Interdomain Repression in the Enhancer Binding Protein NorR

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy.

By

Matthew James Bush

Department of Molecular Microbiology

John Innes Centre

Norwich Research Park, Colney Lane, Norwich, NR4 7UH

September 2011

Acknowledgements

First of all I would like to thank Prof. Ray Dixon (FRS), my supervisor during the period of this research. His advice, guidance and constant encouragement have been invaluable. I would also like to thank past and present members of the Dixon Lab: Dr. Nick Tucker, Dr. Peter Slavny, Dr. Richard Little, Dr. Anton Shmelev and all of the visiting workers that have contributed to my development as a research scientist. Significant thanks go to Nick Tucker whose previous work on NorR has been the precursor to a productive PhD project. Tamaswati Ghosh and Xiaodong Zhang (Imperial College, London) deserve considerable recognition for a successful and thoroughly enjoyable collaboration. I reserve my biggest thanks to Richard who has been largely responsible for my professional development. His expertise and unwavering patience have allowed me to make the most of the last four years in the department of Molecular Microbiology. I would particularly like to thank Richard as well as Jason Terpolilli and Andrzej Tkacz for their friendship and the many laughs we have shared over a cup of tea. I am immensely grateful to those friends and family who have encouraged me throughout my studies; especially my wife, Tor, whose love and support has sustained me each day. This work is for you and our beautiful daughter Elsa.

Abstract

NO (nitric oxide) is an intermediate of respiratory denitrification and is one of the toxic species released by macrophages of the immune system in the defence against invading pathogenic bacteria. In *Escherichia coli*, the expression of the Nitric Oxide (NO) reductase (NorVW) is tightly regulated by NorR, a member of the bacterial Enhancer Binding Protein (bEBP)-family that activates σ^{54}-dependent transcription of the *norVW* genes under conditions of nitrosative stress. Binding of NorR to three conserved enhancer sites upstream of the *norVW* promoter is essential for transcriptional activation and promotes the formation of a stable higher-order nucleoprotein complex. NorR falls into a class of bEBPs that are negatively regulated – the regulatory (GAF) domain represses the activity of the ATPase (AAA+) domain in the absence of NO. NO binds to the non-heme iron centre of the GAF domain, stimulating ATP hydrolysis by the AAA+ domain and establishing an interaction between the activator and σ^{54} that leads to the remodelling of the closed promoter complex. However, the route by which NorR couples signal sensing to substrate remodelling is unknown. Here, the mechanism of interdomain repression in NorR has been investigated by characterising substitutions in the AAA+ domain that bypass repression by the regulatory domain. Most of these substitutions are located in the vicinity of the surface-exposed loops that engage σ^{54} during the ATP hydrolysis cycle or in the highly conserved GAFTGA motif that directly contacts σ^{54}. A combination of genetic and biochemical approaches were used to show that the regulatory domain of NorR is unlikely to control AAA+ activity using previously characterised mechanisms, employed by related bEBPs. Instead, this work identifies a novel mechanism in which the σ^{54}-interaction surface of the AAA+ domain is a target of the GAF-mediated repression mechanism. This hypothesis is further supported by EM-reconstructions of two characterised escape-variants in their on-states, one of which represents the first structure of a bEBP bound to enhancer DNA. In the case of NorR, regulation at the point of σ^{54}-interaction may be linked to the pre-assembly of an inactive hexamer, "poised" at the enhancer sites, enabling the cell to rapidly respond to nitrosative stress.

Publications (see appendix 12.2)

Tucker, N. P., Ghosh, T., Bush, M., Zhang, X., Dixon, R. (2010). Essential roles of three enhancer sites in sigma54-dependent transcription by the nitric oxide sensing regulatory protein NorR. Nuc. Acids Res. **38,** 1182-94

Bush, M., Ghosh, T., Tucker, N., Zhang, X., Dixon, R. (2010). Nitric oxide-responsive interdomain regulation targets the sigma(54) -interaction surface in the enhancer binding protein NorR. Mol Microbiol. **73**, 1278-1288

Bush, M., Ghosh, T., Tucker, N., Zhang, X., Dixon, R. (2011). Transcriptional regulation by the dedicated nitric-oxide sensor, NorR: a route towards NO detoxification. Biochem. Soc. Trans. **39,** 289-293

Ghosh, T.*, Bush, M.*, Sarwick, M., Moal, I., Bates, P., Dixon, R., Zhang, X. The structural basis for enhancer-dependent assembly and activation of the AAA transcriptional activator NorR. *Manuscript in preparation.*

General Abbreviations

σ	Sigma factor
σ^{54}	Alternate sigma factor 54
σ^{70}	Housekeeping sigma factor 70
2-OG	2-oxoglutarate
AAA+	ATPases Associated with various cellular Activities
ACT	Aspartokinase, Chorismate mutase and TyrA
ADP	Adenosine Diphosphate
ADP.AlF$_x$	ADP Aluminium Flouride
ADP.BeF$_x$	ADP Beryllium Flouride
APS	Ammonium Persulphate
ATP	Adenosine Triphosphate
bEBP	bacterial Enhancer Binding Protein
BSA	Bovine Serum Albumin
BsNsrR	*Bacillis subtilis* NsrR
cAMP	cyclic Adenosine Monophosphate
CC	Closed Complex
C domain	Central domain
cGMP	cyclic Guanosine Monophosphate
Cra	Catabolite repressor-activator
CRP	cAMP Receptor Protein
CTD	C-terminal Domain
D domain	DNA-binding domain
DctD	C4-Dicarboxylic Acid Transport Protein D
DmpR	3,4-demethylphenol Catabolism Regulatory Protein
DNA	Deoxyribonucleic Acid
DNIC	Dinitrosyl-iron complex
dNTP	Deoxy Nucleotide Triphosphate
DTT	Dithiothreitol
E	RNA Polymerase (core)
E$\sigma^{54/70}$	RNA Polymerase (sigma factor 54/70 holoenzyme)
EBP	Enhancer Binding Protein
ECF	Extracytoplasmic Function
EDTA	Ethylenediaminetetraacetic acid
EM	Electron Microscopy
EMSA	Electrophoretic Mobility Shift Assay
EPR	Electron Paramagnetic Resonance
ES-MS	Electrospray-Mass Spectrometry
FAD	Flavin Adenine Dinucleotide
Fe-S	Iron-Sulfur
FhlA	Formate-Hydrogen Lyase Activator Protein
FIS	Factor for Inversion Stimulation
FlgR	Flagella Gene Regulator
FMN	Flavin Mononucleotide
FNR	Fumarate and nitrate reductase regulator
Fur	Ferric Uptake Regulator
GAF	cGMP-specific and stimulated phosphodiesterases, Anabaena adenylate cyclases and *E. coli* FhlA

GSNO	*S*-nitrosoglutathione
Hcp	Hybrid cluster protein
Hcy	Homocysteine
HK	Histidine Kinase
Hrp	Hypersensitive response and pathogenicity
HTH	Helix-Turn-Helix
IHF	Integration Host Factor
IPTG	Isopropyl-β-*D*-thiogalactopyranoside
IscR	Iron-Sulfur Cluster Regulator
LB	Luria-Bertani broth
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization-Time Of Flight
MCS	Multiple Cloning Site
NAD	Nicotinamide Adenine Dinucleotide
NifA	Nitrogen Fixation Regulatory Protein A
NgNsrR	*Neisseria gonorrhoeae* NsrR
NMR	Nuclear Magnetic Resonance
Nnr	Nitrate and nitrite reducase regulator
NO	Nitric Oxide
NOD	Nitric Oxide Dioxygenase
NOR	Nitric Oxide Reductase
NorR	Nitric Oxide Reductase Regulator
NorRΔGAF	NorR lacking the N-terminal, Regulatory GAF domain
NOS	Nitric Oxide Synthase
NTD	N-terminal Domain
NTP	Nucleotide Triphosphate
NtrC	Nitrogen Regulatory Protein C
NtrC1	Nitrogen Regulatory Protein C 1
NtrC4	Nitrogen Regulatory Protein C 4
OC	Open Complex
ONPG	Ortho-nitrophenyl-β-galactoside
OPC	Open Promoter Complex
PAGE	Polyacrylamide Gel Electrophoresis
PAS	Per, ARNT and Sim
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDE5	Phosphodiesterase 5
Pfam	Protein Families database
PMF	Proton Motive Force
PSI	Pounds per Square Inch
PspF	Phage Shock Protein F
PFV	Protein Film Voltammetry
PRD	PTS Regulation Domain
R domain	Regulatory domain
R-finger	Arginine-Finger
RNAP	RNA Polymerase (core)
RNAP-σ$^{54/70}$	RNA Polymerase (sigma factor 54/70 holoenzyme)
RNS	Reactive Nitrogen Species
ROO	Rubredoxin:Oxygen Oxidoreductse
ROS	Reactive Oxygen Species
RPM	Revolutions Per Minute

RR	Response Regulator
RT	Room Temperature
SAP	Shrimp Alkaline Phosphatase
SAXS	Small-Angle X-ray Scattering
ScNsrR	*Streptomyces coelicolor* NsrR
SDM	Site Directed Mutagenesis
SDS	Sodium Dodecyl Sulfate
SFM	Scanning Force Microscopy
SMART	Simple Modular Architecture Research Tool
SNAP	*S*-nitroso-*N*-acetylpenicillamine
SNO	*S*-nitrosothiol
SNP	Sodium Nitroprusside
TBE	Tris Borate EDTA
TEMED	N,N,N',N'-Tetramethylethylenediamine
Tris	Tris (hydroxymethyl) aminomethane
TyrR	Tyrosine Regulator
UAS	Upstream Activator Sequence
V4R	Vinyl 4 Reductase
WA	Walker A
WAXS	Wide Angle X-ray Scattering
WB	Walker B
WT	Wild Type
X-gal	5-bromo-4chloro-3-indolyl-B-D-galactopyranoside
XylR	Xylene Catabolism regulatory Protein
ZraR	Zinc Resistance-Associated Protein Regulator

Contents

Index to Tables and Figures

Chapter 4 – Introduction to NorR and the present work

Chapter 5 – Materials and Methods

Chapter 6 - Error-prone PCR mutagenesis of the AAA+ domain in NorR

Chapter 7 - Investigating the escape mechanism of the GAFTGA variant G266D

Chapter 8 - *In vitro* studies of the full-length GAFTGA-variant G266D

Chapter 9 - Investigating the escape mechanism of the Q304E variant

Chapter 10 – General Discussion

Chapter 1 – Nitric Oxide (NO) and the regulation of NO-detoxification

Nitric Oxide (NO) is a highly reactive free-radical gas with wide-ranging roles in both in animals and plants (MacMicking et al. 1997; Bosca et al. 2005; Mur et al. 2008). Depending upon the concentration present it can function either as a signalling molecule e.g. in the regulation of enzymes and transcription factors, or as a potent mediator of cellular toxicity e.g. the antimicrobial response of macrophages. At lower (nanomolar) concentrations, NO functions in signalling. For example, NO causes the stimulation of guanylyl cyclase in smooth muscle cells (Arnold et al. 1977) and as a result the second messenger cyclic GMP (cGMP) is produced, causing muscle relaxation. Consequently, drugs that increase and decrease the NO levels in the body have been developed. Compounds such as sodium nitroprusside (SNP) which release NO are used in the treatment of hypertension to promote vasodilation and a drop in blood pressure (Wang *et al.* 2002; Hanson and Whiteheart 2005). The role of NO in vasodilation has been utilised most famously, in the development of the drug Sildenafil (Viagra). This drug competitively inhibits phosphodiesterase 5 (PDE5) which catalyses the breakdown of cGMP, and is commonly prescribed in the treatment of male impotence (Corbin and Francis 1999). As a small, neutral and lipophilic molecule, NO can rapidly diffuse across cell membranes and at higher (micromolar) concentrations, has toxic effects (Stamler et al. 1992; Lamattina et al. 2003; van Wonderen et al. 2008). The potency of NO as a biological signalling molecule and its toxicity can be largely explained by the single, unpaired electron that renders it highly reactive. It can be rapidly oxidised by the removal of an electron to form the nitrosonium cation (NO^+) or reduced by the addition of an electron to form the nitroxyl anion (NO^-); both are important molecules in the biochemistry of NO (Figure 1.1A). Consequently NO can react with oxygen (O_2), superoxide (O_2^-), nitrogen-species and transition metals leading to the formation of a wide range of derivatives, collectively known as reactive nitrogen species (RNS) (Figure 1.1B) (Lamattina et al. 2003; van Wonderen et al. 2008). NO can react with transition metals such as zinc, iron and copper to form metal-nitrosyl complexes, a reaction utilised in the regulation of certain enzymes and transcription factors. Reaction with the cysteine (thiol) residues of proteins can form *S*-nitrosothiols (SNOs). When superoxide is present, a rapid reaction occurs to form peroxynitrite ($ONOO^-$). This derivative can lead to the formation of a highly reactive, DNA-damaging hydroxyl radical ($OH^\bullet$) and NO_2. Further peroxynitrite induced reactions include the nitration of tyrosine residues to yield Tyr-NO_2 and the oxidation of thiol groups to yield sulfenic and sulfonic acids.
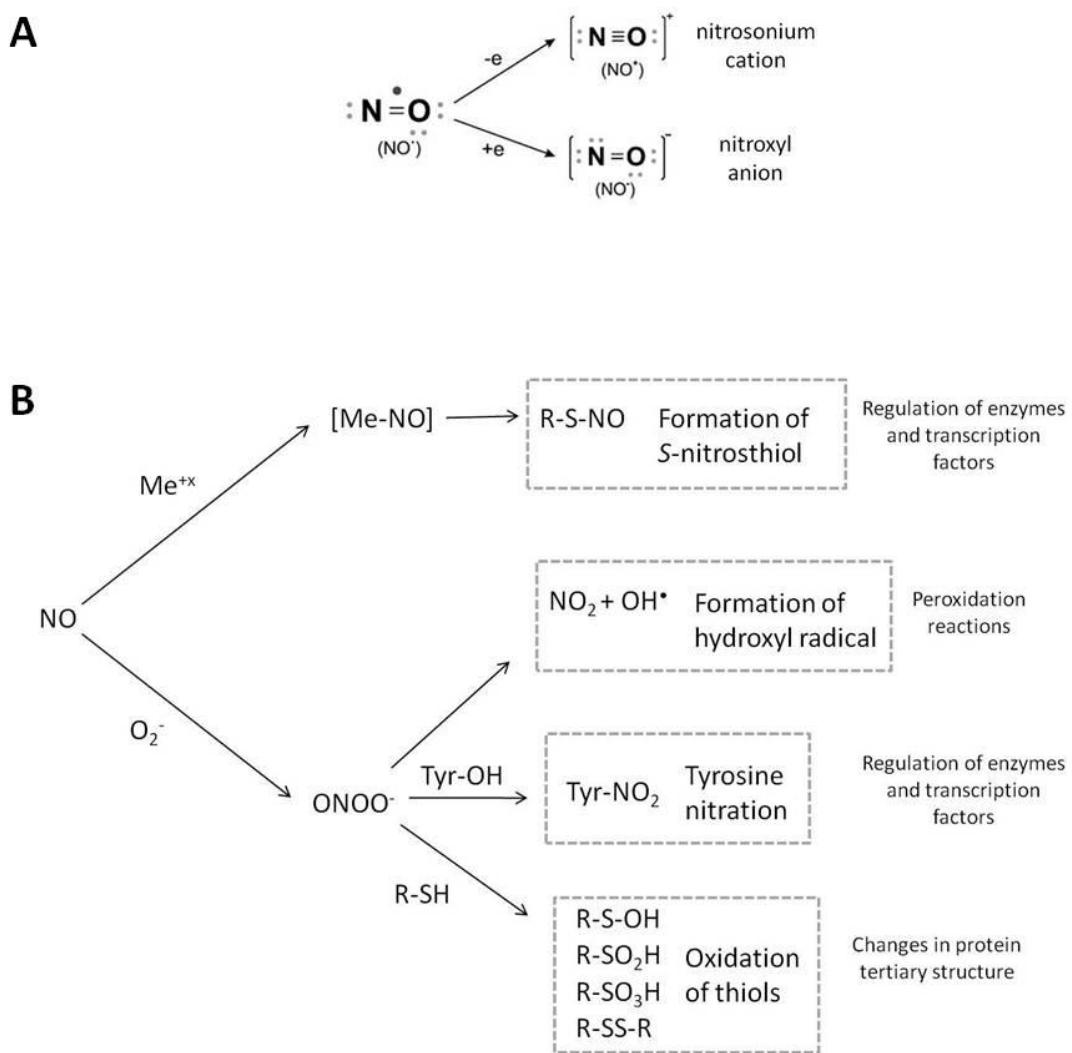
Figure 1.1 - **Chemistry of the interrelated forms of NO and their biological activities**. (**A**) Interconversion of NO forms. NO radical (NO$^\bullet$) is rapidly oxidized by the removal of one electron to give nitrosonium cation (NO$^+$), or reduced by the addition of one electron to form nitroxyl anion (NO$^-$). NO$^+$ and NO$^-$ are important intermediates in the biochemistry of NO. (**B**) Chemical reactions of NO (produced endogenously or released by NO donors). NO can react with transition metals (M^{+x}) such as Zn, Fe and Cu to form metal-nitrosyl complexes. NO$^+$ and NO$^\bullet$ can also nitrosylate thiol groups of cysteines of proteins (R-SNO). NO reacts with superoxide (O$_2^-$) to form the peroxynitrite (ONOO$^-$) derivative that can lead to the formation of hydroxyl radical (OH$^\bullet$) and NO$_2$. Peroxynitrite can also induce the nitration of tyrosine residues to form Tyr-NO$_2$ as well as the oxidation of thiols to produce sulfenic/sulfonic acids (Lamattina et al. 2003).

1.1　Targets of NO and reactive nitrogen species (RNS)

NO may have either a cytostatic or cytotoxic effect, depending on the bacterial species. Evidence indicates that NO/*S*-nitrosothiols can reversibly inhibit DNA replication via a mechanism that involves mobilisation of zinc from metalloproteins (Schapiro et al. 2003). NO itself has also been shown to inhibit respiration by binding to the copper and heme centres in the quinol oxidases cytochrome bo' and bd (Stevanin et al. 2000). In addition, interaction of NO with tyrosine residues has been demonstrated to inhibit the enzyme ribonucleotide reductase which is believed to limit the availability of precursors for the synthesis and repair of DNA (Lepoivre et al. 1991). NO and hydrogen peroxide (H_2O_2) are known to act in synergy in the killing of *E. coli* (Pacelli et al. 1995). Iron chelators reduce the number of double stranded DNA breaks in cultures treated with NO and H_2O_2, indicating a role for the Fenton reaction (the reduction of H_2O_2 by ferrous iron) in DNA damage:

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + OH^\bullet$$

The hydroxyl radical formed in this reaction is highly reactive, causing DNA damage. As a substrate of the Fenton reaction, the role of H_2O_2 is clear but the role of NO in the synergy of killing is less well understood. It has been suggested that interaction of NO with iron-sulfur clusters increases the level of ferrous iron to accelerate the Fenton reaction (Pacelli et al. 1995). Another possible role for NO centres around its inhibition of respiration by binding copper and heme centres in quinol oxidases (Stevanin et al. 2000). It has been suggested that the resulting accumulation of NADH would promote FADH production by an NADH-dependent flavin reductase. FADH would then reduce free ferric iron to the ferrous form, thereby accelerating the Fenton reaction (Woodmansee and Imlay 2002; Woodmansee and Imlay 2003). It is likely that both hypotheses have a role in the NO-dependent damage of DNA. Increasing evidence suggests that proteins that contain iron-sulfur (Fe-S) clusters or mononuclear iron may be the primary target of NO cytotoxicity (Spiro 2007). Well over 200 Fe-S proteins have been identified in *E. coli* with diverse roles including sugar metabolism, amino acids biosynthesis, RNA modification and DNA-synthesis and repair (Johnson et al. 2005; Lill and Muhlenhoff 2006). *In vitro* studies have shown that the Fe-S clusters can be modified by the binding of NO to form dinitrosyl-iron complexes (DNICs) which have a unique Electron Paramagenetic Resonance (EPR) signal at $g = 2.04$ (Drapier 1997; Kennedy et al. 1997; Ding and Demple 2000; Cruz-Ramos et al.

2002; Rogers et al. 2003). Furthermore *in vivo* studies suggest that many iron-sulfur cluster proteins are highly sensitive to NO including aconitase [4Fe-4S] (Gardner et al. 1997), ferredoxin [2Fe-2S] (Rogers and Ding 2001), endonuclease III [4Fe-4S] (Rogers et al. 2003) and the dihyroxy-acid dehydratase (I1vD) [4Fe-4S] (Hyduke et al. 2007; Ren et al. 2008). In addition, a number of transcription factors have also been shown to respond to NO (see below) and it may be that such regulators have evolved to mediate the cellular response to NO cytoxicity.

1.2   Encounter of NO by bacteria

1.2.1   Exogenous NO

In mammals, nitric oxide synthases (NOS) are present in three isoforms (Fang 2004). The eNOS and nNOS isoforms are so-called because of their discovery in endothelial and neuronal tissue respectively (Bredt et al. 1991; Janssens et al. 1992; Sessa et al. 1992). They are constitutively expressed in a range of tissues and are considered low-output enzymes with roles in signal transduction. Production of NO is strongly dependent on calcium (Ca^{2+}) levels. When Ca^{2+} levels increase, a Ca^{2+}-calmodulin complex is formed that binds strongly to eNOS/nNOS to activate NO production. Therefore eNOS and nNOS are collectively termed cNOS. A third isoform, iNOS was first discovered in macrophages and is so-called because its function is independent of calcium signalling. Unlike cNOS, iNOS is not constitutively expressed and only low levels are normally present (Nathan and Xie 1994). Regulation is at the level of transcription (Taylor and Geller 2000); stimulation of receptors and cytokine signalling can trigger signalling cascades that result in large amounts of NO being produced. Macrophages of the immune system are able to express iNOS and produce NO as part of the antimicrobial response. Significantly, lipopolysaccharide (LPS) present on the surface of bacteria can induce the expression of iNOS via the production of cytokines (Nathan 1992; Xie et al. 1994).

NOS enzymes are dimeric (Crane et al. 1998) and catalyse the synthesis of NO from L-arginine via an NωOH-L-arginine intermediate (Figure 1.2). The C-terminal reductase domain contains binding sites for NADPH, FAD and FMN and shares 30-40% similarity with cytochrome P450 reductase (Xie et al. 1994). The oxidation of NADPH releases electrons that are passed between the domains, facilitated by a calmodulin/Ca^{2+} complex. The N-terminal oxidase domain contains the tetrahydobiopterin (BH_4) and heme cofactors.

Oxygen binds the heme and electrons received from the reductase domain drive the two-step oxygenation to NO, releasing L-citrulline as a by-product.

1.2.2 Endogenous NO

NO is an intermediate of denitrification, a process by which denitrifying bacteria convert nitrate into dinitrogen gas (Knowles 1982; Zumft 1997). This anaerobic process is utilised by many bacteria to generate energy and is accomplished by four different types of metalloenzymes in four simple steps (reviewed by (Tavares et al. 2006)) (Figure 1.3). The reduction of nitrate to nitrite is catalysed by a molybdenum-containing nitrate reductase (reviewed by (Richardson et al. 2001)) of which there are two types involved in denitrification. The membrane-bound nitrate reductase (Nar) has three components: NarI and NarH together make up the electron transfer centres whilst NarG contains the active site for the reduction. NarGH is present in the cytoplasm but anchored to the inner cytoplasmic membrane by NarI (Jormakka et al. 2004). The nitrate reductase (Nap) is in contrast located in the periplasm and examples exist in *Desulfovibrio desulfuricans* and *Paracoccus pantotrophus* (Sears et al. 2000; Marietou et al. 2005). The second step in the denitrification pathway involves the reduction of nitrite to nitric oxide (NO) and is facilitated by either the cytochrome cd1 (van Wonderen et al. 2007) or copper-containing nitrite reductases (Nojiri et al. 2009), both of which contain distinct electron transfer and catalytic centres. Nitric Oxide Reductase (NOR) catalyses the two-electron reduction of NO to nitrous oxide (N_2O) (Watmough et al. 2009). This membrane-bound enzyme is found in three forms (cNOR, qNOR and qCuNOR) that have a conserved active site but differ in the number and type of electron transfer centres. The best characterised is the cNOR from *Paracoccus denitrificans* (Field et al. 2008), containing a NorC subunit that receives electrons from cytochrome c or cupredoxin and a binuclear heme iron-non heme iron catalytic NorB subunit. The final step of the complete denitrification pathway is the reduction of N_2O to dinitrogen (N_2), carried out by the perisplasmic multi-copper nitrous oxide reductase (N_2OR) (Haltia et al. 2003). This enzyme contains a binuclear Cu_A centre involved in electron transfer and a multinuclear Cu_Z catalytic centre (Rosenzweig 2000).
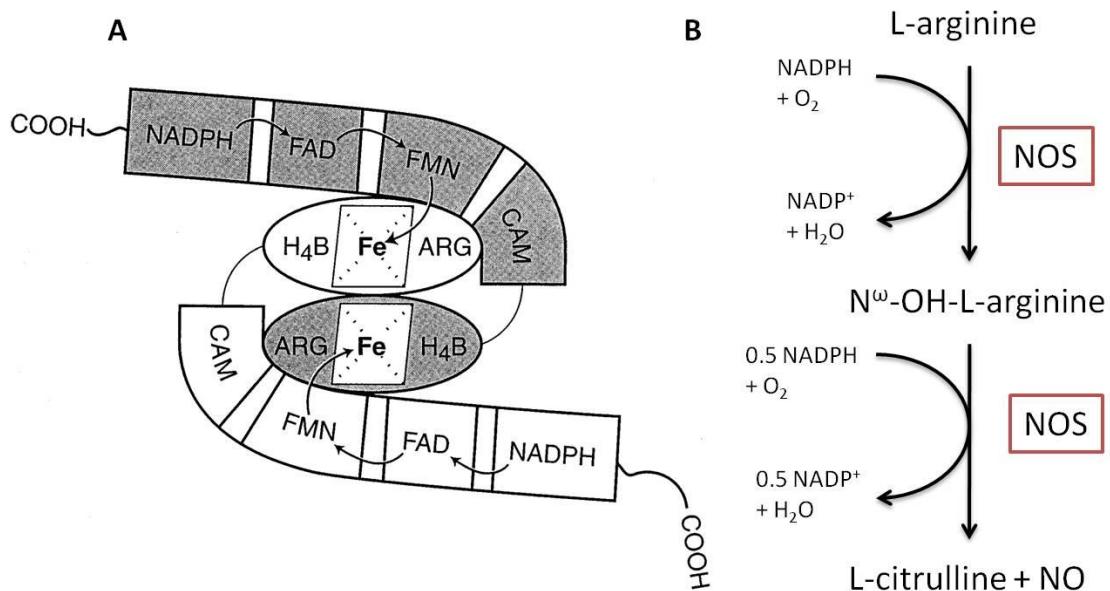
Figure 1.2 – (**A**) Schematic showing the structure of the NO synthase (NOS) enzyme. NOS is a dimer and comprises two domains (shaded and un-shaded). A reductase domain (curved oblong) receives electrons from NADPH and passes them on to the oxygenase domain (oval) of the other monomer via the FAD and FMN cofactors. A calmodulin/Ca^{2+} complex facilitates the transfer of electrons between the two domains. The oxygenase domain contains the tetrahydobiopterin (H_4B) and heme cofactors. Oxygen binds to the heme moiety and the electrons drive the two-step oxygenation of L-arginine to L-citrulline which releases NO. Electron transfer is indicated using arrows and does not occur between the reductase and oxygenase domains of the same subunit (Stuehr et al., 1999). (**B**) Shows the two-step oxygenation of L-arginine to NO via the N^{ω}OH-L-arginine intermediate. L-citrulline is released as a by-product (MacMicking et al. 1997).

NO production has also been observed in non-denitrifying bacteria, including *E. coli* (Ji and Hollocher 1988). Such bacteria are able to carry out steps of the denitrification pathway as part of anaerobic respiration in order to utilise nitrate as a terminal electron acceptor. This process is called respiratory denitrification (Figure 1.3). *E. coli* contains three nitrate reductases, all of which absolutely require molybdenum as a cofactor (Moreno-Vivian et al. 1999). The membrane-bound NarA and NarZ reduce nitrate in the cytoplasm (Berks et al. 1995). Under nitrate-sufficient conditions, NarA is the predominant reductase (Potter et al. 1999), whereas NarZ confers a selective advantage during stationary phase or periods of poor growth (Iobbi-Nivol et al. 1990). Under conditions of nitrate-starvation, the periplasmic nitrate reductase (Nap) is expressed (Stewart et al. 2002). The Nap reductase consists of a number of different subunits; the NapA-NapB complex is believed to receive electrons from the quinol pool via the membrane bound cytochrome NapC, whilst NapG and NapH, but not NapF are also essential for reduction (Brondijk et al. 2002; Nilavongse et al. 2006). Depending on their cellular location, the nitrate reductases in *E. coli* produce nitrite that acts as the substrate for one of two biochemically distinct nitrite reductase enzymes (Wang and Gunsalus 2000). The NirB nitrite reductase is a cytoplasmic siroheme-containing enzyme that uses NADH as an electron donor to reduce nitrite. Together the nitrate reductase NarA/NarZ and the nitrite reductase NirB function in the cytoplasm to reduce nitrate to ammonium. In contrast, the cytochrome C nitrite reductase (NrfA) is located in the periplasm and functions in concert with Nap in respiratory denitrification. NrfA is known to catalyse the six-electron reduction of nitrite to ammonia via NO (Costa et al. 1990) but has more recently been implicated in the detoxification of NO in *E. coli* (Poock et al. 2002; van Wonderen et al. 2008). The function of NrfA as an enzyme in detoxification will be discussed in greater detail later (section 1.4.1).
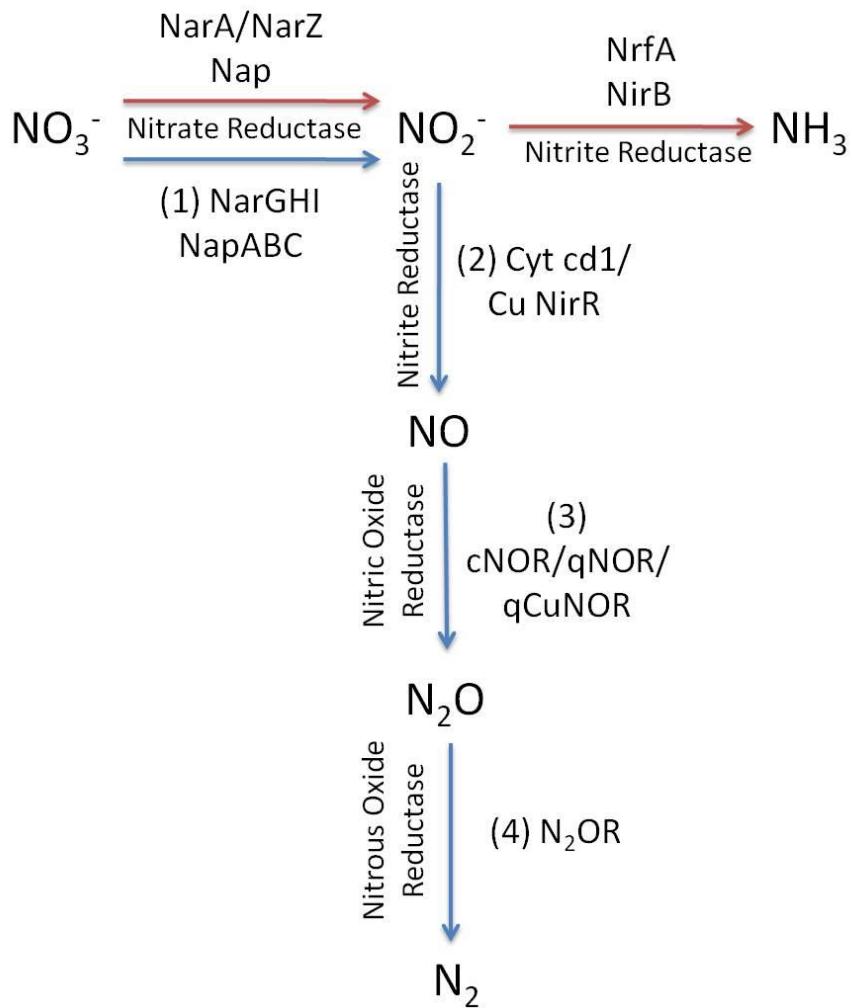
Figure 1.3 – The steps in the denitrification pathway and associated enzymes for *E. coli* (red arrows) and denitrifiers e.g. *Paracoccus denitrificans* (blue arrows). The complete denitrification pathway involves the reduction of nitrate to dinitrogen involving four different types of enzymes: (**1**) nitrate reductase (Nar or Nap), (**2**) nitrite reductase (cytochrome cd1 or copper-containing), (**3**) nitric oxide reductase (cNOR, qNOR or qCuNOR) and (**4**) nitrous oxide reductase (N_2OR) (Tavares et al. 2006). Non-denitrifiers are also able to carry out steps of the pathway under anaerobic conditions, in a process known as respiratory denitrification. Nitrate is reduced to ammonia via nitrate and nitrite reductases in the cytoplasm (NarA/NarZ and NirB) or in the periplasm (Nap and NrfA). The six-electron reduction of nitrite to ammonia by NrfA occurs via a Nitric Oxide intermediate which may be an endogenous source of NO in *E. coli* (Poock et al. 2002; van Wonderen et al. 2008).

