Molecular Microbiology (2015) 95(1), 17–30 ■
doi:10.1111/mmi.12844
First published online 24 November 2014

The structural basis for enhancer-dependent assembly and activation of the AAA transcriptional activator NorR

Matt Bush,[1] Tamaswati Ghosh,[2] Marta Sawicka,[2] Iain H. Moal,[3] Paul A. Bates,[3] Ray Dixon[1]* and Xiaodong Zhang[2]**

[1]*Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK.*

[2]*Centre for Structural Biology, Imperial College London, London SW7 2AZ, UK.*

[3]*Biomolecular Modelling Laboratory, Cancer Research UK London Research Institute, 33 Lincoln's Inn Fields, London WC2A 3LY, UK.*

Summary

σ^{54}-dependent transcription controls a wide range of stress-related genes in bacteria and is tightly regulated. In contrast to σ^{70}, the σ^{54}-RNA polymerase holoenzyme forms a stable closed complex at the promoter site that rarely isomerises into transcriptionally competent open complexes. The conversion into open complexes requires the ATPase activity of activator proteins that bind remotely upstream of the transcriptional start site. These activators belong to the large AAA protein family and the majority of them consist of an N-terminal regulatory domain, a central AAA domain and a C-terminal DNA binding domain. Here we use a functional variant of the NorR activator, a dedicated NO sensor, to provide the first structural and functional characterisation of a full length AAA activator in complex with its enhancer DNA. Our data suggest an inter-dependent and synergistic relationship of all three functional domains and provide an explanation for the dependence of NorR on enhancer DNA. Our results show that NorR readily assembles into higher order oligomers upon enhancer binding, independent of activating signals. Upon inducing signals, the N-terminal regulatory domain relocates to the periphery of the AAA ring. Together our data provide an assembly and activation mechanism for NorR.

Accepted 26 October, 2014. *For correspondence. E-mail ray.dixon@jic.ac.uk; Tel. +44 (0)1603 450747; Fax +44 (0)1603 450778. **For correspondence. E-mail xiaodong.zhang@imperial.ac.uk; Tel. +44 (0)207 594 3151; Fax +44 (0) 207 594 3057.

Introduction

Gene transcription is a fundamental cellular process that is carried out by RNA polymerase (RNAP). The specificity of transcription is partially ensured by sigma factors that recruit RNAP to specific promoter sites. Two classes of sigma factors exist in bacteria, the majority of which belong to the σ^{70} class and can proceed to transcription spontaneously (Buck *et al.*, 2000; Browning and Busby, 2004). In contrast, the major alternative sigma factor, σ^{54}, forms a stable closed complex with RNAP at the promoter site that is transcriptionally incompetent. The activation of σ^{54}-dependent genes requires activator proteins that bind remotely upstream relative to the transcriptional start site. These activator proteins, also called bacterial enhancer-binding proteins (bEBPs), belong to the large AAA protein family and contact σ^{54}-RNAP through DNA looping (Buck *et al.*, 2000). bEBPs usually assemble as hexameric rings in their active forms (reviewed in Rappas *et al.*, 2007) and couple ATPase activity to the isomerisation of σ^{54}-RNAP holoenzyme.

The majority of bEBPs contain an N-terminal regulatory domain that responds to external stimulatory signals, a central ATPase domain that is responsible for ATPase activity and σ^{54}-RNAP interaction and a C-terminal DNA binding domain (DBD) that binds specifically to one or more upstream enhancer sites (Fig. 1A). The N-terminal domain regulates the central AAA domain activity either positively or negatively, often through the control of oligomerisation (reviewed in Ghosh *et al.*, 2010; Bush and Dixon, 2012). Hexamerisation is required both for ATPase activity and for interaction with the holoenzyme bound at the promoter, which leads to the remodelling of the closed complex and transcriptional activation. Two highly conserved surface loops L1 and L2 in the central AAA domain have been shown to interact with σ^{54}-RNAP and undergo nucleotide-dependent conformational changes in a number of well-characterised bEBPs, such as the phage shock protein F (PspF) from *Escherichia coli*, nitrogen regulatory protein C (NtrC) from *Salmonella typhimurium*, and NtrC1 from *Aquifex aerolicus* (De Carlo *et al.*,

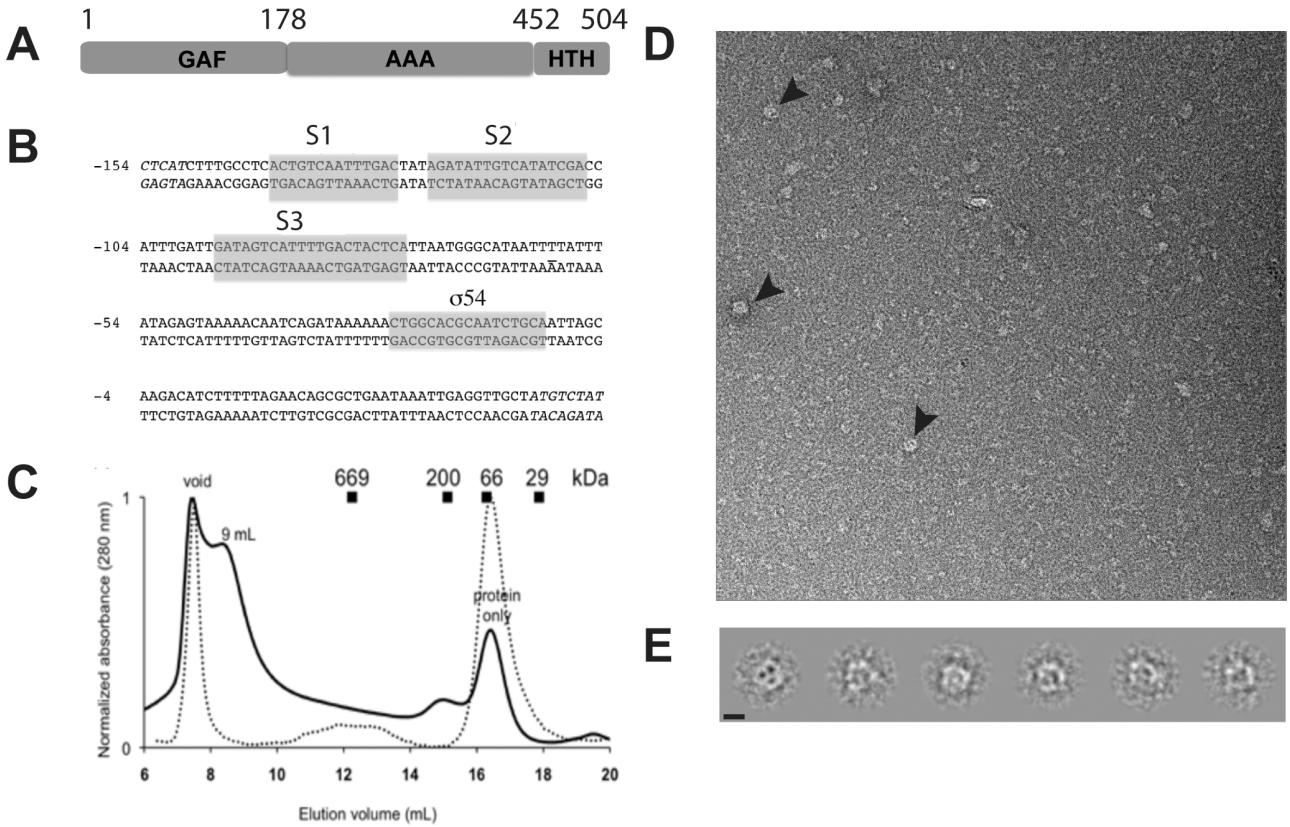

Fig. 1. Enhancer-dependent higher order oligomeric assembly of wild type full length NorR.
A. Domain organisation of NorR protein containing GAF, AAA and HTH DNA binding domains.
B. *norR* enhancer sites relative to the transcription start site. Shaded areas are the σ^{54} promoter and the three NorR binding sites indicated as S1, S2 and S3.
C. Gel filtration chromatography of 3 μM NorR in the absence (*dotted line*) and presence (*solid line*) of 0.4 μM 266 bp dsDNA. The presence of DNA stabilises a higher order oligomeric form of NorR eluted at 9 mL. Corresponding molecular weight of standard globular proteins are indicated at their elution volume.
D. Electron microscopy micrographs of NorR in complex with 266 bp DNA fragment. Ring-shaped higher order oligomeric particles were observed in the presence of DNA (black arrows).
E. Class averages of NorR-DNA showing oligomeric particles. Scale bar = 100 Å.

