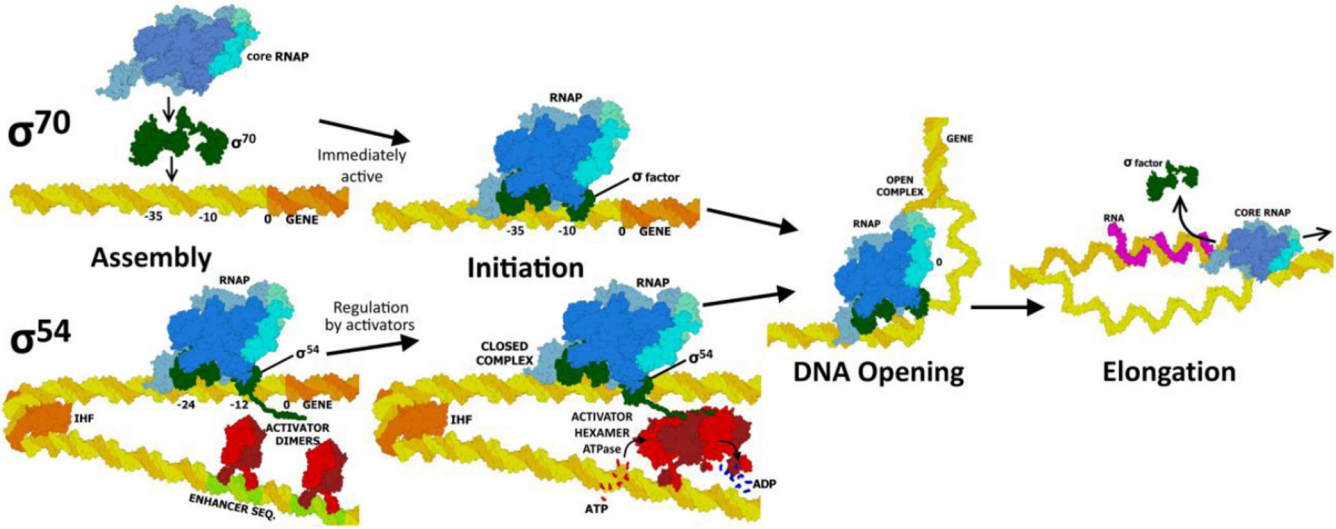


Figure 1. Diagram of transcription initiation mediated by the two classes of bacterial sigma factors

(1) Assembly: sigma factor with RNA polymerase bind upstream of the start site. (2) Initiation: σ^{70} is immediately able to initiate DNA opening, while σ^{54} requires an activation event from one of the transcriptional activators. (3) DNA opening: the RNA polymerase holoenzyme melts DNA. (4) Elongation: the sigma factor can dissociate and core RNA polymerase continues to transcribe RNA using the single stranded DNA template.

