Review

A Perspective on the Enhancer Dependent Bacterial RNA Polymerase

Nan Zhang and Martin Buck *

Department of Life Sciences, Imperial College London, Sir Alexander Fleming Building, Exhibition Road, London SW7 2AZ, UK; E-Mail: nan.zhang@imperial.ac.uk

* Author to whom correspondence should be addressed; E-Mail: m.buck@imperial.ac.uk; Tel.: +44-0-20-7594-5443; Fax: +44-0-20-7504-5419.

Abstract: Here we review recent findings and offer a perspective on how the major variant RNA polymerase of bacteria, which contains the sigma54 factor, functions for regulated gene expression. We consider what gaps exist in our understanding of its genetic, biochemical and biophysical functioning and how they might be addressed.

Keywords: RNA polymerase; Sigma54; transcription; AAA+ ATPase

1. Introduction

Along with the recognition that bacterial RNA polymerases were heterogeneous with respect to their sigma factor content, came the finding that two classes of sigma factor existed in many different types of bacteria [1]. Unlike the major sigma70 class, the sigma54 class of factor was distinctive in being enhancer dependent and relying on a specialised class of transcription activator which used ATP binding and hydrolysis to catalyse the formation of open promoter complexes (RP_O). Closed promoter complexes (RP_C) rarely spontaneously isomerised to open complexes, and so sigma54 dependent systems showed a typical major dependence on cognate activators (reviewed in [2,3]). No hard and fast rules allow the prediction of which bacteria will contain an *rpoN* gene or genes (some organisms have two *rpoN*s) encoding sigma54, although informatics suggests sigma54 use is rooted in the control of cell envelope functions in response to stress [4]. Because the sigma54 factor controls important bacterial stress response genes in pathogenicity and in agriculture [5–9], there is considerable interest

in working out where the dependence arises in biochemical and structural biology terms. Further, what are the advantages the system may have over conventional repression and activation systems used by the sigma70 class of RNA polymerase holoenzymes?

2. The Sigma54 Factor

Work from Sydney Kustu's lab provided the first biochemical evidence that the sigma54 factor (encoded by *rpoN*, also called ntrA) was a dissociable sigma factor which directed RNA polymerase as a holoenzyme to the −12, −24 promoter sequences from which it transcribed [10,11]. Sequence analysis of a range of *rpoN* encoded genes was to provide clear indications that the sigma54 protein was unrelated to the major sigma70 class of sigma factors at the level of primary sequence and most likely its fold and tertiary structure. For the most part this lack of relatedness has been upheld by what limited structural data sets for fragments of sigma54 have been obtained to date. The lack of relatedness suggests that sigma54 may direct a different set of structural transitions in taking RPc to RPo to that orchestrated by sigma70 type factors, rather than just changing the kinetics of the same pathway of conformational change.
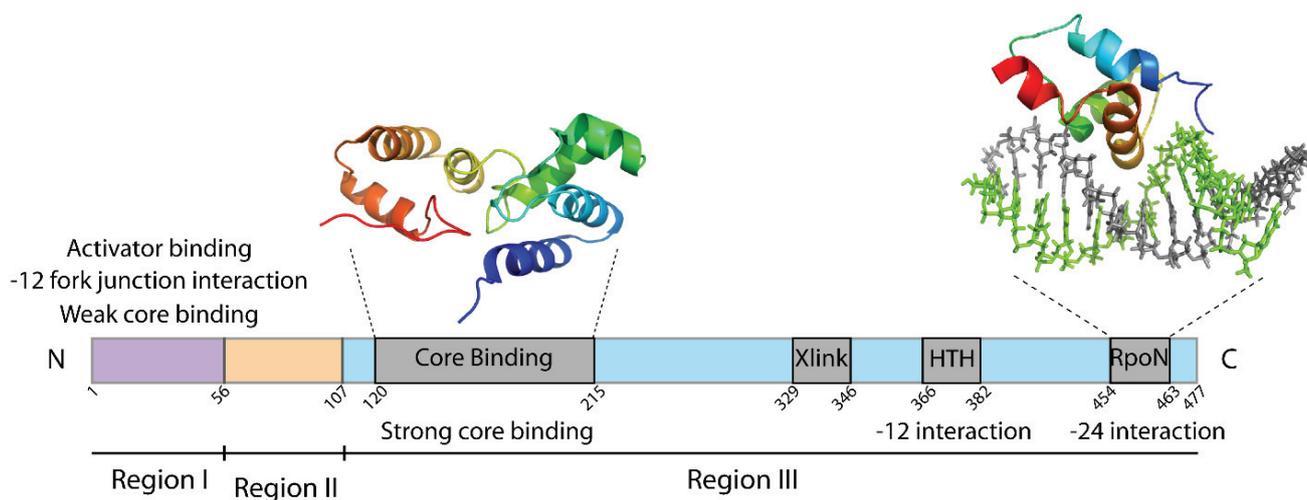