1.3 NO-sensing in bacteria

As part of the response to RNS, bacteria have evolved specific sensor proteins that detect NO and subsequently activate the expression of proteins that mediate the cellular response to nitrosative stress. This is particularly relevant in pathogenic bacteria that encounter NO released by macrophages upon infection of mammalian and plant cells. Many of the NO-sensing proteins and detoxification enzymes are conserved in both pathogenic and non-pathogenic bacteria, suggesting that their evolution may have originally occurred in soil microbes before horizontal gene transfer provided pathogenic bacteria with methods for NO-detoxification (Tucker *et al.* 2010b). Bacterial NO-sensing proteins can be divided into two classes: (i) regulators whose primary function is not to sense NO and (ii) regulators that are solely dedicated to sensing NO. The best example of the first class is the global fumarate and nitrate reductase regulator (FNR) whose primary function is to sense oxygen but which has also been shown to be NO-responsive (Cruz-Ramos et al. 2002). In addition to an FNR orthologue, denitrifying soil bacteria also possess an orthologue of the dedicated NO-sensor nitrate and nitrite reducase regulator (Nnr). For example the denitrifier *P. denitricans* has the Nnr1 protein in addition to an FNR protein (FnrP) that senses both oxygen and NO (Hutchings *et al.* 2002a). The role of the former is to detect NO, activating the transcription of the *nir* and *nor* genes that encode the nitrite and NO reductases respectively. This ensures that nitrate is rapidly converted into nitrous oxide, preventing the accumulation of the toxic intermediates of denitrification (Hutchings et al. 2000; Hutchings and Spiro 2000; Spiro 2007).

1.3.1 Direct NO-sensors in *E. coli*

The role of NO-sensing proteins in the model organism and human pathogen, *Escherichia coli*, will be discussed in this section. A number of regulators in *E. coli* mediate cellular responses to oxygen and reactive oxygen species (ROS) but can additionally respond to NO. The SoxR, IscR and FNR proteins all act as NO-sensors via the Fe-S clusters that they contain. Fe-S proteins are an ancient and important class (Beinert 2000), ubiquitous in nearly all organisms with roles in electron-transfer and catalysis. Their role as redox agents is well understood but in recent years it has emerged that Fe-S cluster proteins can also have regulatory functions (reviewed in (Kiley and Beinert 2003). All Fe-S clusters employ tetrahedral coordination of the iron. The simplest Fe-S cluster is represented by the rubredoxins but more complex clusters exist e.g. [2Fe-2S], [4Fe-4S], and [3Fe-4S] (Kiley and Beinert 2003).

(i)   **SoxR**

SoxR is a member of the MerR-family of regulators and was the first transcription factor shown to contain an Fe-S cluster (reviewed in (Demple *et al.* 2002). Initially, SoxR was implicated in sensing superoxide. Superoxide is a ROS that causes SoxR to activate the transcription of SoxS in *E.* coli, (Demple et al. 2002). SoxS in turn activates the expression of ~45 gene products e.g. superoxide dismutase which catalyse the removal of superoxide and the repair of oxidative stress-induced damage. Electron Paramagnetic Resonance (EPR) spectroscopy indicates that each monomer of the SoxR homodimer contains a [2Fe-2S] cluster that under normal cellular conditions is in its reduced form ($[2Fe-2S]^{1+}$) and unable to activate transcription. Superoxide stress results in the generation of the oxidised form ($[2Fe-2S]^{2+}$) that is competent to activate transcription (Ding and Demple 1997). A membrane-associated NAD(P)H-dependent complex reduces SoxR, returning it to its transcriptionally inactive state, in the absence of superoxide (Koo et al. 2003). More recently, activation of SoxR by NO has been shown to occur via the nitrosylation of the [2Fe-2S] centres to form dinitrosyl-iron complexes (DNICs). Nitrosylated SoxR has a similar transcriptional activity to oxidised SoxR but is short-lived *in vivo*, indicating the presence of mechanisms that repair nitrosylated clusters (Ding and Demple 2000). The nitrosylated [2Fe-2S] cluster of ferredoxin can be repaired *in vitro* by the cysteine desulphurase, IscS (Yang et al. 2002). Despite the identification of SoxR as an NO-sensor, the significance of the SoxRS system in the biological response to NO remains unclear. Recent microarray analyses suggest that SoxR and SoxS only play a minor role in the response to NO (Mukhopadhyay *et al.* 2004; Flatley *et al.* 2005; Justino *et al.* 2005b). However, the inter-relatedness of the different RNS species makes the relative contributions of each molecule difficult to assess and microarray analysis may not detect the subtle effects of NO on the SoxRS regulon (Spiro 2007).

(ii)   **FNR**

FNR is a global regulator that controls the transcription of greater than 100 genes (Guest *et al.* 1996; Kiley and Beinert 1998; Green *et al.* 2001). The primary role of FNR is to sense oxygen via its [4Fe-4S] iron-sulfur cluster and therefore FNR generally controls the expression of proteins that contribute to the anaerobic lifestyle of *E. coli*. In its active form, FNR contains one $[4Fe-4S]^{2+}$ cluster per subunit of the dimer and is competent to bind to DNA. Upon exposure to oxygen, rapid conversion of the cluster from $[4Fe-4S]^{2+}$ to $[2Fe-2S]^{2+}$ prevents dimerisation and therefore DNA-binding. In addition, FNR has been shown

to sense and respond to NO in *Escherichia coli* (Cruz-Ramos et al. 2002; Crack et al. 2008), *Azotobacter vinelandii* (Wu et al. 2000), *Paracoccus denitrificans* (Hutchings *et al.* 2002a) and *Salmonella enterica serovar* Typhimurium (Gilberthorpe and Poole 2008). Here, it is thought that the inactivation of FNR occurs via the formation of monomeric and dimeric dinitrosyl-iron (DNIC) complexes (Cruz-Ramos et al. 2002).

(iii) **IscR**

The iron-sulfur cluster regulator (IscR) contains a [2Fe-2S] cluster and is a member of the Rrf2 family of transcription factors. IscR represses the transcription of the iscRSUA operon (Schwartz *et al.* 2001), the products of which are required for iron-sulfur cluster biogenesis (Zheng *et al.* 1998). Removal of the cluster rather than modification of the redox state has been suggested to modify the activity of IscR by altering the DNA-binding activity of the transcription factor (Schwartz *et al.* 2001). When there is a need for iron-sulfur cluster assembly e.g. under conditions of oxidative stress, the levels of [2Fe-2S] IscR are likely to be low, leading to derepression of the iscRSUA operon. The resulting increase in the expression of the Isc assembly apparatus will eventually lead to higher [2Fe-2S] IscR levels and repression of the operon (Figure 1.4). In this way regulation by IscR is thought to couple iron-sulfur cluster assembly with the cellular requirement for synthesis or repair (Schwartz *et al.* 2001). In the absence of the cluster, IscR has also been shown to directly activate rather than derepress target genes (Figure 1.4). For example under conditions of oxidative stress, the apo form of IscR is thought to activate the expression of the Suf Fe-S cluster assembly system (Yeo *et al.* 2006). Where both the Isc and Suf assembly systems are present (e.g. *E. coli*), the Isc system is thought to a play a primary role in cluster assembly whereas the Suf system acts as a back-up or more specialised system (Takahashi and Tokumoto 2002; Outten *et al.* 2004). In addition to its response to redox status, IscR has been shown to respond to NO (Hyduke *et al.* 2007; Pullan *et al.* 2007; Jones-Carson *et al.* 2008). This corresponds with a proposed role for IscR in the sensing of NO-induced iron-sulfur cluster damage and the subsequent up-regulation of mechanisms for cluster re-synthesis.

(iv) **Fur**

The disruption of iron-sulfur clusters is one of the major toxic effects of NO. In attempt to restore such iron centres, *E. coli* up-regulates proteins involved in the control of intracellular iron levels (reviewed by (Andrews *et al.* 2003). This has been shown to occur

via the inactivation of the iron-responsive ferric uptake regulator (Fur). Indeed a *fur* mutant of *E. coli* has increased susceptibility to NO (D'Autreaux *et al.* 2002). Fur is a global regulator and controls the expression of >90 genes in *E. coli* (Hantke 2001). In contrast to the Fe-S cluster proteins SoxR, FNR and IscR, Fur contains a non-heme ferrous iron centre in its active form that allows it to bind to and prevent expression from Fur-regulated promoters (e.g. those involved in iron sequestration). When the levels of iron are low, the iron centre is lost and Fur is no longer able to repress the transcription of such genes. In addition to IscR (Yeo *et al.* 2006), Fur has been shown to regulate expression of the *suf* operon encoding the secondary, Suf Fe-S cluster assembly system (Figure 1.4). Fur binds to DNA at a site that overlaps the apo-IscR binding-site, preventing IscR-mediated activation of transcription. Iron depletion or oxidative stress leads to derepression through the loss of the Fur iron centre with the concomitant dissociation of the IscR [2Fe-2S] cluster (Lee *et al.* 2008). Derepression at Fur-regulated promoters has also been shown to be induced by the presence of NO which binds to Fur to form a dinitrosyl iron complex (DNIC) (D'Autreaux *et al.* 2002). This presumably increases the intake of Fe^{2+} and the production of the Isc and Suf assembly systems to enable the cell to repair NO-damaged iron containing proteins.

Ferrous iron largely mediates the cytotoxic effects of ROS as hydrogen peroxide leads to the formation of the DNA-damaging hydroxyl radical via the Fenton reaction (Pacelli *et al.* 1995). This reaction is accelerated by superoxide, firstly because the action of superoxide dismutases leads to increased levels of hydrogen peroxide and secondly, because superoxide liberates further ferrous iron due to the attack on the iron-sulfur clusters. The SoxR and OxyR regulators that respond to superoxide and hydrogen peroxide respectively, have additionally been shown to up-regulate the expression of the Fur repressor (Zheng *et al.* 1999). The induction of *fur* transcription by OxyR and SoxRS leads to an increase in Fur protein concentration although the exact consequences of this increase are unclear. It has been suggested that up-regulation may be required to replace Fur damaged by ROS or to prevent the uptake of ferrous iron into the cell, although this seems to contradict the derepression of the Fur regulon under conditions of oxidative stress (D'Autreaux *et al.* 2002). Compartmentalisation of Fe^{3+} by the production of the storage proteins ferritin and bacterioferritin may additionally help to counteract the Fenton reaction in bacteria (Andrews *et al.* 2003). It has been suggested that the ferritin-like protein Dps may have a particularly significant role in the sequestration of iron in the vicinity of DNA
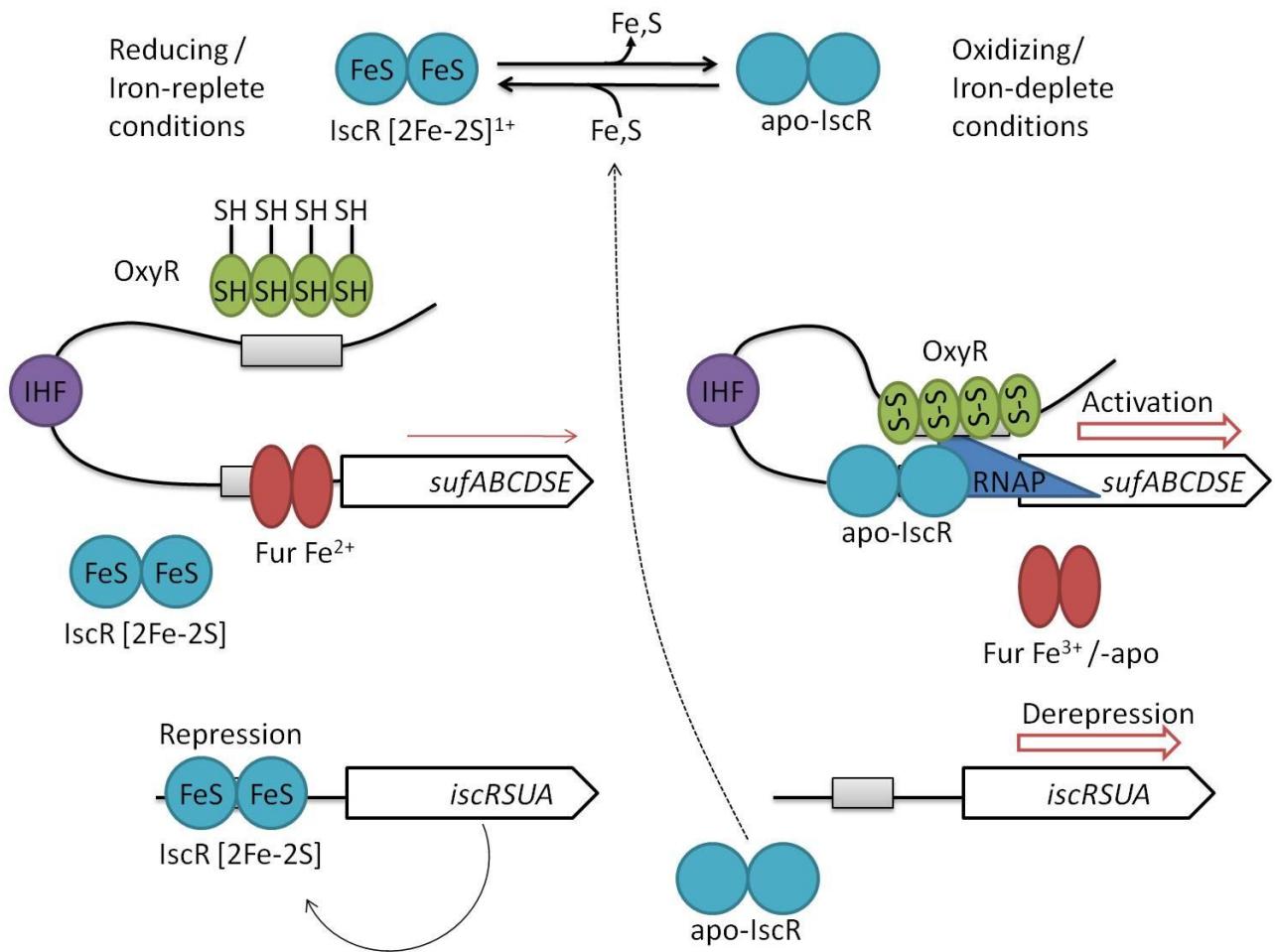
Figure 1.4 – Regulation of the Isc and Suf iron-sulfur cluster assembly systems. Both systems are controlled by a common regulator (IscR), acting in opposite ways to ensure maximal activation of transcription from the *sufABCDSE* and *iscRSUA* promoters under conditions of oxidative stress and iron-depletion. Under reducing conditions, Fur (containing a non-heme iron centre) binds at an overlapping DNA site to prevent binding of the RNA polymerase and apo-IscR-mediated activation of *suf* transcription. Under oxidizing conditions or iron-depletion, the IscR and Fur regulators lose the [2Fe-2S] cluster and ferrous iron centre respectively. This leads to derepression by Fur and activation by apo-IscR. Transcription of *suf* genes further requires OxyR that is activated by peroxide via reversible cysteine oxidation. OxyR-mediated activation is dependent on DNA bending induced by Integration Host Factor (IHF). Activation of transcription from the *isc* promoter occurs under the same conditions but is regulated solely in response to IscR. In contrast to its role in the regulation of the *suf* genes, IscR represses *isc* transcription under reducing/iron-replete conditions when an intact [2Fe-2S] cluster is present. Under oxidizing or iron-deplete conditions, the cluster is lost and derepression of the *isc* system occurs. (Yeo *et al.* 2006). Significantly, IscR and Fur have also been shown to respond to NO. Therefore NO may lead to the expression of the Isc and Suf assembly systems to enable the cell to repair NO-damaged iron containing proteins.

(Almiron *et al.* 1992; Grant *et al.* 1998). In line with this, the transcription of Dps has been shown to be under the control of OxyR (Altuvia *et al.* 1994).

1.3.2   Dedicated NO-sensors in *E.coli*

Two regulators in *E. coli* have been identified that appear solely dedicated to the sensing of NO (Spiro 2007). The bacterial Enhancer Binding Protein (bEBP), NorR, activates the σ^{54}-dependent transcription of genes encoding the flavorubredoxin (NorVW) in response to the direct binding of NO to a mononuclear non-heme iron centre in NorR (Hutchings *et al.* 2002b; Gardner *et al.* 2003; D'Autreaux *et al.* 2005; Tucker *et al.* 2007). σ^{54}-dependent transcription and bEBPs are discussed in greater detail in Chapters 2 and 3 respectively. NorR is the subject of this work and is discussed in depth in Chapter 4. More recently, a second dedicated NO-sensor has been identified. NsrR (reviewed by (Tucker *et al.* 2010b) which like IscR, is a member of the Rrf2 family of transcriptional repressors will now be discussed.

(i)   **NsrR**

First identified in *Nitrosomonas europaea* (Beaumont *et al.* 2004), NsrR is predicted to be encoded in the genomes of most β- and γ-proteobacteria. Notable exceptions include *Pseudomonas aeruginosa* and *Vibrio cholerae* (Rodionov *et al.* 2005). Interestingly in these cases, NorR takes the place of NsrR to control the expression of the NO dioxygenase Hmp. This underlies the importance of a dedicated NO-sensor in the regulation of NO-detoxification mechanisms (Tucker *et al.* 2010b). Microarray studies in *E. coli* identified a number of genes that were NO-inducible in strains that lacked putative NO-responsive regulators (Mukhopadhyay *et al.* 2004; Flatley *et al.* 2005; Justino *et al.* 2005b). It was therefore proposed that *E. coli* contained another NO-sensing regulator, the product of the *yjeB* gene (NsrR). Bioinformatic analysis (Rodionov *et al.* 2005), microarray analysis (Filenko *et al.* 2007) and a genetic approach using transposon mutagenesis (Bodenmiller and Spiro 2006) have been used to show that NsrR negatively regulates the transcription of a number of genes in the absence of nitrosative stress (reviewed by (Tucker *et al.* 2010b). Key members of the NsrR regulon include the *hmp* gene that encodes the flavohaemoglobin (Hmp) and the *nrfA* gene that encodes the periplasmic nitrite reductase, which are enzymes that detoxify NO in *E. coli*. NsrR is also suggested to regulate the expression of the *hcp* and *hcr* genes that encode the Fe-S-containing Hybrid Cluster Protein (Hcp) and its partner reductase, which are suggested to form a hydroxylamine

oxidoreductase converting hydroxylamine to ammonia (Wolfe *et al.* 2002; Cabello *et al.* 2004). However, a role for Hcp as a peroxidase in the ROS-response has also been demonstrated since Hcp can turnover hydrogen peroxide and *hcp* transcription is regulated by OxyR *in vitro* (Almeida *et al.* 2006). It is unclear whether either role represents a biologically relevant function that is repressed by NsrR. It has also been suggested that another RNS-intermediate might represent the true substrate of Hcp and Hcr (Filenko *et al.* 2007). In addition, NsrR has been shown to regulate the expression of the *ytfE* gene that encodes a protein involved in the repair of iron-sulfur clusters (Justino *et al.* 2007; Todorovic *et al.* 2008). Finally, NsrR has been shown to be a negative regulator of motility in *E. coli* (Partridge *et al.* 2009).

Recently, *in vitro* studies of the NsrR proteins from *Streptomyces coelicolor* (Tucker *et al.* 2008), *Bacillis subtilis* (Yukl *et al.* 2008) and *Neisseria gonorrhoeae* (Isabella *et al.* 2009), heterologously expressed and purified in *E. coli,* have shed light on the mechanism by which NsrR might act as an NO-sensor (reviewed by (Tucker *et al.* 2010b). All three studies indicate that NsrR contains a Fe-S cluster but the nature of the cluster was not found to be the same in each case. UV-visible and circular dichroism spectra of *Streptomyces coelicolor* NsrR (ScNsrR) indicate the presence of a [2Fe-2S] cluster (Tucker *et al.* 2008). Significantly this cluster was stable when ScNsrR was purified aerobically, indicating an insensitivity to oxygen that is in line with its role as a dedicated NO-sensor. EPR-spectroscopy indicates that upon exposure to NO, the [2Fe-2S] cluster forms a dinitrosyl iron complex (DNIC). DNA binding assays indicate that the unmodified ScNsrR [2Fe-2S] protein is competent to bind to its promoter targets but that modification by NO prevents the binding to DNA (Tucker *et al.* 2008). Studies of *Neisseria gonorrhoeae* NsrR (NgNsrR), purified aerobically in *E. coli,* are also consistent with the presence of a [2Fe-2S] cluster (Isabella *et al.* 2009). Matrix-Assisted Laser Desorption/Ionization-Time Of Flight (MALDI-TOF) mass spectrometry revealed a higher mass than expected for the apo-protein. Furthermore substitution of one of the three conserved cysteine residues with alanine resulted in a mass that was in line with the expected mass of the apo-form of NgNsrR. In agreement with the ScNsrR data, NgNsrR [2Fe-2S] was able to bind specifically at its target promoter but exposure to NO abolished this activity *in vitro*. In contrast to the *S. coelicolor* and *N. gonorrhoeae* studies, spectroscopy of *B. subtilis* NsrR (BsNsrR) indicates the presence of a [4Fe-4S] rather than a [2Fe-2S] cluster (Yukl *et al.* 2008). Unlike the ScNsrR [2Fe-2S] and NgNsrR [2Fe-2S] proteins, the BsNsrR [4Fe-4S]

protein was not aerobically stable; exposure to oxygen resulted in loss of the cluster with some evidence of the formation of an unstable [2Fe-2S] intermediate. In common with the other studies however, exposure of BsNsrR to NO resulted in the formation of DNIC complexes. Importantly, no studies were carried out to determine the DNA-binding activity of the BsNsrR [4Fe-4S] protein and so it remains unclear whether this form of the protein is biologically relevant. Overall, in the absence of studies in the native hosts, the true nature of the Fe-S clusters in the NsrR proteins from *S. coelicolor, N. gonorrhoeae* and *B. subtilis* remains unclear. It is possible that the anaerobically purified BsNsrR [4Fe-4S] represents the biologically relevant form and that the ScNsrR [2Fe-2S] and NgNsrR [2Fe-2S] proteins form as a result of the breakdown of the cluster during aerobic purification. Initial, unpublished data suggest that when anaerobically purified the ScNsrR protein does contain a [4Fe-4S] cluster, but that exposure to oxygen or NO causes precipitation *in vitro* rather than breakdown to a [2Fe-2S] cluster (Tucker *et al.* 2010b). The demonstration that the binding of ScNsrR [2Fe-2S] and NgNsrR [2Fe-2S] to target promoters is abolished upon exposure to NO, suggests that this form of the protein is biologically relevant (Tucker *et al.* 2010b). An alternative possibility is that NsrR *is* able to respond to oxygen as well as NO and that the [4Fe-4S], [2Fe-2S] and apo-forms of NsrR may all have DNA-binding activity. In a manner analogous to IscR (also a member of the Rrf2 family of transcriptional repressors), NsrR may be able to activate the transcription of genes when the cluster is absent as well as prevent the transcription of a distinct set of genes when the cluster is present (Yeo *et al.* 2006; Tucker *et al.* 2010b). Overall, current evidence supports a model in which NsrR acts as a dedicated NO-sensor via its [2Fe-2S] cluster (Figure 1.5). Exposure of NO, leads to the formation of a DNIC-species leading to the loss of DNA-binding and hence derepression at the promoters of NsrR target genes. In addition to the three conserved cysteine residues, NsrR (like IscR) has a fourth proposed-ligand that is thought to be located in the helix-turn-helix (HTH) motif. This is in agreement with the hypothesis that NO-binding at the Fe-S cluster modifies the DNA-binding affinity of the transcriptional repressor (Tucker *et al.* 2010b).
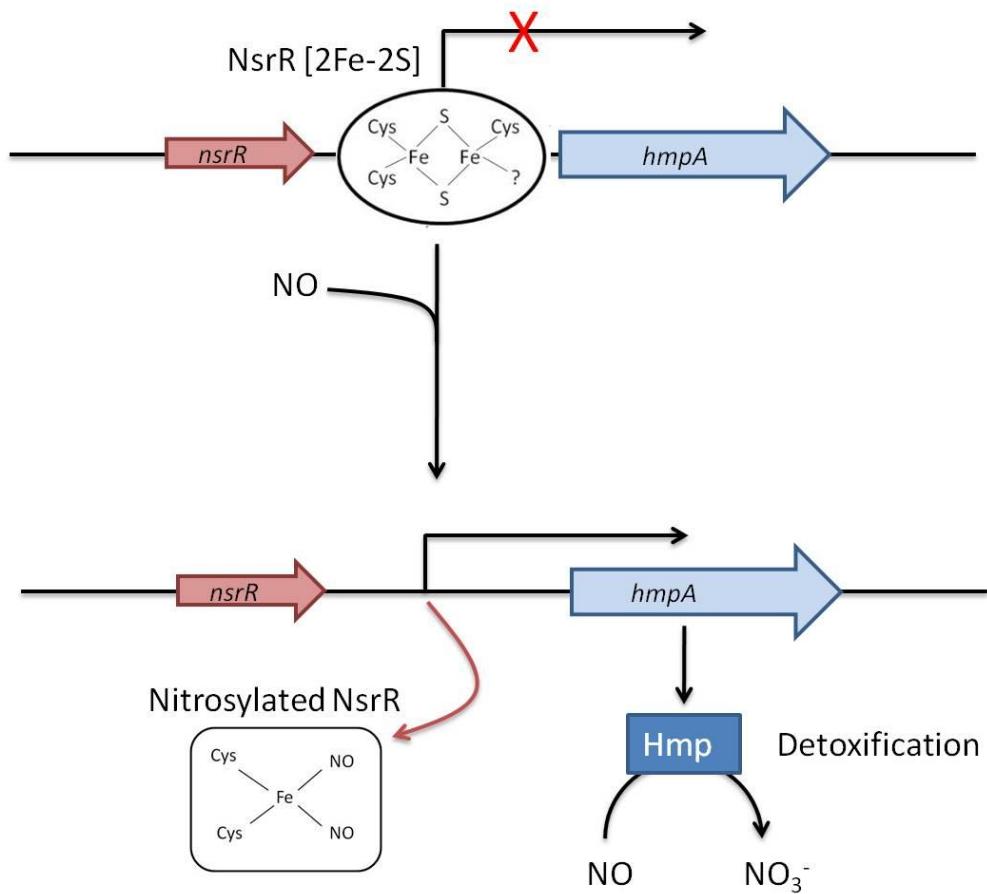
Figure 1.5 – Model of the NsrR regulatory mechanism at a generic *hmp* promoter. NsrR [2Fe-2S] binds in the promoter region of the target gene and prevents the binding of the RNA polymerase holoenzyme. In the presence of NO, a number of dinitrosyl iron (DNIC) complexes can potentially form, leading to the loss of DNA-binding activity and derepression at the target promoter. Only the mononuclear DNIC is shown here. It should also be noted that for clarity, co-regulators such as FNR have been omitted and that *nsrR* is not always genetically associated with *hmp* (Tucker *et al.* 2010b).

1.3.3 The dominant mechanism of NO-sensing by iron-containing proteins

Overall, the majority of NO-sensing mechanisms are based on iron (Figure 1.6). The secondary NO-sensors SoxR, FNR, IscR and the dedicated NO-sensor NsrR all contain iron-sulfur clusters whilst Fur contains a non-heme iron centre. The NO-induced formation of DNICs is irreversible and re-activation of Fe-S cluster-containing proteins therefore requires dissociation of the DNIC and assembly of new iron-sulfur clusters. A number of proteins have been identified as critical for the repair of NO-damaged clusters including the cysteine desulphurase IscS (Yang *et al.* 2002) and a diiron protein YtfE (Justino *et al.* 2007). Both IscS and YtfE in *E. coli* are highly induced upon exposure to NO (Mukhopadhyay *et al.* 2004; Justino *et al.* 2005b; Hyduke *et al.* 2007; Pullan *et al.* 2007). Under aerobic conditions, the levels of these enzymes are not thought to be limiting and it has been suggested that availability of small molecules (ATP, L-cysteine, NADPH/NADH) drive the repair (Ren *et al.* 2008). In contrast to many of the NO-sensing proteins in *E. coli,* the dedicated NO-sensor NorR reversibly binds NO via a non-heme iron centre (D'Autreaux *et al.* 2005) and unlike the other regulators it is not a global regulator, activating transcription from a single promoter (Gardner *et al.* 2003).

1.4 Enzymes of NO detoxification and regulation of their expression

Many of the NO-sensing proteins in *E. coli* directly regulate the transcription of genes that encode enzymes capable of detoxifying nitric oxide (NO). There are at least three enzymes in *E. coli* that directly detoxify NO, employing either NO dioxygenase (NOD) or NO reductase (NOR) activities: the pentaheme periplasmic nitrite reductase (NrfA) that also has NO reductase activity under anaerobic conditions (Poock *et al.* 2002), the flavohaemoglobin (Hmp) that catalyses the oxidation of NO to nitrate in the presence of oxygen (Poole and Hughes 2000; Gardner and Gardner 2002) and the flavorubredoxin/flavodiiron (NorV) protein capable of reducing NO to nitrous oxide in the absence of oxygen (Gardner *et al.* 2002; Gomes *et al.* 2002; D'Autreaux *et al.* 2005). Deletion of any of the genes encoding these enzymes significantly increases the sensitivity of *E. coli* cells to NO grown aerobically or anaerobically. However, all of these enzymes are still not sufficient to protect the cell from NO-cytotoxicity (Mukhopadhyay *et al.* 2004; Justino *et al.* 2005b; Hyduke *et al.* 2007; Pullan *et al.* 2007). These enzymes and the regulation of their expression will now be discussed.
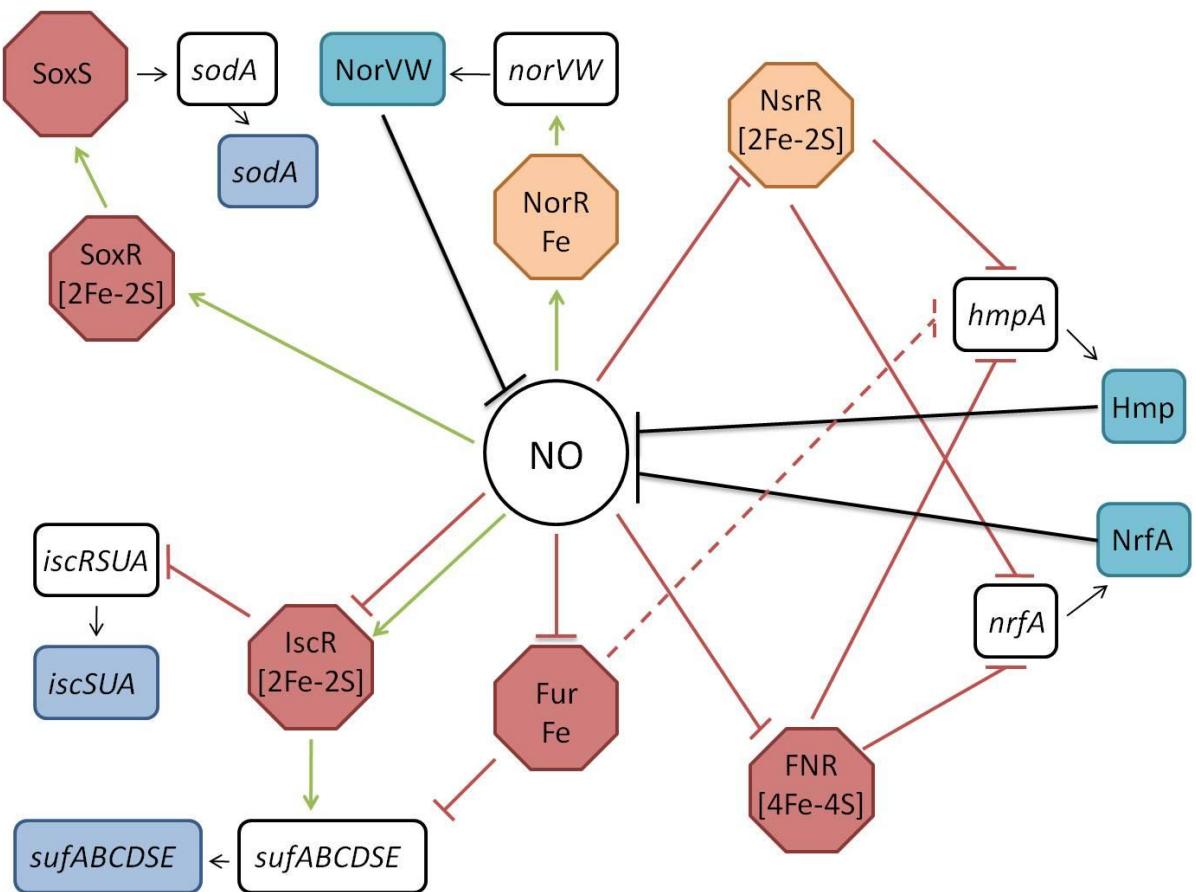
Figure 1.6 – The role of dedicated and secondary NO-sensing proteins as transcription factors. The majority of NO-sensing proteins are secondary sensors (red) i.e. their primary function is in the detection of another signal e.g. O_2 (FNR, IscR, Fur), superoxide (SoxR). The dedicated NO-sensors NorR and NsrR (orange) respond to NO and no other signal (Spiro 2007). The enzymes that detoxify NO in *E. coli* (light blue) are regulated by a variety of NO-sensors. The transcription of the flavorubredoxin, NorV and its redox partner, NorW is regulated by the activator NorR. The transcription of the NO dioxygenase (Hmp) and the periplasmic nitrite reductase (NrfA) is negatively regulated by both the global regulators FNR and NsrR. There is some evidence that the Ferric Uptake Regulator (Fur) negatively regulates the expression of *hmpA* (Hernandez-Urzua *et al.* 2007). Regulation of NrfA and Hmp is more complex than NorR and a number of other non-NO-sensing proteins (not shown) may be involved. Although SoxR primarily senses superoxide, it has also been shown to respond to NO via its [2Fe-2S] cluster (Ding and Demple 2000). SoxR activates SoxS which regulates the transcription of a number of genes including *sodA* that encodes a superoxide dismutase. Under oxidising or iron-deplete conditions, IscR loses its [2Fe-2S] cluster. At the *isc* promoter, this leads to the derepression of transcription. At the *suf* promoter, the apo-IscR protein acts as an activator in concert with the OxyR protein. This activation further requires the derepression of the Fur protein which loses its iron centre under similar conditions (Yeo *et al.* 2006; Lee *et al.* 2008). Transcription of the *isc* and *suf* genes leads to the production of the Isc and Suf iron-sulfur cluster assembly systems. Significantly the IscR and Fur proteins have also been shown to respond to NO (D'Autreaux *et al.* 2002; Hyduke *et al.* 2007; Pullan *et al.* 2007; Jones-Carson *et al.* 2008), in line with a role in the activation of expression of systems that repair damaged Fe-S clusters.