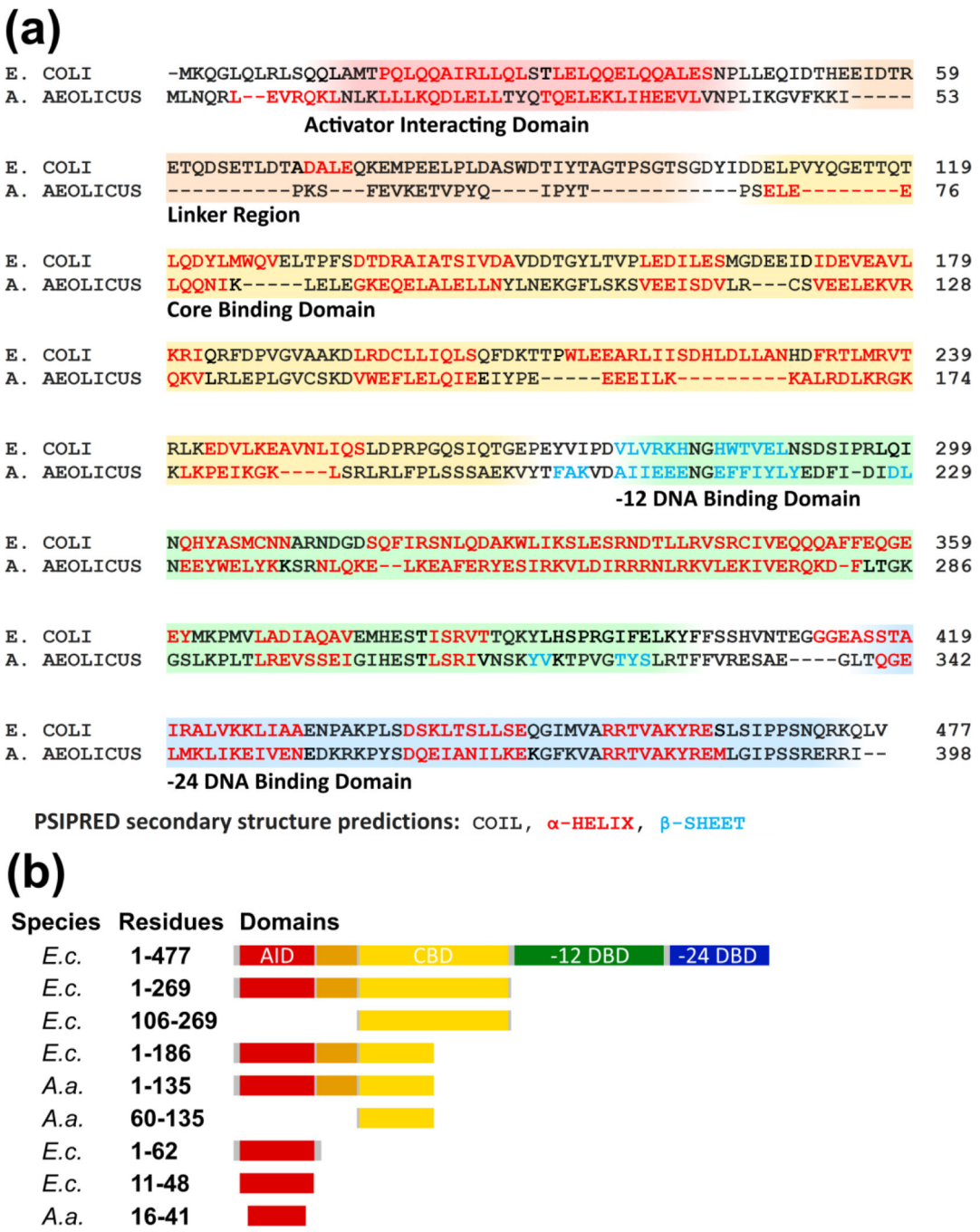


Figure 2. Sequence alignment of *E. coli* σ^{54} and *A. aeolicus* σ^{54}

(a) The predicted secondary structure (black, coil; blue, strand; red, helix) calculated by PSIPRED prediction is shown. Rough domain boundaries for the five functional domains are marked with colored highlights (red, activator interacting domain; orange, linker region; yellow, core binding domain; green, −12 DNA binding domain; blue, −24 DNA binding domain). (b) Domain diagram of all *E.c.* and *A.a.* σ^{54} constructs discussed in this paper.