2006; Rappas *et al.*, 2006; Chen *et al.*, 2010). Changes in these loop conformations have been proposed to drive the remodelling of the σ^{54}-RNAP closed complex, leading to transcriptional activation (Rappas *et al.*, 2005; 2006; De Carlo *et al.*, 2006; Chen *et al.*, 2010).

The nitric oxide (NO)-responsive bEBP, NorR, is required for transcriptional activation of genes that encode a flavorubredoxin that reduces the NO radical under anaerobic conditions (Hutchings *et al.*, 2002). Binding of NO to the iron centre in the N-terminal GAF domain of NorR generates a mononitrosyl complex that releases intra-molecular repression of the adjacent AAA domain (D'Autreaux *et al.*, 2005; Tucker *et al.*, 2008) and deletion of the N-terminal domain renders the activator (NorRΔGAF) constitutively active (D'Autreaux *et al.*, 2005). Our recent study identifies an inter-domain interaction that involves the GAF domain and the afore-mentioned L1 and L2 loops of the AAA domain (Bush *et al.*, 2010). Therefore, it is proposed that the GAF domain negatively regulates the AAA domain by preventing access of the L1 and L2 loops to RNAP-σ^{54} holoenzyme.

One unusual property of NorR is the presence of three enhancer sites upstream of its target promoter, all of which are essential for activity (Fig. 1B). Mutating a single site abolishes its activity *in vivo*. Even in the absence of GAF domain repression, activity still requires the presence of all three enhancer sites (Tucker *et al.*, 2010). This suggests that the GAF domain-mediated repression of NorR activity does not function by preventing oligomer formation, as is the case in NtrC1 and DctD (Lee *et al.*, 2003; Doucleff *et al.*, 2005a). Instead it is likely that enhancer binding *per se* drives the assembly of the NorR hexamer, independent of the binding of NO.

In order to decipher the organisation of bEBPs at their enhancer sites and uncover the NO-dependent activation mechanism in NorR, we investigated the structure and

function of full length NorR, using wild type and a variant that is constitutively active to represent the activated state. The 3D structure of this NorR variant in complex with a 266 bp DNA fragment containing all three upstream activator sequences provides the first structural description of a full length bEBP in complex with its enhancer DNA and reveals an inter-dependent relationship between all three domains of the protein. Functional characterisation of this NorR variant suggests a link between the L1/L2 loops and inter-subunit interactions. Combined with studies of wild-type NorR in the pre-activated state, we provide a mechanism for NorR assembly and activation.

Results

Full length NorR pre-assembles into higher order oligomers upon binding to enhancer DNA

Although the full length NorR protein is competent to bind enhancer DNA, this nucleoprotein complex is inactive with respect to ATP hydrolysis and transcriptional activation in the absence of an NO signal (D'Autreaux *et al.*, 2005). Deletion of the GAF domain results in a constitutively active form of NorR although the activity is strictly dependent on the integrity of all three enhancer sites (Tucker *et al.*, 2010). In light of these findings, we speculated that NorR pre-assembles into a 'pre-activated' hexamer at the enhancer sites, poised to respond rapidly to NO stress. This is in complete contrast to activation mechanisms in other bEBPs such as NtrC and NtrC1 in which the N-terminal receiver domains regulate ATPase activity by controlling higher order ring assembly of AAA subunits. This is also different from other GAF containing σ^{54} activators, such as the NifA-like homolog from *Aquifex aeolicus*. These bEBPs are dimers in their inactive state and form hexameric rings upon phosphorylation by sensor kinases (Ghosh *et al.*, 2010; Batchelor *et al.*, 2013).

To verify that prior to NO-dependent activation, wild type NorR assembles into a hexamer ring upon binding to the three enhancer sites, we characterised the oligomerisation state of NorR in the presence and absence of DNA (Fig. 1). We used analytical gel filtration combined with negative-staining EM to analyse the NorR complex bound to the 266 bp dsDNA fragment containing all three enhancer sites. As shown in Fig. 1C, enhancer DNA shifts the elution peak from 16.5 ml that corresponds to NorR monomers to a higher molecular mass species at 9 ml. In the presence of DNA, there is also a small peak at 15 ml, which could represent a NorR dimer binding to one of the enhancer DNA sites. When visualised using negative stained EM of the peak fractions of 9 ml, higher order oligomeric particles are clearly present and class averages show higher order oligomeric rings (Fig. 1D and E). These data confirm that full length NorR, although inactive, can pre-assemble into higher order oligomers upon binding to the three enhancer sites, in the absence of NO. However, the heterogeneous nature of the particle sizes displayed in the micrographs prevented further 3D structural analysis of wild-type protein.

The Q304E substitution in NorR increases activity in the absence of NO

To explore the inter-domain regulation in NorR, we previously used error-prone PCR to create amino acid substitutions that potentially disrupt the interactions between the GAF and AAA domains (Bush *et al.*, 2010). One such substitution is Q304E, which is located at the end of α4 of the AAA domain (Fig. 2A) and precedes the crucial L2 loop (Fig. 2A and B, Bush *et al.*, 2010). Although Q304 is highly conserved among NorR proteins, interestingly, the equivalent residue is E in PspF and NtrC1 (Fig. 2C). The NorR-Q304E variant activates transcription *in vivo* in the absence of an NO-source and under these conditions this NorR variant has an activity similar to that of the wild-type protein in the 'activated-state' (Bush *et al.*, 2010), suggesting escape from GAF domain-mediated repression (Fig. 2D). However, the presence of NO leads to further induction of NorR activity (Fig. 2D, compare solid and open bars). This implies that the mechanism of inter-domain repression has not been fully disrupted. Removal of the GAF domain in the Q304E variant (Q304EΔGAF) led to constitutive activity *in vivo* as anticipated, at levels similar to that of the truncated derivative of wild-type NorR, lacking the GAF domain (NorRΔGAF) (Fig. 2D). Notably, the activity of the Q304E variant increased significantly in the presence of an NO source compared with the truncated proteins lacking the GAF domain, suggesting that this domain may have a positive role in the activation of the AAA domain once the iron centre is nitrosylated.

In addition to glutamate, an aspartate substitution at position 304 also gave rise to similar activity (Fig. 2D). However, substitutions to asparagine or arginine did not produce functional NorR. Hence, a negatively charged carboxyl group in the side chain of residue 304 may assist NorR to bypass the repressive function of the GAF-domain. In contrast, the Q304A variant had similar properties to wild-type NorR. This is not unexpected, given that the Q304 residue is not thought to directly contact the GAF domain (Bush *et al.*, 2010) and therefore not directly involved in the regulation of AAA domain activity.

The Q304E substitution does not influence enhancer-binding affinity

Since the oligomerisation state and hence the activity of the AAA domain of bEBPs is often controlled by regulatory