Figure 1. Domain organisation of *Escherichia coli* σ⁵⁴. The DNA binding function is mainly associated with σ⁵⁴ Region III, including the DNA crosslinking motif (Xlink), a putative helix-turn-helix motif (HTH) for −12 recognition via E378 and the RpoN box for −24 binding. Structures of two σ⁵⁴ Region III fragments have been obtained (PDB entries: 2K9L and 2O8K).

Sigma54 contains two highly conserved domains (Regions I and III, Figure 1), separated by what may be a flexible linker (Region II). The glutamine- and leucine-rich Region I interacts with activator ATPases, core RNAP and the −12 promoter sequence, thus posing an energy barrier to spontaneous isomerization of RPc to RPo. Region I mutations and deletion often resulted in activator-bypass phenotypes [12,13]. Sigma54 Region II is dispensable for RNAP isomerisation and interaction with DNA. Indeed many bacterial sigma54 proteins do not naturally contain this domain, arguing for its ancillary role. Sigma54 Region III is primarily involved in binding to the promoter DNA at several sites, with the strongest interaction being between the RpoN box and the −24 promoter element. The RpoN

box forms a winged helix-turn-helix (HTH) motif and inserts the recognition helix into the major groove of the −24 promoter sequence [14,15]. Mutations in the RpoN box have been shown to cause a more than 80% reduction in promoter DNA binding affinity [16].

3. Promoter Recognition

Apparently sigma54 RNAP finds promoters by a direct binding route and not sliding, at least *in vitro* [17]. At its cognate promoters the holoenzyme opens up the A:T base pair 3' to the promoter −12 consensus GC sequence and so a repressive fork junction structure with which the Region I of sigma54 interacts is created. Resolution of this structure by the activator ATPases is a part of the enhancer dependent activation of the RP_C to yield an RP_O. Details of the −12 recognition problem and how the local −12 proximal un-stacking is achieved in the stable RP_O from an unstable prior fully stacked DNA RP_C [18] is now a structural biology issue. Genetics implicates a HTH motif E378 residue and specifically also residue L25 in sigma54 in the −12 GC recognition, but the specific amino acid sequences in sigma54 and in core enzyme needed for base un-stacking just downstream of the −12 GC are unknown although may well map to the Region I of sigma54. Similarly DNA melting defective mutants (for RP_O) of sigma54 have not been obtained, in contrast to sigma70, and so quite how sigma54 contributes to opening DNA from −10 to −1 is unknown. However mutants in sigma54 able to allow RP_O formation *in vitro* without activation suggest the barrier to RP_O formation is distinct from determinants that allow formation and maintenance of (at least unstable) forms of RP_O. Although not widely surveyed across promoters, it seems that the sigma54 holoenzyme forms fewer abortive RNA products before elongation than does the sigma70 containing enzyme [19], and so features of sigma70 contributing to the frequency of abortive initiation may not have counterparts in sigma54.

In contrast to a complex recognition of the promoter −12 intimately linked to maintaining RP_C, recognition of the −24 promoter sequence seems relatively well understood, and involves a HTH motif and the RpoN box amino acid sequence, as solved by the Wemmer group by NMR [14]. In contrast to the −35 recognition of sigma70 promoters being mediated by the flap domain of the core RNAP, no such dependence seems so for sigma54 although other roles for the flap domain are suggested by promoter DNA footprinting experiments with a flapless core and sigma54 [20].