1.4.1 Periplasmic nitrite reductase, NrfA

NrfA is known to catalyse the six-electron reduction of nitrite to ammonia via NO (Costa et al., 1990), as part of respiratory denitrifcation in non-dentrifying enterobacteriaceae. In line with this, the 2.5 Å structure of *E. coli* NrfA shows a 52 kDa homodimer (Figure 1.7A), with each monomer consisting of five c-type hemes (Bamford *et al.* 2002; Clarke *et al.* 2008b). An additional electron may be provided by another electron-donating step or from the other monomer in the NrfA homodimer (Clarke *et al.* 2008b). Crystal structures are also available for the NrfA enzymes from *Wolinella succinogenes* (Einsle et al. 2000), *Sulfospirillum deleyianum* (Einsle et al. 1999), *Desulfovibrio desulfuricans* (Cunha *et al.* 2003) and *Desulfovibrio vulgaris* (Rodrigues *et al.* 2006). In each structure, four of the five hemes are attached by the conventional CXXCH motif and have bis-histidine ligands. The catalytic heme (heme 1) however, is attached via a novel CXXCK sequence motif with the lysine (K126 in *E. coli*) coordinating the heme on the proximal side and a water (hydroxide) molecule or substrate on the distal side. There are four other highly conserved residues in the active site (R106, Y216, H264 and Q263 in *E. coli*) that provide a positive environment for the heme centre and act as potential proton donors (Figure 1.7B). The glutamine is positioned 8 Å away from the heme and coordinates an essential calcium ion. Together the glutamine and calcium ion are thought to increase substrate affinity by supporting a network of hydrogen-bonded water molecules (Clarke *et al.* 2008a). A second calcium ion has been suggested to have a structural role (Cunha *et al.* 2003). It is proposed that the input of electrons occurs at heme 2; there are then two potential routes for electron transfer to the active site. Either electrons move to the nearest catalytic heme via heme 3 (Figure 1.7C) or across the protomer-protomer interface via heme 5 to reach heme 1 of the adjacent monomer (Clarke *et al.* 2008b).

The ability of NrfA to act as an NO-reductase *in vitro* was first shown in 1990 (Costa *et al.* 1990). Since then the contribution of NrfA as an NO reductase (NOR) has been assessed *in vivo* through comparative studies of *E. coli* wild-type and mutant strains deficient for the periplasmic nitrite reductase (Poock *et al.* 2002). Wild-type cells (deficient for the other nitrite reductase, NirB) were calculated to have an *in vivo* turnover by NrfA of 390 NO s^{-1} anaerobically, whilst no activity was detected in the mutant strain. This indicates that NrfA can act as an NO reductase as well as functioning in the six-electron reduction of nitrite to ammonia. Furthermore Nrf^- cells were more sensitive than Nrf^+ cells to treatment with either NO or nitrosating agent *S*-nitroso-*N*-acetylpenicillamine (SNAP). This suggests a

role for NrfA in the detoxification of NO. Recently, protein film voltammetry (PFV) has shown that the NrfA protein has a higher NO reductase activity than either of the dedicated NO reductases flavohaemoglobin and flavorubredoxin (van Wonderen *et al.* 2008). The authors suggest a role for the periplasmic nitrite reductase in pathogenic bacteria as part of the first line of defence against exogenously encountered NO. The site of NO reduction is likely to be near to the periplasmic face of the membrane, which would help to detoxify the toxic free radical before it enters the cell (Poock *et al.* 2002). This is anticipated to allow the flavorubredoxin, NorV to efficiently detoxify any remaining NO under anaerobic conditions. Whilst studies demonstrate that NrfA confers the capacity for NO reduction and detoxification, there is limited data to support an *in vivo* role for NrfA in protection against NO toxicity. The physiological role for NrfA as a NOR therefore remains uncertain.

(i)   **Transcriptional regulation of *nrfA* expression**

Regulation of *nrfA* is complex with expression controlled by at least three regulators and two nucleoproteins (Figure 1.8B) (Browning *et al.* 2002; Browning *et al.* 2005) . Under anaerobic conditions, the global regulator FNR binds to a site centred at position -41.5 relative to the *nrf* promoter and activates expression through interaction with RNA polymerase (RNAP). An *fnr* mutant is unable to produce NO, most likely due to a lack of *nrf* expression, providing evidence for the production of NO by NrfA and its regulation by FNR (Corker and Poole 2003). The FNR protein is sufficient to activate transcription under anaerobic conditions (reviewed by (Guest *et al.* 1996) but FNR-dependent activation is blocked by the binding of the nucleoproteins Integration Host Factor (IHF) and Factor for Inversion Stimulation (FIS) downstream of the *nrf* promoter. This repression integrates two further signals into the regulation of *nrf* expression. Firstly, FNR-dependent activation is reliant upon NarL or NarP, two homologous response regulators controlled by the NarX and NarQ sensor kinases respectively (Rabin and Stewart 1993; Darwin and Stewart 1996). Phosphorylated NarL/P disrupts the binding of the DNA-bending IHF upstream of the promoter, in response to nitrate or nitrite (Browning *et al.* 2002).   Secondly, the nucleoprotein, FIS, functions in catabolite repression to ensure that transcription only occurs under nutrient-limiting conditions (Browning *et al.* 2005). This is in contrast to the FNR-dependent activation of *nir* expression where the catabolite repressor-activator (Cra) protein ensures that transcription only occurs under nutrient-rich conditions (Figure 1.8A) (Tyson *et al.* 1997). This difference in the regulation at the *nrf* and *nir* promoters ensures
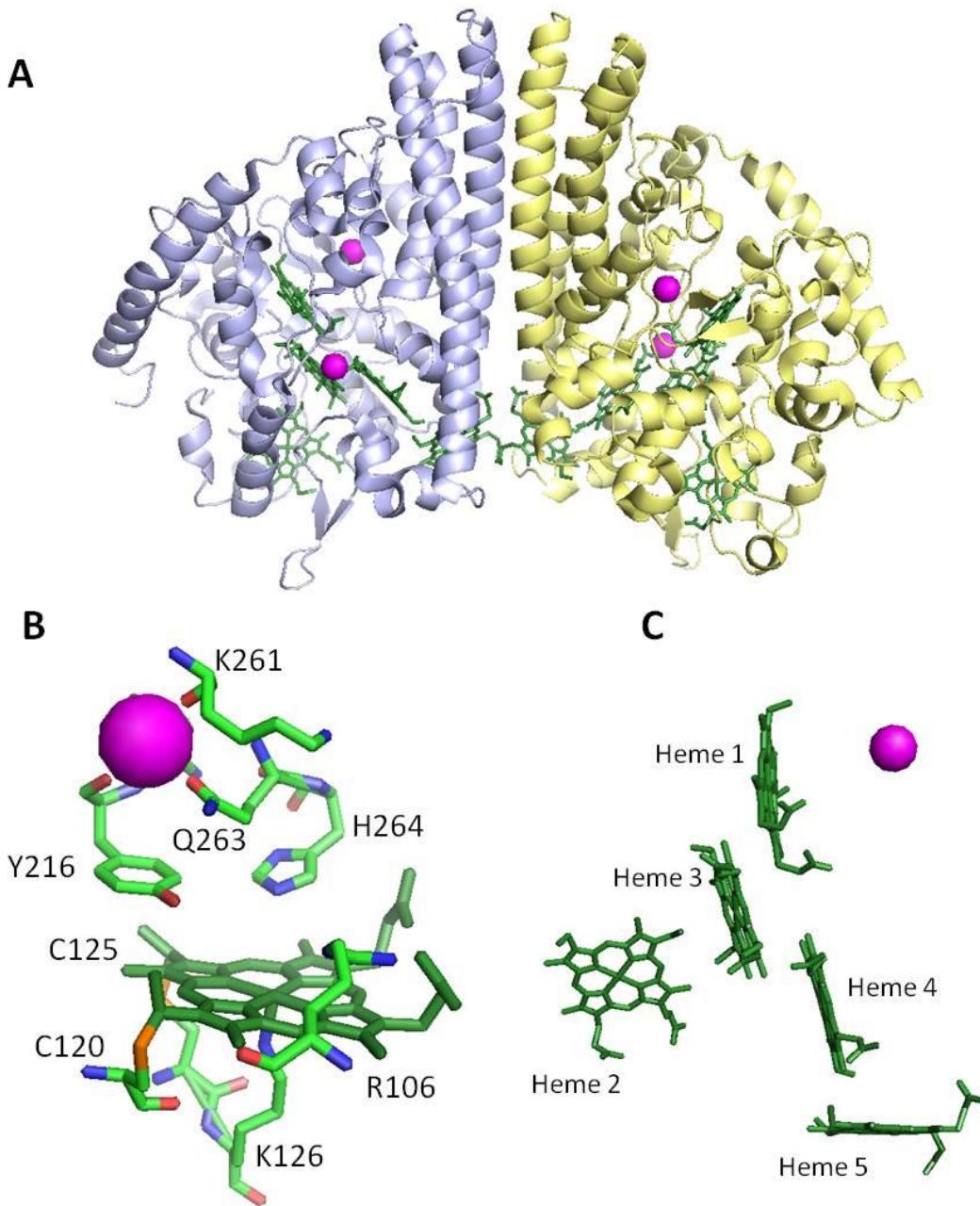
Figure 1.7 - Overall structure of *E. coli* NrfA. (**A**) The NrfA dimer, showing arrangement of monomers (pale blue and pale yellow) about a 2-fold crystallographic axis PDB ID: 1GU6 (Bamford *et al.*, 2002). Heme groups are shown in green and calcium ions in magenta. (**B**) The ligand environment of the active site heme (heme 1). Five conserved residues (R106, K126, Y216, H264 and Q263) provide a positive environment around the active site and may act as proton donors. Once again the catalytic calcium ion is shown in magenta. (**C**) The arrangement of the 5 hemes in an NrfA monomer. Heme 1 is the catalytic heme and the catalytic calcium iron is shown in magenta. The hemes are numbered according to the order of attachment to CXXC motifs on the polypeptide chain. The orientation of the hemes is similar to that shown in (**A**).

that the cytoplasmic and periplasmic nitrite reductases, NirB and NrfA respectively, are expressed under different conditions. In agreement with this, the expression of the *nrf* and *nir* operons at varying nitrate and nitrite levels suggest that NrfA has a physiological role where nitrate (or nitrite) is limiting in the cell environment, whereas NirB functions when the substrate is in excess (Wang and Gunsalus 2000). Recently, yet another regulator was identified as having a role in controlling expression from the *nrf* promoter. There is residual activation by nitrite even in a *narP-narL* double mutant (Rabin and Stewart 1993) and subsequently NsrR was confirmed as a fifth regulator. Anaerobic expression of *nrfA* was increased in an NsrR mutant, identifying NsrR as a negative regulator of *nrfA* expression and repressor titration experiments were fully consistent with this conclusion (Filenko *et al.* 2007). The NsrR protein binds at an upstream site and represses transcription in the absence of nitric oxide (NO) (Figure 1.8B) (Browning *et al.* 2010). It is interesting to note that both enzymes involved in the periplasmic pathway for respiratory denitrification (Nap and Nrf) but not the cytoplasmic pathway (Nar and Nir) are members of the NsrR regulon (Filenko *et al.* 2007).

1.4.2   Flavohaemoglobin, Hmp

The flavohaemoglobin (Hmp), encoded by the *hmp* gene, is an NO dioxygenase (NOD) that converts NO to nitrate at the expense of oxygen and NAD(P)H (Gardner *et al.* 1998b; Gardner *et al.* 2000b; Poole and Hughes 2000; Gardner and Gardner 2002):

$$2NO + 2O_2 + NAD(P)H \rightarrow 2NO_3^- + NAD(P)^+ + H^+$$
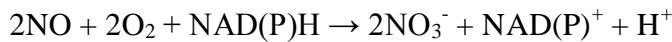
This reaction occurs via a two-electron flavoenzyme mechanism (Gardner *et al.* 2000b). The overall fold of *E. coli* Hmp consists of a heart-shaped structure in which the N-terminal globin domain, the FAD-binding domain and the C-terminal NAD binding domains are clearly distinguished (Figure 1.9A; (Ilari *et al.* 2002). Initially NADH reduces bound FAD and the heme iron is then reduced by $FADH_2$. Oxygen binds to the reduced flavohaemoglobin and the oxygenated Hmp dioxygenates NO to form nitrate at a rate of 240 NO s^{-1} (Figure 1.9B) (Gardner *et al.* 2000a; Gardner *et al.* 2000b). The requirement of oxygen in this reaction correlates with the protection of the citric acid cycle enzyme, aconitase by Hmp under aerobic conditions (Gardner *et al.* 1998a). At low concentrations of oxygen (microaerobic conditions) it has been suggested that Hmp catalyses a similar
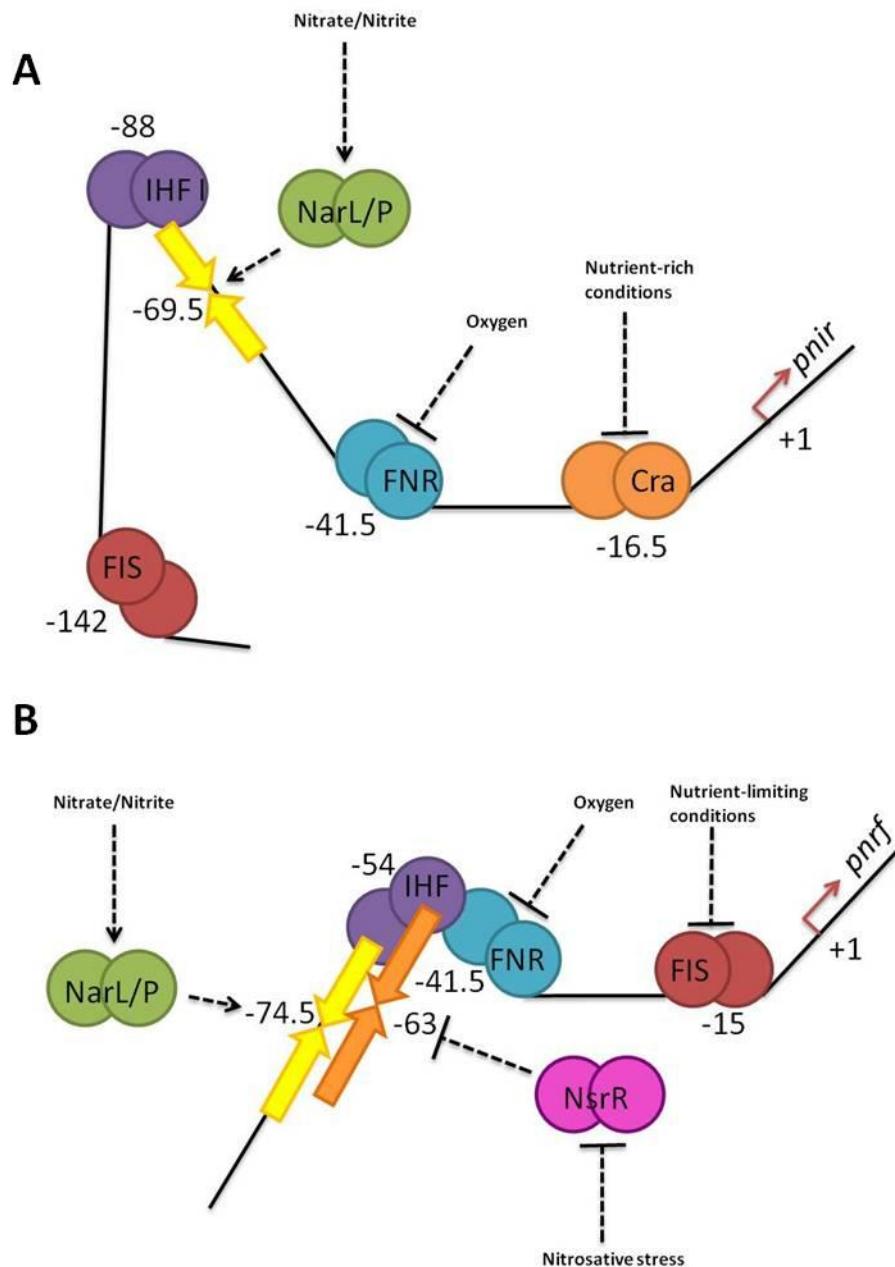
Figure 1.8 - Control of bacterial transcription at the *pnir* and *pnrf* promoters by activators and repressors. (**A**) FNR (blue)-dependent activation of transcription at the *nir*-promoter is repressed by the binding of IHF (purple) and FIS (red) at the -88 and -142 positions respectively (Browning *et al.* 2000). Under nutrient-limiting conditions, the Cra protein (orange) binds to a site centred on position -16.5 and represses transcription (Tyson *et al.* 1997). The binding of NarL/P (green) to a site centred on position -69.5 (yellow arrows), in response to nitrate or nitrite, counteracts the effects of FIS and IHF and enables transcription providing the Cra protein is not bound. (**B**) FNR (blue)-dependent activation of transcription at the *nrf*-promoter is repressed by the binding of IHF (purple) and FIS (red) at positions -54 and -15 respectively. Provided that FIS does not bind and in the presence of nitrate or nitrite, the NarL/P (green) activator binds at a site centred on position -74.5 (yellow arrows) to displace IHF and enable FNR-mediated activation of transcription (Browning *et al.* 2005). Unlike at the *nir*-promoter, FIS binds downstream of the FNR-binding site and functions instead of the Cra protein in catabolite repression; transcription is prevented under nutrient-rich conditions. The binding of the NsrR (pink) repressor to a site that overlaps both the IHF and NarL/P sites (orange arrows) provides a further level of regulation (Ogura and Wilkinson 2001; Hanson and Whiteheart 2005; Rodionov *et al.* 2005; Browning *et al.* 2010). In the presence of nitric oxide (NO), NsrR cannot bind, enabling transcription at the promoter if the other criteria for activation are also met.

42

conversion of NO to nitrate via a nitroxyl-oxidation reaction. Here, NO binds the heme before reaction with oxygen but this reaction is approximately 50% slower than the conventional nitric oxide dioxygenase (NOD) reaction (Hausladen *et al.* 2001). In addition, there is some evidence that Hmp functions under anaerobic conditions. *hmp* mutants are sensitive to NO both aerobically and anaerobically (Gardner *et al.* 1998a), suggesting that a function for flavohaemoglobin might exist in the absence of oxygen. It has been shown that Hmp can act as an NO reductase to produce nitrous oxide (N_2O) but this gives a turnover of only 0.24 NO s^{-1} (Figure 1.9C; (Kim *et al.* 1999), bringing into question the physiological significance of this activity. Comparative studies in aerobically and anaerobically grown *E. coli* cells reveal that Hmp is an efficient NO dioxygenase but exhibits little NO reductase activity. Furthermore, the NOR activity of Hmp did not confer any protection to the NO-sensitive aconitase (Gardner and Gardner 2002). Significantly, transcriptome profiling indicated that the addition of the NO donor *S*-nitrosoglutathione (GSNO) up-regulated *hmp* only under aerobic and not anaerobic conditions (Flatley *et al.* 2005).

(i)   **Anaerobic Regulation of *hmp* expression**

Under anaerobic conditions, regulation of the *hmp* gene is facilitated by the global regulator FNR. Expression of an *hmp-LacZ* fusion increased three- to four- fold in an FNR-deficient mutant, indicating that FNR negatively regulates *hmp* expression (Poole *et al.* 1996). Indeed, inspection of the *hmp* promoter reveals an FNR binding site centred at position +5 (Cruz-Ramos *et al.* 2002). Although FNR primarily senses oxygen via its [4Fe-4S] cluster to regulate the anaerobic lifestyle of *E. coli*, EPR spectroscopy has revealed that it can also respond to NO (Cruz-Ramos *et al.* 2002). A mechanism of derepression has subsequently been proposed in which NO binds to the [4Fe-4S] clusters of FNR dimers, generating dinitrosyl-iron (DNIC) complexes that are no longer able to bind to the FNR-box at the *hmp* promoter (Figure 1.10) (Cruz-Ramos *et al.* 2002). The derepression of transcription under anaerobic conditions in response to NO could support a minor role for the flavohaemoglobin as an NO reductase. When only the dioxygenase (NOD) activity of Hmp is considered, it is unclear how the NO-sensing function of oxygen-sensitive FNR would contribute to the control of *hmp* expression.
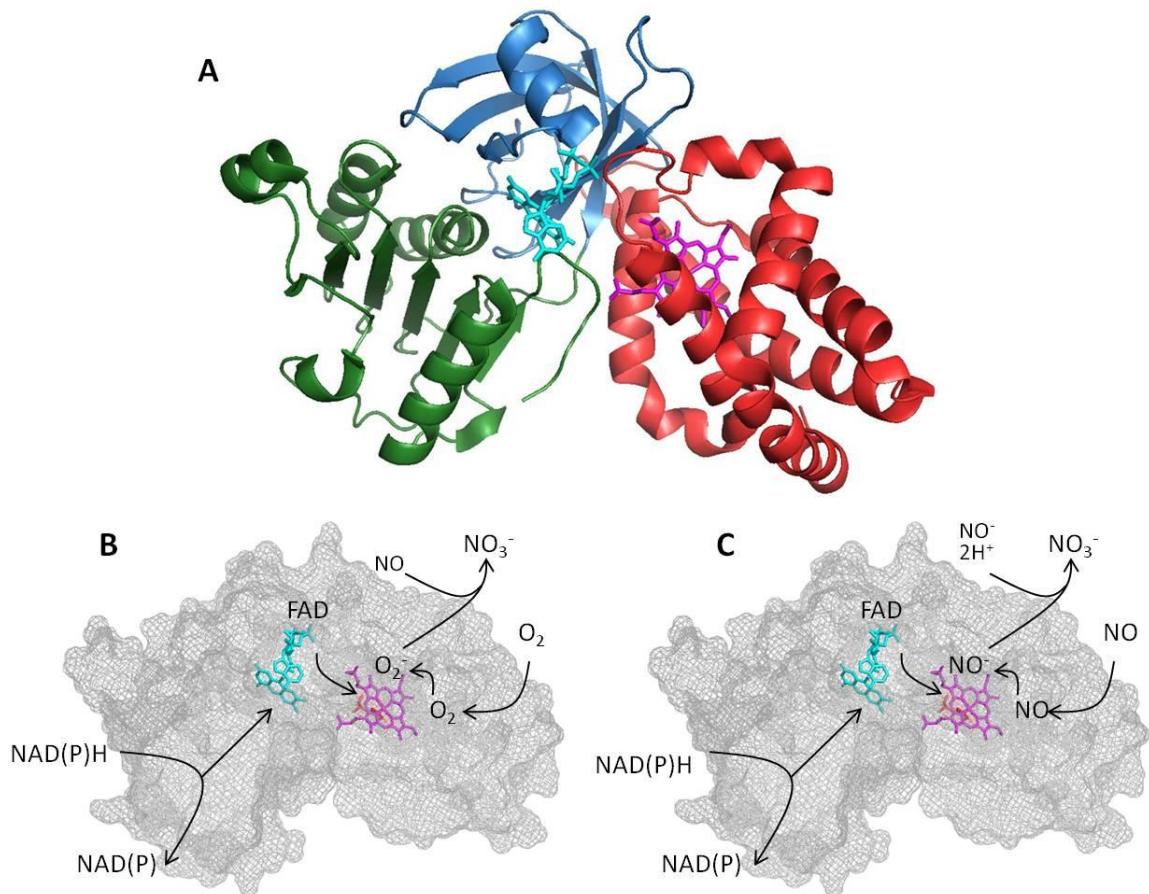
Figure 1.9 – (**A**) **Overall structure of *E. coli* ferric unliganded flavohaemoglobin** (PDB ID: 1GVH).The heart-shaped structure is positioned with the flavin binding domain at the upper apex (sky blue), the globin domain on the lower right side (red), and the NAD-binding domain on the lower left side (green). FAD is shown by a cyan-stick model and the heme group by a magenta stick model. NAD(P)H is not shown bound at the NAD-binding domain (Ilari *et al.* 2002). (**B** and **C**) Electron transfer and ligand reduction in Hmp. The heart-shaped structure (Ilari *et al.* 2002) is shown as a surface-mesh with the FAD and heme emphasised by stick-models. Electrons from NAD(P)H are passed to FAD (cyan) and thence to the single heme (magenta). His-85 is the proximal heme ligand (red); O_2 and other ligands bind in the enlarged distal heme pocket (Poole and Hughes 2000). (**B**) **Aerobic (NOD) activity of Hmp.** O_2 is reduced at the ferrous heme to superoxide. O_2 occupancy at the heme appears to persist during the oxygenation reaction yielding nitrate, probably via a bound peroxynitrite intermediate species (not shown). The alternative nitroxyl-oxidation reaction is not illustrated. (**C**) **Anaerobic (NOR) activity of Hmp.** NO reacts with the ferrous heme to give a nitrosyl species. Electron transfer from heme gives NO^-. N_2O formation may occur via a dimeric species (not shown). In both reactions, the ferric heme resulting from electron transfer to the ligand can be re-reduced in the presence of NAD(P)H via FAD.
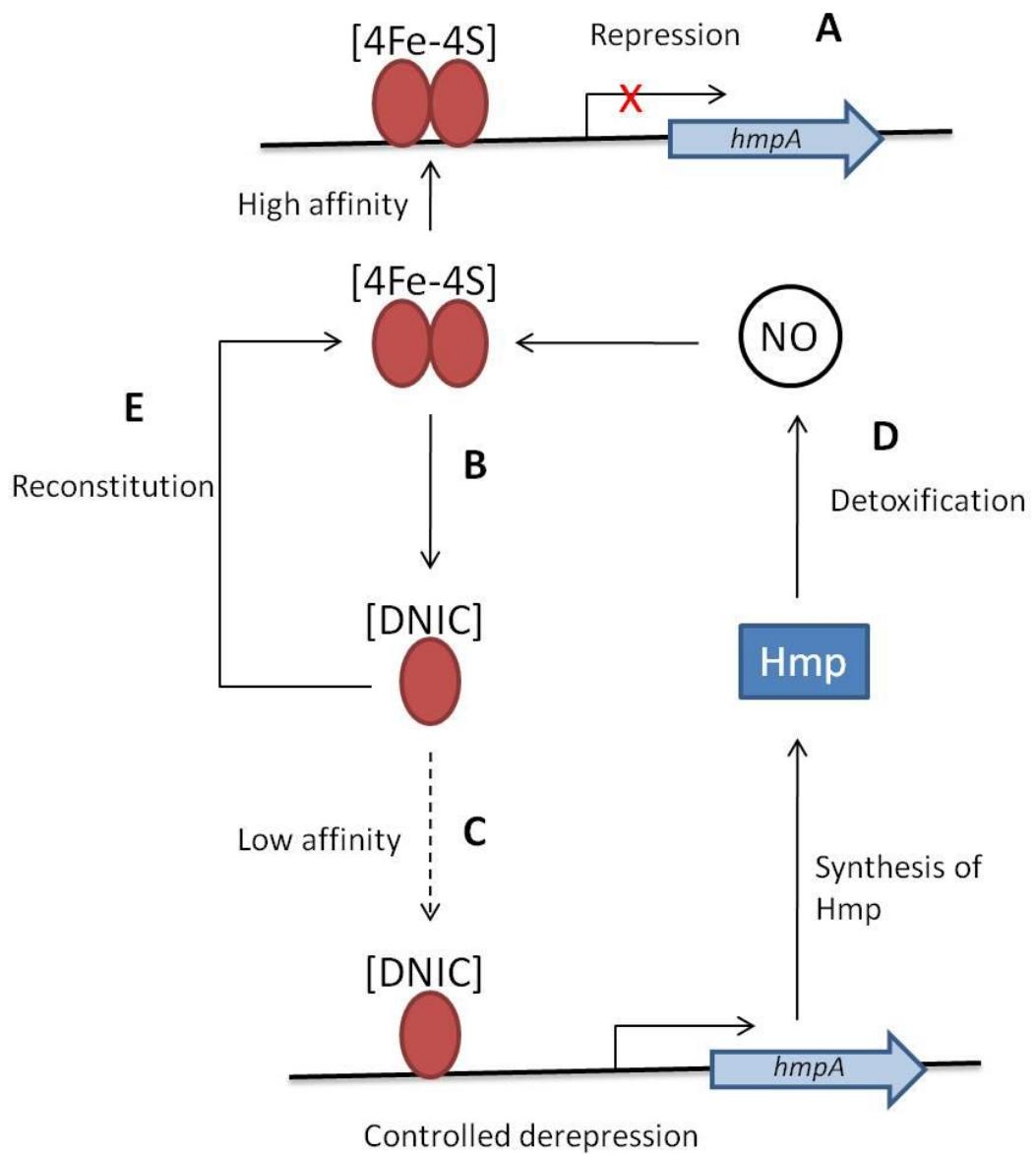
Figure 1.10 - Anaerobic regulation of the *E. coli hmp* gene by the NO-responsive regulator, FNR. FNR forms are represented by red ovals. The $[4Fe-4S]^{2+}$ and DNIC Fe-S clusters are labelled. The *hmp* promoter, NO and Hmp molecules are also shown. (**A**) Under anaerobic condititions, FNR (which is a dimer in the $[4Fe-4S]^{2+}$ form) binds to the *hmp* promoter and prevents expression of *hmp*. (**B**) FNR senses NO by reaction with the $[4Fe-4S]^{2+}$ cluster to generate a dinitrosyl iron (DNIC) complex. This is likely to cause FNR to dissociate into a monomer. (**C**) NO-treated FNR binds the *hmp* promoter with a lower affinity to avoid sudden and complete derepression of the *hmp* gene. (**D**) Eventually depression is achieved and Hmp is synthesized. Hmp can then act to detoxify NO. (**E**) When the NO concentration is reduced, Fe-S cluster-repairing mechanisms can reconstitute the NO sensor FNR. (Cruz-Ramos *et al*. 2002).

(ii)   Regulation of *hmp* expression by NsrR

The *hmpA* gene has already been shown to be subject to regulation by FNR (Cruz-Ramos et al., 2002) and the confirmation that Hmp is part of the NsrR regulon increases the complexity of regulation at the *hmp* promoter (Rodionov *et al.* 2005; Bodenmiller and Spiro 2006; Filenko *et al.* 2007). As mentioned previously, NsrR is a negative regulator of *hmp* expression, and is released from the *hmp* promoter as a result of NO binding to the probable [2Fe-2S] cluster to form DNICs (Figure 1.5; (Tucker *et al.* 2010b). The recent discovery of NsrR means that the relative roles of FNR and NsrR in the negative regulation of *hmp* expression are unknown (Spiro 2007). The binding sites for FNR and NsrR overlap (Bodenmiller and Spiro 2006) and it is not known whether the two regulators can bind simultaneously. There is some evidence *in vitro* to suggest that the NsrR proteins from *S. coelicolor* and *N. gonorrhoeae* are aerobically stable (Tucker *et al.* 2008; Isabella *et al.* 2009). Therefore NsrR may regulate *hmp* expression under aerobic conditions to ensure that transcription only occurs in response to NO. In addition, binding sites for the MetR regulator have been identified in the intergenic region between the *hmp* promoter and the divergently transcribed *glyA* gene (Figure 1.5) (Lorenz and Stauffer 1995). It has been suggested that MetR (the regulator of methionine biosynthesis in *E. coli*) regulates the transcription of both genes depending upon the availability of the homocysteine (Hcy) cofactor (Membrillo-Hernandez *et al.* 1998). When the levels of Hcy are high, MetR-Hcy induces transcription from the *glyA* promoter. RNS apparently act to deplete the pool of available Hcy, allowing MetR to instead promote *hmp* transcription. Activation of *hmp* is lost in a *metR* mutant, evidence in support of this mechanism (Membrillo-Hernandez *et al.* 1998) but MetR-dependent regulation has not been shown for NO itself and methionine biosynthesis genes do not respond to NO under anaerobic conditions (Justino *et al.* 2005b). This mechanism of regulation remains unclear therefore, especially as Mukhopadhyay et al. showed no role for MetR in the regulation of *hmp* (Mukhopadhyay *et al.* 2004). Together, NsrR (and possibly MetR) may function aerobically to ensure that Hmp is only produced in response to nitrosative stress whilst FNR ensures no functional enzyme is present in the absence of oxygen and NO (Figure 1.11). Recent evidence suggests Fur modestly represses expression at the *hmp* promoter in *E. coli* (Hernandez-Urzua *et al.* 2007) but it is not clear whether this is a direct effect; no strongly predicted Fur binding sites have been identified at *hmp* promoters.
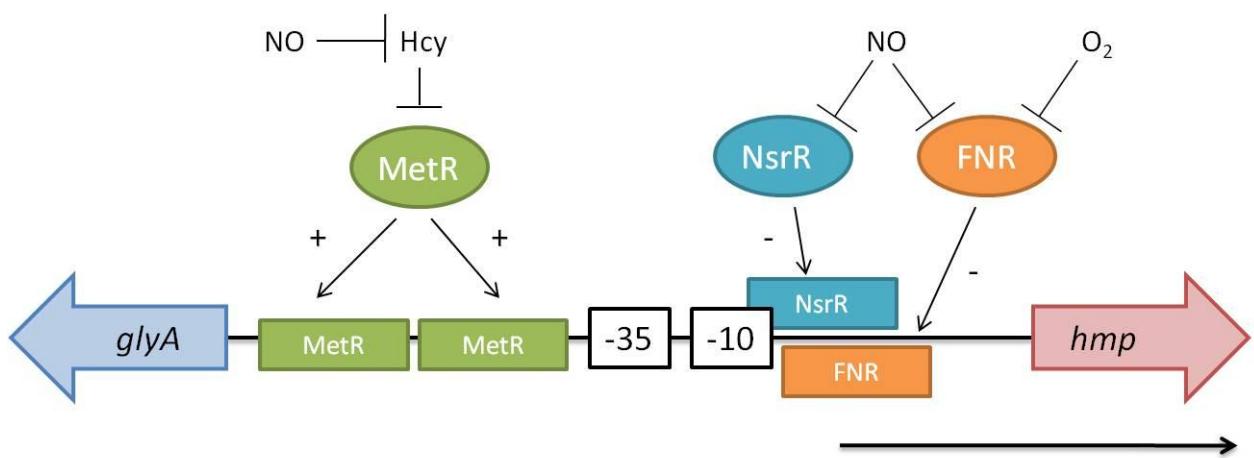
Figure 1.11 - Regulation of flavohaemoglobin (Hmp) expression in *E. coli* (Spiro 2007). Schematic (not to scale) of the intergenic region between the divergently transcribed *glyA* and *hmp* genes. The MetR protein and corresponding binding sites are shown in green, FNR and its corresponding binding site in orange and NsrR and its binding site in blue. The FNR and MetR binding sits have been experimentally verified (Lorenz and Stauffer 1995; Cruz-Ramos *et al.* 2002); the NsrR binding site is inferred from genetic and bioinformatic studies (Lupas and Martin 2002; Bodenmiller and Spiro 2006).