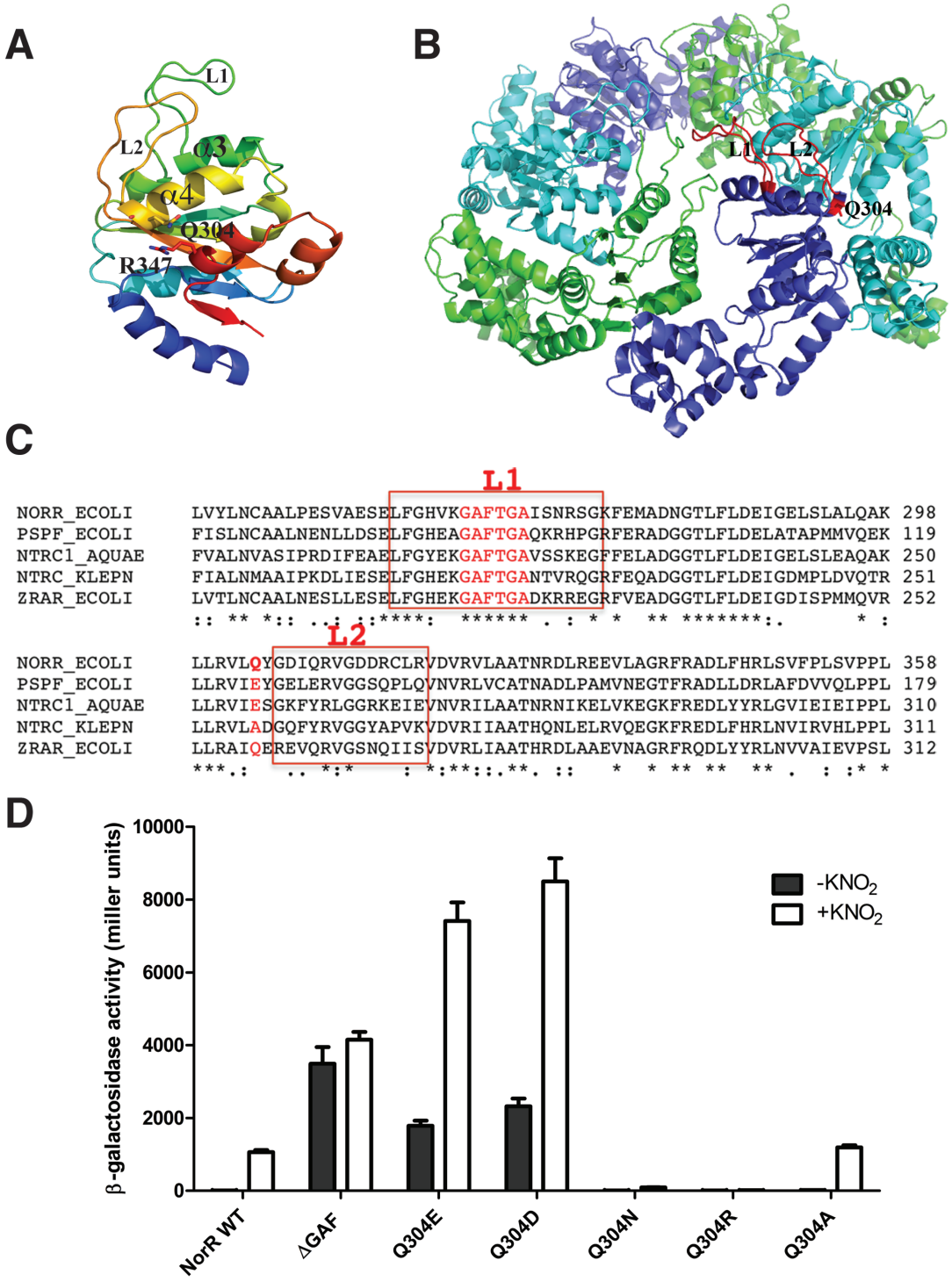

Fig. 2. Characterisation of the NorR-Q304E variant *in vivo*.
A. Cartoon representation of NorR AAA domain (modelled on NtrC1, pdbcode 1NY6) with L1, L2 loops and residue Q304 labelled.
B. Hexameric assembly based on ZraR showing L1 and L2 loops (1OJL).
C. Sequence alignment of bEBPs highlighting L1, L2 loops as well as Q304 and the GAFTGA motif (red).
D. Transcriptional activation by NorR-Q304 variants *in vivo* as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x-axis. 'NorR' refers to the wild-type protein and 'NorRΔGAF' refers to the truncated form lacking the GAF domain (residues 1–170). Cultures were grown either in the absence (black bars) or presence (white bars) of 4 mM potassium nitrite, which induces endogenous NO production. Error-bars show the standard error of the three replicates carried out for each condition.

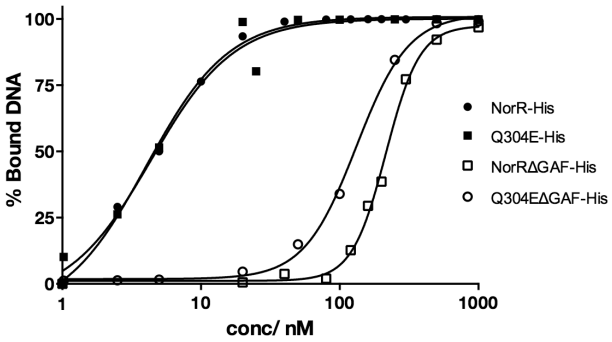

Fig. 3. DNA binding by the Q304E variant. Binding activity of the Q304E-His (closed squares) and Q304EΔGAF-His (open circles) variants compared to NorR-His (closed circles) and NorRΔGAF-His (open squares) to the 266 bp DNA fragment as determined by EMSA. The percentage of fully shifted DNA was quantified using a Fujix BAS 1000 phosphoimager.

domains (Ghosh *et al.*, 2010), we questioned whether the NorR-Q304E substitution might bypass the repressive function of the GAF domain by altering the assembly of higher order oligomers. Since binding of NorR to enhancer sites is essential for the formation of stable oligomers and enhancer DNA appears to be a key ligand in the activation of NorR as a transcription factor (Tucker *et al.*, 2010), we first investigated whether the Q304E mutation influences DNA binding. For this and subsequent biochemical experiments we utilised an N-terminal histidine tag as an aid to protein purification. The presence of this tag does not significantly affect the activity of either wild-type NorR or the Q304EΔGAF variant *in vivo* (see Fig. S1). Although the histidine tag apparently reduces the *in vivo* activity of Q304E, the ratio between the uninduced and NO-stimulated activities is similar to that of the non-tagged protein (Fig. S1). Electrophoretic mobility shift assays (EMSA) were performed to measure the binding of wild-type and the Q304E variant to enhancer DNA in either their full length or truncated forms lacking the regulatory GAF domain (ΔGAF). The affinity of NorR and NorRΔGAF for a 266 bp DNA fragment, containing all three enhancer sites, was not significantly influenced by the presence of the Q304E mutation (Fig. 3). Both NorR-His wild-type and the Q304E-His variant bound to the 266 bp fragment with a dissociation constant (Kd) of approximately 5 nM. Therefore, the Q304E variant does not alter the affinity of enhancer DNA binding. However, both NorRΔGAF-His and Q304EΔGAF-His exhibited decreased affinity for the 266 bp DNA fragment, which was estimated to be in the 150–200 nM range. Thus, the full-length proteins (both wild-type and variant) have a greater affinity for enhancer DNA than the truncated forms lacking the GAF domain (Fig. 3). Therefore the GAF domain appears to influence the DNA binding affinity of NorR.

The Q304E variant exhibits enhancer-independent activation in vivo *and DNA-independent ATPase activity* in vitro

Our previous analysis has demonstrated that all three enhancer sites upstream of the *norV* promoter are required both for transcriptional activation by wild-type NorR *in vivo* and to activate the ATPase activity of NorR *in vitro* (Tucker *et al.*, 2010). To examine the role of individual enhancer sites *in vivo*, we introduced NorR and its variants into *E. coli* strains containing mutant enhancers that prevent binding to individual enhancer sites (Tucker *et al.*, 2010). Consistent with previous results, transcriptional activation by either wild-type NorR or NorRΔGAF was significantly diminished when the consensus sequence of either binding site 1 (S1), site 2 (S2) or site 3 (S3) was altered (Fig. 4A). This was also observed with the Q304E variant under non-inducing conditions, indicating that the escape phenotype observed for this variant in the absence of an NO source is dependent on the presence of the three enhancer sites. Surprisingly however, when endogenous NO was present, transcriptional activation by the Q304E variant was not significantly affected by the presence of the enhancer mutations. This implies that when intramolecular repression by the GAF domain is fully relieved through the formation of the mononitrosyl iron complex, the Q304E variant can form a functional oligomer in the absence of at least one of the enhancer sites. Consistent with this observation, in the absence of the GAF domain, the Q304EΔGAF variant was competent to activate transcription in the S2 and S3 enhancer mutant strains, irrespective of the presence of a source of NO (Fig. 4A).