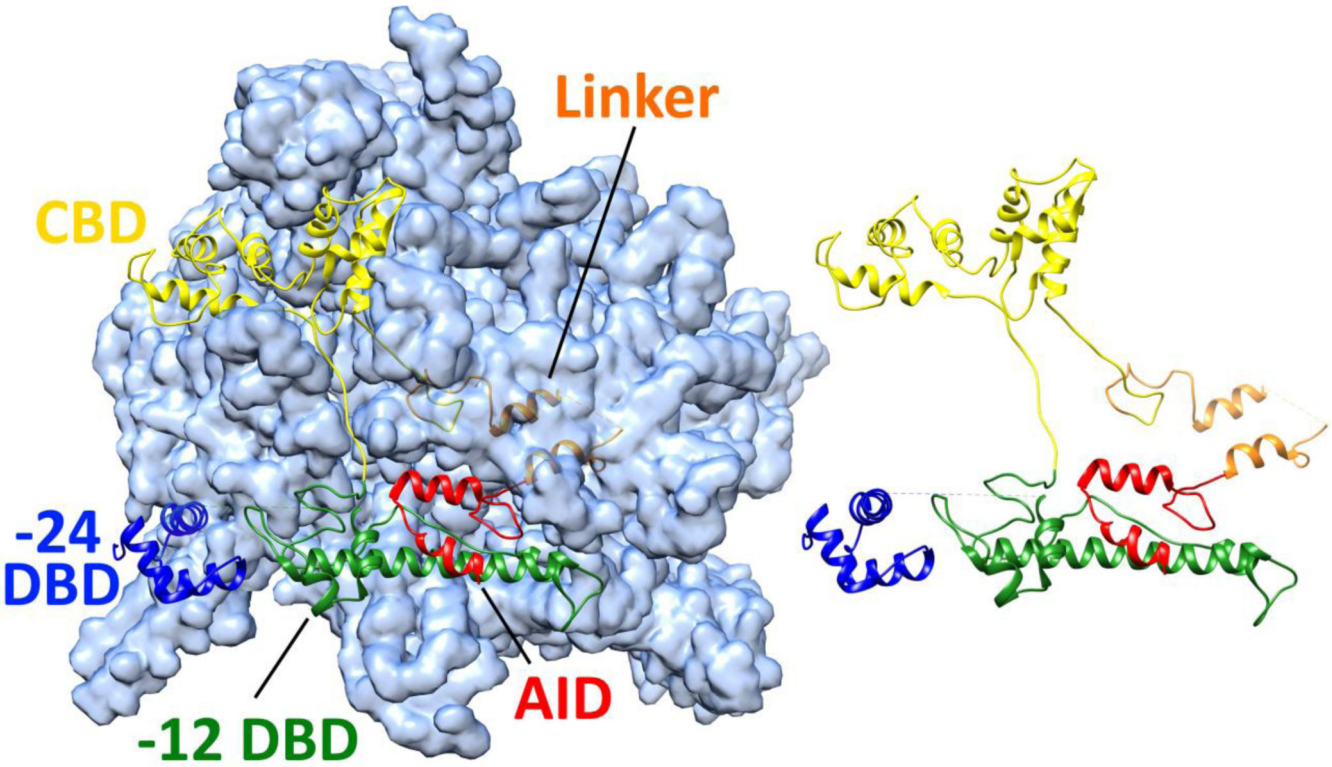


Figure 3. Crystal structure of σ^{54} RNAP holoenzyme

RNAP subunits α, α, β, β', and ω are shown with a surface (transparent light blue). The domains of σ^{54} are shown as ribbons bound to the RNAP, from N- to C-terminus the AID (red), linker region (orange), CBD (yellow), −12 DBD (green), and −24 DBD (blue). On the right is the same structure showing only σ^{54} with the core RNAP surface model removed. Residues could not be assigned before the AID (M1 to A14), within the linker region (D71 to D80), and between the −12 and −24 DBDs (Q387 to A415). PDB ID: 5BYH [36].

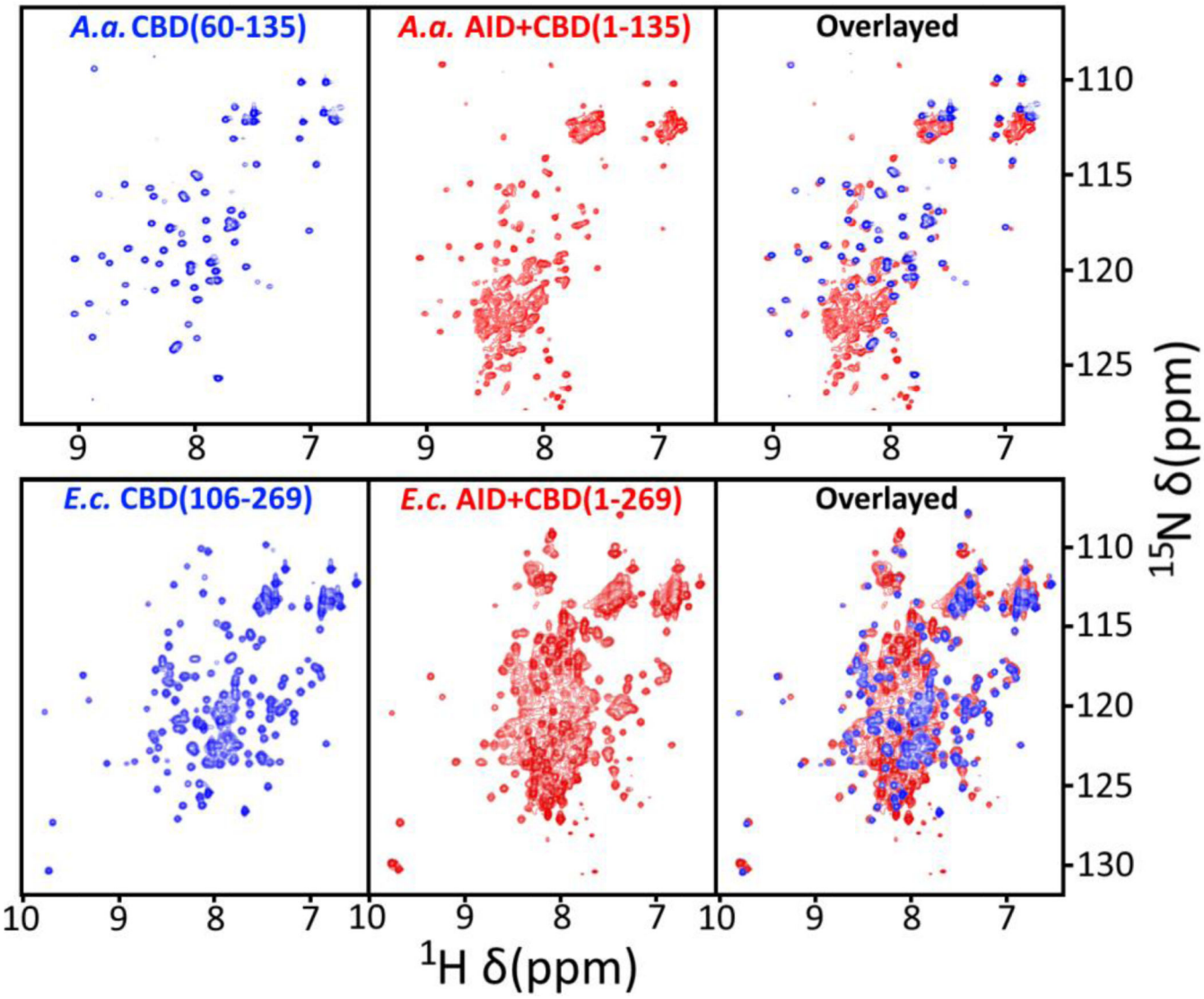


Figure 4. HSQCs of the core binding domain and activator interacting domain

Comparison of ^{1}H-^{15}N HSQCs from the core binding domain (left, blue) and the activator interacting domain plus the core binding domain (red, middle) of *A aeolicus* (top) and *E. coli* (bottom) σ^{54}. In both species, overlaying the HSQC of the core binding domain alone with a construct containing both the activator interacting domain and core binding domain (right) reveals unaccounted for peaks corresponding to the unstructured activator interacting domain and linker domain.