4. Activator Remodel of the Sigma54 Holoenzyme

Sigma54 activators (such as NtrC and PspF) belonging to the AAA+ protein family assemble into hexamers and fuel the rearrangement of RP_C to RP_O. Cryo-EM reconstitutions carried out by the Zhang lab revealed up to three sigma54 activators within a hexamer could directly contact the RP_C structure for isomerisation [21]. These contacts were made asymmetrically via the GAFTGA loop one motifs to sigma54 Region I and the upstream −30 promoter region [22], and were possibly accompanied by the splitting of the hexameric ATPase ring in order to exert directional forces [23]. A bridging density within the holoenzyme was observed to physically block the DNA loading channel formed between the β and β' subunits prior to activation (Figure 2A). This density was assigned to a part of sigma54 Region I and it relocated downstream towards the +1 site in the intermediate complex (RP_I) when the ADP-AlF$_X$ hydrolysis analogue was added. The DNA melting site was misaligned with the DNA loading site in the RP_I in the proposed model, which could constitute an ancillary inhibitory mechanism.

The Stockley and Tuma labs further addressed sigma54 domain movements in relation to promoter DNA and ATP hydrolysis by smFRET analysis (Figure 2B, [24]). Sigma54 Region I moved by approximately 9 Å towards the leading edge of the −10 to −1 transcription bubble in the RP_I when activated with ADP-AlF$_x$. This downstream movement upon activation fully agreed with the Cryo-EM observations and potentially correlated with blockage removal by the "power stroke" action of activator ATPases. Once ATP was fully hydrolysed, Region I would retract slightly upstream, possibly to accompany the DNA loading event. In contrast, Region III remained rather static with respect to the −24 promoter sequence along the activation pathway for making RP_O from RP_C.
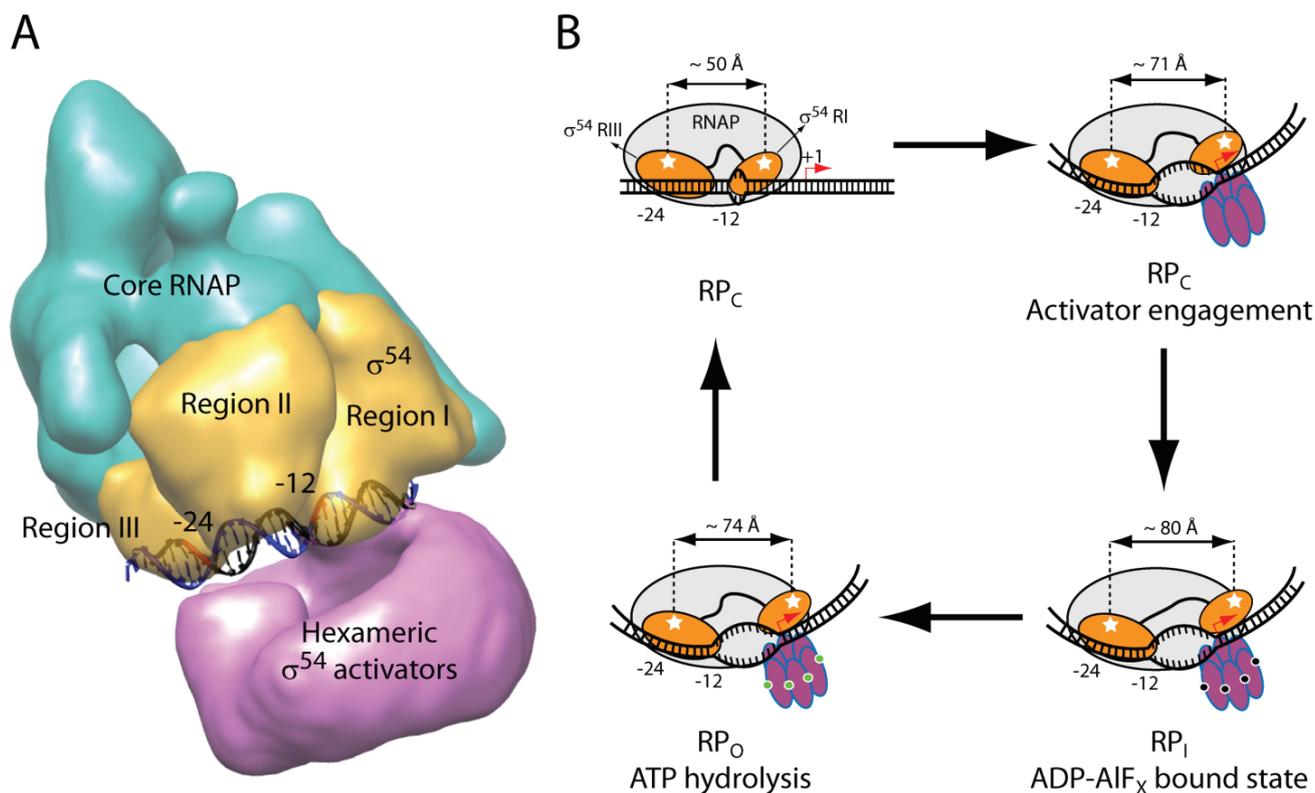

Figure 2. Domain architecture and the proposed mechanism of activation. (**A**) Cryo-EM reconstitution of the Eσ^{54}-PspF-ADP-AlF$_x$ with the promoter DNA modelled in [21]; (**B**) Domain movements of σ^{54} Regions I and III during the transcription activation cycle (modified from [24]). The white stars depict fluorophores used in the smFRET experiments.