To further explore enhancer dependency, ATPase activity assays were performed *in vitro*. The wild-type form of NorR does not exhibit ATPase activity unless the ferrous iron centre in the GAF domain is activated by NO to form the mononitrosyl iron complex (D'Autreaux *et al.*, 2005). To enable comparisons under non-inducing conditions, we compared the ATPase activity of the Q304E variant with that of NorRΔGAF. As demonstrated previously this truncated protein exhibits activity in the absence of NO that is strongly stimulated by the presence of enhancer DNA (Fig. 4B). Likewise, the ATPase activity of the NorR-Q304E variant was stimulated by the presence of enhancer DNA. However, compared to NorRΔGAF, the cooperativity decreased, resulting in lower activity, particularly at low protein concentrations (Fig. 4 compare panels B and C). The concentration dependent behaviour of NorRΔGAF implies that self-association of the subunits is required for maximum activity. Hence, although NorR-Q304E retains the requirement for enhancer DNA for maximal ATPase activity, the effect of the enhancer sites is significantly reduced. Interestingly, full-length NorR-Q304E also exhibited ATPase activity in the absence of enhancer DNA. A

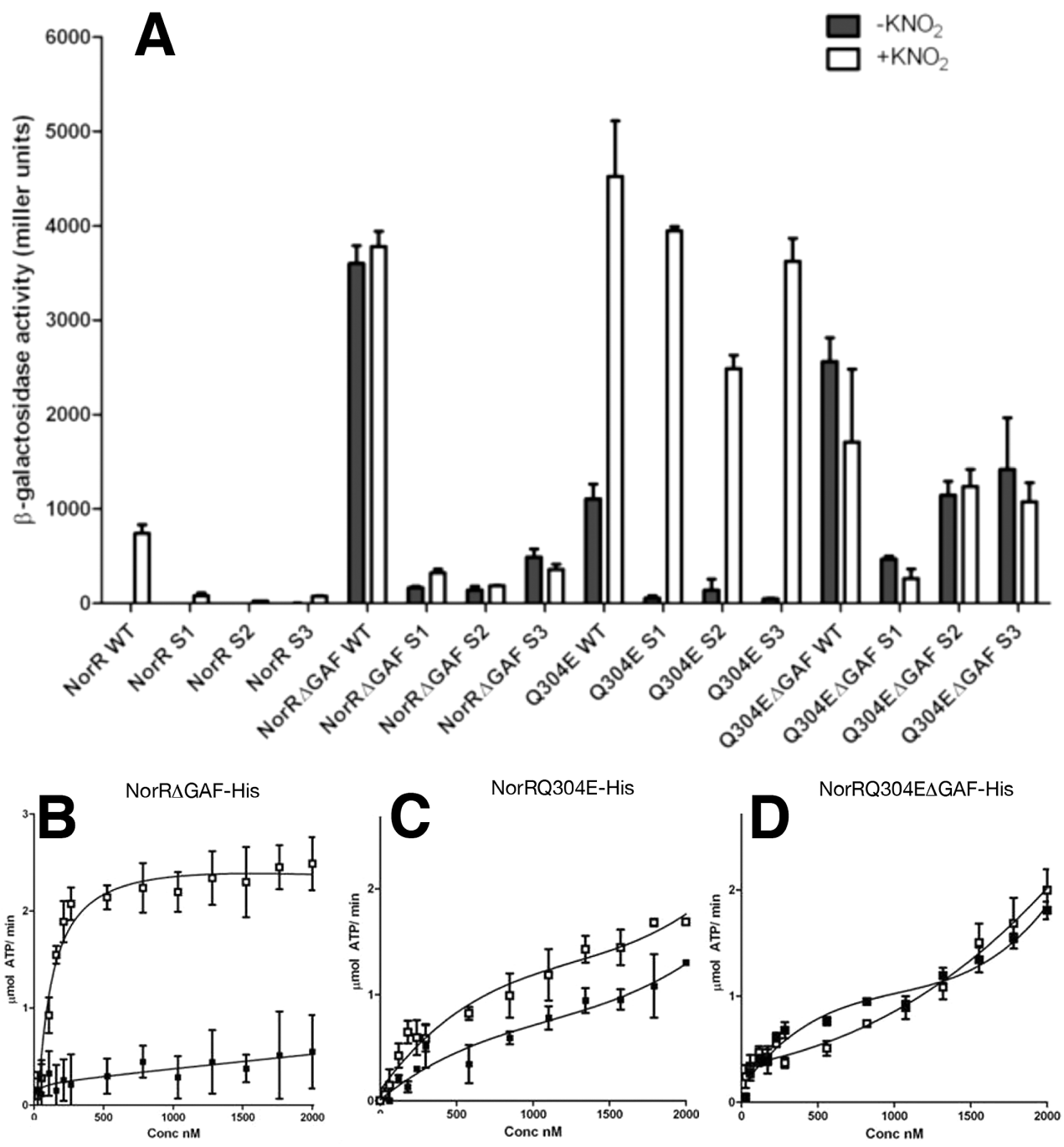

Fig. 4. *In vivo* and *in vitro* activities in the absence and presence of enhancer DNA.
A. Comparison of *in vivo* transcriptional activation by NorR and NorR-Q304E in *E. coli* strains with mutations in each of the NorR enhancer sites. NorR constructs were transformed into strains of *E. coli* with either three wild-type (WT) NorR binding sites (GT-(N7)-AC) or with one of three NorR binding sites (S1, S2, S3) altered to GG-(N7)-CC. Cultures were grown either in the absence (black bars) or presence (white bars) of 4 mM potassium nitrite. Error-bars show the standard error of the three replicates carried out for each condition.
B. ATPase activity of NorRΔGAF-His, in response to protein concentration and in the absence (closed squares) or presence (open squares) of enhancer DNA, C. same as B. for Q304E-His and D. for Q304EΔGAF-His. A 266 bp DNA fragment was used at a final concentration of 5 nM. Data are shown as the mean from at least two experiments.

similar level of enhancer-independent ATPase activity was observed with Q304EΔGAF (Fig. 4D). Therefore, the Q304E substitution appears to facilitate enhancer-independent ATPase activity, particularly in the absence of the GAF domain, but results in a decreased rate of ATP hydrolysis at low protein concentrations. To negate any possibility of ATPase activity from contaminants, a D286A substitution at the Walker B motif was combined with the Q304E mutation. As expected, this Walker B mutation significantly reduced the ATPase activities of the full length and truncated forms of the Q304E protein (see Fig. S2). Overall, these results suggest that the Q304E substitution increases the potential for NorR to form a functional oligomer in the absence of enhancer DNA, particularly when intramolecular repression by the GAF domain is relieved.

Negative stained EM studies of the full-length NorR-Q304E variant in the presence of enhancer DNA

In order to understand the molecular properties of the Q304E variant we carried out structural studies of this protein in complex with the 266 bp DNA fragment containing all three enhancer sites. The variant protein has the advantage that it forms a stable oligomeric complex in the presence of enhancer DNA (Fig. 5A), which most likely represents the 'activated-form' of NorR since the mutant protein has considerable activity in the absence of NO (Fig. 2D). We used negative stained electron microscopy for improved image contrast and analysed 5700 single particles using multivariant statistics analysis and classification followed by angular reconstitution implemented in Imagic (van Heel *et al*., 1996). Initial eigenimages show 6-fold symmetry, consistent with hexameric arrangement of AAA domains (Fig. S3). Although DNA is likely to be asymmetric in this complex, in order to reveal the relative arrangement between GAF, AAA and DBD of NorR and the general location of DNA relative to NorR protein, we applied 3 or 6-fold symmetry to subsequent image processing in order to increase signal-to-noise ratios and to improve the reconstruction, given that the AAA proteins are likely to be hexameric at the resolution achievable by negative stained electron microscopy.

The electron microscopy 3D reconstruction with 3-fold symmetry at 22 Å resolution is composed of two-stacked rings (Fig. 5 and Fig. S3). The top face of the ring structure (Fig. 5C) is consistent with six distinct subunits arranged around the symmetry axis, and measures 155 Å in diameter. When viewed from the side, the top half of the molecule appears dome-like and measures 54 Å in height (Fig. 5C). Six distinct density regions curve downwards to contact the bottom ring (Fig. 5C). These dimensions are comparable with 3D reconstructions of other hexameric AAA proteins, including that of p97 (ND1 and D2 rings measure 145–170 Å and ~120 Å in diameter, respectively) (Beuron *et al*., 2003) and MCM helicase (130–155 Å wide double hexamer) (Costa *et al*., 2006). We have therefore assigned the top ring to be the NorR hexamer. The other face of the structure (Fig. 5, C-*bottom view* and D-*III*) appears rounded, particularly the outer density, so that no clear hexameric pattern can be detected. The bottom ring measures 140 Å in diameter, with a very wide (94 Å) central opening (Fig. 5DIII).