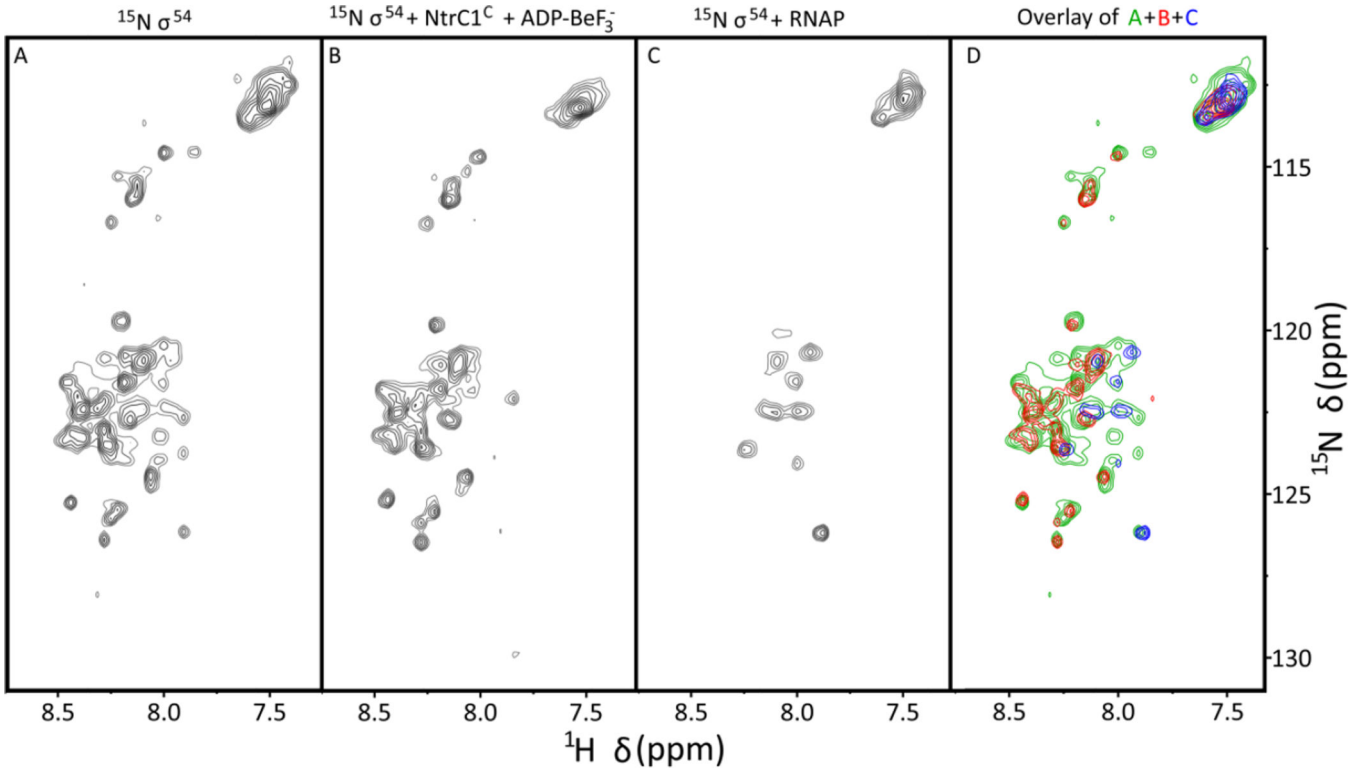


Figure 5. HSQCs showing binding of the activator interacting domain to NtrC^{C} and RNAP

^{1}H-^{15}N HSQCs at high contour levels showing the *E.c.* σ^{54}(1-477) construct with most well-dispersed peaks corresponding to folded domains below the cutoff (A). These peaks correspond to the AID and linker and are almost all gone in the *E.c.* σ^{54}ΔAID(106-477) spectrum (not shown). Some of the σ^{54} AID peaks broaden in the presence of NtrC1^{C} and ADP-BeF$_3^-$ (B) and RNA polymerase (C) indicating that the AID binds to NtrC1^{C} in the ATP state and the core RNA polymerase as part of the *E.c.* σ^{54}-RNA polymerase complex. An overlay (D) of the spectra of free σ^{54} (green), σ ^{54} bound to NtrC1^{C} and ADP-BeF$_3^-$ (red), and σ^{54} bound to RNAP (blue) suggests that the activator and RNAP may interact with different regions of the AID, presumably the N-terminal and C-terminal helices respectively.