Overall it is clear that activation for making RP_O requires a repositioning of the sigma54 Region I, which occurs through the use of at least two GAFTGA loops of the activator ATPases that bind there. Whether the activator binds a rare state of the holoenzyme in RP_C, or drives the formation of an entirely new functional state is not yet clear. However, the R336A sigma54 variant phenocopies the action of the binding of the activator to RP_C in that the Region I has moved compared to that in RP_C, and so the barrier to passing from RP_C to at least an RP_I if not RP_O is not so large. Release of the activators from the sigma54 Region I may be key in passing to RP_O from RP_I, as may a contact of the activator with upstream promoter DNA. Kinetic studies on the *glnAP2* promoter using COSMO methodologies indicates that the transition from RP_C to transcript generation takes around ninety seconds, and activators interact with both RP_C and RP_O [18]. These single molecule studies also lend support to the idea that

an initial form of RP_C exists without the DNA base unstacking at −12 [18]. Given the heterogeneous nature of nucleotide bound states of the activators, that the hexameric activators may assume an opened form and use a subset of the six possible sigma-contacting mobile surface features [22,25]. Their asymmetric functioning seems a necessary part of the transcription activation mechanism and may reflect the asymmetric target of the σ^{54} within the RP_C. A processive functioning of the activators may not be necessary for activation, because the activator bound RP_I has been shown to possess RP_O like characteristics [26]. However quite how and when DNA opening occurs for making the RP_O with sigma54 is not known, and could involve other functions of the activators ATPase rather than a direct sigma DNA melting activity. Notably when double stranded DNA is outside of the RNAP in RP_I, extensive cross links can be made with the bound activator, which might guide dsDNA into the holoenzyme for making RP_O, or indeed act on it directly for making RP_O [21]. An interaction of DNA with the activators ATPase domain has been observed [27]. Further, activator binding to RP_C (enabled by the use of non-hydrolysable nucleotide analogues such as $ADP-AlF_x$) when promoter DNA from −10 to −1 is unpaired allows transcription initiation, suggesting the holoenzyme can accept ssDNA and then open up the start site to allow RNA synthesis, without the need for ATP hydrolysis by the activator protein [26].

5. Signaling for Sigma54 Dependent Transcription

Genes under control of the sigma54 are required under specific stress conditions or when a particular C or N source is available. Classical examples include the NtrC dependent activation of genes in response to nitrogen limitation, the use of XylR in expressing genes for catabolism of aromatic hydrocarbons and PspF in activating the psp response for inner membrane stress. Although all contain the critical AAA+ domain needed for remodelling the closed promoter complex, the activators are all signaled to via a range of N-terminal regulatory domains. These fall into a wide range of different classes ranging from the response regulator type aspartic acid phosphorylation target in two component members, to GAF domains in NifA and CARF domains in RtcRs (reviewed in [2]). In many cases the exact ligand interacting with these N-terminal regulatory domains remains unknown.

6. Conclusions and a Final Perspective

Detailed structural insights into RP_C, RP_I and RP_O are now needed to work out quite what it is in structural terms that establishes the barrier to making RP_O, and how RP_C and RP_O interconvert. Comparisons to the use of the sigma70 Region I as a place keeper for DNA and that of sigma54 Region I as a target for activation, maintaining the repressive fork junction structure and escaping the action of certain phage inhibitors of RNAP activity such as T7 phage gp2 will be one valuable outcome from such studies. Similarly, knowing what the functional state the core RNAP is in when bound to sigma54 will provide insights into the importance of the sigma-core interface in gene control and potentially antibiotic action. For example in the presence of sigma54 is the core enzyme in a catalytically competent state, and can smFRET data be reconciled with RNAP clamp opening and clamp closing, and the processive closed state of the core enzyme? Do the sigma54 and its activators take the RNAP down a new set of conformationally distinct changes for making RP_O, or is it simply altering the kinetics of a single pathway for making the RP_O from the RP_C?