A prominent feature of our reconstruction is that the central channel spanning the entire length (85 Å) of the NorR molecule is not hollow, unlike structures of other AAA proteins (Beuron *et al*., 2003; De Carlo *et al*., 2006; Costa *et al*., 2008; Ahuja *et al*., 2009). Six self-associating density lobes, arranged around the symmetry axis, sit below the top hexameric ring and just above the bottom ring (Fig. 5C, cut open view). This small ring-like mass has a diameter of 85 Å and a 26 Å wide central pore at the top and narrows at the bottom. Clear density continues above and below these lobes connecting them to the top and bottom faces of the molecule (Fig. 5C). Comparing our 3D map to that of the activated full-length NtrC bound to ADP.AlF$_x$ (Fig. 6, EMD-1218) (De Carlo *et al*., 2006), the only available structure of a full-length bEBP, shows that the overall architecture of the top-half of the NorR molecule is similar to that of the NtrC protein (see Fig. 6). However the NtrC structure is somewhat flat, with small protrusions above, whereas the NorR protein displays a dome-shaped ring assembly (side views). The small ring-like density found at the bottom of the hexameric ring in our model corresponds well to the DNA-binding (D) domain containing density in the NtrC structure (Fig. 6, *side views*). In the NorR reconstruction the small ring density sits further below the main ring, unlike the NtrC D domains that pack tightly against the ATPase domain ring in the ADP.AlF$_x$ bound state (Fig. 6). Therefore, we propose that the top half of the map displaying the clear hexameric ring density comprises the AAA domain and the N-terminal GAF domain. The smaller ring found in the middle possibly holds the DNA binding domains. The circular density at the bottom of our EM map (Fig. 5C bottom view and Fig. 5DIII) remains unaccounted for by NorR and was therefore assigned to the dsDNA encircling the hexameric ring. Consistent with this assignment, the diameter of the ring density measures ~25 Å, in agreement with the diameter of dsDNA.

Modelling and docking of NorR domains and dsDNA into the EM reconstruction

To obtain the relative orientation of the three NorR domains and dsDNA, we obtained homology models of the GAF, AAA and DNA-binding domains of NorR. In addition, a 100 bp circular DNA was constructed. The models were

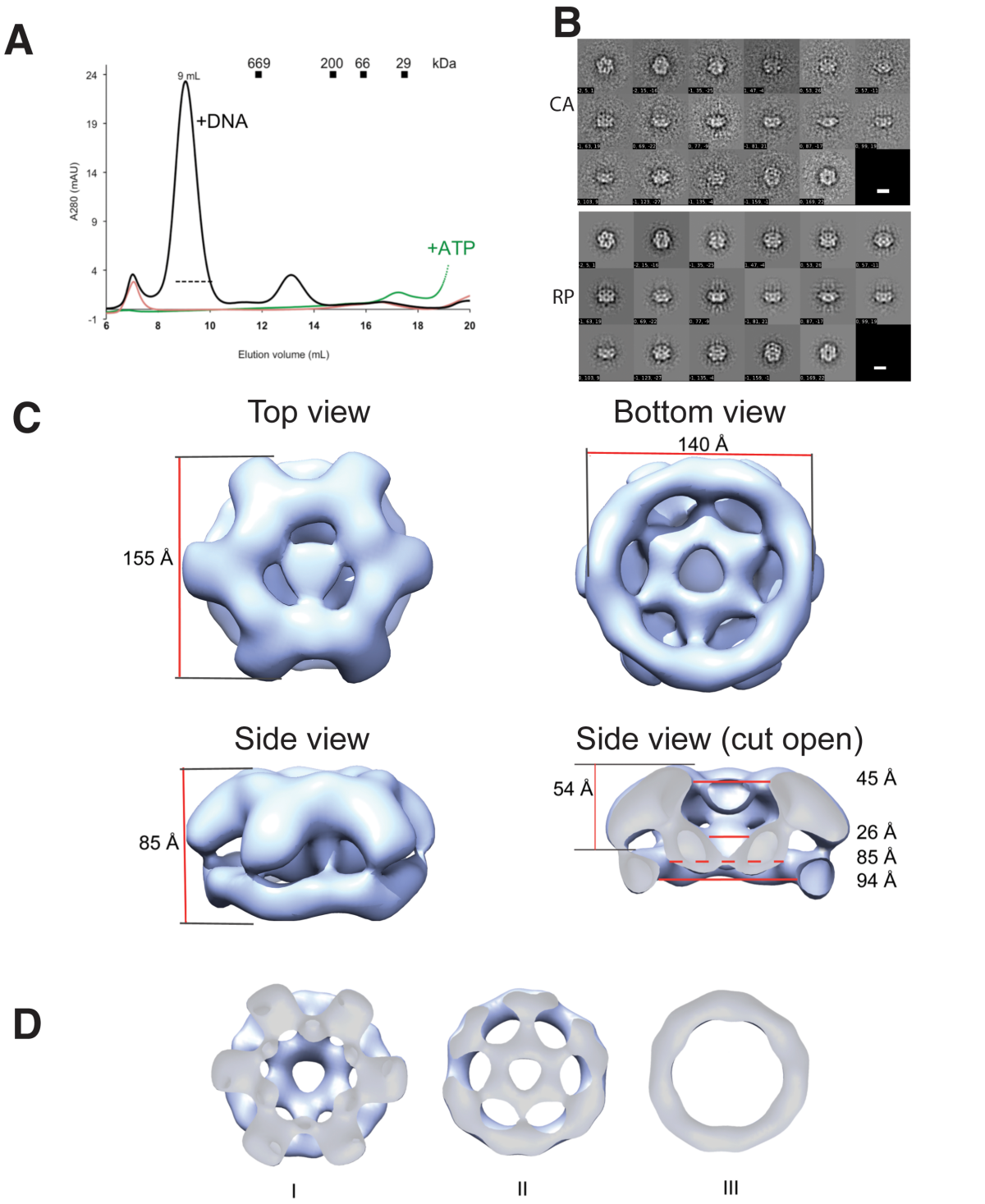

Fig. 5. 3D-reconstruction of NorR-Q304E bound to the 266 bp DNA fragment containing all three enhancer sites.
A. Gel filtration analysis of NorR-Q304E variant in complex with 266 bp DNA containing all three enhancer sites. A high molecular weight peak at 9 mL dominates, indicating higher order oligomer formation upon enhancer binding.
B. Class averages (CA) and reprojections (RP) of the 3D reconstruction along the same euler angles.
C. Surface representations of the 3D reconstruction shown in different orientations. The overall dimensions for the complex are given. A side-view has been cut-open to reveal the central chamber spanning the entire length of the molecule.
D. A view from the top to the bottom face of the nucleoprotein complex, along the symmetry axis. The protein monomers assemble into a hexameric ring around a wide central channel (I), with a clear asymmetric ring on the opposite face of the complex (III). Six distinct density regions lie just below the hexameric ring with clear connections to the bottom (II) rings in the structure.

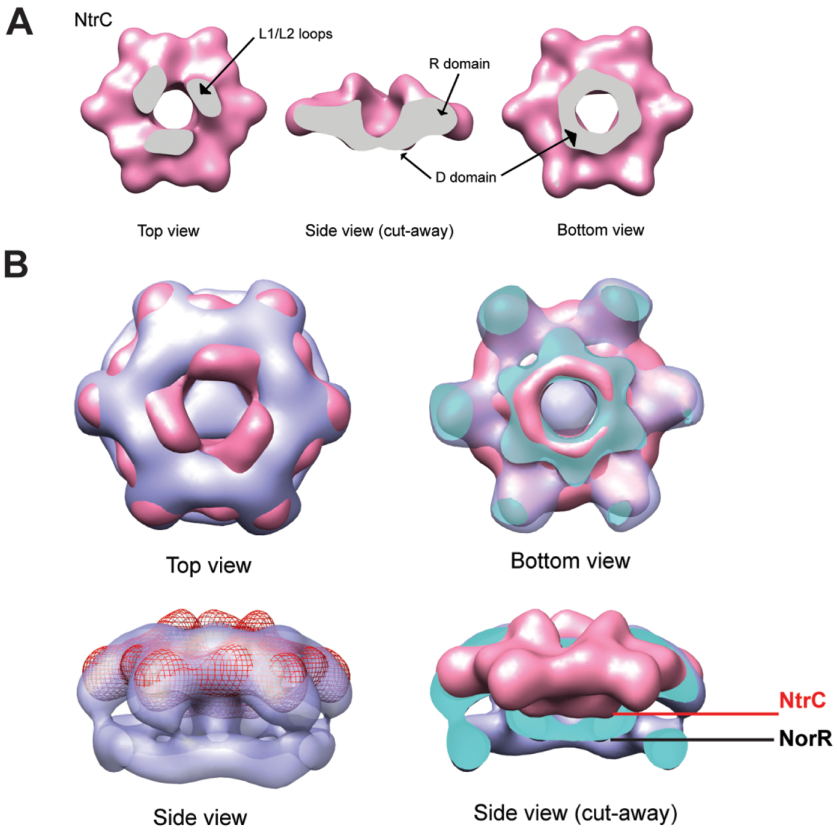

Fig. 6. Superimposed negative-stain EM maps of activated full-length NorR(Q304E) – DNA complex and NtrC-ADP.AlFx.
A. A surface representation of the 28 Å structure of NtrC (EMD 1218; De Carlo *et al.*, 2006) shown in different orientations. Positions of the N-terminal receiver (R) domain, the C-terminal DNA-binding (D) domains and the surface exposed L1/L2 loops of the AAA+ domain in the hexameric ring structure are indicated.
B. The superimposed maps (NtrC map is in pink or red mesh and NorR in blue) are shown in top, bottom and side view orientations. The NorR map was filtered to ~28 Å, and the top rings of the two maps were aligned in Chimera. Also shown is a cut-open side view (surface caps are in green) of the NorR mutant map to highlight the architectural similarity shared between the top-half of the NorR molecule and the NtrC structure.