1.4.3   Flavorubredoxin, NorV

The flavorubredoxin (NorV) is an NO reductase (NOR) capable of reducing NO to nitrous oxide under anaerobic conditions (Gardner *et al.* 2002; Gomes *et al.* 2002; D'Autreaux *et al.* 2005). NorV is a member of the A-type flavoproteins which are characterised by two core domains, revealed in the well-studied rubredoxin:oxygen oxidoreductse (ROO) from *Desulfovibrio gigas* (Figure 1.12; (Frazao *et al.* 2000; Gomes *et al.* 2000). The N-terminal region consists of a metallo-β-lactamase-like domain containing the non-heme di-iron active site. A flavodoxin-like domain contains a FMN moiety that is well positioned for electron transfer and receives electrons from a rubredoxin protein to reduce oxygen to water (Chen *et al.* 1993). NorV proteins in enterobacteria contain a third C-terminal domain, pointing to the fact that such proteins are a fusion of the flavoprotein and the rubredoxin redox partner. NorV is therefore known as a flavorubredoxin. The NorV protein has been shown to act as a nitric oxide reductase (NOR) with a turnover of around 15 mol NO-mol NorV^{-1} s^{-1} (Gomes *et al.* 2002). Immediately downstream of the *norV* gene in *E. coli* and co-transcribed with it, is the *norW* gene encoding the NADH:flavorubredoxin oxidoreductase (Gomes *et al.* 2000), which acts as the electron donor for NO reduction. It has been shown that NorW is required for maximal NorV activity both *in vivo* (Gardner *et al.* 2002) and *in vitro* (Gomes *et al.* 2002). There are two possible mechanisms for reduction of NO to nitrous oxide by NorV (Gardner *et al.* 2002). The diferrous centre could initially bind two NO molecules, which are each reduced to form two nitroxyl anions (HNO) with a concomitant oxidation to form a diferric centre. The nitroxyl anions are suggested to combine to form N_2O and water. Alternatively, the diferrous centre could initially bind a single NO molecule which is reduced by two electrons to produce a Fe^{3+}-O-Fe^{3+}-NO_2^- (H+) species. Reaction of this intermediate with a second NO molecule is suggested to form the product N_2O.
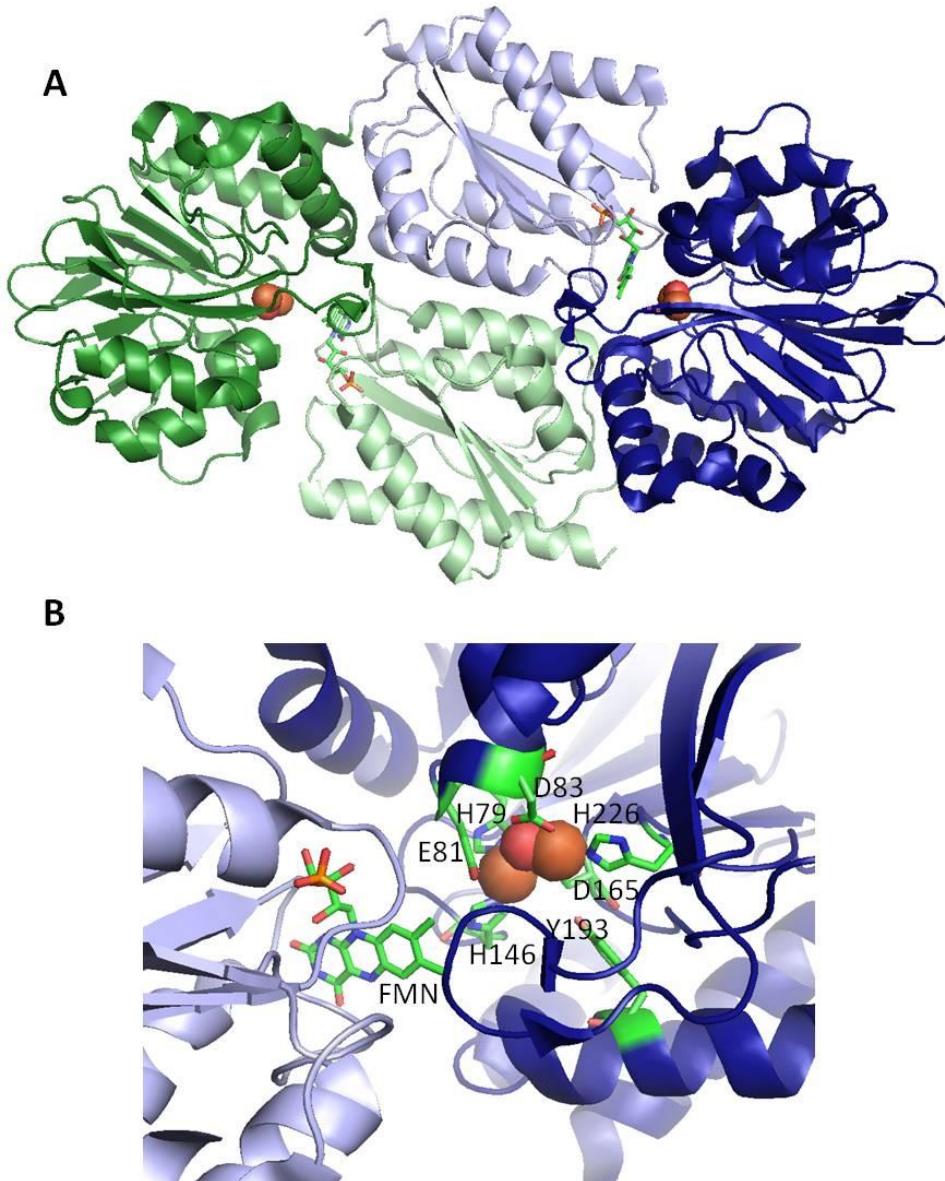
Figure 1.12 - Structure of *Desulfovibrio gigas* rubredoxin:oxygen oxidoreductse (ROO). PDB ID: 1E5D (Frazao *et al.* 2000). (**A**) – Modular structure of the dimeric ROO enzyme (monomers in blue and green). The two core domains are shown; the metallo-β-lactamase-like domain (dark blue or dark green) containing the non-heme di-ferrous active site (Fe^{2+}-O-Fe^{2+}; iron =orange spheres; oxygen = red) and the flavodoxin-like (light blue or light green) domain that contains the FMN moiety (stick model). *E. coli* NorV has an additional C-terminal rubredoxin-like domain (Frazao *et al.* 2000). (**B**) di-iron centre of ROO. *E. coli* flavorubredoxin (NorV) has 34% sequence similarity with *D. gigas* ROO and the conserved residues at the di-iron centre are labelled (H146, E81, H79, D83, H226, D165 and Y193 in *D. gigas*). Reduction of NO is proposed to occur at the di-iron centre since NO binds the di-ferrous centre (Fe^{2+}-O-Fe^{2+}) to form a ferric-nitroxyl species, detectable by EPR (Gomes *et al.* 2000). The proximal Y193 residue is thought form a hydrogen bond interaction with bound NO. The location of the low potential FMN isoalloxazine ring is shown; it is positioned to reduce the di-ferric centre by two electrons for cycles of NO reduction. The coordination of FMN is not shown (Gardner *et al.* 2002).

(i)   Transcriptional regulation of *norVW* expression

Divergently transcribed from the *norVW* genes is the *norR* gene, located 187bp upstream of the *norV* start site (Tucker *et al.* 2005). The NorR protein is a bacterial Enhancer Binding Protein (bEBP) of the sigma 54 (σ^{54}) class and has been shown to activate transcription of *norV* and *norW* in response to NO (Hutchings *et al.* 2002b; Gardner *et al.* 2003). The role of bEBPs in the initiation of transcription in bacteria will be discussed in Chapters 2 and 3. As the subject of this work, the bEBP NorR will be discussed in greater depth in Chapter 4.

1.5   Overview

Given the toxicity of NO at higher concentrations, it is easy to understand why bacteria including *E. coli* have evolved to use enzymes with NO reductase (NOR) and dioxygenase (NOD) activities (Figure 1.13A). The combination of the aerobic NODs and anaerobic NORs allow detoxification of NO across the physiological range of oxygen concentrations (Gardner *et al.* 1998a; Gardner and Gardner 2002; Gardner *et al.* 2002). Indeed, both NorV (a NOR) and Hmp (a NOD) show evidence of activity under microaerobic conditions (Gardner *et al.* 2002) that may ensure complete protection against nitrosative stress as bacteria make the transition between aerobic and anaerobic states. Together such enzymes help to detoxify NO before it reaches lethal concentrations. However where bacteria encounter higher levels of NO production (e.g. released by macrophages of the mammalian immune system), NO-mediated damage will undoubtedly occur. There exist therefore, various systems in bacteria to repair damage to proteins and to DNA. Since the production of detoxification enzymes and repair systems is energetically expensive to the cell, their expression is tightly regulated by a number of regulators (Figure 1.13B). This regulation is commonly mediated by iron and Fe-S cluster-containing proteins that act as direct NO-sensors to activate (or derepress) transcription of key genes only under conditions of nitrosative stress (Figure 1.6). Since many pathogenic bacteria encounter NO as part of the immune response, inhibition of the detoxifying NODs and NORs and/or their regulators may represent a novel therapy to enhance the host response or the action of antibiotics in pathogenic infections (Gardner *et al.* 2002).
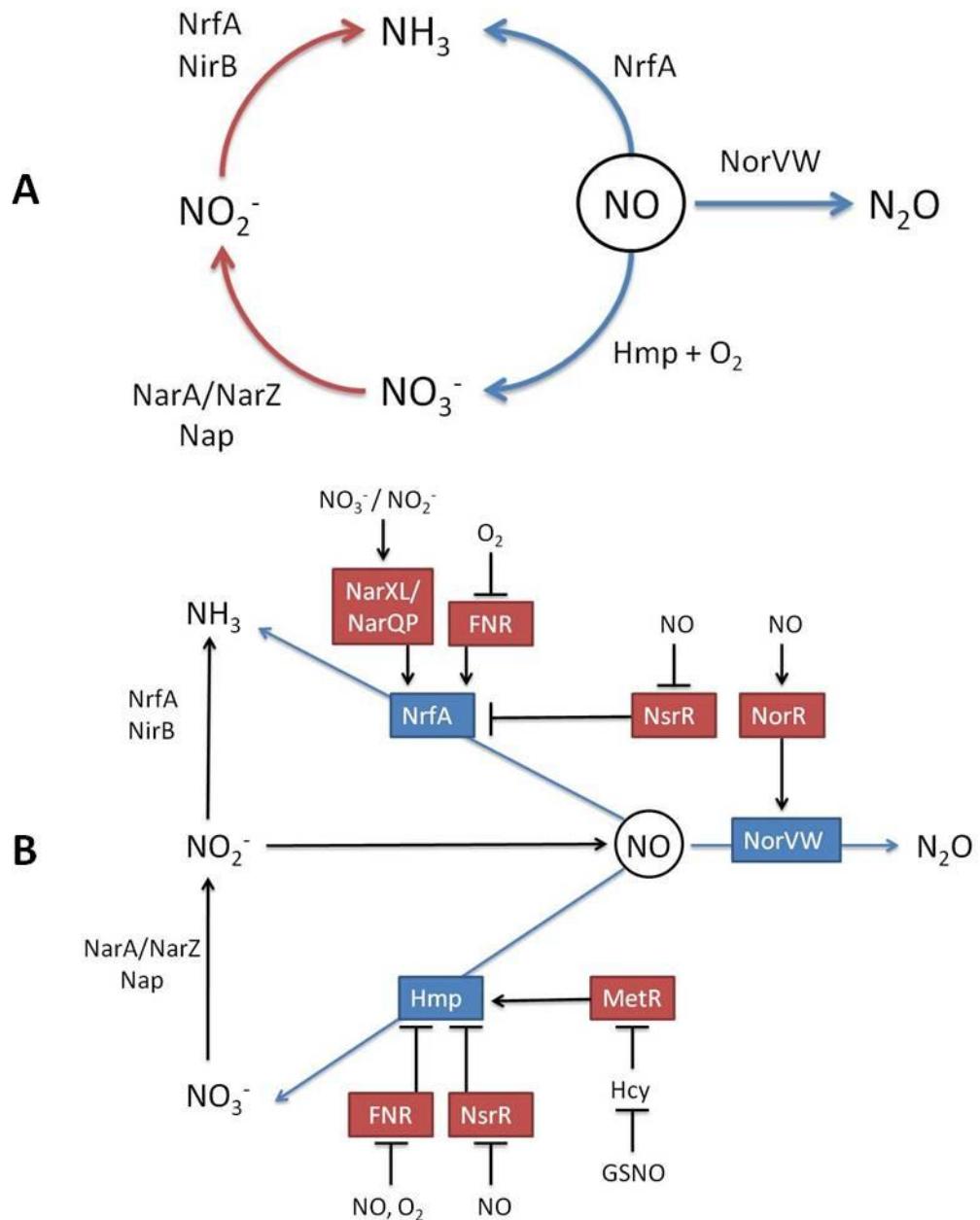
Figure 1.13 – (**A**) **Pathways for NO-formation and metabolism in *E. coli*.** Under anaerobic conditions, *E. coli* reduces nitrate to ammonium in the cytoplasm (using the nitrate reductase NarA/NarZ and the nitrite reductase NirB) and the periplasm (using the nitrate reductase Nap and the nitrite reductase NrfA). NrfA has also been suggested to act as an NO reductase (NOR) in the detoxification of NO. In the absence of oxygen, NO can be reduced to nitrous oxide by the NOR flavorubredoxin (NorVW). In the presence of oxygen the NO dioxygenase (NOD) Hmp oxidises NO to nitrate. Reactions utilised in respiratory denitrification are shown by red arrows whilst detoxification is shown by blue arrows (Spiro 2006). (**B**) **Regulation of NO-detoxification enzymes.** The NORs NrfA and NorVW and the NOD Hmp are shown in blue. Regulators of these enzymes are shown in red along with their signals. Positive regulation is denoted by arrows, negative regulation is donated by perpendicular lines. GSNO = *S*-nitrosoglutathione; Hcy = homocysteine (Spiro 2007).

Chapter 2 – Role of sigma factors in the initiation of transcription in bacteria

2.1   The core RNA Polymerase

In contrast to eukaryotes and archaea, the transcription of all genes in bacteria is dependent on a single form of the core RNA polymerase (RNAP) enzyme (E). This multi-subunit polymerase (~400 kDa) has a subunit composition of $\alpha_2\beta\beta'\omega$ (Figure 2.1). Structural studies of the bacterial core enzyme reveal an overall "crab-claw" shape with the large β and β' subunits (~150 and 155 kDa respectively) forming the two "pincers", either side of the active site (Zhang *et al.* 1999; Murakami *et al.* 2002a; Murakami *et al.* 2002b). This is a feature also present in the archaeal RNAP (Hirata *et al.* 2008) and eukaryotic RNAPs (Fu *et al.* 1999; Cramer *et al.* 2000; Cramer *et al.* 2001) which show a similar overall structure. Each of the two identical α-subunits (37 kDa) consists of two domains that are connected via a ~20 amino acid flexible linker (Blatter *et al.* 1994). The ~26 kDa N-terminal domain (αNTD) dimerises to facilitate the assembly of the β and β' subunits. The ~9 kDa C-terminal domain (αCTD) binds to promoter DNA and interacts with a diverse range of activators to modulate the level of transcription (Gourse *et al.* 2000; Browning and Busby 2004). Lastly, the small (91-residue, 11 kDa) ω subunit (Mathew and Chatterji 2006) assists in the folding and incorporation of the β' subunit in the final stage of the core RNAP assembly (Ghosh *et al.* 2010). In addition to its structural role, the ω subunit has been shown to play a functional role in the stringent response since in the absence of ω, RNAP cannot respond to the effector ppGpp (Vrentas *et al.* 2005).

2.2   Initiation of transcription

The first step in the initiation of transcription in bacteria is promoter recognition (Busby and Ebright 1994). Although the core enzyme (E) contains all the catalytic elements required for transcription, it can only bind to promoter sequences weakly and in a largely non-specific manner (Paget and Helmann 2003). The major promoter-RNAP contact is facilitated by the αCTD domains which bind to the UP element, a ~20 bp sequence upstream of the transcriptional start site of some promoters (Ross *et al.* 2001). In order to initiate transcription of specific genes, the core RNAP (E) associates with a sixth subunit, the sigma factor (σ) to yield the RNAP holoenzyme ($\alpha2\beta\beta'\omega\sigma$, E$\sigma$). σ factors confer the specificity of RNAP for different promoters and are often associated with the expression of sub-sets of genes required for different cellular activities (Buck *et al.* 2000). The σ factor directs the binding of the core enzyme via interaction with three different promoter elements. σ domain 2.4 recognises the conserved -10 sequence whilst σ domain 4.2
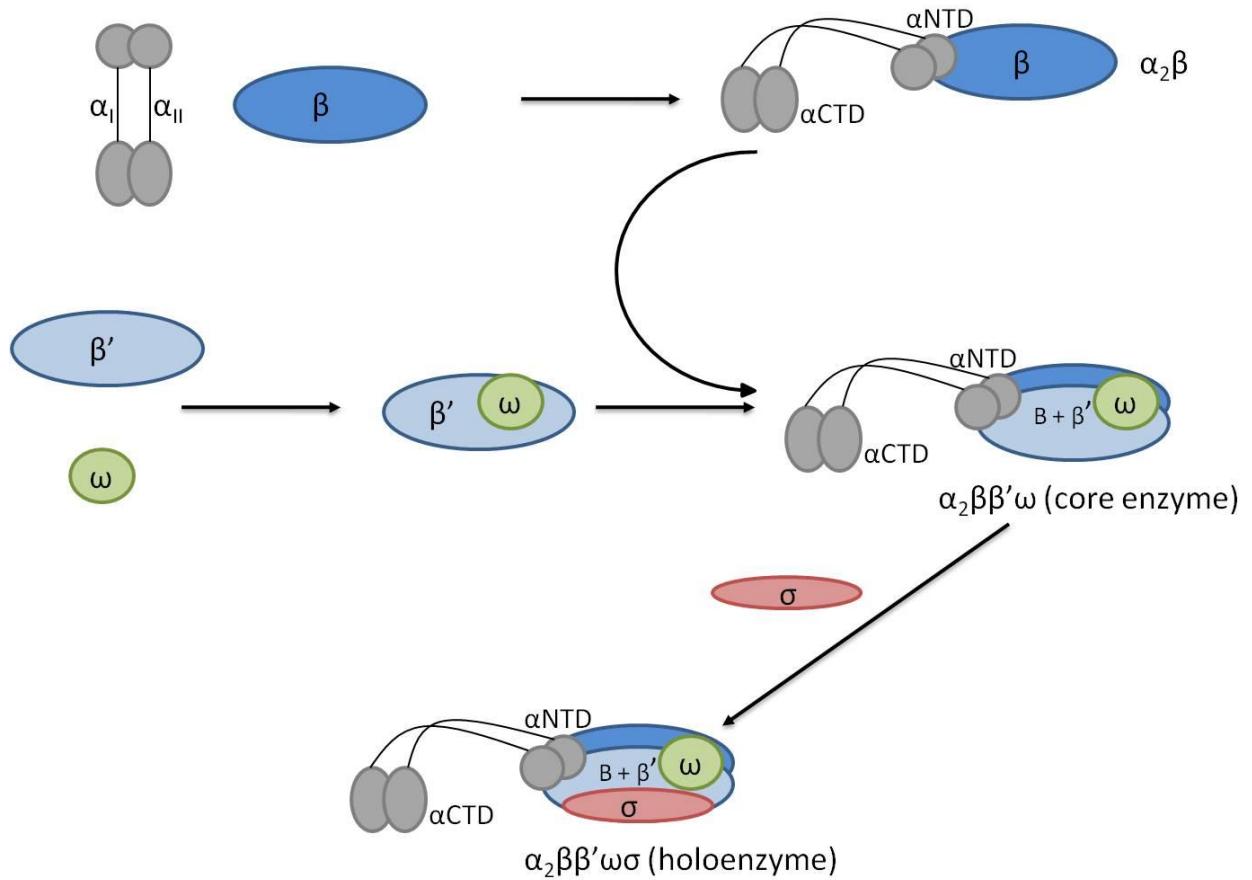
Figure 2.1 - Schematic showing the assembly of the bacterial RNA polymerase holoenzyme. In the first step, the two identical α subunits (37 kDa) dimerise. Each α subunit is composed of a ~26 kDa N-terminal domain (αNTD) and a ~9 kDa C-terminal domain (αCTD), connected via a flexible ~20 amino acid linker. Dimerisation of α provides the scaffold for the assembly of the core enzyme. The ~150 kDa β subunit binds to α_2. The 11 kDa ω subunit facilitates the folding of the ~ 155 kDa β' subunit and together β' and ω bind to form the core enzyme composed of $\alpha_2\beta\beta'\omega$. In order to bind to and transcribe from specific promoters, the core enzyme must associate with one of the available σ-factors in the cell to form the holoenzyme ($\alpha_2\beta\beta'\omega\sigma$).

recognises the conserved -35 sequence (Campbell *et al.* 2002; Murakami *et al.* 2002a). The extended -10 element, located immediately upstream of the -10 sequence of some promoters , is 3-4 bp in length and recognised by domain 3 of the σ factor (Barne *et al.* 1997; Murakami *et al.* 2002a; Sanderson *et al.* 2003).

Once the polymerase has bound to the promoter, open complex formation must occur to open up the double-stranded DNA for transcription (deHaseth *et al.* 1998). During isomerisation, unwinding of the duplex DNA is initiated at the -10 promoter element (Guo *et al.* 2000). A conserved threonine residue in σ (T429 in *E. coli*) destabilises the stacking of the conserved adenine base at the -11 position, causing it to "flip" out and stack against a tyrosine residue in region 2.3 of σ (Schroeder *et al.* 2007; Schroeder *et al.* 2008; Schroeder *et al.* 2009). Subsequently, the template strand is loaded into the RNAP active site and synthesis of an RNA chain begins with initial short-chain products being formed and released in a process known as abortive initiation. During this time, the holoenzyme moves forward but the contacts between the trailing edge of Eσ and the -35 promoter element remain intact. This causes the DNA strands of the -10 promoter element to be extruded from the main DNA channel in a process called DNA scrunching (Kapanidis *et al.* 2006; Revyakin *et al.* 2006). It has been proposed that the release of the energy stored in the "scrunched" intermediate is responsible for the disruption of Eσ-promoter contacts that leads to promoter clearance. Eventually, the sigma factor may be released (Mooney *et al.* 2005) and the core enzyme is able to participate in the elongation process, catalysing the addition of complementary nucleotides to form an extending RNA chain before termination occurs to complete the transcription process (Figure 2.2).

2.3 Regulation by Sigma factors

Since their discovery (Burgess *et al.* 1969) it has become clear that the role of the sigma factor is central to the ability of the RNA polymerase enzyme to carry out transcription. Not only do sigma factors direct the binding of the polymerase to specific promoters, they also inhibit non-specific initiation, mediate the isomerisation of the closed promoter complex and regulate the process of promoter escape (Buck *et al.* 2000). The most abundant σ factor in *E. coli* is σ^{70} (RpoD). Discovered in 1968 (Burgess *et al.* 1969), the 70 kDa protein is often referred to as the housekeeping factor as it transcribes the majority of genes in growing cells. Although some bacteria contain only one σ factor e.g. *Mycoplasma genitalium*, the majority of bacteria contain multiple σ factors. For example, the
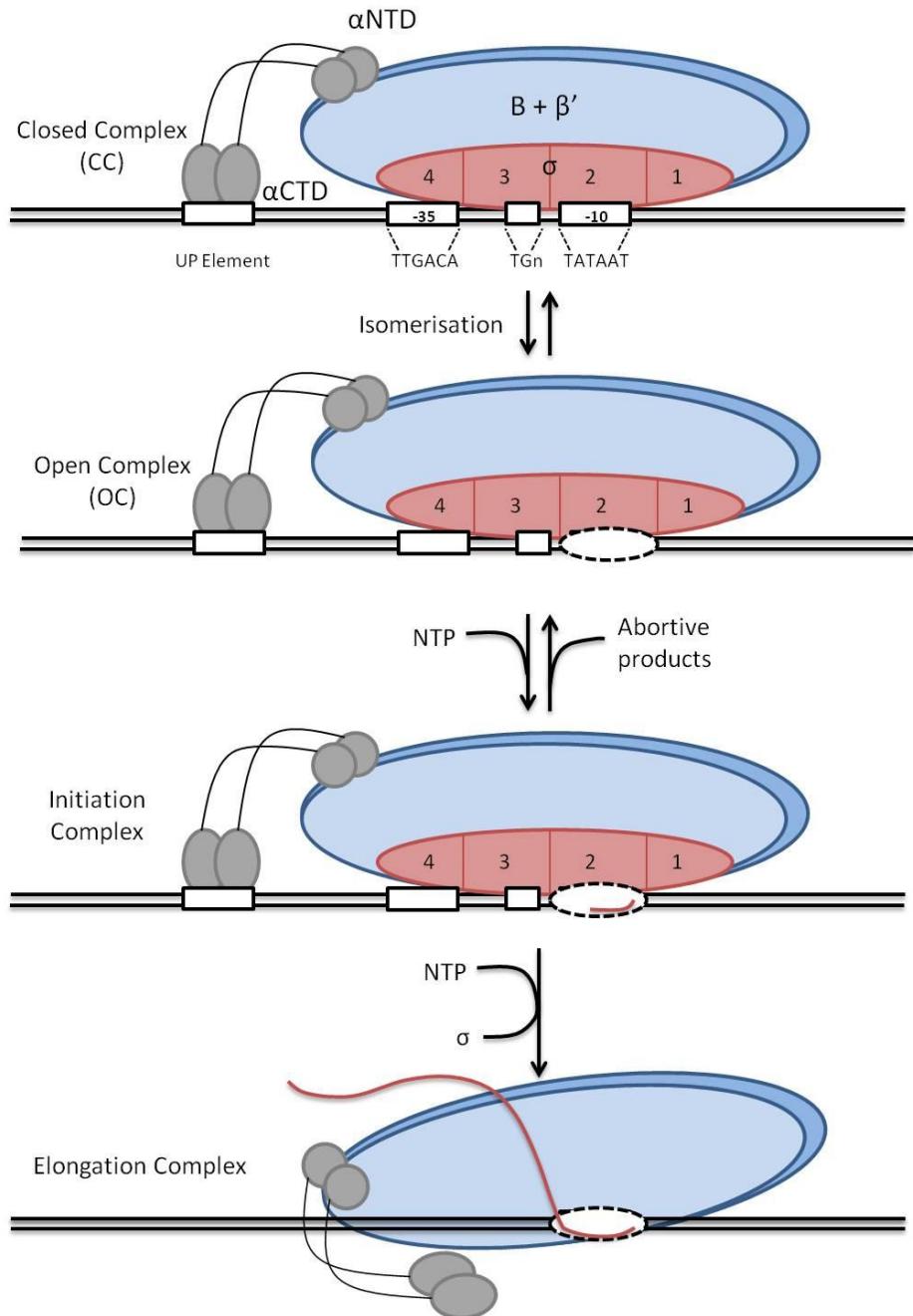
Figure 2.2 - Schematic of the pathway of bacterial transcription initiation (σ^{70}-family dependent). The σ factor directs the RNA polymerase holoenzyme (ω subunit not shown) to bind at the -10 (recognised by σ region 2.4) and -35 (recognised by σ region 4.2) promoter elements. The -10 and -35 consensus sequence is shown for the housekeeping σ^{70}. An additional interaction can form between the extended -10 element (if present) and σ domain 3 at some promoters. The core enzyme itself can interact with the UP element (if present), a ~20 bp sequence upstream of the -35 element that is recognised by the αCTD domains. The promoter recognition sequences as well as the different subunits of the holoenzyme are labelled. After promoter binding, the closed complex (CC) isomerises to form the open complex (OC): the double stranded DNA "melts" or unwinds to form the transcription "bubble". The initiating complex then forms and RNA synthesis begins using the single-stranded DNA as a template. Initially, small abortive RNA products are produced (abortive initiation) before the transcribing complex moves into the elongation phase. σ may be released and the mRNA molecule is formed as the RNA chain length increases. Figure adapted from (Browning and Busby 2004).

actinomycete *Streptomyces coelicolor* which has a complex life cycle and is highly adapted to living in soil environments, contains 63 alternative sigma factors (Gruber and Gross 2003). A vast range of regulatory strategies exist in bacteria to control the level of these alternative σ factors in the cell. Of particular importance is the production of a diverse range of anti-σ factors that bind and sequester their corresponding sigma factors in response to environmental cues (Helmann 1999). In addition to the housekeeping σ^{70}, *E. coli* contains six alternative σ factors that are up-regulated in response to various environmental conditions or stresses (Ishihama 2000): σ^{38} (RpoS), σ^{32} (RpoH), σ^{28} (RpoF), σ^{19} (FecI), σ^{24} (RpoE) and σ^{54} (RpoN). As these σ factors accumulate, they compete with σ^{70} for binding to the core RNA polymerase. In this way the bacterial cell is able to channel a greater proportion of its energy and resources towards the stress response. For example, in response to the stress induced from heat shock, the σ^{32} (RpoH) and σ^{24} (RpoE) σ factors bind core RNAP to activate the expression of genes that assist the cell in surviving increases in temperature (Yura and Nakahigashi 1999; Raivio and Silhavy 2001). Based on sequence similarity, the recognised promoter consensus sequence and their mode of activation, the bacterial σ factors can be placed in two distinct groups: the σ^{70}-family and the σ^{54} family (Wosten 1998).

The sigma factors σ^{38}, σ^{32}, σ^{28}, σ^{19}, σ^{24} as well as the housekeeping σ^{70} factor are all members of the σ^{70}-class. Each member binds to conserved -10 and -35 promoter elements, although the consensus sequences and spacing differs for each sigma-factor. The $E\sigma^{70}$ holoenzyme recognises and binds to the consensus sequences TTGACA at -35 and TATAAT at -10 and the spacing between these sequences is crucial for expression (Harley and Reynolds 1987). Sequence comparison of the σ^{70}-family of σ factors reveals four conserved helical domains (σ_1, σ_2, σ_3 and σ_4), connected by flexible linkers (Figure 2.3A). Each domain can be divided into functionally distinct sub-regions that have roles in promoter recognition, core binding and isomerisation (Murakami and Darst 2003; Haugen *et al.* 2008).
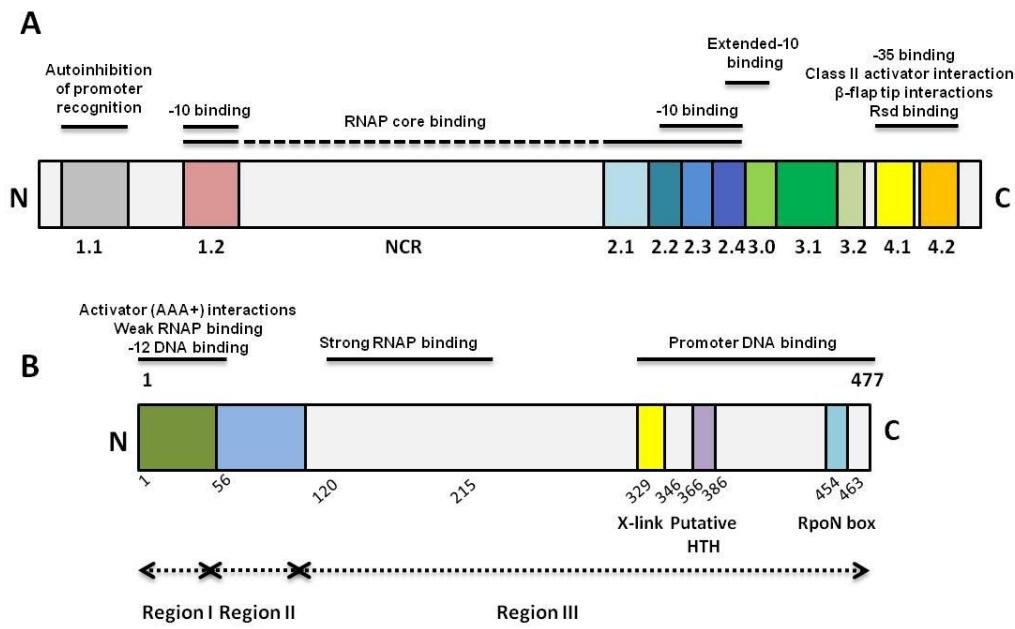
Figure 2.3 – The σ⁷⁰ and σ⁵⁴ sigma factors. Adapted from (Ghosh *et al.* 2010). (**A**) Schematic showing the sequence and conserved structural elements of the σ⁷⁰ factor. Suggested roles for the structural elements are indicated above. (**B**) Domain organisation of σ⁵⁴. *E. coli* σ⁵⁴ (residues 1-477) consists of 3 regions (I to III). DNA binding motifs include the DNA cross-linking (XLINK) region, the helix-turn-helix (HTH) motif and the RpoN box, all present at the C-terminus. Region I interacts with the activator of transcription. Region II is often acidic and occasionally absent. The location of the main core RNAP binding determinants (residues 120-215) is shown (Buck *et al.* 2000).