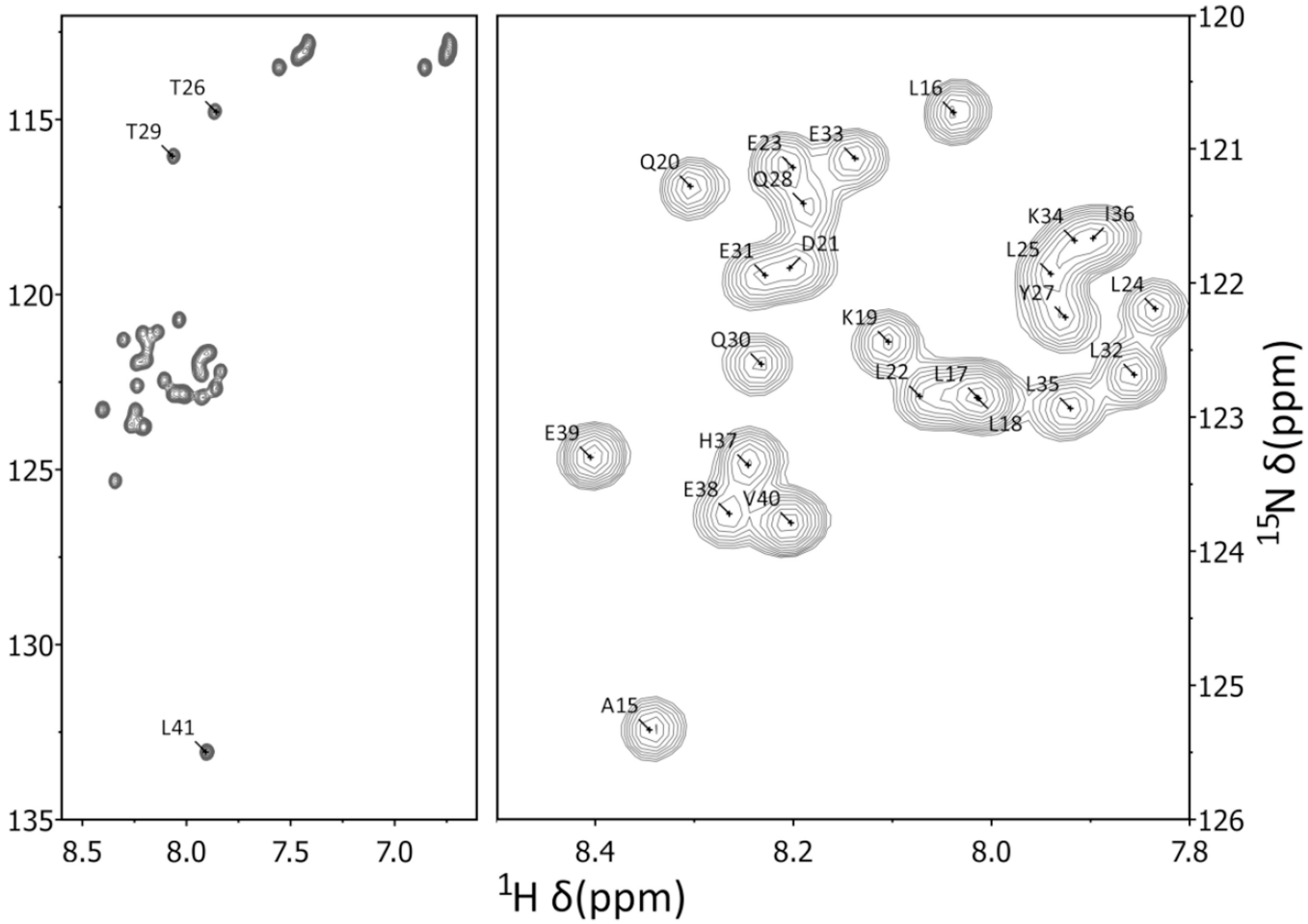


Figure 6. Chemical shift assignments of the minimal activator interacting domain AID(16-41)

Amide assignments of the AID(16-41) construct obtained from 3D ^{15}N-resolved [^{1}H-^{1}H]-NOESY data. A tall spectrum (left) shows the assignments of the outlying peaks for T26, T29 and the C-terminal residue L41. A zoomed in spectrum (right) shows the assignments of the remaining peaks.