then manually fitted into the EM density map. The monomeric models of NorR GAF and AAA domains were manually fitted as rigid bodies into the hexameric ring density (top half of our map), such that the AAA domain occupied the central density of the ring (Fig. 7A). In this orientation the L1/L2 loops are surface exposed (Fig. 7A and D) and can project outwards in an ATP-dependent fashion to contact σ^{54}, as seen in the cryo-EM model of the PspF AAA

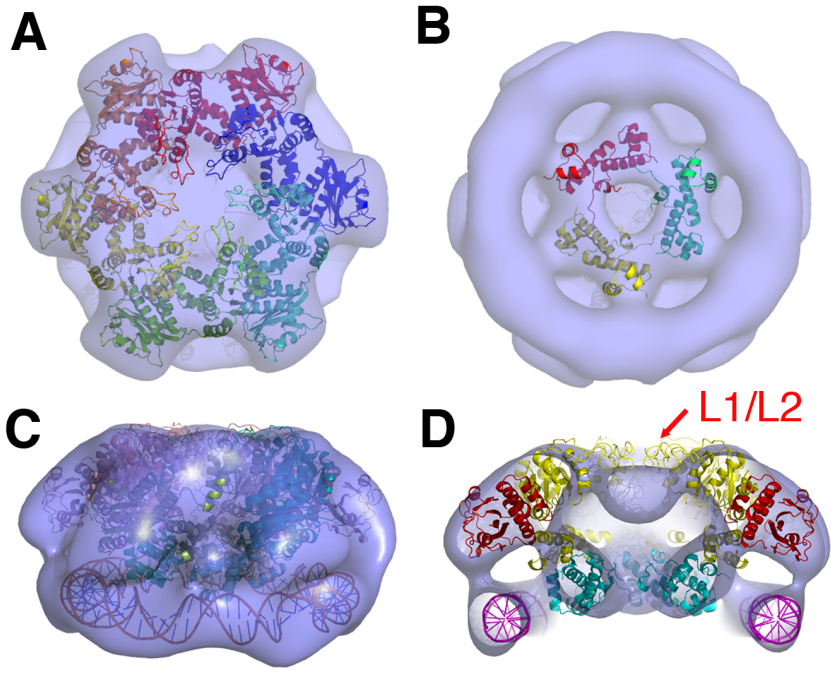

Fig. 7. Fitting and assigning NorR domains into the EM density.
A. Top view of the six homology models of AAA and GAF domains fitted into the EM density.
B. Three pairs of DNA binding domains fitted into the inner density viewed from the bottom.
C. A 100 bp circular dsDNA fitted into the density below the NorR hexamer.
D. Cut-open view of AAA (yellow), GAF (red) and DNA binding domains (cyan) fitted into the EM density.

hexamer-σ^{54} complex formed in the presence of the transition state analogue, ADP.AlF$_x$ (Rappas *et al*., 2005). However, in the current EM reconstruction, we do not observe protruding densities for L1/L2 loops such as those observed in NtrC-ADP.AlF$_x$ complex (Fig. S3). This is consistent with the fact that L1/L2 loops are proposed to project outwards in the presence of ATP or ADP.AlF$_x$ (Rappas *et al*., 2006). The GAF domain fitted into the arm-like density that originates at the periphery of the central density and then curves down to contact the rounded ring density occupying the bottom half of the map (Figs 7A and 7D). The DNA binding domain was fitted manually into the small density lobe located below the AAA ring density (Fig. 7B, 7D). Since DNA binding domains have been shown to form dimers and palindromic sequence exists in all three enhancer sites (Tucker *et al*., 2004), we therefore fitted a dimer of DNA binding domains from NtrC4 as a unit (PDB code 4FTH). Three pairs of the dimer account well for the small ring density found below the AAA hexamer ring plane (Fig. 7B). A hexamer model of full-length NorR was subsequently generated by applying 6-fold symmetry for the GAF and AAA domain models and 3-fold symmetry for DNA binding domains. Together, the AAA and GAF domains account well for the dome-like shape of the top half of the Q304E-DNA reconstruction (Fig. 7A and C). A circular dsDNA model fitted nicely in the bottom half of the map (Fig. 7). However, the exact DNA path is unknown as the imposed symmetry eliminates any differences between the DNA density relative to the hexameric protein. To improve the fitted protein model, the manually docked model underwent two stages of refinement, first generating related structures by optimising the agreement with the EM density and other physical constraints, then by energy minimisation including charge and shape complementarity, and finally clustering and selecting the final structure by cluster size, structural integrity and the complementarity of the inter-domain interactions. Compared to the initial fitted model, the correlation coefficient between the synthetic density and the cryo-EM density improved from 0.44 to 0.70.

Our reconstruction and model show that in the 'activated-state' of NorR, the GAF domain is located at the periphery of the AAA domain while the DNA binding domain is located below the AAA domain, occupying the space protected by the GAF domains. Interestingly, dsDNA is not located directly below the DNA binding domain, but just below the outer rim of the NorR hexamer, almost immediately below the GAF domains (Fig. 7D). The whole assembly is a very tightly organised complex, with the GAF domain contacting both the AAA domain and the DNA binding domain, as well as potentially the DNA itself. Likewise, the DNA binding domain is connected to the AAA domain and the GAF domain as well as the DNA. As noted above, NorRΔGAF showed reduced affinity for the enhancer sites, consistent with a model in which the GAF domain may contribute to DNA binding (Fig. 3A).

Discussion

The NorR oligomer is stabilised by encircling DNA

One unusual property of NorR compared to other bEBPs is the absolute requirement of binding to all three enhancer sites for oligomerisation and other functionalities. Our data here show that dsDNA containing the three NorR-binding sites encircle the NorR hexameric ring. Wrapping of dsDNA around the oligomeric complex would therefore stabilise the hexameric assembly of the ATPase domains. This is in line with previous negative-stained EM and biochemical results that showed hexameric ring formation and maximal ATP hydrolysis by NorRΔGAF requires the presence of the 266 bp DNA fragment carrying the three enhancers (Tucker *et al*., 2010). Based on our reconstruction we estimated that a minimum of 350 Å or ~100 bp DNA is required to encircle the ring. This is consistent with our previous findings that a shorter 66 bp fragment, which contains the three enhancer sites, is insufficient to stabilise a hexameric ring of NorRΔGAF, which further translates to the inability of small DNA fragments containing a single NorR-binding site to stimulate the ATPase activity (Tucker *et al*., 2010). The extra 160 bp DNA that is not wrapped around the NorR hexamer is not visible in our reconstruction due to the flexible nature of free DNA, which would have been averaged out during image processing.

The enhanced activity of Q304E variant is due to partial escape of GAF domain repression as well as enhancer-independent ATPase activity

The escape phenotype of the Q304E variant is different to that of substitutions we previously characterised in the highly conserved GAFTGA motif located in the L1 loop, which completely bypass repression mediated by the GAF domain. GAFTGA motifs are located on the surface exposed L1 loops on the top of the AAA ring and are responsible for interaction with σ^{54}-RNAP (Figs 2C and 7D). The properties of these GAFTGA substitutions suggested that repression by the GAF domain targets the surface exposed LI/L2 loops and hence prevents access of σ^{54}-RNAP to this interaction surface (Bush *et al*., 2010). Our structural model suggests that the GAF domains in the ATPase-competent Q304E-DNA hexamer are located at the periphery of the AAA ring, making the L1/L2 loops accessible to σ^{54}-RNAP, resulting in active transcription in the absence of NO signal (Fig. 7D). Our EM studies

confirm that DNA binding *per se* increases the tendency of wild-type NorR to form higher order oligomers. This results in stimulation of ATPase activity, suggesting that DNA binding induces conformational changes that promote oligomerisation and catalytic activity (Bush *et al*., 2010; Tucker *et al*., 2010). It is possible that the Q304E mutation itself promotes such a conformation, hence explaining the reduced dependence on enhancer DNA for ATPase activity and *in vivo* activity. The equivalent residue to Q304 (Fig. 2A) in other bEBPs such as NtrC1 and PspF is glutamic acid rather than glutamine (Fig. 2C), which is well placed to interact with one of the R fingers (R168 in PspF and R347 in NorR) (Rappas *et al*., 2006; Chen *et al*., 2010). R finger residues play important roles in oligomerisation and catalytic activity, as mutating the equivalent residue in PspF (R168) to Ala results in constitutive hexamers, which are defective in ATPase activity (Schumacher *et al*., 2004; Rappas *et al*., 2005). It is therefore highly plausible that the Q304 residue, which is highly conserved in NorR homologs, may play important roles in regulating ATPase activity and formation of hexamers through the control of the R finger residue.