2.4 σ^{54}-dependent transcription

Unlike the σ^{70}-family, the σ^{54}-family of sigma-factors contains just a single member, σ^{54}, which shows little sequence similarity to the σ^{70}-class (Merrick 1993; Buck *et al.* 2000). Although members of both families associate with the same core polymerase enzyme, the resulting holoenzymes activate transcription by entirely different mechanisms. Both the σ^{70}-type and the alternate σ^{54} factors form holoenzyme-promoter complexes with a default closed and non-productive form (Guo *et al.* 2000) (Figure 2.4). However, the requirements for the formation of an open promoter complex differ. σ^{54} binds to different consensus sequences that are more strongly conserved than σ^{70}. Binding occurs at the GG -24 and TGC -12 elements (Morett and Buck 1989) that are part of the wider sequence YTGGCACGrNNNTTGCW (highly conserved = upper case, weakly conserved = lower case, N = non-conserved, Y = pyrimidines, R = purines, W is A or T, (Barrios *et al.* 1999). The -12 element is critical for σ^{54}-dependent transcription; changes at the -12 position of the DNA or substitution of amino acids of the protein, to destroy interaction with this element causes deregulated transcription (Guo *et al.* 1999; Wang *et al.* 1999). It appears that interaction of σ^{54} with the -12/-11 fork junction prevents binding of the holoenzyme to the non-template strand, a key step in the DNA melting process (Guo *et al.* 2000). So the σ^{54} holoenzyme binds promoter sequences tightly in such a way that isomerisation is not spontaneous; the holoenzyme is transcriptionally silent (Cannon *et al.* 2003; Cannon *et al.* 2004). Therefore, initiation of σ^{54}-dependent transcription is unique in that it requires an activator of the AAA+ (ATPases Associated with various cellular Activities) class that couples the energy produced from ATP hydrolysis to changes in the structure of the σ^{54} subunit and the fork junction. The inhibitory interaction at the -12/-11 position is removed so that the thermodynamic and kinetic barriers that restrict open complex formation in the σ^{54} system are overcome (Cannon 2000; Guo *et al.* 2000; Cannon *et al.* 2001; Zhang *et al.* 2002). In contrast, the σ^{70} holoenzyme binds to the consensus -10 (TATAAT) and -35 (TTGACA) sequences to form an energetically unfavourable closed complex (CC) that is readily converted into an open complex (OC) without a requirement for activators (Guo *et al.* 2000).
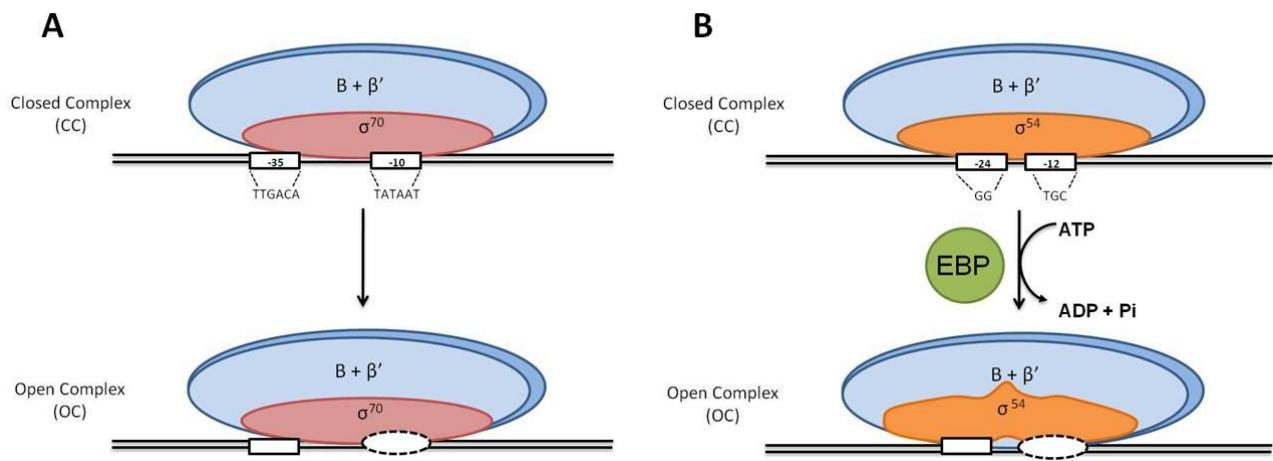
Figure 2.4 - Activation of bacterial transcription by the RNAP-σ^{70} (A) and RNAP-σ^{54} (B) holoenzymes. The σ^{70} factor directs the binding of polymerase to the consensus -10 (TATAAT) and -35 (TTGACA) sequences to form an energetically unfavourable closed complex (CC) that is readily converted into an open complex (OC) to initiate transcription. In contrast, the σ^{54} factor directs the binding of RNAP to conserved -12 (GG) and -24 (GC) promoter elements that are part of the wider consensus YTGGCACGrNNNTTGCW (Barrios et al., 1999). This forms an energetically favourable closed complex (CC) that rarely isomerises into the open complex (OC). In order to form the transcription "bubble", a specialised activator (an Enhancer Binding Protein, EBP) must bind and use the energy from ATP hydrolysis to re-model the holoenzyme.

Since the regulation of σ^{54} and σ^{70}-dependent transcription is so different, it is pertinent to examine the evolutionary advantages of regulating the transcription of genes through the activation of the alternative σ factor rather than by regulating the pool of available σ^{70}. Due to the requirement of an activator, transcription is tightly regulated and σ^{54}-dependent transcription occurs primarily in response to cellular and extracellular signals that regulate the activity of the AAA+ protein. As a result, the activation of transcription occurs rapidly and specifically. This is important since σ^{54} commonly binds to the promoters of genes associated with the bacterial stress response. For example in response to the presence of Nitric Oxide (NO), NorR activates transcription from the *norV* promoter, leading to the expression of the flavorubredoxin (NorV) that functions in NO detoxification (Hutchings *et al.* 2002b). In addition to regulation of transcription via signal-sensing bEBPs, the output from σ^{54}-dependent promoters can be controlled in response to global regulatory signals (Shingler 2010). This regulation could be mediated by factors that counteract the binding of either the holoenzyme at the promoter or of the activator (exclusion). In *Pseudomonas* sp. strain ADP, the bEBP NtrC has been shown to activate σ^{54}-dependent transcription of the *atzR* gene from solution (Porrua *et al.* 2009). The AtzR protein acts as a regulator of cyanuric acid metabolism but also autoregulates its own expression by binding to a site that overlaps the *atzR* promoter to inhibit the formation of the closed complex by the σ^{54}-RNA polymerase (Porrua *et al.* 2009). In *Klebsiella aerogenes*, NtrC activates transcription from the σ^{54}-*nac* promoter under nitrogen-limiting conditions. In this case, Nac negatively autoregulates its own expression by a mechanism of anti-activation; it binds within the intergenic region (between the upstream NtrC binding sequences and the promoter) to prevent a productive interaction between the activator and the holoenzyme (Feng *et al.* 1995). Both the mechanisms of exclusion and anti-activation seem to be utilised in regulating the activation of transcription by DctD at the σ^{54}-*dctA* promoter. cAMP-bound CRP (cAMP receptor protein) is able to bind to two sites that overlap the binding sites of the bEBP but is also able to directly interact with the holoenzyme to inhibit transcription (Wang *et al.* 1998). *In vitro* data suggest that this cAMP-CRP-bound holoenzyme forms an alternative closed complex that slowly converts into one that is subject to bEBP-activation. Therefore, in some cases the activation of σ^{54}-dependent transcription by bEBPs can be modulated in response to regulators that communicate global signals.

2.5   σ^{54} domain architecture

The σ^{54} factor, encoded by the *rpoN* (*ntrA*) gene, is composed of three regions, based on function (Figure 2.3B) (Cannon *et al.* 1997; Buck *et al.* 2000; Bordes *et al.* 2003; Ghosh *et al.* 2010). Region I (residues 1-56 in *E .coli*) is a glutamine and leucine rich sequence that represents the regulatory domain of σ^{54}, binding to the -12 promoter element to form the RNAP-σ^{54} regulatory centre (Wigneshweraraj *et al.* 2001). This nucleoprotein structure prevents spontaneous open complex formation (Wang *et al.* 1995; Syed and Gralla 1998; Cannon *et al.* 1999b). Direct binding of region I to the central, ATP-hydrolysing domains of the activators PspF and NifA has been shown in the presence of nucleotide (Chaney *et al.* 2001; Bordes *et al.* 2003) and biochemical studies show that deletion of region I bypasses the requirement for an activator when pre-melted DNA is used (Guo and Gralla 1998; Chaney and Buck 1999). This indicates that this region is the target of the AAA+ activator in σ^{54}-dependent transcription. Region II (residues 57-107 in *E. coli*) is variable in amino acid composition and length, ranging from 26 residues in *Rhodobacter capsulatus* to 110 residues in *Bradyrhizobium japonicum* but can be characterised by the predominance of acidic residues (Southern and Merrick 2000). This region is not essential for σ^{54}-dependent transcription; region II is virtually absent in some bacterial species e.g. *Bacillus subtilis* (Buck *et al.* 2000). However, deletions in *Klebsiella pneumoniae* σ^{54} region II significantly impair the activity of holoenzyme in open complex formation (Southern and Merrick 2000). Current evidence suggests region II has roles in DNA binding (Cannon *et al.* 1999a) and DNA melting (Wong and Gralla 1992). The C-terminal Region III (residues 108-477 in *E. coli*) is well conserved, containing the major determinants for binding to promoter DNA (residues 329-463 in *E. coli*) (Buck *et al.* 2000; Wigneshweraraj *et al.* 2002; Burrows *et al.* 2003; Burrows *et al.* 2004). These determinants include a DNA cross-linking motif (residues 329-346 in *E. coli*) (Cannon *et al.* 1994), and a RpoN box (residues 454-463 in *E. coli*), (Taylor *et al.* 1996). Recently, NMR-based structural studies of the C-terminal region of *Aquifex aeolicus* σ^{54} bound to promoter DNA have revealed a helix-turn-helix (HTH) motif in which the RpoN box forms the recognition helix that binds to the -24 promoter element (Doucleff *et al.* 2005b; Doucleff *et al.* 2007). An additional HTH motif, previously suggested to bind to the -12 promoter element is present outside of the RpoN box (residues 366-386 in *E. coli*) (Merrick and Chambers 1992; Buck *et al.* 2000). As well as containing promoter-binding determinants, region III of σ^{54} also contains determinants for core RNAP binding (residues 120-215 in *E. coli*) (Gallegos and Buck 1999; Hsieh *et al.* 1999; Buck *et al.* 2000).

2.6   The structural basis for activator-dependence

Although high-resolution crystal structures have been determined for σ^{70}-family members (Campbell *et al.* 2002; Murakami *et al.* 2002b; Jain *et al.* 2004; Sorenson *et al.* 2004) the alternate sigma-factor (σ^{54}) has only yielded low-resolution small-angle x-ray scattering and cryo-Electron Microscopy (cryo-EM) structures (Svergun *et al.* 2000; Rappas *et al.* 2005; Bose *et al.* 2008b). Recently, cryo-EM has revealed structural features of the RNAP-σ^{54} holoenzyme that explain the stability of the closed complex and therefore the need for an activator (Bose *et al.* 2008b; Ghosh *et al.* 2010). Reconstructions of RNAP-σ^{54} in the presence and absence of an activator protein have identified three distinct structural regions (named D1, D2 and D3 by Bose et al.) of σ^{54}, each positioned on the β' side and on the upstream face of the core RNA polymerase (Bose *et al.* 2008b). The D1 region likely represents the core RNAP binding domain (residues 120-215 in *E. coli*) of region III and is located at the tip of the β' subunit, well positioned to contact the β' coiled-coil motif which is the binding site of σ-factors in the core enzyme (Young *et al.* 2001). This is consistent with Nuclear Magentic Resonance (NMR)-studies of the core-binding region of *Aquifex aeolicus* σ^{54} in which one surface is negatively charged and predicted to interact with the coiled-coil motif of β' (Hong *et al.* 2009). Bose et al., attribute the density of the D3 region to the DNA-binding domain of region III that includes the RpoN box. Of particular importance is the presence of a strong bridging density (Db), connecting the two "pincers" of the polymerase enzyme that is more pronounced in RNAP-σ^{54} compared to an RNAP-σ^{70} structure. Since the Db region correlates with the -12 position of the promoter DNA, it has been proposed that this connecting density (attributed to region I), obstructs the loading of DNA into the active site channel of the core enzyme. Therefore the presence of the bridging density (Db) could explain why the σ^{54}-holoenzyme forms an energetically favourable closed complex unlike the σ^{70}-holoenzyme which spontaneously isomerises into an open form. Comparison of the RNAP-σ^{54} reconstruction with one also in the presence of an activator and nucleotide transition-state analogue reveals a significant conformational change in region I upon activator binding that is coupled to a slide of the DNA towards the active site of the polymerase (Figure 2.5) (Bose *et al.* 2008b). Indeed, hydroxyl radical footprinting and photo-crosslinking have demonstrated that the ATPase domain of the activator is within 12 Å of the -12 promoter element during open complex formation (Burrows *et al.* 2004). This is consistent with a role for the activator in remodelling the nucleoprotein regulatory centre (Chaney *et al.* 2001; Wigneshweraraj *et al.* 2001; Bordes *et al.* 2003). Based on studies of σ^{54} (residues 69-198) from *Aquifex aeolicus*

it has also been suggested that the hydrophobic interface between the N-terminal and C-terminal sub-domains of the core binding region (D1 in the Bose et al. reconstruction) might be disrupted upon activator binding to region I of σ^{54}, paving the way for conformational changes in the holoenzyme and the isomerisation of the closed complex (Hong *et al.* 2009). Overall, structural and biochemical studies have revealed three main roles for the activator in σ^{54}-dependent transcription (Bose *et al.* 2008b; Ghosh *et al.* 2010). First, the activator must stimulate DNA melting at the -12 promoter element since activator-bypass mutants only function in the presence of pre-melted DNA (Guo and Gralla 1998; Chaney *et al.* 2001). Secondly, the activator must remodel region I of the σ factor that physically blocks the loading of the DNA into the active site (Bose *et al.* 2008b). Thirdly, the activator must cause the repositioning of the DNA-binding domains of σ^{54} downstream since the -12 promoter element at which DNA melting originates is located too far upstream from the active site of the core enzyme for elongation to proceed (Bose *et al.* 2008b). The properties of such activators will now be discussed, in Chapter 3.
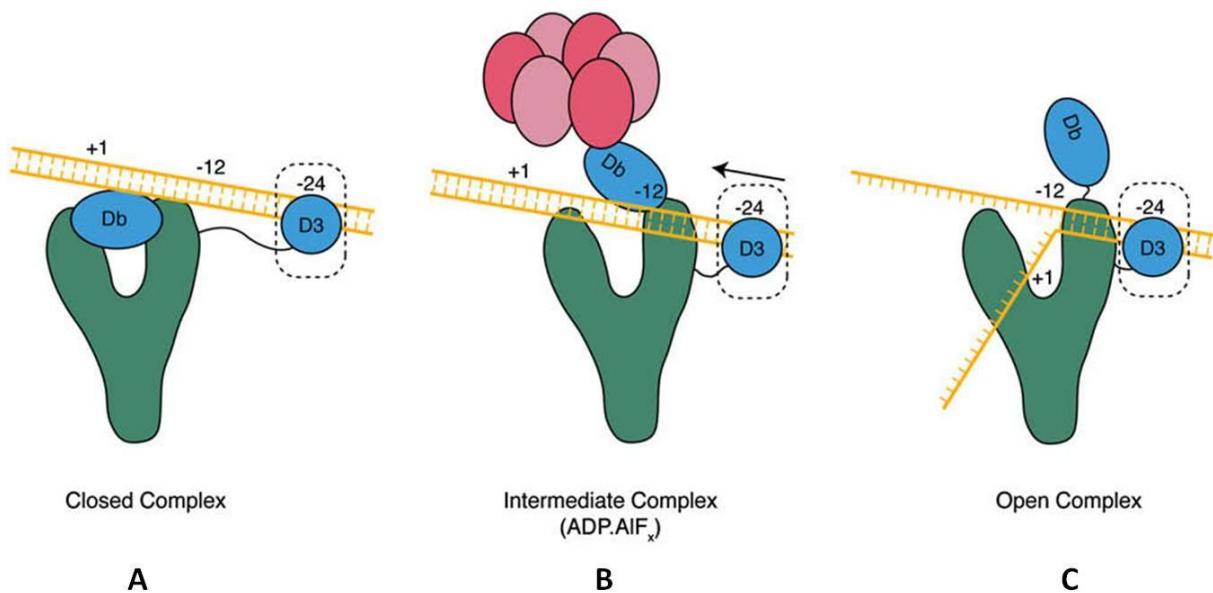
Figure 2.5 - A schematic representation of the proposed relative positions and movements of σ^{54} domains and promoter DNA in the closed (A), intermediate (B), and open (C) complexes (Bose *et al.* 2008b). The DNA-binding region (D3 density) and Region I (Db density) of σ^{54} are shown in blue with the core enzyme in green. The bEBP/activator is shown as a hexamer in red and the promoter DNA is shown in orange. Cryo-EM structures of the RNAP-σ^{54} holoenzyme indicate that the "Db density" is in close proximity to the -12 position of the promoter DNA. Therefore it has been proposed that Region I of σ^{54} prevents the initiation of transcription by obstructing the loading of DNA into the active site channel of the core RNAP. The activator (i) causes the melting of DNA at the -12 position, (ii) interacts with Region I to relocate the Db density and (iii) results in the downstream movement of the DNA-binding (D3) region of σ^{54}, bringing the origin of DNA melting (-12) near to the active site.

Chapter 3 - The AAA+ Enhancer Binding Proteins (EBPs)

AAA+ (ATPases Associated with various cellular Activities) proteins are universal in living organisms, functioning as molecular machines to convert the chemical energy stored in ATP into a mechanical energy that can be used in various cellular processes (Neuwald *et al.* 1999; Ogura and Wilkinson 2001; Lupas and Martin 2002; Hanson and Whiteheart 2005; Tucker and Sallai 2007). In contrast to σ^{70}-dependent transcription, σ^{54}-dependent transcription absolutely requires the presence of an activator of the AAA+ class that couples the energy generated from ATP hydrolysis to the isomerisation of the RNAP-σ^{54} closed complex (Schumacher *et al.* 2004). Such activators typically bind at sites 80-150 bp upstream of the promoter, known as upstream activator sequences (UASs) or enhancer sites (Figure 3.1A). This is similar to the binding of eukaryotic enhancer binding proteins (EBPs) and so activators of σ^{54}-dependent transcription are referred to as bacterial enhancer binding proteins (bEBPs) (reviewed in (Zhang *et al.* 2002; Schumacher *et al.* 2006; Rappas *et al.* 2007; Wigneshweraraj *et al.* 2008). Despite only sharing 20% sequence similarity with other AAA+ proteins, bEBPs are sufficiently similar in structure and function to be classified as members (Neuwald *et al.* 1999). Since the activator binds so far upstream of the transcriptional start site, DNA must bend between the enhancers and the promoter site in order for it to interact with the RNAP-σ^{54} holoenzyme (Figure 3.1B) (Zhang *et al.* 2002; Wigneshweraraj *et al.* 2005; Schumacher *et al.* 2006). Such DNA looping has been visualised by electron microscopy (EM) (Su *et al.* 1990) and scanning force microscopy (SFM) (Rippe *et al.* 1997). DNA looping is often aided by the integration host factor (IHF), a small heterodimeric protein which binds between promoter and enhancer sites to bend the DNA up to 180° (Yang and Nash 1989; Hoover *et al.* 1990; Arfin *et al.* 2000). Since correct interfacing between the bEBP and holoenzyme is crucial for the activation process, the phasing of the IHF-binding site relative to the promoter is important (Claverie-Martin and Magasanik 1992; Huo *et al.* 2006). In addition to ensuring efficient activator-holoenzyme contact, IHF-induced changes in DNA topology contribute to the specificity and efficiency of activation (Perez-Martin and De Lorenzo 1995; Dworkin *et al.* 1997). For example, at the *pspA-E* promoter, IHF has been shown to mediate architectural changes that aid the binding of the bEBP PspF and increase promoter output (Jovanovic and Model 1997). Once DNA bending has been induced, the bEBP utilises nucleotide triphosphate (NTP) hydrolysis to drive conformational rearrangements in the holoenzyme that promote the transition of the closed to an open complex (Figure 3.1C) (Schumacher *et al.* 2004; Rappas *et al.* 2006). Although ATP binding to a bEBP

65
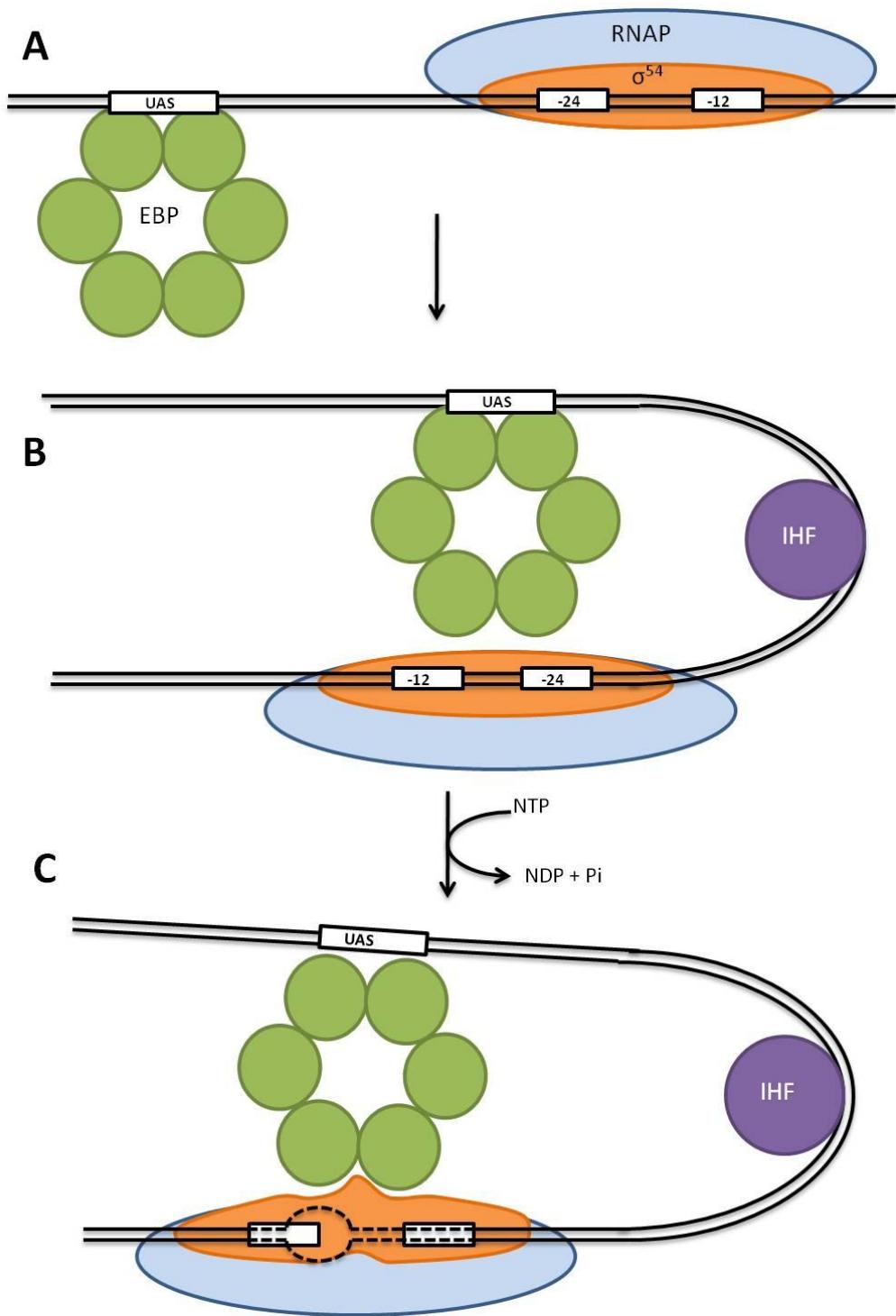
Figure 3.1 - The activation of bacterial transcription initiation (σ^{54}-family dependent). The σ factor directs the RNA polymerase holoenzyme to bind at the -12 and -24 promoter elements. (**A**) Interaction of the bacterial Enhancer Binding Protein (bEBP) with the σ^{54}-RNAP holoenzyme is dependent on the binding of the activator (shown as an oligomer in green) to Upstream Activator Sequences (UAS), 80-150 bp upstream of the transcriptional start site. (**B**) DNA looping occurs; often facilitated by other proteins such as Integration Host Factor (IHF), enabling bEBP-σ^{54} interactions. In σ^{54}-dependent transcription, the closed complex does not spontaneously undergo isomerisation (melting of the double stranded DNA). (**C**) Nucleotide hydrolysis by the activator promotes re-modelling of the closed complex through a series of protein-protein and protein-DNA interactions that promote the formation of the open complex (Buck *et al.* 2000)

dimer has been demonstrated (Rombel *et al.* 1999), oligomerisation is required to stimulate full ATPase activity (Wikstrom *et al.* 2001; Rappas *et al.* 2005). Well characterised examples of bEBPs include the nitrogen regulatory protein C (NtrC), C4-dicarboxylic acid transport protein D (DctD), the nitrogen fixation regulatory protein (NifA), the phage shock protein F (PspF), the xylene catabolism regulatory protein (XylR) and 3,4-demethylphenol catabolism regulatory protein (DmpR) (Table 3.1) (Studholme and Dixon 2003).

3.1   Domain architecture

In common with their eukaryotic counterparts, bEBPs are modular proteins and in general consist of three domains (Studholme and Dixon 2003; Schumacher *et al.* 2006). The N-terminal regulatory (R) domain has a role in signal perception and modulates the activity of the bEBP. The central AAA+ domain (C) is responsible for ATP-hydrolysis in the bEBP; it is indispensible and often sufficient to activate σ^{54}-dependent transcription (Berger *et al.* 1995; Jovanovic *et al.* 1999; Wikstrom *et al.* 2001; Xu *et al.* 2004a). Lastly, the C-terminal DNA binding domain (D) contains a helix-turn-helix (HTH) motif (Pelton et al., 1999; Sallai et al., 2005) that enables specific UAS/enhancer site recognition (Xu and Hoover 2001). However, not all activators of σ^{54}-dependent transcription consist of each of the three domains (Table 3.1 and Figure 3.2). Whilst the presence of the central domain is conserved, some bEBPs lack either the regulatory domain (i.e. consist of C+D) or the DNA-binding domain (i.e. consist of R+C) (Hutcheson *et al.* 2001; Elderkin *et al.* 2002; Brahmachary *et al.* 2004). In addition, the regulatory domains do not share common homology and contain a variety of sensory motifs depending on the signal that is detected (Table 3.1) (Studholme and Dixon 2003). Consequently, the bEBP family has been divided into five groups (I-V) based on the organisation of the three domains that are present (Wigneshweraraj *et al.* 2005). The structure and function of these three domains will now be discussed in more detail.

bEBP	Full-name	Organism	Function (target)	Regulatory domain (s)	Stimuli/ mode of regulation	Regulation of AAA+ activity	Reference
PspF	phage shock protein F	*Escherichia coli*	Phage-shock response (Eσ^{54})	absent	interaction *in trans* with PspA	Negative (by PspA)	(Elderkin *et al.* 2002; Elderkin *et al.* 2005; Joly *et al.* 2009)
NifA	nitrogen fixation regulatory protein	*Azotobacter vinelandii*	Regulation of Nitrogenase (Eσ^{54})	GAF domain	1. Small-ligand binding (2-oxoglutarate) 2. *in trans* in response to NifL	1. Positive (by 2-OG) 2. Negative (by NifL)	(Martinez-Argudo *et al.* 2004b; Martinez-Argudo *et al.* 2004c)
NtrC	nitrogen regulatory protein C	*Salmonella enterica*	Regulation of Glutamine Synthetase (and other functions of nitrogen metabolism) (Eσ^{54})	Response Regulator (RR)	Phosphorylation by sensor-kinase NtrB	Positive	(Weiss *et al.* 1991; Austin and Dixon 1992; Doucleff *et al.* 2005a; De Carlo *et al.* 2006)
NtrC1	nitrogen regulatory protein C 1	*Aquifex aeolicus*	homolog of NtrC (Eσ^{54})	Response Regulator (RR)	Phosphorylation	Negative (R domain releases inhibition of C domain)	(Lee *et al.* 2003; Doucleff *et al.* 2005a)
NtrC4	nitrogen regulatory protein C 4	Aquifex aeolicus	homolog of NtrC (Eσ^{54})	Response Regulator (RR)	Phosphorylation	Negative (R domain releases inhibition of C domain)	(Batchelor *et al.* 2008; Batchelor *et al.* 2009)
XylR	xylene catabolism regulatory protein	*Pseudomonas putida*	Regulation of oxidative transformation of toluene, *m*-xylene and *p*-xylene (Eσ^{54})	V4R domain	Small-ligand binding (aromatic effectors)	Negative (R domain releases inhibition of C domain)	(Fernandez *et al.* 1995; Perez-Martin and Lorenzo 1995; Perez-Martin and de Lorenzo 1996)
DctD	C4-dicarboxylic acid transport protein D	*Sinorhizobium meliloti*	Regulation of C4-dicarboxylic acid transport (Eσ^{54})	Response Regulator (RR)	Phosphorylation	Negative (R domain releases inhibition of C domain)	(Xu *et al.* 2004a; Xu *et al.* 2004b; Doucleff *et al.* 2005a)
DmpR	3,4-demethylphenol catabolism regulatory protein	*Pseudomonas CF600*	Regulation of (methyl)phenol catabolism (Eσ^{54})	V4R domain	Small-ligand binding (aromatic effectors)	Negative (R domain releases inhibition of C domain)	(Shingler and Moore 1994; Shingler and Pavel 1995; Wikstrom *et al.* 2001)
NorR	nitric oxide reductase regulator	*Escherichia coli*	Regulation of NO-detoxification system (Eσ^{54})	GAF domain	Small-ligand binding (Nitric Oxide)	Negative (R domain releases inhibition of C domain)	This work; (Hutchings *et al.* 2002b; D'Autreaux *et al.* 2005; Tucker *et al.* 2007)
ZraR	zinc resistance-associated protein regulator	*Salmonella typhimurium*	Regulation of heavy-metal tolerance system (Eσ^{54})	Response Regulator (RR)	Phosphorylation by sensor-kinase ZraS	Unknown	(Leonhartsberger *et al.* 2001; Sallai and Tucker 2005)
FhlA	formate-hydrogen lyase activator protein	*Escherichia coli*	Regulates formate-hydrogen lyase (which oxidises formic acid) (Eσ^{54})	2 x GAF domains	Small-ligand binding (Formate)	Negative	(Schlensog *et al.* 1994; Hopper and Bock 1995)
TyrR	tyrosine regulator	*Escherichia coli*	Regulation of aromatic amino acid biosynthesis and transport (Eσ^{70})	PAS and ACT domains	Small-ligand binding (Phenylalanine, tyrosine or tryptophan) in presence of ATP	ATPase-inactive (activates or represses Eσ^{70} promoters)	(Pittard and Davidson 1991; Pittard *et al.* 2005)
HrpR/S	hypersensitive reaction pathogenicity island regulator	*Pseudomonas syringae*	Regulation of *hrp-hrc* pathogenicity island (Eσ^{54})	absent	interaction with apparent sensor-kinase (HrpS)	Negative (HrpS regulated by HrpV; HrpR regulated by HrpG)	(Preston *et al.* 1998; Schuster and Grimm 2000; Hutcheson *et al.* 2001; Jovanovic *et al.* 2011)
FlgR	flagella gene regulator	*Helicobacter pylori*	Regulates class III flagella genes. Only σ^{54} activator in organism – activates from solution (no D domain) (Eσ^{54})	Response Regulator (RR)	Phosphorylation by sensor-kinase FlgS	Negative (release of repression)	(Spohn and Scarlato 1999; Brahmachary *et al.* 2004)

Table 3.1 – The function and regulation of selected examples of bEBPs
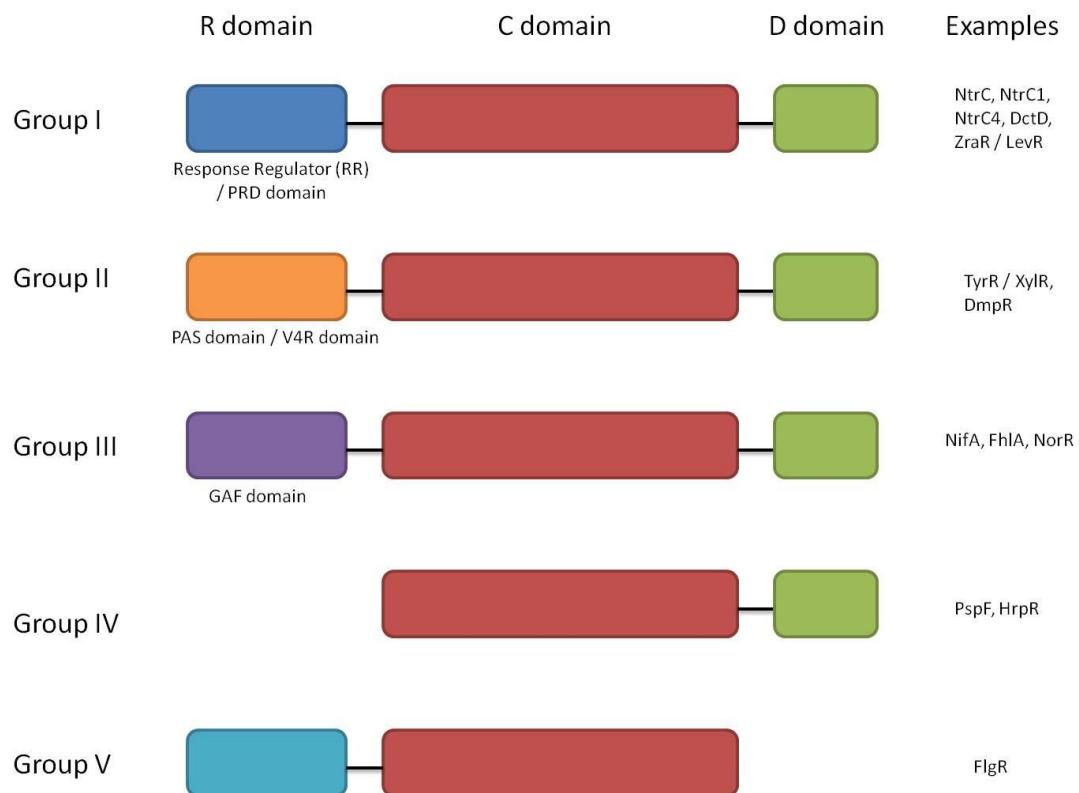
Figure 3.2– Domain architecture of the five groups (I-V) of bEBPs (Wigneshweraraj *et al.* 2005). The central AAA+ domain (C, Red) is highly conserved and absolutely essential for σ^{54}-dependent transcription. The C-terminal DNA-binding domain (D, Green) consists of a helix-turn-helix (HTH) motif that directs the bEBP to specific UAS/enhancer binding sites and in some bEBPs is absent (Group V). The N-terminal regulatory domain (R) is not well conserved between members of the bEBP family. Different sensory domains are present depending on the environmental signal to be detected but in some bEBPs it is absent (Group IV). Group I bEBPs contain a response regulator (RR) or PTS regulation (PRD) domains (Blue). Group II bEBPs contain Per, ARNT, and Sim (PAS) domains or a V4R (vinyl 4 reductase) domain (Orange). Group III bEBPs contain a cGMP-specific and stimulated phosphodiesterases, Anabaena adenylate cyclases and *E. coli* FhlA (GAF) domain (Purple).