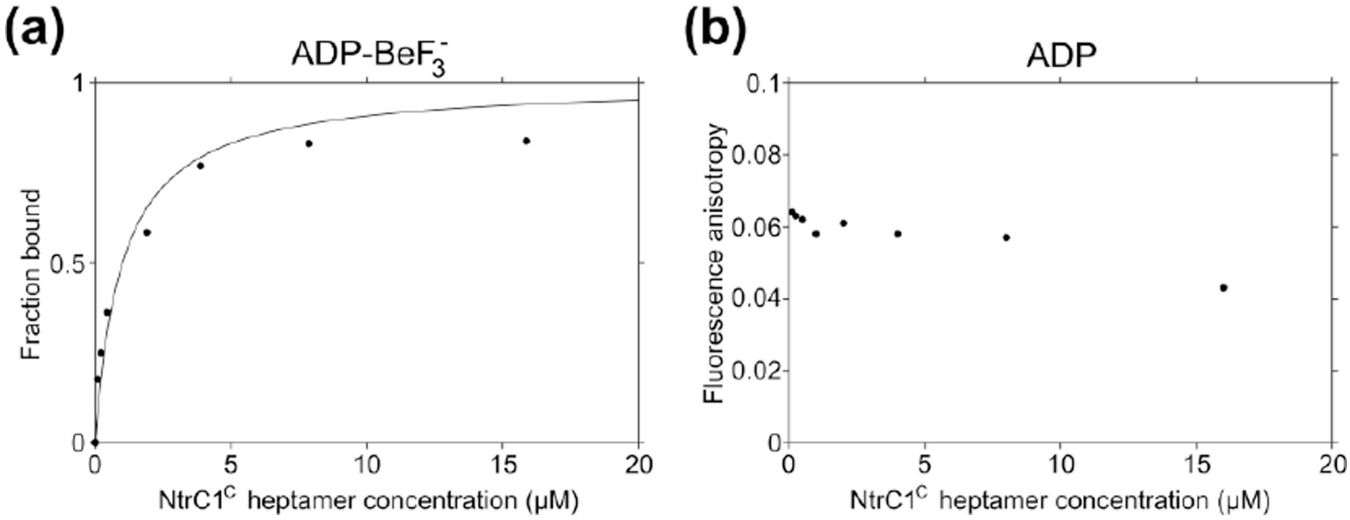


Figure 7. Binding constant of the AID to NtrC1^{C} determined by fluorescence anisotropy

(a) Fluorescence anisotropy data for 0.16 μM Alexa488-labeled AID(16-41) with 500 μM ADP-BeF_3^- and increasing concentrations of NtrC1^{C} heptamer. The fraction bound is determined by the measured fluorescence anisotropy relative to its maximum value with high excess NtrC1^{C} (fully bound AID(16-41)) and its minimum value with no NtrC1^{C} (fully unbound AID(16-41)). The K_d of the binding is the concentration where half of the peptide is bound, which occurs at 1 μM in the presence of ADP-BeF_3^-. (b) Fluorescence anisotropy data when 500 μM ADP is used instead of ADP-BeF_3^-. No binding is detected in the presence of ADP only.