Interestingly, although in the absence of NO, Q304E exhibited similar levels of activity to that of wild type protein in the 'activated-form', the addition of inducing agent further stimulated transcriptional activation by the Q304E variant, resulting in a 'hyperactive' state (Fig. 2D). This would suggest that NO interactions with the GAF domain result in a different conformational state of the Q304E variant that is favourable for its activities. This 'hyperactive' state appears to tolerate at least one defective enhancer site (Fig. 4A), implying that this altered conformation can compensate for the lack of enhancer binding.

Multiple roles for the GAF domain

Our previous studies indicate that the GAF domain regulates NorR activity by blocking the access of the L1 and L2 loops to the RNAP-σ^{54} holoenzyme (Bush *et al*., 2010). Interaction of NO with the iron centre in the GAF domain may induce changes that relocate this regulatory domain, releasing the inhibition on the AAA domain. In the absence of NO, the Q304E variant has similar activity to that of wild-type NorR when NO is present (Fig. 2D), suggesting that Q304E likely represents the 'activated-form' of the wild type protein.

In addition to the repressive functionality of the GAF domain, our structure of the Q304E-DNA complex indicates that the GAF domain interacts with the DNA binding domain and potentially with the enhancer DNA itself. This implies that the GAF domain could play a role in controlling the affinity of NorR for its DNA target sites. Indeed, NorRΔGAF has reduced affinity for DNA compared to full length NorR. This may help to ensure enhancer dependency by preventing runaway assembly of the hexamer off the DNA and consequent *cis* activation of non-cognate σ^{54}-promoters. These inter-dependent interactions emphasise the highly synergistic domain functionality that has evolved amongst the bEBP family.

A model for NorR assembly and activation

The GAF domain-mediated mechanism of AAA repression in NorR does not involve prevention of hexamer formation as in the case of other bEBPs such as NtrC1, DctD or other GAF containing NifA-like homologs (Lee *et al*., 2003; Doucleff *et al*., 2005b; Batchelor *et al*., 2013). Instead, it is likely that the GAF domain of NorR negatively regulates the AAA hexamer by preventing access of the L1 and L2 loops to σ^{54}-RNAP. This repression mechanism might also serve to lock the loops in a restrained conformation that feeds back to the nucleotide-binding site to prevent ATP hydrolysis. This model of repression allows for the pre-assembly of a hexameric bEBP, which in the case of NorR, may allow the cell to rapidly respond to NO-induced stress.

Our study confirms that enhancer DNA is bound to the opposite face of the AAA domain from its σ^{54}-interaction surface (Fig. 7D) although the exact path of the DNA is unclear. Earlier mutagenesis data identified G266, the second G in the signature GAFTGA motif (Fig. 2C), as being involved in the interaction with the regulatory GAF domain in the 'pre-activated' state. This suggests that prior to activation, the GAF domain is located above the AAA ring. Our studies on full length wild-type NorR suggest that prior to activation, upon binding to enhancers, NorR forms higher order oligomers, most likely hexamers, which are stabilised by DNA wrapping. However, NorR remains inactive by blocking the interaction surface of σ^{54}-RNAP via the GAF domains. Upon binding to NO signal, the GAF domains relocate downwards to the side of the AAA ring, making the L1/L2 loops of the AAA domains accessible for σ^{54} interaction and competent to carry out nucleotide-dependent remodelling. Our results thus provide a conceptual model for NorR activation. However, the exact molecular basis for this activation awaits further high resolution structural information on the 'pre-activated' and 'activated-form' of NorR.

Experimental procedures

Assaying NorR activity in vivo

Transcriptional activation by NorR *in vivo* was measured by introducing wild-type and mutant plasmids into MH1003 (Hutchings *et al*., 2002; D'Autreaux *et al*., 2005). Cultures were prepared as described previously (D'Autreaux *et al*., 2005). Levels of expression of the *norV-lacZ* fusion were then

determined by assaying β-galactosidase activity as previously described (Tucker *et al.*, 2008).

Protein purification

Non his-tagged NorR was over-expressed and purified as described previously (D'Autreaux *et al.*, 2005). His-tagged full-length and GAF domain deleted forms of wild-type and Q304E substituted NorR (Bush *et al.*, 2010), with an additional N-terminal TEV cleavable His-tag, were overexpressed from pET-M11 (EMBL) but with the NcoI site altered to an NdeI site, to allow easy cloning of the *norR* sequence. The proteins were purified by nickel affinity chromatography and gel filtration (Bush *et al.*, 2010).

ATPase assays

ATPase activities were measured using the coupled assay as described previously (Tucker *et al.*, 2010).

Analytical gel filtration

Gel filtration chromatography of His-tagged NorR-Q304E and Q304EΔGAF proteins alone and in complex with a 266 bp DNA fragment, containing all three enhancer sites, were performed using a Superose 6 column (10 × 300mm, 24 ml) as described previously (Tucker *et al.*, 2010). The DNA fragment was generated by PCR as described previously (Tucker *et al.*, 2010).

Negative-stain electron microscopy and image processing

To prepare the NorR-Q304E nucleoprotein complex, a solution containing 2 μM NorR-Q304E protein and 0.6 μM 266 bp dsDNA fragment (molar ratio of 12:1 protein monomer: DNA) in buffer containing 10 mM Tris-HCl pH 8.5, 50 mM NaCl, 8 mM $MgCl_2$ was incubated for 2 min at 37°C. The samples were then cooled at room temperature for 5 min and left on ice. 2 μl of the samples were applied on the glow-discharged grids and stained with 2% uranyl acetate. Data were collected on the Phillips CM200 FEG electron microscope operating at 200 kV, at a magnification of 50 000×. Micrographs were recorded directly on a 4k × 4k CCD camera (TVIPS, Germany), giving a pixel size of 1.76 Å on the specimen scale, which was then coarsened to 3.52 Å/pixel.

Single particles were selected interactively within the program Boxer (within EMAN), with poor quality particles or 'junk' discarded manually. The dataset was then extracted from the coarsened micrographs using the particle coordinates. Contrast-transfer-function (CTF) correction was applied to images extracted into 512 × 512 pixel boxes, before reducing the box size to 128 × 128 pixels. The complete picked dataset contained 5700 particles. All further image processing was done using the IMAGIC-5 software (van Heel *et al.*, 2000). The raw particles were initially band-pass filtered with a low and high frequency cutoffs of 150 Å and 10 Å, respectively, centred (without using reference images) and subjected to multivariate statistical analysis (MSA), hierarchical ascendant classification and averaging (van Heel *et al.*, 2000). The best classums that represent different projections were used as references for multi-reference alignment (MRA). A brute force alignment program (program written and kindly lent by Dr. Timothy Grant, CBEM, Imperial College) that uses a modified version of the MRA algorithm in IMAGIC-5 was used following alignment. Inspection of eigenimages after MSA on translationally aligned particles indicated the presence of a sixfold symmetry component. This confirms that NorR-Q304E assembles into a hexamer upon enhancer DNA binding. In order to obtain the organisation of NorR hexamer and the relative position of DNA, a preliminary 3D reconstruction was calculated with a three-fold rotational symmetry to reflect three enhancer sites although the 266 bp DNA should impose some degree of asymmetry in the complex. After iterative refinement, a stable 3D reconstruction was obtained. Non-symmetrised data processing was carried out independently and this reconstruction shows similar features although at a lower resolution due to the lack of symmetry. There are some asymmetric features in the non-symmetrised reconstruction but the details are insufficient to provide meaningful interpretations of the differences between different protomers and we therefore focus our discussions on the domain relationship between NorR and DNA which are consistent between the 3-fold and non-symmetrised reconstructions.