3.2　Role of the central AAA+ (C) domain

The central (AAA+) domain is responsible for nucleotide binding and hydrolysis, oligomerisation and σ^{54} contact. It is indispensable and often sufficient for transcriptional activation both *in vitro* and *in vivo* (Berger et al. 1995; Jovanovic et al. 1999; Wikstrom et al. 2001; Xu et al. 2004a). The AAA+ domain is the most conserved of the three domains and has been divided into seven conserved regions, C1-C7 (Figure 3.3) (Morett and Segovia 1993; Osuna *et al.* 1997).

3.2.1　Conserved regions of the C domain

Crystal structures of the PspF, NtrC and ZraR bEBP AAA+ domains (Lee *et al.* 2003; Rappas *et al.* 2005; Sallai and Tucker 2005) reveal an α/β subdomain, followed by a smaller α-helical subdomain, characteristic of all AAA+ proteins (Zhang *et al.* 2002; Rappas *et al.* 2007). The nucleotide binding site is located in the cleft between these subdomains and between two adjacent protomers (Figure 3.7) (Rappas *et al.* 2007). AAA+ domains are characterised by the Walker A and Walker B motifs that have roles in nucleotide binding and hydrolysis (Walker *et al.* 1982; Hanson and Whiteheart 2005). Walker A is in the C1 region and forms a P-loop with the consensus GxxxxGK[T/S] that interacts with the phosphates of ATP (Saraste *et al.* 1990). The requirement for the Walker A motif has been shown in a number of bEBPs including PspF (Schumacher *et al.* 2004) and NtrC (Rombel *et al.* 1999). In *Pseudomonas putida* XylR, the G268N substitution abolishes ATP-binding and hydrolysis (Perez-Martin and de Lorenzo 1996). Likewise, the Walker B motif of the C4 region has a consensus hhhhDE (h = any hydrophobic amino acid) and has been shown to be required for nucleotide hydrolysis. Mutagenesis of the key aspartate residue suggests a role in coordination of Mg^{2+}, required for ATP hydrolysis (Rombel *et al.* 1999; Schumacher *et al.* 2004) and this residue is also thought to activate water for nucleophilic attack of the γ-phosphate. Another common feature of AAA+ proteins is the presence of the Sensor I and Sensor II motifs, present in the conserved regions C6 and C7 respectively (Schumacher *et al.* 2006). Sensor I residues are located within a loop with the side chain of threonine in between the Walker A (WA) and Walker B (WB) motifs. This threonine residue has been implicated in coupling nucleotide hydrolysis to conformational change (Schumacher *et al.* 2007). Sensor II residues within the C7 region are located in the third helix of the α-helical subdomain and have been implicated in nucleotide base binding. I226 in PspF has been suggested to be involved in this function whilst the adjacent arginine residue points towards the γ-phosphate and may

be involved in hydrolysis (Schumacher *et al.* 2006). Indeed R227 in PspF was suggested to have a role in Mg^{2+} coordination (Zhang *et al.* 2002; Rappas *et al.* 2006). Members of the AAA+ superfamily also contain one or two arginine residues (R-fingers) that have been implicated in inter-subunit catalysis and nucleotide sensing (Ogura and Wilkinson 2001; Lupas and Martin 2002; Hanson and Whiteheart 2005). In accordance with this, bEBPs contain two potential R-fingers that together with the catalytically important Sensor II residues are located at the protomer interface. In PspF, the predicted R-finger R168 has been shown to be required for ATP-hydrolysis but not for ATP-binding (Schumacher *et al.* 2004) whilst the same is observed in NtrC for the second R-finger R294 (R162 in PspF) (Rombel *et al.* 1998; Rombel *et al.* 1999). These phenotypes reflect the observation that the ATPase active site is formed at the interface between adjacent protomers to "share" the catalytic arginine(s). The structure of the AAA+ protein p97 shows the catalytic arginine protruding from one protomer into the catalytic site of the adjacent protomer, contacting the γ-phosphate of ATP (Zhang *et al.* 2000). Indeed the recent publication of the crystal-structure of ATP-bound NtrC1 indicates that the first R-finger (R299 in NtrC1) engages with the γ-phosphate (Chen *et al.* 2010). Therefore oligomerisation is essential for the bEBP's ability to hydrolyse ATP.

3.2.2 Structural elements of AAA+ domains specific to bEBPs

The bEBP-subfamily of AAA+ domains contains specific structural features that enable nucleotide-dependent interactions with σ^{54} (Figure 3.3) (Zhang *et al.* 2002; Schumacher *et al.* 2006). Most conserved amongst these is the GAFTGA motif (conserved region C3), which forms a loop on the surface of the AAA+ domain that contacts σ^{54} during the ATP hydrolysis cycle (Bordes *et al.* 2003). Crystal structures of the NtrC1 and ZraR AAA+ domains (Lee *et al.* 2003; Sallai and Tucker 2005) show that the GAFTGA motif is located at the α/β subdomain surface at the tip of loop 1 (L1) which is inserted into helix 3 (H3). Although the surface-exposed loops appear to point towards the central pore of the oligomeric rings, the GAFTGA motifs are not in an extended conformation. Stable interaction between the GAFTGA loop and σ^{54} has only been observed in studies that use the ATP-transition state analogue, ADP.AlF$_x$. Cryo-electron microscopy (Cryo-EM) studies of the PspF central domain in complex with σ^{54} and ADP.AlF$_x$ reveals a hexameric bEBP ring in contact with monomeric σ^{54} (Rappas *et al.* 2005). Significantly, the reconstruction reveals connecting electron densities between the bEBP and σ^{54}. Fitting of the PspF AAA+ crystal structure into the 3D-reconstruction confirms that the GAFTGA-
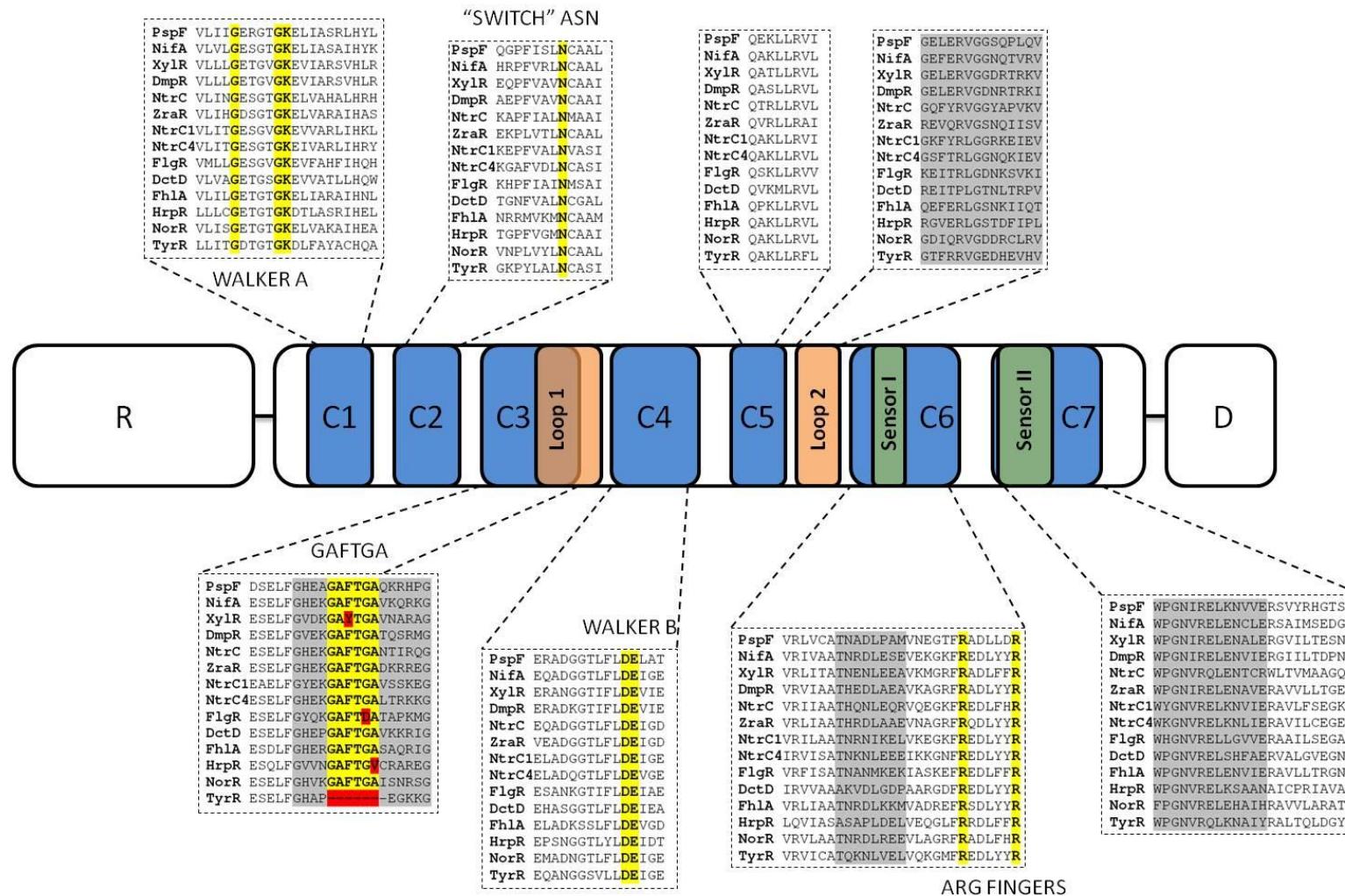
Figure 3.3 - Domain map and sequence alignment of the conserved regions of bEBP AAA+ domains (C1-C7; (Morett and Segovia 1993). The conserved regions are based on a structure-based sequence alignment (Schumacher *et al.* 2006). Key residues (Walker A, "Switch" Asn, GAFTGA, Walker B and R-fingers) are highlighted in yellow and non-consensus sequences in the alignments are highlighted in red. The locations of Loop 1, Loop 2, Sensor I and Sensor II motifs are indicated with their sequences highlighted in grey. Alignments were conducted using ClustalW (www.ebi.ac.uk/clustalw/*)* using the sequences from UniProtKB/Swiss-Prot (http://www.expasy.ch/): PspF (*E. coli*), NifA (*A. vinelandii*), XylR (*P. putida*), DmpR (*Pseudomonas* sp.), NtrC (*E. coli*), ZraR (*E. coli*), NtrC1 (*A. aeolicus*), NtrC4 (*A. aeolicus*), FlgR (*H. pylori*), DctD (*S. meliloti*), FhlA (*E. coli*), HrpR (*P. syringae*), NorR (*E. coli*), TyrR (*E. coli*). The R (regulatory) and D (DNA binding) domains are also illustrated, though not to scale.

containing L1, assisted by loop 2 (L2) mediate this interaction. Therefore, it is likely these conserved motifs enable nucleotide-dependent σ^{54}-interaction to initiate the transition of the closed complex. In accordance with this, the GAFTGA motif has been shown to be critical for open complex formation (Zhang *et al.* 2002). The effect of substituting residues of the GAFTGA motif has been studied in the bEBPs NtrC (North *et al.* 1996; Li *et al.* 1999; Yan and Kustu 1999), DctD (Wang *et al.* 1997), NifA (Gonzalez *et al.* 1998), DmpR (Wikstrom *et al.* 2001) and PspF (Chaney *et al.* 2001; Bordes *et al.* 2003; Bordes *et al.* 2004; Dago *et al.* 2007; Zhang *et al.* 2009). There seems to be an absolute requirement for an intact GAFTGA motif; mutation of any one of the six amino acids has severe effects on the bEBP's ability to hydrolyse ATP, contact σ^{54} or activate transcription (Table 3.2). Other bEBP-like proteins that lack this motif such as the *E. coli* TyrR protein and NtrC from *Rhodobacter capsulatus* are unable to activate σ^{54}-dependent transcription (Bowman and Kranz 1998; Poggio *et al.* 2002; Pittard *et al.* 2005). In the case of TyrR, the α-helical residues that precede the GAFTGA motif in other bEBPs (ESELFGHEK) are thought to couple ATP hydrolysis to effects on transcription at σ^{70}-dependent promoters (Kwok *et al.* 1995).

3.2.3 Coupling ATP hydrolysis to the activation of transcription

ATP hydrolysis is coupled to open complex formation via conformational changes in the AAA+ domain that ultimately lead to the relocation of the GAFTGA-containing L1, and the L2 loops. A number of crystal structures of bEBP AAA+ domains have been reported providing detailed information about the nucleotide binding pocket and other key determinants (Lee *et al.* 2003; Rappas *et al.* 2005; Sallai and Tucker 2005; Chen *et al.* 2010). In order to examine the structure of the AAA+ domain at the discrete stages of the nucleotide cycle, a variety of structural and biochemical techniques have been employed (Bose *et al.* 2008a). This includes the soaking of transiently stable crystals of the PspF C domain (1-275) in the presence of different nucleotides to obtain various nucleotide-bound structures (Rappas *et al.* 2006). High resolution crystal structures of NtrC1 bound to both ATP and ADP have also been published (Lee *et al.* 2003; Chen *et al.* 2010). Lower resolution techniques that include Cryo- EM and SAXS (small-angle X-ray scattering)/WAXS (wide angle X-ray scattering) used in conjunction with nucleotide analogues have provided information about the larger, macromolecular conformational changes that occur in the bEBP as ATP is hydrolysed (Rappas *et al.* 2005; De Carlo *et al.* 2006; Chen *et al.* 2007; Chen *et al.* 2008). However, caution should be taken in the

analysis of all such structures since the coordination of ATP hydrolysis between bEBP-protomers is thought to involve different nucleotide-bound states (Joly *et al.* 2006).

(i) Studies of the ground-state of nucleotide hydrolysis

The ATP-bound "ground" state of the nucleotide cycle has been examined by soaking crystals of $PspF_{1-275}$ with ATP, either in the absence of Mg^{2+} to prevent hydrolysis or by using a hydrolysis-defective variant of the bEBP (Figure 3.4A) (Buck *et al.* 2006; Rappas *et al.* 2006; Bose *et al.* 2008a). Structures reveal the residues responsible for binding ATP within the nucleotide binding pocket. The Walker B glutamate (E108 in PspF) senses the γ-phosphate of ATP and forms a strong interaction with a nearby highly conserved asparagine (N64 in PspF). The adjacent aspartate (D107 in PspF) coordinates the position of a water molecule for nucelophilic attack of this γ-phosphate (Rappas *et al.* 2006). In NtrC1, the conserved R-finger from the adjacent protomer (R299 in NtrC1, R168 in PspF) appears to engage with the γ-phosphate, stabilising the binding of ATP (Chen *et al.* 2010). Significantly, the ATP-bound structures of PspF and NtrC1 indicate that the L1 and L2 loops are in a raised conformation, consistent with low-resolution SAXS-derived structures using the ground-state analogue $ADP-BeF_x$ in NtrC1 (Rappas *et al.* 2006; Chen *et al.* 2007; Chen *et al.* 2010). Biochemical experiments confirm that both PspF and NtrC1 can establish contact with σ^{54} in the presence of this ground-state analogue (Chen *et al.* 2007; Bose *et al.* 2008a). Taken together this data indicates that the binding and sensing of nucleotide causes significant conformational changes in the bEBP AAA+ domain. The combined data to date suggest that the GAFTGA-containing L1 loop, assisted by L2 is released to make an initial, unstable interaction with $E\sigma^{54}$.

(ii) Studies of the transition state of nucleotide hydrolysis

Studies using the transition state analogue $ADP.AlF_x$ have established a clear role for the GAFTGA-motif of loop 1 in contacting Region I of σ^{54} prior to the remodelling of the closed complex (Chaney *et al.* 2001; Bordes *et al.* 2003; Cannon *et al.* 2003). Most significantly, a stable interaction between the GAFTGA-containing L1 loop and σ^{54} is formed in the presence of the transition state analogue, as revealed by the fitting of the $PspF_{1-275}$ crystal structure into the Cryo-EM reconstruction of the activator in complex with σ^{54} (Rappas *et al.* 2005; Rappas *et al.* 2006). In agreement with this, low-resolution structures of NtrC and NtrC1 bound to $ADP.AlF_x$, reveal electron density above the plane of the oligomeric bEBP corresponding to a raised position of the L1 and L2 loops (De

Carlo *et al.* 2006; Chen *et al.* 2007). Significantly, the ADP.AlF$_x$ complex of NtrC1 was shown to be more stable than the ADP-BeF$_x$ complex, indicating that ATP-hydrolysis strengthens the unstable interaction between the bEBP and σ^{54} that forms in the ground state (Chen *et al.* 2007). Similar results have been obtained for the bEBP PspF (Burrows *et al.* 2009). In addition, this form of the activator is able to initiate the early stages of promoter "melting" (Burrows *et al.* 2004), identifying it as a true intermediate in the transition of the closed complex. However, the "trapped" forms of PspF are largely defective for nucleotide hydrolysis and incapable of activating transcription (Chaney *et al.* 2001; Burrows *et al.* 2009) suggesting that continuation of the cycle of nucleotide-driven conformational changes is vital for formation of the open complex.

(iii) Studies of the post-nucleotide hydrolysis state

The "post-hydrolysis" state has been extensively studied through the publication of crystal-structures of various bEBPs in their ADP-bound forms (Lee *et al.* 2003; Rappas *et al.* 2005; Sallai and Tucker 2005; Rappas *et al.* 2006). Comparison of the post-hydrolysis state with ground- and transition- state structures has shed light on the conformational rearrangements that occur upon ATP hydrolysis. Structures of PspF (Figure 3.4B) indicate that the release of the γ-phosphate to form ADP causes a 90° rotation of the glutamate side chain of the Walker B motif (E108 in PspF). As a result, the interaction between the glutamate and the conserved asparagine (N64 in PspF) is broken resulting in the Walker B glutamate interacting instead with the Sensor I threonine residue (T148 in PspF), via a water molecule (Figure 3.4C). Communication of the altered position of Walker B by the asparagine residue leads to conformational changes in helix 3 (H3) and helix 4 (H4) that are translated to loop 1 (L1) and loop 2 (L2) via strategically placed residues in the central domain. The functional significance of these key residues has now been assessed and it has emerged that both intra- and inter-subunit interactions have a role in modulating the conformation of the σ^{54}-interaction surface (Figure 3.5) (Joly and Buck 2010). The structure of ADP-bound PspF$_{1-275}$ suggests that the Walker B-asparagine "switch" results in the disruption of an interaction between the R131 residue of the L2 loop and the E97 residue of H3 (Figure 3.5A) (Rappas *et al.* 2006). Single substitutions at these positions abolish ATPase activity and σ^{54}-interaction whilst residue swap experiments (R131E/E97R) partially restore activity, demonstrating the importance of this polar interaction (Joly and Buck 2010). Following this, the E97 residue is proposed to interact with R91 whilst the R131 residue is thought to contact the L1 loop residue E81. Variants

of R91 in PspF were unaltered for ATPase activity and although they showed only a slight decrease in the ability to contact σ[54], they were significantly less able to isomerise σ[54], indicating that this residue is important for substrate remodelling after the initial interaction (Joly and Buck 2010). This residue does not show high conservation in the bEBP subfamily of AAA+ proteins but in many cases, the adjacent residue may serve a similar function. Together these new interactions result in the compaction of the loops down towards the surface of the AAA+ domain enabling σ[54] relocation, crucial to the conversion from the closed to the open complex (Rappas *et al.* 2006; Chen *et al.* 2007; Bose *et al.* 2008a).

Interestingly, it has now been shown that inter-subunit (*in trans*) interactions confer cooperativity in nucleotide-dependent substrate remodelling of PspF (Figure 3.5B) (Joly and Buck 2010). In the ATP-bound state, the E130 residue at the base of L2 is proposed to interact *in trans* with the R98 of H3 from the adjacent subunit. Upon release of the γ-phosphate, the E130 residue is expected to interact instead with R95 of H3 also from the adjacent protomer. Substitutions made at these positions cause the uncoupling of ATPase activity and substrate remodelling since these variants are able to oligomerise and hydrolyse ATP but are not competent to contact σ[54]. Therefore the "switching" of this *in trans* interaction between protomers is thought to contribute to the coordination of the L1 and L2 loops during nucleotide hydrolysis (Joly and Buck 2010). These residues do not show strict conservation in the bEBP subfamily but similar interactions may play a significant role in the coupling of ATP hydrolysis to substrate remodelling. For example in NtrC1, the equivalent residues (F226, L229 and Y261) are thought to facilitate *in trans* hydrophobic contracts (Chen *et al.* 2010).

(iv) **The Walker B-asparagine "switch"**

Analysis of active site structures reveal that the glutamate "switch" residue of the Walker B motif (E108 in PspF) is a common feature of hydrolysis in the majority of AAA+ proteins (Zhang and Wigley 2008). Whereas substitutions of the adjacent Walker B aspartate have severe effects on various aspects of PspF activity, in line with a key role for this residue in ATP hydrolysis, substitution of the adjacent glutamate (E108 in PspF) only has moderate effects upon activity. Such variants have been used to effectively study intermediate states en route to open complex formation and confirm the pivotal role of the
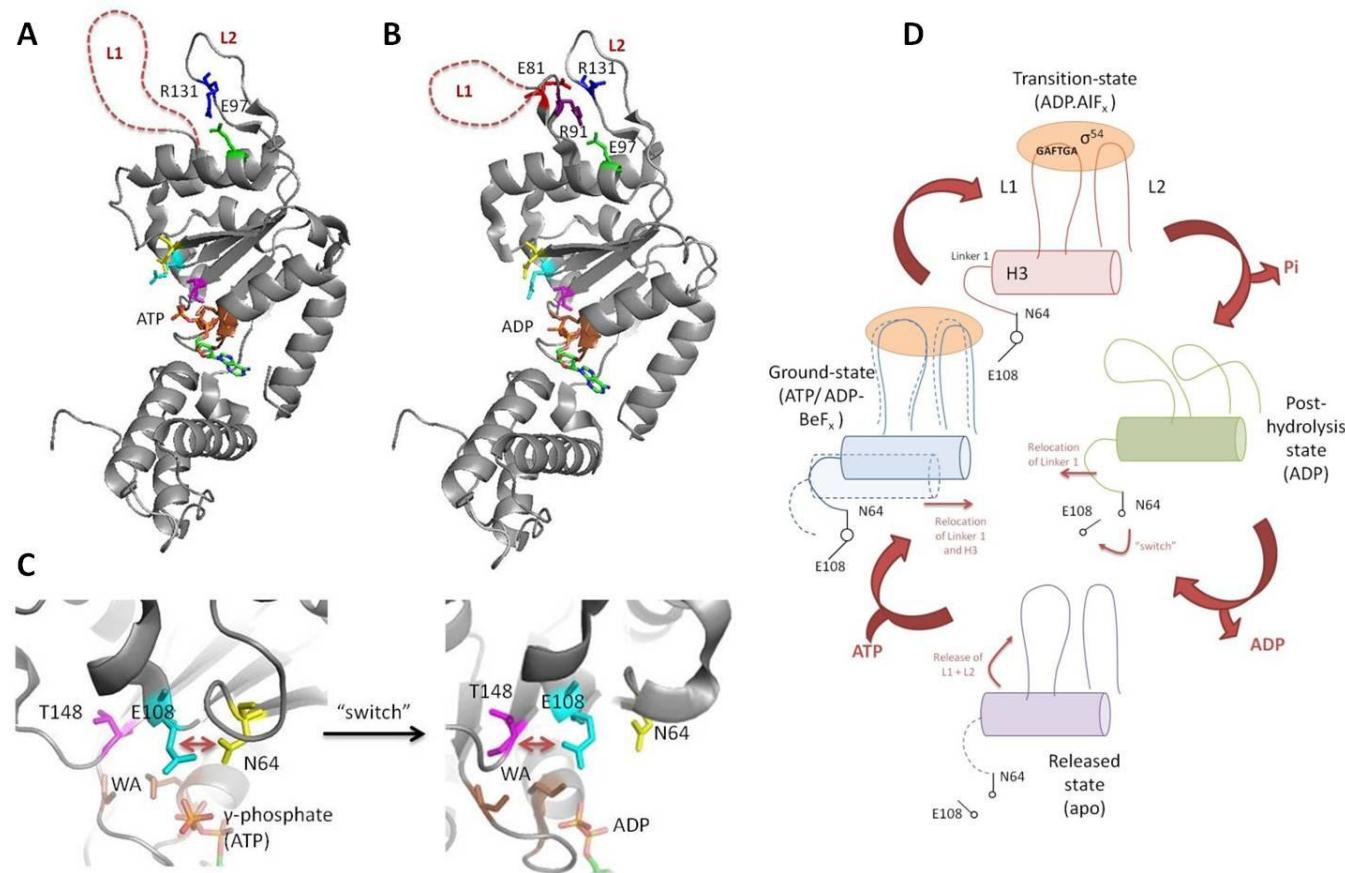
Figure 3.4 – Structure of monomeric PspF$_{1-275}$ bound to ATP (**A**) (PDB code 2C96/2C9C) **and ADP** (**B**) (PDB code 2C98/2C9C). Important motifs are highlighted: Walker A (brown), Walker B/E108 (cyan), Sensor I/T148 (magenta), "Switch" N64 (yellow), E97 (bright green), E81 (red), R131 (blue), R91 (purple). Loop 1 (L1) and loop 2 (L2) are labelled with the non-resolved fold of L1 indicated using a red-dotted line. (**C**) **The switching mechanism of the Walker B E108 residue**. In the ATP-bound state, E108 interacts with N64 (indicated by double-headed arrow). In the ADP-bound state there is a 90° rotation in the N64 side chain so that it interacts with the Sensor I residue T148 (indicated by double-headed arrow). Upon ATP-hydrolysis, the absence of the γ-phosphate induces the "switch", resulting in the E97 residue breaking the contact with R131 of L2. R131 is then thought to contact E81 whilst E97 contacts R91 in PspF. This switch in the interactions causes the compaction of the loops downwards allowing for σ^{54}-relocation. (**D**) **Summary of the nucleotide-driven conformational changes that occur during ATP-hydrolysis, as proposed in PspF** (Rappas *et al.* 2006). The ground- (blue), transition- (red), post-hydrolysis- (green) and released (purple) states are indicated. For simplicity, only the "switch" interactions are shown with the associated re-locations of linker 1, helix 3 (H3) and the L1/L2 loops.
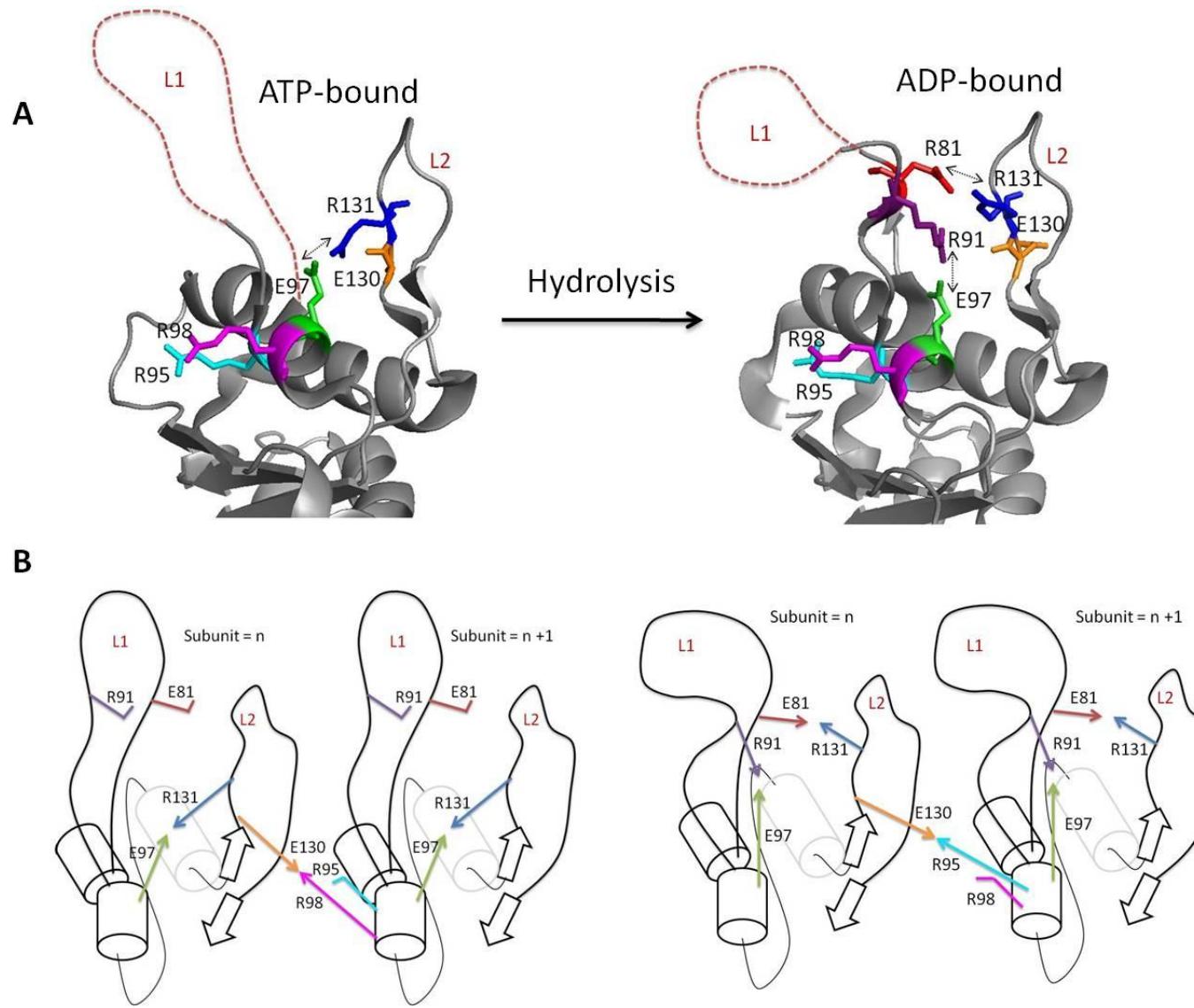
Figure 3.5 – *in cis* and *in trans* interactions predicted to form during nucleotide hydrolysis in PspF. Interactions *in cis* centre around the E97 (green) residue which interacts with either R131 (blue) (ATP-state) or R91 (purple) (ADP-state) depending on the position of the Walker B-asparagine "switch". Upon γ-phosphate release, E97 breaks its interaction with R131, allowing R131 to instead contact R81 (red). These new interactions result in the compaction of the L1 and L2 loops down towards the surface of the AAA+ domain enabling σ[54] relocation. Interactions *in cis* centre around the E130 residue (orange) (subunit n) which contacts the R98 residue (magenta) (subunit n+1) in the ATP-bound state but interacts with the R95 residue (cyan) (subunit n+1) in the ADP-bound state. (**A**) Region of the crystal structure of PspF$_{1-275}$ in ATP-bound form (PDB 2C96) and ADP-bound form (PDB 2C98) showing the location of the key residues involved in inter- and intra- subunit interactions. Interactions occurring within the protomer are indicated by double-headed arrows. (**B**) Schematic showing the *in cis* and *in trans* interactions that occur before and after γ-phosphate release. Adapted from (Joly and Buck 2010). Interactions are indicated by arrows and residue colours correspond to those in (**A**).