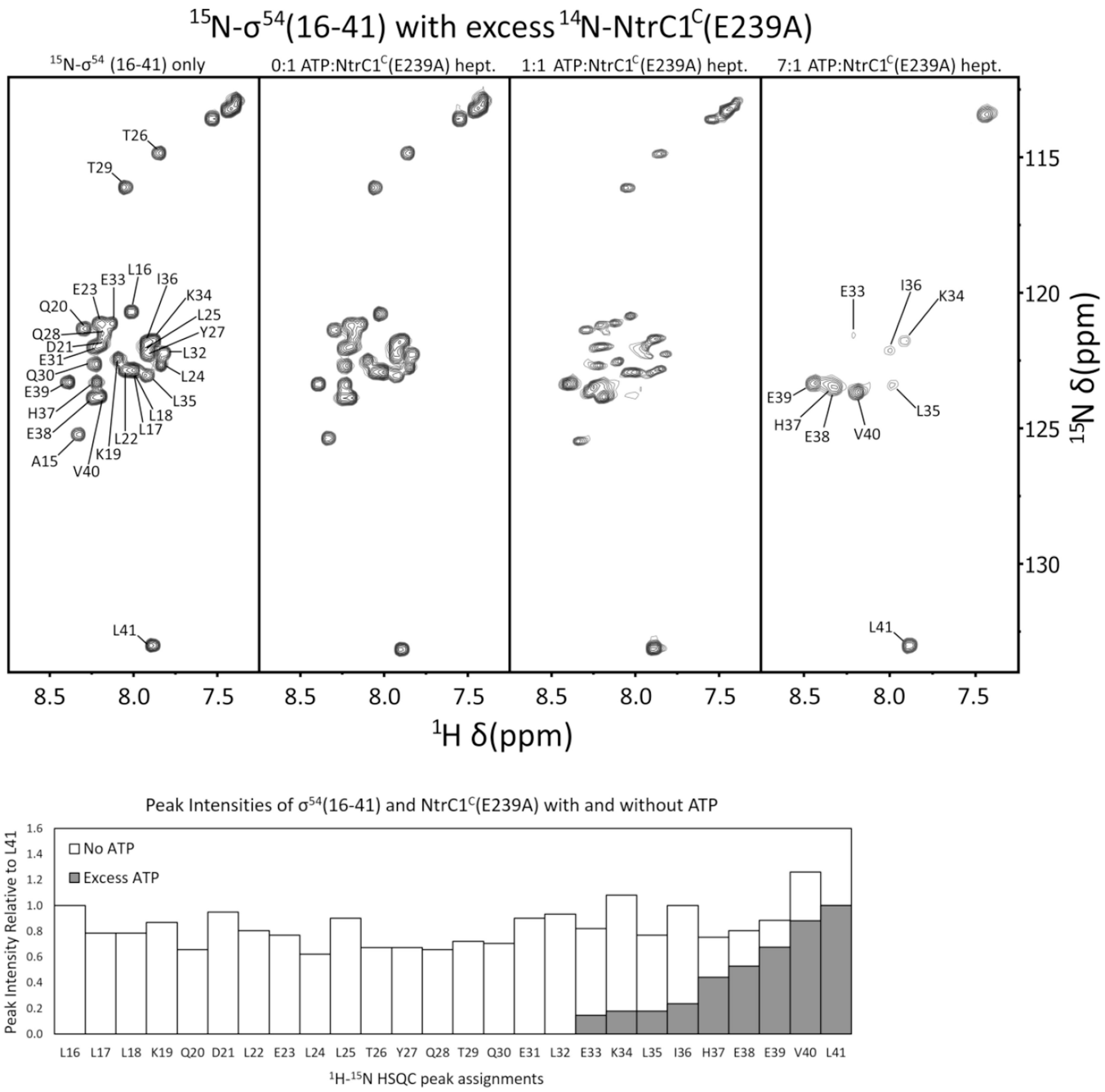


Figure 8. Peak broadening in the HSQC of the AID upon activator binding

^{1}H-^{15}N HSQC of *A.a.* ^{15}N-σ^{54}AID(16-41) and ^{14}N-NtrC1^{C}(E239A) heptamer without nucleotide (left) with 1:1 ATP:NtrC1^{C}(E239A) heptamer (middle) and with 7:1 ATP:NtrC1^{C}(E239A) heptamer (right) showing the disappearance of most AID peaks when bound to NtrC1^{C}(E239A) trapped in the ATP state. The remaining AID peaks at high ATP concentrations are localized to the second predicted helix and correspond to E38, E39, V40, and L41, with broad peaks present for K34, L35 and I36. One remaining glutamine side chain peak likely corresponds to Q30. Peak intensities of each residue ^{15}N-σ^{54}AID(16-41)

with ^{14}N-NtrC1^{C}(E239A) in the presence (gray) and absence (white) of ATP are shown below.