Docking and refinement of the high resolution structural models into the EM density

Models of individual NorR domains were generated based on crystal structures of *A. aeolicus* NtrC1 AAA+ domain bound to ADP (PDB code: 1NY6), DNA-binding domain of ZraR (PDB code: 1OJL, chain A) and GAF domain from Acinetobacter phosphoenolpyruvate-protein phosphotransferase (PDB code: 3CI6); NorR AAA+ domain (residues 190–417) shares 48% sequence identity with $NtrC1^C$ (residues 142–369). Each domain was initially fitted individually into the EM density as a rigid body.

To improve the fitted model, the manually docked model underwent two stages of refinement. In the first stage, a modified version of the Integrative Modelling Platform package [IMP_REF] was used to generate 10,000 structures optimally consistent with the EM density map and a number of other constraints. First, the model was coarse grained as one sphere per residue. Monte Carlo simulated annealing was used to optimise the position and orientation of the GAF, α/β, α-helical and DNA binding domains, all of which were allowed to move independently as rigid bodies. To reduce the dimensionality of the search, C_3 symmetry was imposed such that only two adjacent AAA monomers were adjusted with the remainder reconstructed via C_3 and C_3^2 symmetry operations. The DNA, which already fitted well within the density map, was kept fixed. The agreement of the models with the cryo-EM data was determined as the overlap score of their synthetic density map, at 20 Å resolution, with the experimental density, as described by Lakser *et al.* [MULTIFIT_REF]. This term was weighted with a factor of 10 000. Hinges connecting the GAF and α/β domains, and between the α/β and α-helical domains, were modelled as harmonic springs beyond an upper bound C_α-C_α distance of 5 Å ($k = 2.0$ Å^{-1}). Further, an excluded volume restraint was included to prevent clashes; interpen-

etration of residue spheres was penalised with a harmonic potential (k = 5.0 Å^{-1}). Connectivity restraints were included to ensure contact between the DNA and the DNA binding domains, as well as between adjacent α/β domains; connected domain pairs which do not have at least one residue centre within 10 Å of the other were penalised with a harmonic penalty (k = 0.1 Å^{-1}). For each Monte Carlo step, each domain was moved in a random direction by an amount taken from a uniform distribution in the range of [0, 1] Ångstroms, and rotated around a random axis by an angle taken from the uniform distribution of [0, 0.6] radians. The move was accepted or rejected according to the Metropolis criterion, with pseudo-temperature decreasing linearly from 50 to 0 over 5000 iterations. The algorithm was run 10 000 times. The cross-correlation of the synthetic EM densities improved from and initial value of 0.44 to typically 0.65–0.70. In the second stage of refinement, each structure was reverted back to atomic resolution and extensively minimised using the CHARMM27 force field [CHARMM_REF]; 1000 steps of steepest descent to remove clashes and 1000 steps conjugate gradient for finer tuning. The structural integrity of the models and the physical complementarity between the domains was calculated by evaluating the energy of each structure using the DFire statistical pair potential [DFIRE_REF]. Finally, the lowest energy 1000 structures were clustered at 5 Å resolution. A single large cluster of 834 structures was found, indicating a consensus between the low energy models. Of this cluster, the lowest energy structure was selected as the final model.

Acknowledgements

We would like to thank Suhail Islam for providing the circular dsDNA model. TG is funded by Imperial College London and Overseas Research Scholarship. MB and RD were supported by grants BB/J004561 and BB/D009588 from the Biotechnology and Biological Sciences Research Council and by the John Innes Foundation. This work was also in part supported by BBSRC grant BB/H012249/1 to XZ.

References

Ahuja, D., Rathi, A.V., Greer, A.E., Chen, X.S., and Pipas, J.M. (2009) A structure-guided mutational analysis of simian virus 40 large T antigen: identification of surface residues required for viral replication and transformation. *J Virol* **83:** 8781–8788.

Batchelor, J.D., Lee, P.S., Wang, A.C., Doucleff, M., and Wemmer, D.E. (2013) Structural Mechanism of GAF-Regulated sigma(54) Activators from Aquifex aeolicus. *J Mol Biol* **425:** 156–170.

Beuron, F., Flynn, T.C., Ma, J., Kondo, H., Zhang, X., and Freemont, P.S. (2003) Motions and negative cooperativity between p97 domains revealed by cryo-electron microscopy and quantised elastic deformational model. *J Mol Biol* **327:** 619–629.

Browning, D.F., and Busby, S.J. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2:** 57–65.

Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y., and Gralla, J.D. (2000) The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* **182:** 4129–4136.

Bush, M., and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. *Microbiol Molecul Biol Rev* **76:** 497–529.

Bush, M., Ghosh, T., Tucker, N., Zhang, X., and Dixon, R. (2010) Nitric oxide-responsive interdomain regulation targets the sigma54-interaction surface in the enhancer binding protein NorR. *Mol Microbiol* **77:** 1278–1288.

Chen, B., Sysoeva, T.A., Chowdhury, S., Guo, L., De Carlo, S., Hanson, J.A., *et al*. (2010) Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase. *Structure* **18:** 1420–1430.

Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A., and Onesti, S. (2006) Structural basis of the Methanothermobacter thermautotrophicus MCM helicase activity. *Nucleic Acids Res* **34:** 5829–5838.

Costa, A., van Duinen, G., Medagli, B., Chong, J., Sakakibara, N., Kelman, Z., *et al*. (2008) Cryo-electron microscopy reveals a novel DNA-binding site on the MCM helicase. *EMBO J* **27:** 2250–2258.

D'Autreaux, B., Tucker, N.P., Dixon, R., and Spiro, S. (2005) A non-haem iron centre in the transcription factor NorR senses nitric oxide. *Nature* **437:** 769–772.

De Carlo, S., Chen, B., Hoover, T.R., Kondrashkina, E., Nogales, E., and Nixon, B.T. (2006) The structural basis for regulated assembly and function of the transcriptional activator NtrC. *Genes Dev* **20:** 1485–1495.

Doucleff, M., Chen, B., Maris, A.E., Wemmer, D.E., Kondrashkina, E., and Nixon, B.T. (2005a) Negative regulation of AAA+ ATPase assembly by two component receiver domains: a transcription activation mechanism that is conserved in mesophilic and extremely hyperthermophilic bacteria. *J Mol Biol* **353:** 242–255.

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

Ghosh, T., Bose, D., and Zhang, X. (2010) Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiol Rev* **34:** 611–627.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* **116:** 17–24.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., *et al*. (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* **33:** 307–369.

Hutchings, M.I., Mandhana, N., and Spiro, S. (2002) The NorR protein of Escherichia coli activates expression of the flavorubredoxin gene norV in response to reactive nitrogen species. *J Bacteriol* **184:** 4640–4643.

Lee, S.Y., De La Torre, A., Yan, D., Kustu, S., Nixon, B.T., and Wemmer, D.E. (2003) Regulation of the transcriptional activator NtrC1: structural studies of the regulatory and AAA+ ATPase domains. *Genes Dev* **17:** 2552–2563.

Rappas, M., Schumacher, J., Beuron, F., Niwa, H., Bordes, P., Wigneshweraraj, S., *et al*. (2005) Structural insights into the activity of enhancer-binding proteins. *Science* **307:** 1972–1975.

Rappas, M., Schumacher, J., Niwa, H., Buck, M., and Zhang, X. (2006) Structural basis of the nucleotide driven confor-

mational changes in the AAA+ domain of transcription activator PspF. *J Mol Biol* **357:** 481–492.
Rappas, M., Bose, D., and Zhang, X. (2007) Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. *Curr Opin Struct Biol* **17:** 110–116.
Schumacher, J., Zhang, X., Jones, S., Bordes, P., and Buck, M. (2004) ATP-dependent transcriptional activation by bacterial PspF AAA+ protein. *J Mol Biol* **338:** 863–875.
Tucker, N.P., D'Autreaux, B., Studholme, D.J., Spiro, S., and Dixon, R. (2004) DNA binding activity of the Escherichia coli nitric oxide sensor NorR suggests a conserved target sequence in diverse proteobacteria. *J Bacteriol* **186:** 6656–6660.
Tucker, N.P., D'Autreaux, B., Yousafzai, F.K., Fairhurst, S.A., Spiro, S., and Dixon, R. (2008) Analysis of the nitric oxide-sensing non-heme iron center in the NorR regulatory protein. *J Biol Chem* **283:** 908–918.
Tucker, N.P., Ghosh, T., Bush, M., Zhang, X., and Dixon, R. (2010) Essential roles of three enhancer sites in sigma54-dependent transcription by the nitric oxide sensing regulatory protein NorR. *Nucleic Acids Res* **38:** 1182–1194.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.