Walker B glutamate in transmitting nucleotide-dependent conformational changes (Joly *et al.* 2007). Sequence alignments of bEBPs indicate that the glutamate switch-interacting asparagine (N64 in PspF) is strictly conserved, in agreement with the hypothesis that this residue plays a key role in these specialised activators (Joly *et al.* 2009). In PspF, variants of the N64 residue have altered oligomeric states and ATPase activities, suggesting that the conserved asparagine plays a significant role in the organisation of the active site. In line with this, the position of Mg^{2+}-bound ATP, and the water molecule involved in nucleophilic attack of the γ-phosphate correspond with a catalytic role of the asparagine in ATP hydrolysis (Rappas *et al.* 2006). Significantly, in the absence of the glutamate or asparagine side chains (i.e. in the variants E108A or N64A), PspF was shown to form a stable complex with σ^{54}, indicating that the interaction between the activator and σ^{54} is not strictly dependent on the Walker B glutamate or asparagine residues (Joly *et al.* 2007; Joly *et al.* 2009). However, the N64A variant was significantly less able to form open complexes than the wild-type activator. This defect was suppressed when pre-melted DNA was used, suggesting that the conserved asparagine is also required for the melting of the promoter DNA and the associated loading of the template into the RNAP active site (Joly *et al.* 2009). Despite the Walker B-interacting asparagine not being present in all members of the AAA+ family, the distance between the glutamate and this residue seems to be conserved. Consequently, it has been suggested that the positioning of these two residues is important for their communication with each other and for forming a fully functional catalytic site. In the case of the specialised bEBP family, the communication helps to control the positioning of the GAFTGA-containing L1 loop. In AAA+ proteins that do not contain the GAFTGA insertion, it is likely that the interaction between these residues controls similar nucleotide-dependent conformational changes that regulate functionality in the oligomeric ring (Joly *et al.* 2009).

(v) **R-finger directed "switch"**

More recently, the publication of a high resolution crystal structure of the ATP-bound NtrC1 central domain (NtrC^C) and comparison with the ADP-bound form, has identified an alternative mechanism for coupling hydrolysis to substrate remodelling (Lee *et al.* 2003; Buck and Hoover 2010; Chen *et al.* 2010). In order to examine the configuration of an ATP-bound activator, the Walker B glutamate was replaced with alanine, allowing NtrC1 to bind but not to turnover the nucleotide. This led to the first structure of a bEBP in which the highly conserved arginine (R)-finger (R299 in NtrC1) is seen to contact the γ-

phosphate (Figure 3.6B). This is in contrast to the ATP-bound structure of monomeric PspF$_{1-275}$, where the alternative R-finger R162 (R293 in NtrC1) rather than R168 (R299 in NtrC1) is predicted to be in close proximity to the γ-phosphate of the adjacent protomer (Figure 3.8B; Rappas et al., 2006). Comparison of the ATP- and ADP-bound NtrC1 structures indicates that the engagement of the γ-phosphate by the R299 R-finger stimulates a rearrangement of interaction networks in the same protomer (*in cis*), at the R-finger side of the inter-protomer interface. It is proposed that interaction of this R-finger with the γ-phosphate causes helical distortions that ultimately lead to the transition of the L1 and L2 loops to a raised conformation. The K250 residue appears to be particularly important in this transition; the side chain exists in distinct environments in the two nucleotide-bound states (Chen et al. 2010). This model is in stark contrast to prior studies in PspF which indicate that the Walker B-asparagine "switch" is responsible for modulating the conformation of the σ^{54}-interaction surface in the protomer on the Walker AB side of the inter-protomer interface (Joly et al. 2007; Joly et al. 2008a). However, caution should be taken when interpreting the role of the R299 residue in the ATP-bound structure of the E239A NtrC1 variant. Comparison of the wild-type (ADP-bound) and E239A (ATP-bound) variant structures suggests that the Walker B substitution prevents an interaction between residue 239 and the other possible R-finger (R293) (Figure 3.6). R293 instead interacts with E242 which appears to cause the displacement of R299, bringing it into closer proximity with the ATP γ-phosphate. Furthermore, compared to the ATP-bound structure of PspF (Figure 3.8B), the Walker B aspartate appears to be too distant from ATP for hydrolysis to occur. Therefore the role of the R299 residue in coupling hydrolysis to substrate remodelling remains unclear. Although work in NtrC1 suggests a role for R299 (R168 in PspF) in γ-phosphate sensing, studies in PspF suggests that R293 (R162 in PspF) is the "true" R-finger.

(vi) **A model of the nucleotide-driven conformational change**

Overall, structural and biochemical studies have shown that the coupling of ATP-hydrolysis to open complex formation by Eσ^{54} is dependent upon conformational changes in the AAA+ domain (Figure 3.4D) (Rappas et al. 2006; Bose et al. 2008a; Chen et al. 2010). These changes centre on the "sensing" of the γ- phosphate by either the asparagine "switch" residue that detects changes in the Walker B motif upon ATP hydrolysis (Rappas et al. 2006) or by the highly conserved R-finger of the adjacent protomer (Chen et al. 2010). Depending upon the stage in the ATP-hydrolysis cycle and the position of the
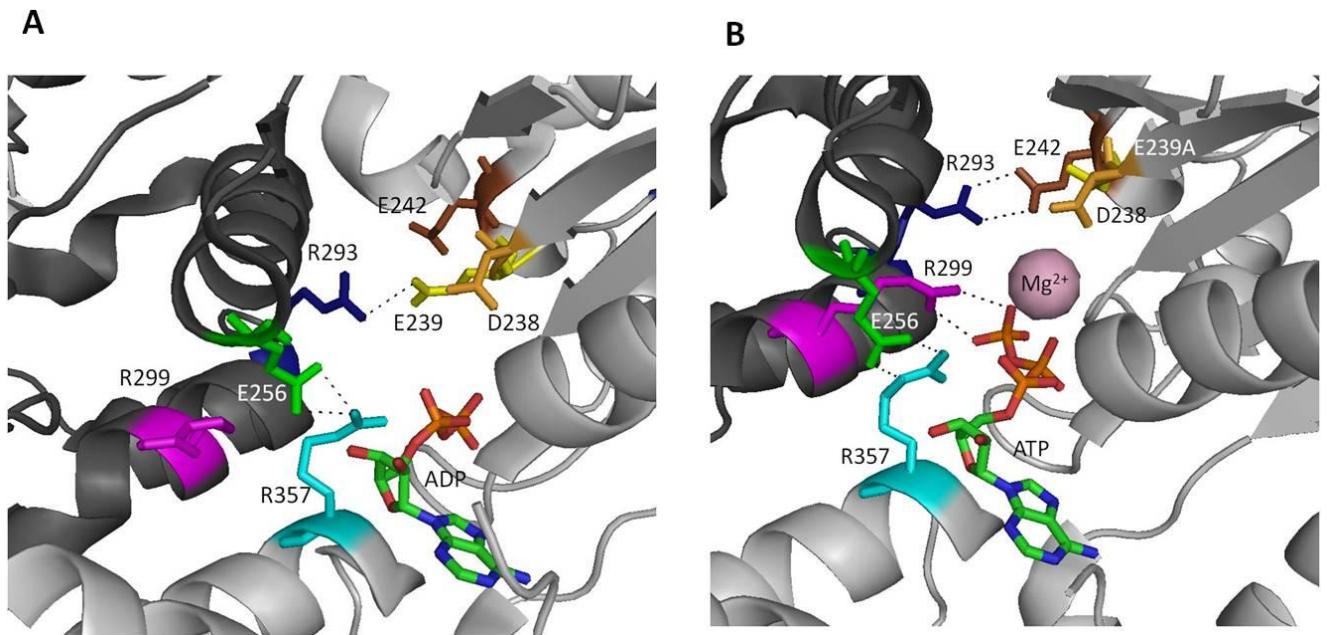
Figure 3.6 - **Comparison of the ADP-bound and ATP-bound structures of the AAA+ domains of NtrC1 and NtrC1 (E339A) respectively**. In each structure the proposed R-fingers are indicated (R293; dark blue and R299; magenta). The Sensor II arginine (R357) is in cyan. The Walker B "DE" residues (D238 and E239) are in orange and yellow respectively. The E242 (brown) and E256 (green) residues may form inter-protomer interactions (**A**) The ADP-bound structure of NtrC1 AAA+ (PDB ID: 1NY6 B-C interface; (Lee *et al.* 2003)). Work in PspF suggests that R293 Is the real R-finger and in this structure is shown to interact with the Walker B glutamate residue. In both ADP and ATP-bound structures, E256 forms an inter-protomer contact with the Sensor II arginine (R357). (**B**) The ATP-bound structure of NtrC1 (E239A) AAA+ (PDB ID: 3MOE; (Chen *et al.* 2010)). The magnesium ion is shown as a light pink sphere. To achieve this structure the Walker B glutamate was substituted for an alanine to prevent the turnover of ATP. Consequently, the possible R-finger R293 instead interacts with the E242 residue rather than the Walker B residue from the adjacent protomer. This appears to result in the displacement of the R299 residue, shifting it onto the other side of E256. Therefore caution should be taken with the suggestion that R299 has a direct role in γ-phosphate sensing. In agreement with this, the Walker B aspartate appears to be too far from the γ-phosphate for ATP hydrolysis to occur (compare ATP-bound NtrC1 [Figure 3.6B] with ATP-bound PspF [Figure 3.8B]).

"switch" or R-finger, the GAFTGA-containing L1 and L2 loops adopt varying conformations (Chen *et al.* 2010; Joly and Buck 2010). Upon nucleotide binding, the loops are in an extended conformation and the GAFTGA motif forms an unstable interaction with σ^{54}. ATP hydrolysis strengthens this interaction and remodelling of the holoenzyme can then occur to enable open complex formation. Upon phosphate release, rearrangement of both *in cis* and *in trans* interactions cause the loops to disengage with the σ factor. The cycle can then start-over with the exchange of ADP for ATP. The GAFTGA motif thus performs a crucial role in the 'power stroke' of bEBPs in coupling ATP hydrolysis to conformational rearrangements of the σ^{54}-RNA polymerase. The role of the individual residues within this essential motif will be discussed shortly (section 3.2.4)

(vii) **The coordination of ATP hydrolysis**

Structural studies of the activator bound to different nucleotides have helped establish the conformational changes within the AAA+ domain that couple nucleotide hydrolysis to σ^{54}-contact (Rappas *et al.* 2006). The ADP.AlF$_x$-bound cryo-EM structure of PspF$_{1-275}$ indicates that not all protomers within the AAA+ hexamer contact σ^{54} during the transition state of ATP hydrolysis, indicating assymmetry (Rappas *et al.* 2005). In line with this, a number of hetero-hexameric AAA+ proteins exist (e.g. eukaryotic MCM2-7) comprising up to six different proteins, strongly suggesting that each subunit may have a distinct role in the activity of the hexamer (Forsburg 2004; Bochman and Schwacha 2008). Studies in the homo-hexameric bEBP, PspF, have subsequently confirmed that an asymmetric configuration is a key requirement for open complex formation (Joly and Buck 2011). The introduction of the GAFTGA substitution T86A into single-chain forms of PspF with two or three subunits, allowed Joly and Buck to examine the minimal requirements for σ^{54}-contact and substrate remodelling. Such a substitution was shown to uncouple the ATPase and oligomerisation activities of the bEBP from its ability to contact and remodel σ^{54}. Results show that the minimal configuration for stable interaction with σ^{54} is two adjacent functional subunits, revealing that more than one GAFTGA-containing L1 loop is likely to contact σ^{54} at the point of ATP-hydrolysis (Joly and Buck 2011). However, two wild-type subunits, opposite relative to each other in the ring give rise to a hexamer that is markedly less able to interact with σ^{54} or form open complex. This result strongly suggests that asymmetry in the hexamer is important for its ability to contact and remodel σ^{54}, which is itself asymmetrical. This is in contrast to the previously proposed model of homotropic

control in which the requirement of three subdomains for the formation of the catalytic site leads to ring-imposed symmetry constraints (Schumacher *et al.* 2008). Here, it is anticipated that alternating nucleotide states (ATP, ADP and unbound) around the ring produce a two-fold symmetric conformation, suggested to be important for effective ATP hydrolysis at opposite sites in the hexamer (Figure 3.8A). Significantly, efficient open complex formation requires at least one more additional subunit, revealing that the minimal requirements for stable interaction with σ^{54} are different to those for substrate remodelling (Joly and Buck 2011).

The finding that only a subset of the GAFTGA-containing L1 loops are required to mediate σ^{54}-interaction and open complex formation in turn, reveals that only a subset of ATP-hydrolysis sites are required for bEBP activity (Joly and Buck 2011). In support of this, work carried out to determine how ATP hydrolysis between protomers is coordinated has suggested that ATPase activity in PspF is partially sequential (Joly *et al.* 2006). In the AAA+-family of proteins, two models of nucleotide occupancy may explain how hydrolysis is coordinated (Figure 3.7) (Ades 2006). Homogeneous nucleotide occupancy has been observed in a number of AAA+ protein crystal structures (Lenzen *et al.* 1998; Zhang *et al.* 2000; Gai *et al.* 2004), supporting a model of concerted/synchronised hydrolysis (Figure 3.7D) in which subunits of the AAA+ ring simultaneously hydrolyse ATP. Other AAA+ structures show mixed nucleotide occupancy with ATP, ADP or no nucleotide bound at the catalytic sites between subunits (Bochtler *et al.* 2000; Wang *et al.* 2001). This supports a model of sequential or rotational hydrolysis (Figure 3.7C) in which heterogenous nucleotide occupancy is coordinated between protomers. Studies using the bEBP PspF reveal that either ATP or ADP stimulates oligomerisation of the activator and that physiological ADP concentrations stimulate the ability of the protein to hydrolyse ATP. This suggests that in PspF, ADP binding: (i) promotes formation of the stable hexamer and (ii) causes structural changes leading to increased ATPase in adjacent protomers. Furthermore, where non-optimal binding of nucleotide occurs there are negative homotropic effects (Schumacher *et al.* 2008). High ATP concentrations, at which every catalytic site in the hexamer is likely to be in the ATP-bound form, inhibit ATP hydrolysis and the activation of transcription *in vitro*. Taken together these data strongly suggest that heterogeneous nucleotide occupancy, coordinated between protomers in the hexameric ring plays a crucial role in the activation of σ^{54}-dependent transcription by bEBPs. Therefore, the concerted/synchronised model of hydrolysis can be discounted

(Figure 3.7C). Because of the "ATP-inhibition" and "ADP-stimulation" of PspF ATPase activity, nucleotide hydrolysis in bEBPs is unlikely to be a stochastic process, in which each catalytic site is independent (Figure 3.7A). Rather, the arrangement of the hydrolysis site and *in vitro* data suggest that cooperativity exists between protomers of the hexamer (Joly *et al.* 2006). Indeed R-finger residues have been shown to function *in trans*, coordinating the bound nucleotide in the adjacent subunit of the hexamer (Zhang *et al.* 2002; Briggs *et al.* 2008; Greenleaf *et al.* 2008; Chen *et al.* 2010). Furthermore, recent mutagenesis in PspF has identified non R-finger residues involved in inter-protomer interactions that help to coordinate the position of the L1 and L2 loops during ATP hydrolysis (Joly and Buck 2010). Subsequently, a model of sequential or rotational hydrolysis is favoured (Figure 3.7B/C). These mechanisms would create an asymmetry in the exposure of GAFTGA motifs in the hexamer that has been shown to be important for the mechanical action of the activator (Joly *et al.* 2006; Joly and Buck 2011). However the exact number of nucleotides bound and the conformation of individual protomers at the discrete steps of hydrolysis still remains unclear. The generation of high-resolution crystal structures of mixed nucleotide-bound hexamers would therefore be beneficial.
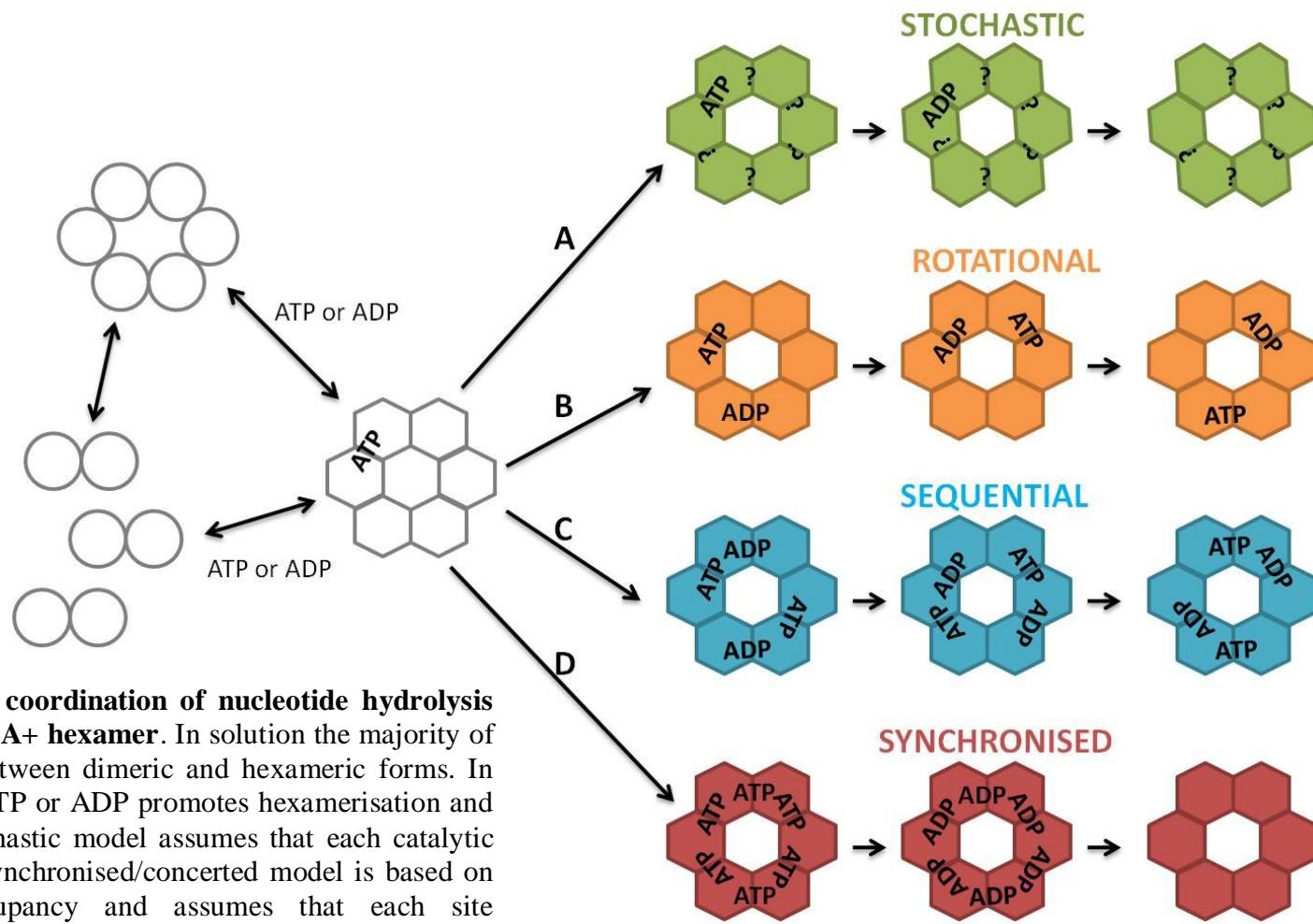
Figure 3.7 – Models for the coordination of nucleotide hydrolysis between protomers in the AAA+ hexamer. In solution the majority of bEBPs exist in equilibrium between dimeric and hexameric forms. In the case of PspF, binding of ATP or ADP promotes hexamerisation and ATP hydrolysis. (**A**) The stochastic model assumes that each catalytic site is independent. (**D**) The synchronised/concerted model is based on homogenous nucleotide occupancy and assumes that each site simultaneously hydrolyses ATP. Data in PspF suggests that ATP hydrolysis occurs either via the rotational (**B**) or sequential (**C**) mechanisms that utilise heterogeneous nucleotide occupancy. Both models are based on cooperativity between protomers in the hexamer. (Joly *et al.* 2006)

85
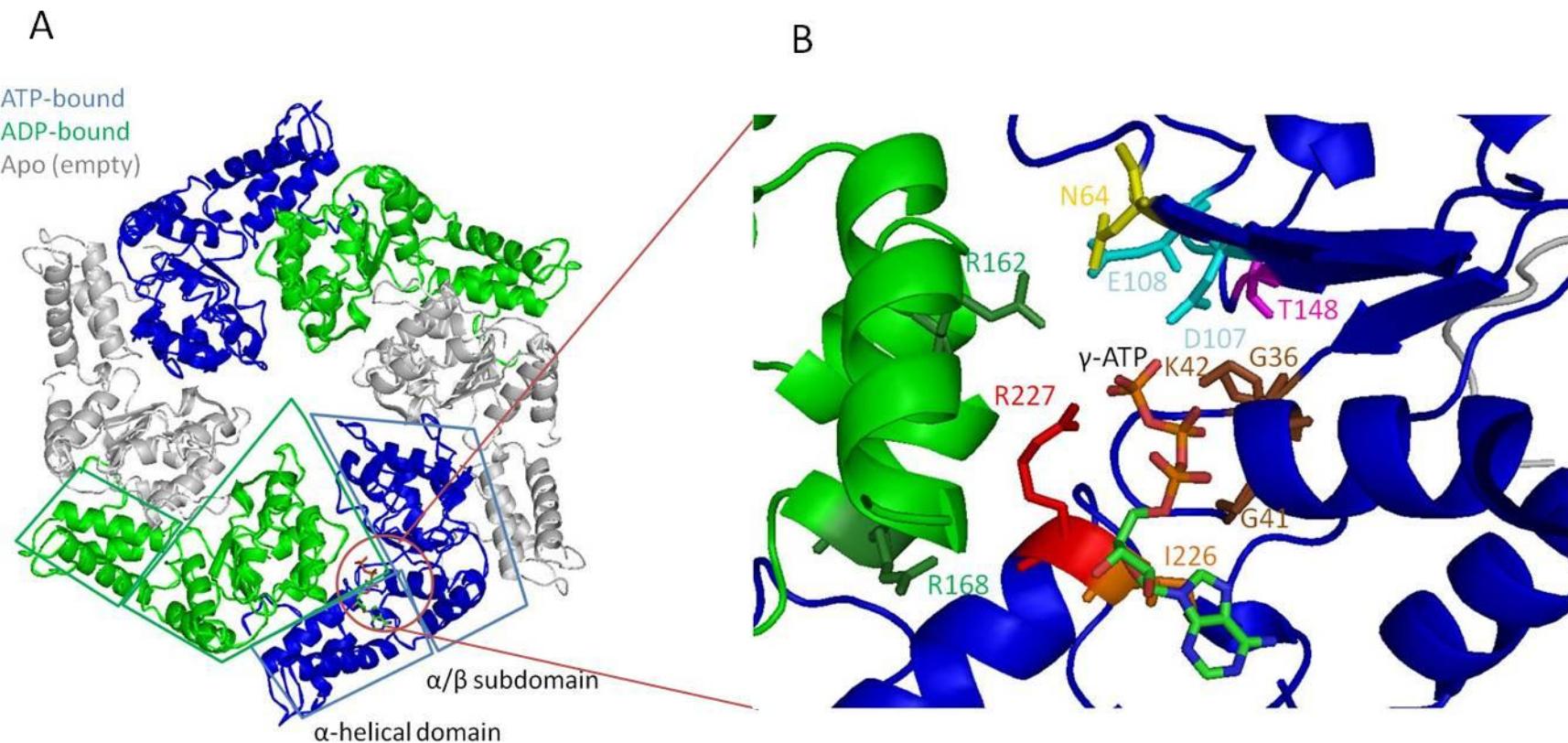
Figure 3.8 – Model of homotropic coordinated ATP hydrolysis between heterogenously occupied subunits. (**A**) Model of hexameric PspF built from the structure of monomeric PspF$_{1-275}$ (PDB 2C96) by Xiaodong Zhang, Imperial College, showing the α/β and α-helical subdomains of two adjacent protomers in the AAA+ ring. Subunits proposed to be bound by ATP at any one time are in blue, ADP-bound in green and unbound (apo) in grey. (**B**) Close-up of the nucleotide hydrolysis site in PspF in the ATP-bound state. The Walker A (G36, G41, K42) residues are labelled in brown, Walker B (D107 and E108) residues are labelled in cyan, Sensor I threonine (T148) in magenta, conserved asparagine (N64) in yellow, Sensor II arginine in red, Sensor II isoleucine in orange and the putative *in trans* R-fingers (R162, R168) in green. The location of ATP and its γ-phosphate is also indicated. Walker A forms a P-loop that interacts with the phosphates of ATP. The Walker B aspartate has a role in coordination of Mg^{2+} and the glutamate residue is thought to activate water for nucleophilic attack of the γ-phosphate. The conserved asparagine functions in the hydrolysis-dependent "switch" (Rappas *et al.* 2006). The Sensor I threonine residue (T148) has been implicated in coupling nucleotide hydrolysis to conformational change (Rappas *et al.* 2006). Sensor II residues are located in the third helix of the α-helical subdomain. I226 in PspF has been implicated in nucleotide base binding whilst the R227 residue points towards the γ-phosphate and may have a role in Mg^{2+} coordination (Schumacher *et al.* 2006). The R-fingers that have been implicated in inter-subunit catalysis and nucleotide sensing (Ogura and Wilkinson 2001; Lupas and Martin 2002; Hanson and Whiteheart 2005).

3.2.4 The GAFTGA motif

Sequence alignments of bEBP AAA+ domains indicate a very high level of conservation for the GAFTGA motif (Figure 3.3) (Zhang *et al.* 2009), reflecting its importance in σ^{54}-dependent transcription. The first glycine residue of the motif appears to be absolutely conserved although has not been widely studied in bEBPs (Table 3.2). In NtrC, random mutagenesis identified the G215V mutation that abolished *in vivo* and *in vitro* transcriptional activation (Li *et al.* 1999). In DctD, substitution of the equivalent residue (G220D) also produced an inactive protein *in vivo* (Wang *et al.* 1997). Likewise, substitution of the second amino acid of the GAFTGA motif in the bEBPs NtrC, NifA and DctD, gave rise to variants that were unable to activate transcription (North *et al.* 1996; Wang *et al.* 1997; Gonzalez *et al.* 1998). However, these variants showed little or no reduction in ATPase activity, suggesting that whilst this residue may be required for contacting σ^{54}, it does not communicate with the ATP-hydrolysis machinery of the AAA+ domain. Interestingly, in the case of the A216C variant, the defect in transcriptional activation is relieved in a form of NtrC that is incapable of binding DNA (Yan and Kustu 1999). Therefore, it appears that in the A216C mutant-version of NtrC, DNA binding prevents the bEBP from contacting σ^{54}, suggesting a relationship between DNA binding and AAA+ function.

The role of the conserved theronine of the GAFTGA motif in transcriptional activation is better understood (Table 3.2). Substitution of the T218 residue of NtrC, the T308 residue of NifA and the T223 residue of DctD abolishes the ability of the bEBP to activate transcription (Wang *et al.* 1997; Gonzalez *et al.* 1998; Li *et al.* 1999). The same is true for the T85A and T85V substitutions in PspF but T86S remains partially active (Chaney *et al.* 2001). Strong evidence for a direct interaction between the threonine residue of the GAFTGA-motif and Region I of σ^{54} was provided through the identification of substitutions (e.g. G4L in Region I) that specifically suppress the defects of the partially active T86S variant (Chaney *et al.* 2001; Bordes *et al.* 2004; Dago *et al.* 2007). Significantly, the G4L substitution allows the T86S variant to interact with but not activate the Eσ^{54} (G4L) complex when the promoter DNA is "pre-melted". This supports a role for the GAFTGA motif in the "sensing" of the promoter DNA conformation downstream of the -10 position. It has been suggested that communication of this information to σ^{54} via Region I might allow Eσ^{54} to establish contact with single-stranded DNA, required for open complex formation (Dago *et al.* 2007).

The role of the phenylalanine of the GAFTGA motif in transcriptional activation by the bEBP has also been extensively studied (Table 3.2). Mutagenesis of this residue in NtrC (F217), NifA (F307), DctD (F222), DmpR (F312) and PspF (F85) produced bEBPs that failed to activate transcription (Wang *et al.* 1997; Gonzalez *et al.* 1998; Li *et al.* 1999; Wikstrom *et al.* 2001; Bordes *et al.* 2003). The exception is the F307Y variant of NifA which retained 20% of its activity *in vivo* (Gonzalez *et al.* 1998). Indeed, 16 of 248 bEBPs identified in an alignment have a naturally occurring tyrosine at this position (Zhang *et al.* 2009). This indicates that an aromatic ring here is absolutely required for transcriptional activation by the bEBP. To further investigate the role of the phenylanine of the GAFTGA motif in σ^{54}-dependent transcription, the F85 residue of PspF was systematically substituted with 10 other amino acid residues and the functionality of the resulting variants assessed *in vitro* (Zhang *et al.* 2009). Each of the substitutions rendered the bEBP unable to activate transcription from the *nifH* promoter. The F85H, F85I, F85W, F85L, F85C and F85Q variants retained the ability to hydrolyse ATP, explained by their ability to self-associate. However, they were unable to interact with σ^{54} to form "trapped" complexes in the presence of ADP.AlF$_x$, indicating that this residue is critical for bEBP-σ^{54} contact. Since some variants e.g. F85Q showed a significant decrease in ATPase activity, it was suggested that the F85 residue communicates with the ATP hydrolysis site. In contrast, the F85A, F85E and F85R variants showed <10% of ATPase activity compared to the wild-type protein and gel filtration indicates that this was due to a defect in their ability to form higher order oligomers. This suggests that there is a structural and functional link between the phenylalanine and the distant interface of self-association. PspF F85Y was the only variant able to interact with σ^{54} to form the ADP.AlF$_x$ "trapped" complex. However, it too was unable to activate transcription and this can be explained by the inability of the protein to form activator-σ^{54}-DNA complexes using promoter DNA probes with a mismatch at the -11/-12 position (Bordes *et al.* 2004; Zhang *et al.* 2009). Importantly, the G4L substitution in Region I of σ^{54} (Dago *et al.* 2007) can rescue this σ^{54}-DNA interaction defect (Zhang *et al.* 2009). Therefore the conserved phenylalanine of the GAFTGA motif plays a role in "sensing" the conformation of DNA at the -12 promoter position, in agreement with the model based on recent Cryo-EM reconstructions (Bose *et al.* 2008b). Overall, studies using different bEBPs suggest that the conserved phenylalanine has multiple, interrelated roles during transcriptional activation. The conserved threonine has been shown to contact Region I of σ^{54} and the adjacent phenylalanine is also critical for this interaction. Previous studies suggest that F85 stabilises loop 1 (L1)–Region I interactions indirectly through the

positioning of T86 (Bordes *et al.* 2003) and therefore its major role is likely to be in contacting the promoter DNA rather than σ^{54}.

The role of the second glycine of the GAFTGA motif has been less well studied. The G219K variant of NtrC showed only a 50% reduction in activity but failed to initiate open complex formation. Surprisingly, G219K showed improved DNA binding properties (North *et al.* 1996). In contrast, the G219C variant was competent to activate transcription with an increased ATPase activity that is most likely due to increased oligomerisation. However, as is the case for the A216C variant of NtrC, binding to enhancer DNA, prevents the bEBP from activating transcription. The ability of the variant to activate transcription is restored in a form of the bEBP defective for DNA-binding (Yan and Kustu 1999). Taken together, the data suggests that in NtrC, there exists communication between the second glycine of the GAFTGA motif and the C-terminal DNA-binding domain.