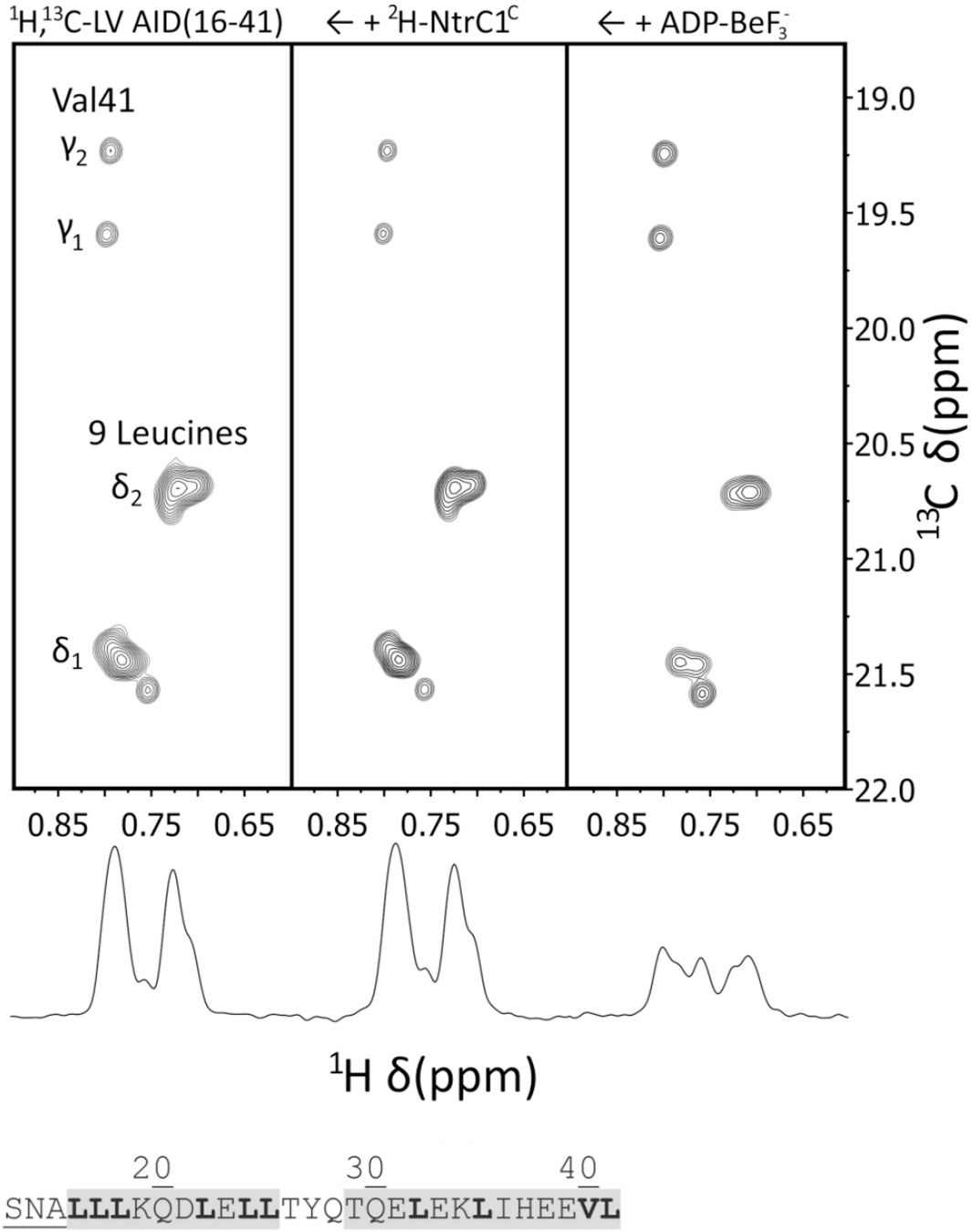


Figure 9. Methyl-TROSY HMQC spectra of AID(16-41) bound to NtrC1^{C}

2D ^{1}H-^{13}C methyl-TROSY HMQC spectra of AID(16-41) with ^{1}H-^{13}C labeling of δ_1 or δ_2 methyls of leucine and γ_1 or γ_2 methyls of valine and ^{2}H and ^{12}C labeling of all other carbons and hydrogens. Integration of all leucine and valine side chain methyls matches the expected number of nine leucines per one valine in the AID(16-41) alone (left). The spectrum is unaffected by addition of excess ^{2}H-NtrC1^{C} (middle) but integration indicates six leucines are broadened by the addition of ^{2}H-NtrC1^{C} and ADP-BeF$_3^-$ (right) leaving three leucines and one valine. The sum projection (below) also reflects the broadening of

some leucine peaks when ADP-BeF_3^- is present. The sequence of the AID peptide with its two predicted α-helices highlighted in gray is shown below. The nine leucines and one valine are in bold. The three underlined residues “SNA” in front of the first helix are the result of TEV cleavage and are not part of the true *A.a.* σ^{54} sequence.

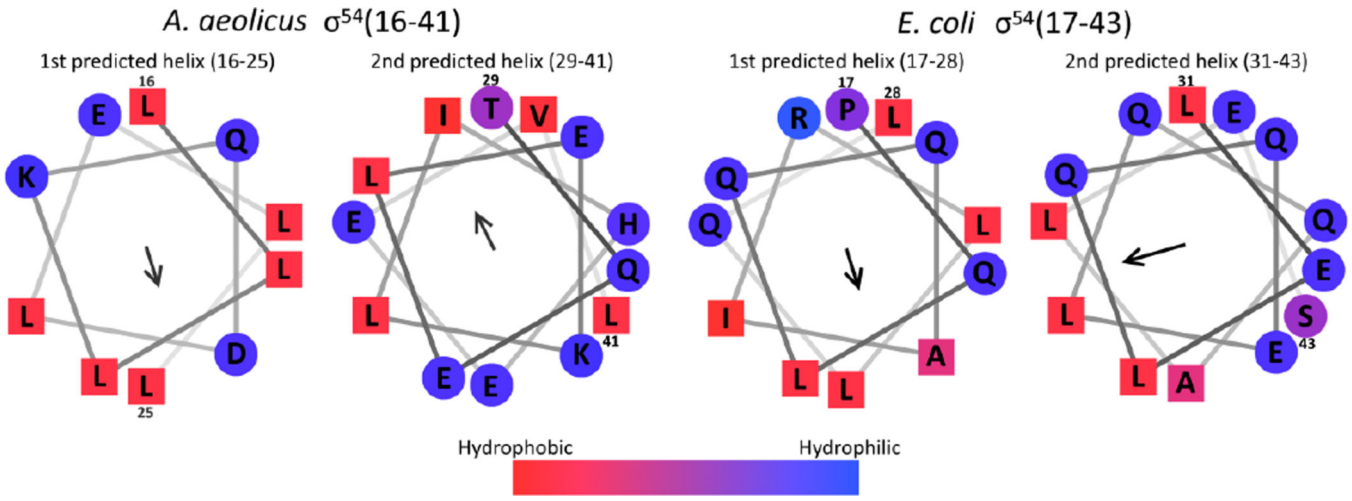


Figure 10. Helical wheel diagrams of predicted secondary structure in the AID

Helical wheel diagrams of the two PSIPRED predicted helices of the activator interacting domain in *A. aeolicus* (left) and *E. coli* (right). Residues are colored by their Kyte-Doolittle hydrophobicity score (more red, more hydrophobic; more blue, more hydrophilic) with hydrophobic residues shown as squares and hydrophilic residues shown as circles [55]. Amphipathic nature of the predicted helices is indicated by the arrows pointing towards their hydrophobic side.