Quite where sigma54 came from remains a mystery-it cannot be rooted and seems to have no obvious structural counterpart (although detailed structural analysis may overturn this view, if not for the full protein then for some of its domains). Genomics methods and RNAseq studies will now no doubt offer us perspectives on why have sigma54 at all-is it a relic of a transcription repression mechanism, can it (or has it?) evolve in some cases to activator independence, and does it have repressive functionality alongside its gene activation responsiveness? Currently ChipSeq and RNAseq work with sigma54 and its holoenzyme to define its regulon suggests complexity in the roles of sigma54 [28,29]. For example some RP$_C$s seem not to be served by cognate activators, and amongst these a subset seem to repress transcription by other forms of RNAP holoenzyme. In combination with genomic approaches to better understand roles of sigma54 binding sites on bacterial chromosomes, sufficient knowledge of sigma 54 and sigma70 structure-function relationships will allow us to make orthogonal non-native chimeric sigma factors-to tackle issues of promoter use in synthetic biology and an increased insularity from the bacterial chassis used in such work. Since sigma54 is served by cognate activators of the AAA+ protein family, detailed atomic structures of their co-complexes with RP$_C$ will help work out how such AAA+ proteins remodel their targets. Lastly, the full role of sigma54 in bacterial physiology is unknown. Many of its activators respond to signals which are poorly characterised, yet underpin important processes such as RNA repair (the *rtcBA* genes activated by RtcR) and envelope maintenance (the *psp* operon activated by PspF). As details of how genes turn on and off in individual cells through transient changes in, for example, local DNA superhelicity to yield the so called bursty behaviours of gene transcription revealed by time series studies, one might wonder if the sigma54 factor with its coupled ATPase overlays some special features to such time series.

Acknowledgments

Author Contributions

Nan Zhang and Martin Buck equally contributed to the writing.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Merrick, M.J. In a class of its own—The RNA polymerase sigma factor sigma 54 (sigma N). *Mol. Microbiol.* **1993**, *10*, 903–909.
2. Bush, M.; Dixon, R. The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 497–529.
3. Joly, N.; Zhang, N.; Buck, M.; Zhang, X. Coupling AAA protein function to regulated gene expression. *Biochim. Biophys. Acta* **2012**, *1823*, 108–116.

4. Francke, C.; Groot Kormelink, T.; Hagemeijer, Y.; Overmars, L.; Sluijter, V.; Moezelaar, R.; Siezen, R.J. Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics* **2011**, doi:10.1186/1471-2164-12-385.

5. Dixon, R.; Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* **2004**, *2*, 621–631.

6. Fisher, M.A.; Grimm, D.; Henion, A.K.; Elias, A.F.; Stewart, P.E.; Rosa, P.A.; Gherardini, F.C. *Borrelia burgdorferi* sigma54 is required for mammalian infection and vector transmission but not for tick colonization. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5162–5167.

7. Fouts, D.E.; Abramovitch, R.B.; Alfano, J.R.; Baldo, A.M.; Buell, C.R.; Cartinhour, S.; Chatterjee, A.K.; D'Ascenzo, M.; Gwinn, M.L.; Lazarowitz, S.G.; *et al.* Genomewide identification of *Pseudomonas syringae pv. Tomato DC3000* promoters controlled by the *hrpL* alternative sigma factor. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2275–2280.

8. Studholme, D.J.; Buck, M. The biology of enhancer-dependent transcriptional regulation in bacteria: Insights from genome sequences. *FEMS Microbiol. Lett.* **2000**, *186*, 1–9.

9. Wolfe, A.J.; Millikan, D.S.; Campbell, J.M.; Visick, K.L. *Vibrio fischeri* sigma54 controls motility, biofilm formation, luminescence, and colonization. *Appl. Environ. Microbiol.* **2004**, *70*, 2520–2524.

10. Kustu, S.; Santero, E.; Keener, J.; Popham, D.; Weiss, D. Expression of sigma 54 (NtrA)-dependent genes is probably united by a common mechanism. *Microbiol. Rev.* **1989**, *53*, 367–376.

11. Popham, D.L.; Szeto, D.; Keener, J.; Kustu, S. Function of a bacterial activator protein that binds to transcriptional enhancers. *Science* **1989**, *243*, 629–635.

12. Cannon, W.; Chaney, M.; Buck, M. Characterisation of holoenzyme lacking sigman regions I and II. *Nucleic Acids Res.* **1999**, *27*, 2478–2486.

13. Syed, A.; Gralla, J.D. Isolation and properties of enhancer-bypass mutants of sigma 54. *Mol. Microbiol.* **1997**, *23*, 987–995.

14. Doucleff, M.; Malak, L.T.; Pelton, J.G.; Wemmer, D.E. The C-terminal RpoN domain of sigma54 forms an unpredicted helix-turn-helix motif similar to domains of sigma70. *J. Biol. Chem.* **2005**, *280*, 41530–41536.

15. Doucleff, M.; Pelton, J.G.; Lee, P.S.; Nixon, B.T.; Wemmer, D.E. Structural basis of DNA recognition by the alternative sigma-factor, sigma54. *J. Mol. Biol.* **2007**, *369*, 1070–1078.

16. Wang, L.; Gralla, J.D. Roles for the C-terminal region of sigma 54 in transcriptional silencing and DNA binding. *J. Biol. Chem.* **2001**, *276*, 8979–8986.

17. Friedman, L.J.; Mumm, J.P.; Gelles, J. RNA polymerase approaches its promoter without long-range sliding along DNA. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9740–9745.

18. Friedman, L.J.; Gelles, J. Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell* **2012**, *148*, 679–689.

19. Tintut, Y.; Wang, J.T.; Gralla, J.D. Abortive cycling and the release of polymerase for elongation at the sigma 54-dependent *glnAP2* promoter. *J. Biol. Chem.* **1995**, *270*, 24392–24398.

20. Burrows, P.C.; Severinov, K.; Ishihama, A.; Buck, M.; Wigneshweraraj, S.R. Mapping sigma 54-RNA polymerase interactions at the −24 consensus promoter element. *J. Biol. Chem.* **2003**, *278*, 29728–29743.

21. Bose, D.; Pape, T.; Burrows, P.C.; Rappas, M.; Wigneshweraraj, S.R.; Buck, M.; Zhang, X. Organization of an activator-bound RNA polymerase holoenzyme. *Mol. Cell* **2008**, *32*, 337–346.

22. Zhang, N.; Joly, N.; Buck, M. A common feature from different subunits of a homomeric AAA+ protein contacts three spatially distinct transcription elements. *Nucleic Acids Res.* **2012**, *40*, 9139–9152.

23. Sysoeva, T.A.; Chowdhury, S.; Guo, L.; Nixon, B.T. Nucleotide-induced asymmetry within ATPase activator ring drives sigma54-RNAP interaction and ATP hydrolysis. *Genes Dev.* **2013**, *27*, 2500–2511.

24. Sharma, A.; Leach, R.N.; Gell, C.; Zhang, N.; Burrows, P.C.; Shepherd, D.A.; Wigneshweraraj, S.; Smith, D.A.; Zhang, X.; Buck, M.; *et al.* Domain movements of the enhancer-dependent sigma factor drive DNA delivery into the RNA polymerase active site: Insights from single molecule studies. *Nucleic Acids Res.* **2014**, *42*, 5177–5190.

25. Zhang, N.; Gordiyenko, Y.; Joly, N.; Lawton, E.; Robinson, C.V.; Buck, M. Subunit dynamics and nucleotide-dependent asymmetry of an AAA+ transcription complex. *J. Mol. Biol.* **2014**, *426*, 71–83.

26. Burrows, P.C.; Joly, N.; Buck, M. A prehydrolysis state of an AAA+ ATPase supports transcription activation of an enhancer-dependent RNA polymerase. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9376–9381.

27. Lawton, E.; Jovanovic, M.; Joly, N.; Waite, C.; Zhang, N.; Wang, B.; Burrows, P.; Buck, M. Determination of the self-association residues within a homomeric and a heteromeric AAA+ enhancer binding protein. *J. Mol. Biol.* **2014**, *426*, 1692–1710.

28. Brown, D.R.; Barton, G.; Pan, Z.; Buck, M.; Wigneshweraraj, S. Nitrogen stress response and stringent response are coupled in *Escherichia coli*. *Nat. Commun.* **2014**, doi:10.1038/ncomms5115.

29. Zhao, K.; Liu, M.; Burgess, R.R. Promoter and regulon analysis of nitrogen assimilation factor, sigma54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. *Nucleic Acids Res.* **2010**, *38*, 1273–1283.