In common with the other residues of the GAFTGA motif, studies in NtrC, NifA, DctD and DmpR reveal that the second alanine residue is critical for σ^{54}-dependent transcription. The A220T and A220V variants of NtrC (North *et al.* 1996; Li *et al.* 1999), the A310N, A310D and A310G variants of NifA (Gonzalez *et al.* 1998), the A225T variant of DctD (Wang *et al.* 1997) and the A315T variant of DmpR (Wikstrom *et al.* 2001) all fail to activate transcription. Only the A310S variant of NifA exhibits activity, although this is less than 20% compared to the wild-type protein (Gonzalez *et al.* 1998). Many of the variants of this position are still able to hydrolyse ATP and it is therefore likely that these substitutions destabilise the Loop1-Region I contact that forms at the regulatory centre of the closed complex.

bEBP	Residue	NorR	Substitution	Comments	Reference
NtrC	G215	G262	V	Fails to activate transcription *in vivo* or *in vitro*	(Li *et al.* 1999)
DctD	G220	G262	D	Fails to activate transcription *in vivo*	(Wang *et al.* 1997)
NtrC	A216	A263	V	Fails to activate transcription despite little reduction in ATPase activity	(North *et al.* 1996)
NtrC	A216	A263	C	Sufficient ATPase. Increased oligomerisation state. Binds enhancer but activity perturbed by DNA	(Yan and Kustu 1999)
NifA	A306	A263	D/N	Fails to activate transcription *in vivo*	(Gonzalez *et al.* 1998)
DctD	A221	A263	V/D	Fails to activate transcription *in vivo*	(Wang *et al.* 1997)
NtrC	F217	F264	L	Fails to activate transcription *in vivo* or *in vitro*	(Li *et al.* 1999)
NifA	F307	F264	A/L/I/P/R/H/N	Fails to activate transcription *in vivo*	(Gonzalez *et al.* 1998)
NifA	F307	F264	Y	<20% activity *in vivo*	(Gonzalez *et al.* 1998)
DctD	F222	F264	L	Fails to activate transcription *in vivo/in vitro*. ATPase 13% of WTΔNTD	(Wang *et al.* 1997)
DmpR	F312	F264	L	Fails to activate transcription *in vivo*. ATPase 75-85% of WTΔNTD (ATP), 23-25% WTΔNTD (dATP).	(Wikstrom *et al.* 2001)
PspF	F85	F264	A/E/R	<1% activity WT– *in vitro* transcription assays. Does not form ADP.AlFx-dependent trapped complex. Decreased ATPase activity. Defective for oligomerisation	(Bordes *et al.* 2003) (Zhang *et al.* 2009)
PspF	F85	F264	C	<1% activity WT– *in vitro* transcription assays. Does not form ADP.AlFx-dependent trapped complex. Decreased ATPase activity. Oligomerises in the presence/absence nucleotide	(Zhang *et al.* 2009)
PspF	F85	F264	H/I/W	<1% activity WT– *in vitro* transcription assays. Does not form ADP.AlFx-dependent trapped complex. WT ATPase activity. Nucleotide-dependent oligomerisation (~WT)	(Zhang *et al.* 2009)
PspF	F85	F264	L/Q	<1% activity WT– *in vitro* transcription assays. Does not form ADP.AlFx-dependent trapped complex. Decreased ATPase activity. Nucleotide-dependent oligomerisation (~WT)	(Zhang *et al.* 2009)
PspF	F85	F264	Y	<1% activity WT– *in vitro* transcription assays. Forms ADP.AlFx-dependent trapped complex. Decreased ATPase activity. Oligomerises in the presence/absence nucleotide. Cannot form activator-DNA-σ^{54} complex (phenotype rescued by G4L substitution in σ^{54})	(Zhang *et al.* 2009)
NtrC	T218	T265	A/N	Fails to activate transcription *in vivo/in vitro*.	(Li *et al.* 1999)
NifA	T308	T265	A/L/M/P/R/V/G /C/S	Fails to activate transcription *in vivo*	(Gonzalez *et al.* 1998)
DctD	T223	T265	I	Fails to activate transcription *in vivo/in vitro*. ATPase 123% of WTΔNTD	(Wang *et al.* 1997)
DctD	T223	T265	A	Fails to activate transcription. Significant ATPase retained	(Wang *et al.* 1997)
PspF	T86	T265	A	<1% activity WT– *in vitro* transcription assays. Wild-type ATPase activity Does not form ADP.AlFx-dependent trapped complex	(Chaney *et al.* 2001) (Bordes *et al.* 2003)
PspF	T86	T265	S	52% activity WT – *in vitro* transcription assays. *In vivo* ~ 25% WT Region I G4L σ^{54} substitution restores transcription activation activity Wild-type ATPase activity. Forms ADP.AlFx-dependent trapped complex	(Chaney *et al.* 2001) (Bordes *et al.* 2003) (Bordes *et al.* 2004) (Dago *et al.* 2007)
PspF	T86	T265	V	<1% activity WT– in vitro transcription assays. Wild-type ATPase activity Does not form ADP.AlFx-dependent trapped complex	(Chaney *et al.* 2001)
NtrC	G219	G266	K	50% ATPase but fails to activate transcription. Improved DNA binding	(North *et al.* 1996)
NtrC	G219	G266	C	Increased ATPase due to increased oligomerisation Binding of enhancer DNA *prevents* transcription	(Yan and Kustu 1999)
NtrC	A220	A267	T	Fails to activate transcription despite little reduction in ATPase activity	(North *et al.* 1996)
NtrC	A220	A267	V	Does not activate transcription *in vivo* but shows "hyperactivity" at low concentrations *in vitro*	(Li *et al.* 1999)
NifA	A310	A267	S	<20% activity *in vivo*	(Gonzalez *et al.* 1998)
NifA	A310	A267	N/D/G	Fails to activate transcription *in vivo*	(Gonzalez *et al.* 1998)
DctD	A225	A267	T	Fails to activate transcription *in vivo/in vitro*. Reduced ATPase	(Wang *et al.* 1997)
DmpR	A315	A267	T	Fails to activate transcription *in vivo*. Wild-type ATPase activity.	(Wikstrom *et al.* 2001)

Table 3.2 – Substitutions made in related bEBPs within the highly conserved GAFTGA motif. WT = wild-type; NTD = N-terminal domain

3.2.5 Oligomerisation of the bEBP

Since the nucleotide hydrolysis site is formed through interactions between residues of adjacent protomers (Figure 3.8), oligomerisation of the AAA+ domain is required to form a bEBP that is competent to activate transcription (Zhang *et al.* 2002; Rappas *et al.* 2007). bEBPs are typically dimeric in their inactive state with the dimerisation determinants existing in either the N-terminal regulatory (R) domain as is the case for DctD (Meyer *et al.* 2001; Park *et al.* 2002) or C-terminal DNA-binding (D) domains, as shown for NtrC (Pelton *et al.* 1999; Hastings *et al.* 2003). In response to a stimulatory signal, the oligomerisation of the bEBP is then facilitated through interactions between the central AAA+ (C) domains.

Recently the exact functional oligomeric state of bEBPs has become a matter for debate. The first structure determined for an activator of σ^{54}-dependent transcription was of the isolated ADP-bound ATPase domain of the NtrC1 protein from the extreme thermophile *Aquifex aeolicus* (Lee *et al.* 2003). The 3.1 Å structure revealed a heptameric ring with a height and diameter of 40 Å and 124 Å respectively. The recently published crystal structure of the ATP-bound form of NtrC1 confirms a heptameric arrangement and furthermore negative-stain EM of the Walker B mutant-derivative (that can bind but not hydrolyse nucleotide) shows it is competent to form a complex with σ^{54} (Buck and Hoover 2010; Chen *et al.* 2010). However, other bEBP structures have revealed hexameric arrangements (Table 3.3). When the crystal structure of the isolated ATPase domain of PspF ($PspF_{1-275}$) is fitted into the Cryo-EM structure of the activator in complex with σ^{54}, the electron density can accommodate six monomers (Rappas *et al.* 2005). Indeed, electrospray-mass spectrometry (ES-MS) shows that six monomers of $PspF_{1-275}$ form a complex with monomeric σ^{54}, consistent with bEBPs functioning as hexamers (Ogura and Wilkinson 2001; Lupas and Martin 2002). The 3 Å X-ray structure of the zinc-responsive ZraR from *Salmonella tyohimurium*, lacking the N-terminal regulatory domain also reveals a hexameric arrangement (Sallai and Tucker 2005). These more recent studies have called into question whether the heptameric configuration of NtrC1 (Lee *et al.* 2003) represents a physiologically relevant form of the bEBP. In the absence of the C-terminal (DNA binding) and N-terminal (regulatory) domains, it has remained unclear as to whether the higher order oligomer is a heptamer or has another stoichiometry. In addition, the odd-number of subunits does not match up with the dimeric arrangement of the receiver domains in both active and inactive forms (Lee *et al.* 2003; Doucleff *et al.* 2005a). To

resolve the effect that different domains have on the oligomerisation state of σ^{54}-dependent activators, a more thorough analysis was conducted using different domain combinations of the NtrC4 protein from *Aquifex aeolicus* (Batchelor *et al.* 2009). As is the case for NtrC1, the ATPase activity of NtrC4 is subject to negative regulation. The assembly of the active oligomer is repressed by the receiver domain and phosphorylation is likely to remove this repression (Batchelor *et al.* 2008). Unlike NtrC1, NtrC4 has a partially disrupted receiver–AAA+ domain interface and can assemble into active oligomers at high protein concentrations independent of phosphorylation (Batchelor *et al.* 2008). ES-MS experiments show that full-length NtrC4 (NtrC4^{RCD}) and activated NtrC4, lacking the DNA binding domain (NtrC4^{RC}) form hexamers. In contrast, the isolated ATPase domain (NtrC4^{C}), non-activated NtrC4, lacking the DNA binding domain (NtrC4^{RC}) and NtrC4 lacking the regulatory domain (NtrC4^{CD}) all form heptamers. A heptameric arrangement for the central ATPase domain in isolation is consistent with the heptamer observed when this domain of NtrC1 is crystallised (Lee *et al.* 2003). Therefore it seems that for the extreme thermophile *Aquifex aeolicus*, a heptamer is the most stable arrangement for the AAA+ domain in the absence of regulatory and DNA-binding domains. This is in contrast to the central domain of PspF which forms hexamers when in isolation (Rappas *et al.* 2005) and so despite the high conservation of AAA+ domains, the PspF and NtrC1/NtrC4 central domains must have some differences. Interestingly, when the regulatory domain of NtrC4 is absent a heptamer is formed but when present and activated, hexamerisation occurs. Since the activated receiver domain stabilizes the hexameric form of NtrC4, it appears that an intermediate mechanism of regulation exists; somewhere between the negative mechanism of NtrC1/DctD and positive mechanism of NtrC (Batchelor *et al.* 2008; Batchelor *et al.* 2009). Overall, studies examining the oligomeric state of NtrC4 have conclusively shown that truncated or non-activated proteins may have a propensity to exhibit altered stoichiometries. In the crystal structure of ZraR, it has been suggested that the AAA+ domain is held in a hexameric configuration by the DNA-binding domains which are dimeric in nature (Sallai and Tucker 2005). Therefore one explanation is that in the absence of this domain, dimerisation determinants do not exist to prevent the heptamerisation of NtrC1 and NtrC4. Due to the difficulty in activating full-length NtrC1, the oligomeric state of this construct cannot be assessed but the similarity of NtrC4 and NtrC1 suggests that NtrC1 would form hexamers. In addition to bEBPs, heptameric arrangements have also been frequently observed in other proteins of the AAA+ family (Table 3.3) such as MCM (Yu *et al.* 2002; Costa *et al.* 2006b; Costa *et al.* 2006a) RuvB

(Miyata *et al.* 2000), ClpB (Kim *et al.* 2000; Akoev *et al.* 2004), magnesium chelatase (Reid *et al.* 2003), HslU (Rohrwild *et al.* 1997), Lon (Stahlberg *et al.* 1999) and the C-terminal domain of p97 (Davies *et al.* 2008). Significantly, hexamers have also been observed for each of these proteins (Miyata *et al.* 2000; Zhang *et al.* 2000; Yu *et al.* 2002; Willows *et al.* 2004; Costa *et al.* 2006b; Costa *et al.* 2006a; Park *et al.* 2006). Various explanations exist for the existence of two different isoforms of the same AAA+ protein. In the case of RuvB from *Thermus thermophilus* and the MCM protein from *Methanothermobacter thermoautotrophicus,* heptamers form in the absence of DNA but hexamerisation occurs when DNA is present (Miyata *et al.* 2000; Yu *et al.* 2002; Costa *et al.* 2006b; Costa *et al.* 2006a). It is possible that the heptamer facilitates the loading of DNA into the central channel of the protein ring before the loss of a single subunit results in the hexamer (Yu *et al.* 2002). The bacterial protein-disaggregating chaperone, ClpB forms heptamers in the absence of nucleotide but undergoes rearrangements to form hexamers when ATP or ADP binds (Kim *et al.* 2000; Akoev *et al.* 2004). This implies that during the ATP-hydrolysis cycle (as ATP binds, is hydrolysed and ADP is released), there is "switching" between hexameric and heptameric states. This partial ring-dissociation has been suggested to facilitate the "prying apart" of aggregated substrates (Akoev *et al.* 2004). In the case of the ATPase HslU, rings of 7-fold and 6-fold symmetry have apparently been observed under the same conditions (Rohrwild *et al.* 1997). Here it is unclear as to whether the heptameric form of the protein is competent to associate with the partner protease, HslV. However, it has been suggested that the symmetry mismatch between a heptameric HslU and a hexameric HslV may facilitate loading of the substrate into the proteolytic chamber, as has been suggested for ClpAP (Kessel *et al.* 1995). Unlike the ATP-dependent proteases HslUV and ClpAP, the Lon protease combines both proteolytic and ATPase functions within a single subunit. The Lon protease from *Saccharomyces cerevisiae* has been reported to consist of seven flexible subunits (Stahlberg *et al.* 1999), possibly reflecting the requirement for "mismatch symmetry" of the HslUV and ClpAP systems. However electron microscopy of the Lon protease from *Escherichia coli* indicates a hexameric arrangement (Park *et al.* 2006). Likewise, the magnesium chelatase subunit BchI from the proteobacterium *Rhodobacter capsulatus* has 6-fold symmetry (Willows *et al.* 2004) but the equivalent subunit in the cyanobacterium *Synechocystis* has 7-fold symmetry (Reid *et al.* 2003). This suggests that it may also be possible for the same AAA+ protein to have different oligomeric states in different organisms or evolutionary groups. Finally, proteins may have heptameric AAA+ domains that are hexameric when expressed

as part of the intact protein, as has been shown for the bEBP NtrC4 (Batchelor *et al.* 2009). For example, when crystallised, the C-terminal (D2) domain of p97 reveals a 7-fold symmetry (Davies *et al.* 2008) but Cryo-EM studies of the full-length form indicate that in fact the protein has a hexameric arrangement (Zhang *et al.* 2000).

3.3 Role of the amino-terminal domain

3.3.1 Signal sensing

Many bEBPs contain an N-terminal or regulatory (R) domain that responds to various environmental signals and regulates the activity of the central AAA+ domain as a result (Figure 3.9) (Schumacher *et al.* 2006). There are three main ways in which activation of the bEBP occurs in response to an environmental cue: (1) phosphorylation, (2) ligand binding and (3) protein-protein interaction (Table 3.1). Depending on the method of bEBP activation, different domains are found in the regulatory region of the protein (Studholme and Dixon 2003). Many bEBPs are part of two-component systems that couple an external stimulus to an internal response (Stock *et al.* 2000). Such systems are commonly composed of a histidine protein kinase (HK) with a conserved kinase core domain and a response regulator protein (RR) with a conserved regulatory domain. Extracellular stimuli are sensed by the HK to modulate its activity in phosphotransfer. The HK translates a phosphoryl group to a conserved aspartate in the RR (a reaction catalysed by the RR) and the phosphorylated RR is able to activate a downstream effector domain that elicits a specific response in the bacterial cell. The bEBPs NtrC, NtrC1, NtrC4, DctD, ZraR and FlgR all have RR domains that are phosphorylated by specific sensor kinases. The best-studied RR in this group is the bEBP NtrC which is phosphorylated at the conserved D54 residue by the sensor kinase NtrB in response to the nitrogen status of the cell (Reitzer 2003). Briefly, the phosphorylation cascade is controlled by the uridylyltransferase (GlnD) which transmits the nitrogen status to the NtrB protein via the PII protein GlnB. Under nitrogen-limiting conditions, NtrB phosphorylates NtrC, activating it as a bEBP. Under nitrogen-excess conditions, the phosphatase activity of NtrB prevents NtrC activation (reviewed in (Dixon and Kahn 2004) The NtrC1 and NtrC4 bEBPs are both classed as NtrC-family members based on the high amino acid similarity (59%) but the cognate HKs and the signals controlling the two-component systems have not yet been identified (Deckert *et al.* 1998; Lee *et al.* 2003). In response to phosphorylation by the HK DctB, the DctD RR in *Sinorhizobium meliloti* and *Rhizobium leguminosarum* activates the expression of DctA (Yurgel *et al.* 2000; Meyer *et al.* 2001), a transport protein that allows bacteria to

Protein	Class	Function	From	Technique	Oligomeric structure(s)		Reference
PspF[C]	bEBP/ AAA+	Transcriptional activator that regulates phage-shock response	*E.coli*	Electron Microscopy	Hexamer		(Rappas *et al.* 2005)
NtrC[RCD]	bEBP/ AAA+	Transcriptional activator that regulates nitrogen metabolism	*Salmonella tyohimurium*	SAXS/WAXS (small- angle and wide- angle X-ray scattering) and Electron Microscopy	Hexamer (activated)		(De Carlo *et al.* 2006)
ZraR[CD]	bEBP/ AAA+	Zinc-responsive transcriptional activator	*Salmonella tyohimurium*	X-ray crystallization	Hexamer		(Sallai and Tucker 2005)
NtrC1[C]	bEBP/ AAA+	Homologue of NtrC (*Salmonella enterica*)	*Aquifex aeolicus*	X-ray crystallization	Heptamer (ADP-bound and ATP-bound)		(Lee *et al.* 2003; Chen *et al.* 2010)
NtrC4 (RC, C, CD)	bEBP/ AAA+	Homologue of NtrC (*Salmonella enterica*)	*Aquifex aeolicus*	Electrospray ionization mss spectrometry (ES-MS)	Full-length and activated-RC domain proteins hexameric	Isolated ATPase, unactivated-RC and CD proteins heptameric	(Batchelor *et al.* 2009)
RuvB	AAA+	ATP-dependent motor for branch migration in homologous recombination	*Thermus thermophilus*	Electron Microscopy	Hexamer in presence of dsDNA	Heptamer in absence of dsDNA	(Miyata *et al.* 2000)
MCM	AAA+	orthologue of eukaryotic replicative helicase candidate	*Methano-thermobacter thermo-qutotrophicus*	Electron Microscopy	Hexamer in presence of dsDNA	Heptamer in absence of dsDNA and presence of nucleotide	(Yu *et al.* 2002; Costa *et al.* 2006b; Costa *et al.* 2006a)
ClpB	AAA+	Chaperone in protein-disaggregating machinery	*E.coli*	Sedimentation equilibrium/ sedimentation velocity and Electron Microscopy	Hexamer predominant in presence of ATPγS and ADP	Heptamer predominant under low ionic strength conditions	(Kim *et al.* 2000; Akoev *et al.* 2004)
HslU	AAA+	Part of HslUV two-component protease	*E.coli*	Electron Microscopy; scanning transmission electron microscopy (STEM) of cross-linked protein; cross-linking/EMSA	Mixture of hexameric and heptameric rings		(Rohrwild *et al.* 1997)
Lon protease	AAA+	ATP-dependent protease	*E.coli* and *Saccharomyces cerevisiae*	Electron Microscopy	Hexamer (*E.coli*)	Heptamer (*Saccharomyces*)	(Stahlberg *et al.* 1999; Park *et al.* 2006)
Magenisum chelatase	AAA+	Drives insertion of Mg^{2+} into protoporphyrin (chlorophyll biosynthesis)	*Synechocystis* (cyano-bacterium) and Rhodobacter capsulatus (proteo-bacterium)	Electron Microscopy	Hexamer (Rhodobacter)	Heptamer (Synechocystis)	(Reid *et al.* 2003; Willows *et al.* 2004)
p97	AAA+	Homotypic membrane fusion	Mammalian	X-ray crystallization Electron Microscopy	Hexamer (full-length)	Heptamer (C-terminal D2 only)	(Zhang *et al.* 2000; Davies *et al.* 2008)

Table 3.3 – The formation of hexamers and heptamers by AAA+ proteins and bEBPs of the AAA+ protein family

use C4-dicarboxylic acids to grow as free-living cells or to power nitrogen fixation within symbiotic bacteroids (Park *et al.* 2002). The ZraR RR is phosphorylated by the HK ZraS in response to high Zn^{2+} concentrations and the activation of the effector domain of the bEBP results in the expression of a periplasmic Zn^{2+}-binding protein, ZraP (Leonhartsberger *et al.* 2001). Finally in *Helicobacter pylori*, the sensor HK FlgS, phosphorylates the RR FlgR, activating it as a bEBP and leading to the transcription of genes required for flagella biosynthesis (Spohn and Scarlato 1999; Brahmachary *et al.* 2004).

Other σ^{54}-activators have a regulatory domain that binds small effector molecules. Direct binding of aromatic compounds to a vinyl 4 reductase (V4R) domain activates the bEBPs DmpR and XylR (Perez-Martin and Lorenzo 1995; Shingler and Pavel 1995; Shingler 1996). In response to small-ligand binding, DmpR activates the expression of the *dmp* operon that encodes enzymes involved in the catabolism of phenol and methylphenols (Shingler *et al.* 1989; Shingler *et al.* 1993; Shingler and Moore 1994). XylR binds to toluene, *m*-xylene and *p*-xylene to activate transcription at the Pu promoter of the TOL plasmid allowing *Pseudomonas putida* to grow on toluene and related hydrocarbons (Delgado *et al.* 1995) N-terminal domain swaps between the XylR and DmpR proteins confirm that the specificity of the response is conferred by the regulatory domain (Shingler and Moore 1994). Another domain, the GAF (cGMP-specific and stimulated phosphodiesterases, Anabaena adenylate cyclases and *E. coli* FhlA) domain is a member of a large and diverse domain family that is found in all kingdoms of life (Aravind and Ponting 1997). These domains have a range of functions; they are found in diverse evolutionary and functional contexts. FhlA contains two GAF domains that bind formate to activate the transcription of the formate hydrogen lyase system (Hopper and Bock 1995). NorR contains a single GAF domain which binds NO to activate transcription of the *norV* and *norW* genes, enabling NO detoxification and is the subject of this work (Hutchings *et al.* 2002b; D'Autreaux *et al.* 2005). The crystal structure of the *Saccharomyces cerevisiae* YKG9 GAF domain reveals it is structurally similar to another class of ubiquitous amino-terminal signalling domain (Ho *et al.* 2000). The Per, ARNT and Sim (PAS) domain (Ponting and Aravind 1997) often detects signals via a bound cofactor such as heme or flavin (Taylor and Zhulin 1999) and like the GAF domain, is present in many proteins in eukaryotes, bacteria and archaea (Zhulin *et al.* 1997). Although the PAS and GAF domains share little sequence similarity they have similar structures and may share a common ancestor (Ponting and Aravind 1997; Ho *et al.* 2000). Finally, the aspartokinase,

chorismate mutase and TyrA (ACT) domain is common in metabolic enzymes that are regulated by amino acid concentration. The bEBP, TyrR contains both a PAS domain and an ACT domain, and facilitates the activation or repression of transcription of genes involved in aromatic amino acid biosynthesis and transport, although it is not an activator of σ^{54}-dependent transcription (Pittard and Davidson 1991; Pittard *et al.* 2005; Verger *et al.* 2007). The ACT domain is most likely the binding site for the aromatic amino acids tyrosine, phenylalanine or tryptophan whereas the PAS domain has been suggested to have a role in contacting the αCTD of RNAP (Pittard *et al.* 2005).

A further group of bEBPs regulate the activity of the AAA+ domain through protein-protein interaction with another protein called an anti-activator. In nitrogen fixing *Azotobacter Vinelandii*, the bEBP NifA is bound by the anti-activator NifL to prevent transcription of *nif* genes under conditions that are inappropriate for nitrogen fixation (Money *et al.* 2001; Martinez-Argudo *et al.* 2004b; Martinez-Argudo *et al.* 2004c). In addition, the NifA protein also contains other sensing domains. Binding of 2- oxoglutarate (2-OG) to the GAF domain of NifA antagonises the influence of adenosine nucleotides on the NifL-NifA interaction to ensure that the bEBP is not inhibited by NifL under nitrogen-fixing conditions (Little and Dixon 2003; Martinez-Argudo *et al.* 2004a).

Some bEBPS such as PspF, lack an N-terminal regulatory domain altogether (Studholme and Dixon 2003). PspF is instead negatively regulated by PspA *in trans* (Dworkin *et al.* 2000; Elderkin *et al.* 2002). Initially PspA was suggested to be an escaped regulatory domain of PspF but phylogenetic analysis has placed PspF in a distinct clade of response regulators (Studholme and Dixon 2003). The *pspABCDE* operon encodes several proteins that help maintain membrane integrity. It has been suggested that upon proton motive force (PMF) dissipation, PspB and PspC act as positive regulators of transcription by binding PspA and relieving inhibition of PspF. This enables PspF to activate the σ^{54}-dependent transcription from the *psp* promoter (Model *et al.* 1997; Adams *et al.* 2003; Darwin 2005). Like PspF, the related HrpR and HrpS bEBPs both lack an N-terminal regulatory domain. The HrpR/HrpS system regulates transcription from the *hrp* (hypersensitive response and pathogenicity) gene cluster (Schuster and Grimm 2000; Hutcheson *et al.* 2001) that encodes plant pathogenicity genes including components and effectors of a type III secretion pathway in *Psuedomonas syringae* (Alfano and Collmer 1997; Collmer *et al.* 2000). The extracytoplasmic function (ECF) σ-factor (Missiakas and Raina 1998), HrpL, is

the primary transcription factor that controls the expression of the *hrp* gene cluster. HrpR and HrpS have been shown to interact, forming a stable heteromeric complex and activate the σ^{54}-dependent transcription of *hrpL* (Hutcheson *et al.* 2001). In a manner analogous to the regulation of the PspF protein, in the absence of a regulatory (R) domain, HrpS (but not HrpR) is specfically bound by another protein, HrpV, to repress the activity of the heterohexamer (Jovanovic *et al.* 2011). However, there is no apparent homology between HrpV and PspA (Preston *et al.* 1998; Studholme and Dixon 2003).

3.3.2 Controlling the activity of the central (C) AAA+ domain

(i) **Negative regulation as a dominant mechanism of control**

As discussed, the N-terminal regulatory (R) domain allows the bEBP to regulate transcription at σ^{54}-dependent promoters in response to an environmental cue. Various R domains exist that function in two-component signalling or to detect the binding of small ligands. However, bEBPs have developed different methods for the transduction of this signal from the R domain (the site of detection) to the enzymatic C domain. Generally, this transduction can be subject to (1) positive control or (2) negative control (Shingler 1996). Assaying the activity of the bEBP (i.e. ATP hydrolysis, oligomerisation or activation of transcription) in a truncated form that lacks the N-terminal R domain has become the standard method for determining the mechanism of control. The first indication that the R domains of bEBPs may function in the repression of AAA+ domain activity came from the identification of semi-constitutive variants of XylR and DmpR that had substitutions in either the R domain, C domain or the interdomain linker (often referred to as the B-linker, Q-linker or L_1-linker) (Delgado *et al.* 1995; Fernandez *et al.* 1995; Shingler and Pavel 1995). Furthermore, N-terminally truncated forms of the proteins that lacked the R domain exhibited constitutively active phenotypes *in vivo*, indicating that the C domain is subject to repression by the R domain. For DmpR, constitutive ATPase activity *in vitro* was demonstrated in the absence of this negative control (Fernandez *et al.* 1995; Shingler and Pavel 1995). These observations led to a model of interdomain repression in which the R-domain represses the ATPase activity of the C-domain in the absence of a small molecule. Ligand binding to the R domain is then expected to cause derepression, allowing the bEBP to hydrolyse ATP and activate transcription (Figure 3.10A).
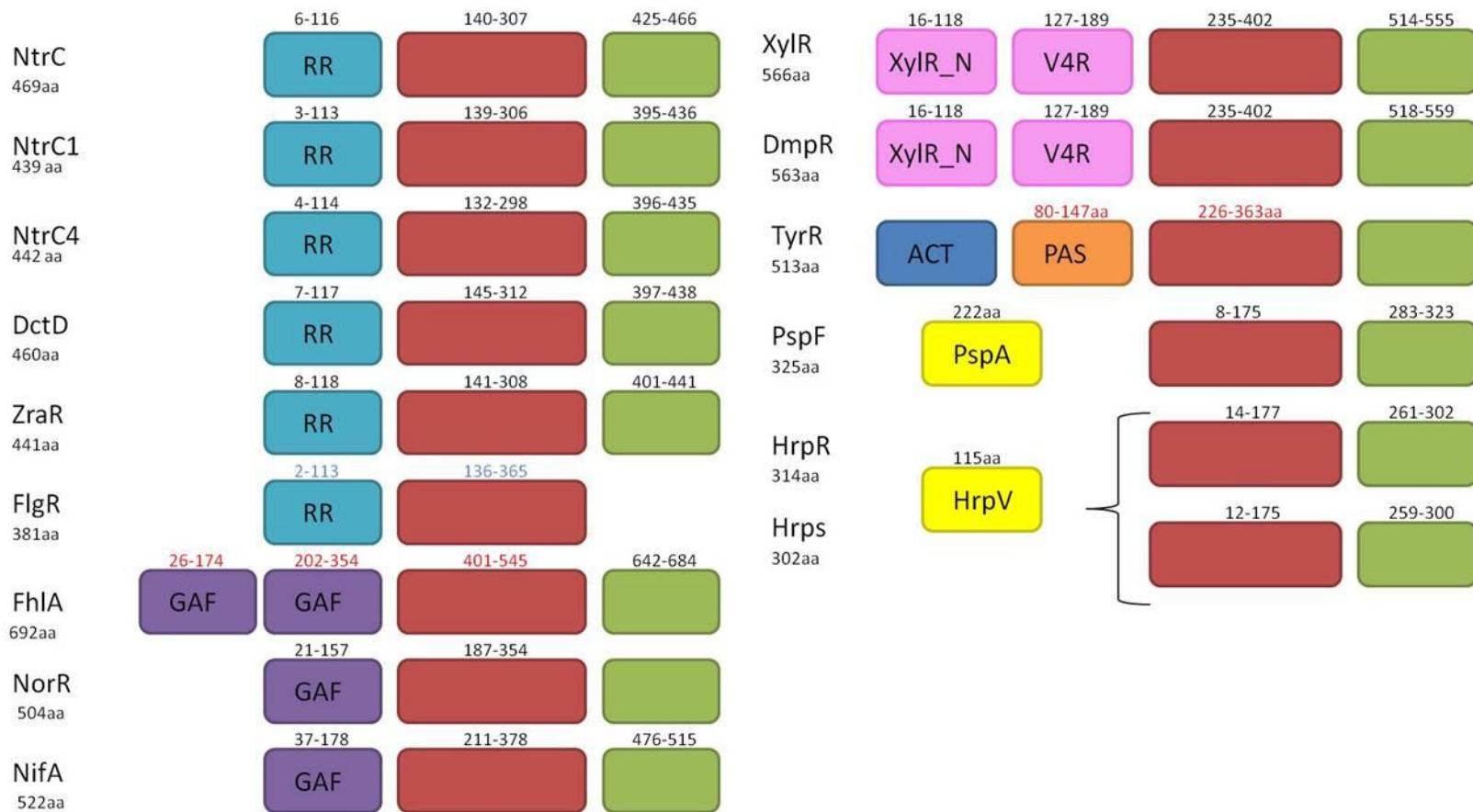
Figure 3.9 – Domain architecture of bEBPs indicating the type of regulatory (R) domain present. The different domains are coloured: The central (C) domain in red, the DNA-domain (D) if present, in green, response-regulator (RR) in light blue, cGMP-specific and stimulated phosphodiesterases, Anabaena adenylate cyclases and *E. coli* FhlA (GAF) in purple, vinyl 4 reductase (V4R) and XylR_N (found next to V4R domains) in pink, aspartokinase, chorismate mutase and TyrA (ACT) in dark blue, Per, ARNT and Sim (PAS) in orange and *in trans* regulators in yellow. HrpR and HrpS are co-activators of transcription and therefore grouped together. Where available, the predicted domain boundaries are indicated above the domain representation in black (Pfam) or red (SMART). The size of the gene product is also indicated, according to the UniProtKB database. Sequences for NtrC1 and NtrC4 were from *Aquifex aeolicus*; PspF, NorR, FhlA and TyrR were from *Escherichia coli*; NtrC and ZraR from *Salmonella typhimuriu*; XylR from *Pseudomonas putida*; DctD from *Sinorhizobium meliloti*; DmpR from *Pseudomonas CF600*; NifA from *Azotobacter vinelandii*; HrpR/S from *Pseudomonas syringae* and FlgR from *Helicobacter pylori*.

In a similar manner, NorR, the subject of this work, is thought to undergo derepression in response to the binding of NO to the R domain (D'Autreaux *et al.* 2005). More recently, this mechanism of negative control has been identified in the response regulator bEBPs DctD and NtrC1 (Lee *et al.* 2003). Here, phosphorylation and not effector binding, relieves interdomain repression (Lee *et al.* 2003; Doucleff *et al.* 2005a). Removal of the R domain and L_1-linker in DctD produced an active protein without the need for phosphorylation (Gu *et al.* 1994). Similar results were obtained with NtrC1 where activity was repressed both *in vivo* and *in vitro* in the presence of the R domain and L_1-linker but derepressed in their absence (Lee *et al.* 2003). The founder member of the σ^{54}-dependent class of transcription factors, NtrC, in contrast is positively regulated (Figure 3.10B). Deletion of the R domain to give a form of the activator that can no longer be phosphorylated by NtrB results in a constitutively inactive form of the protein, indicating that the AAA+ domain is subject to positive regulation (Drummond *et al.* 1990; Weiss *et al.* 1991). Here the phosphorylation of the R domain has a genuine stimulatory, rather than a derepressive function. Therefore, despite sharing ~60% sequence similarity, the NtrC and NtrC1 bEBPs have evolved entirely different mechanisms of regulation. In the absence of a transcriptional assay, it is not known whether ZraR is subject to positive or negative regulation but on the basis of structural similarities, it is likely to belong to the NtrC-subgroup (Sallai and Tucker 2005). Overall, the relative advantages of protein-protein interaction, phosphorylation and effector binding as control mechanisms are not understood but it seems that whatever the mechanism of sensing, negative regulation is the dominant mechanism of control (Shingler 1996).

(ii)   **Functions of the C domain targeted by the R domain**

In order for the output of the bEBP to be regulated, the sensory domain must respond to the detection of an environmental or metabolic signal by controlling the activity of the AAA+ domain that is indispensible and often sufficient for σ^{54}-dependent transcription (Berger *et al.* 1995; Jovanovic *et al.* 1999; Wikstrom *et al.* 2001; Xu *et al.* 2004a). Irrespective of whether the R domain regulates the C domain positively or negatively, it has been shown to target two different aspects of AAA+ activity: (1) the oligomerisation of the AAA+ domain and (2) the ATPase activity of the AAA+ domain. These targets will now be considered.
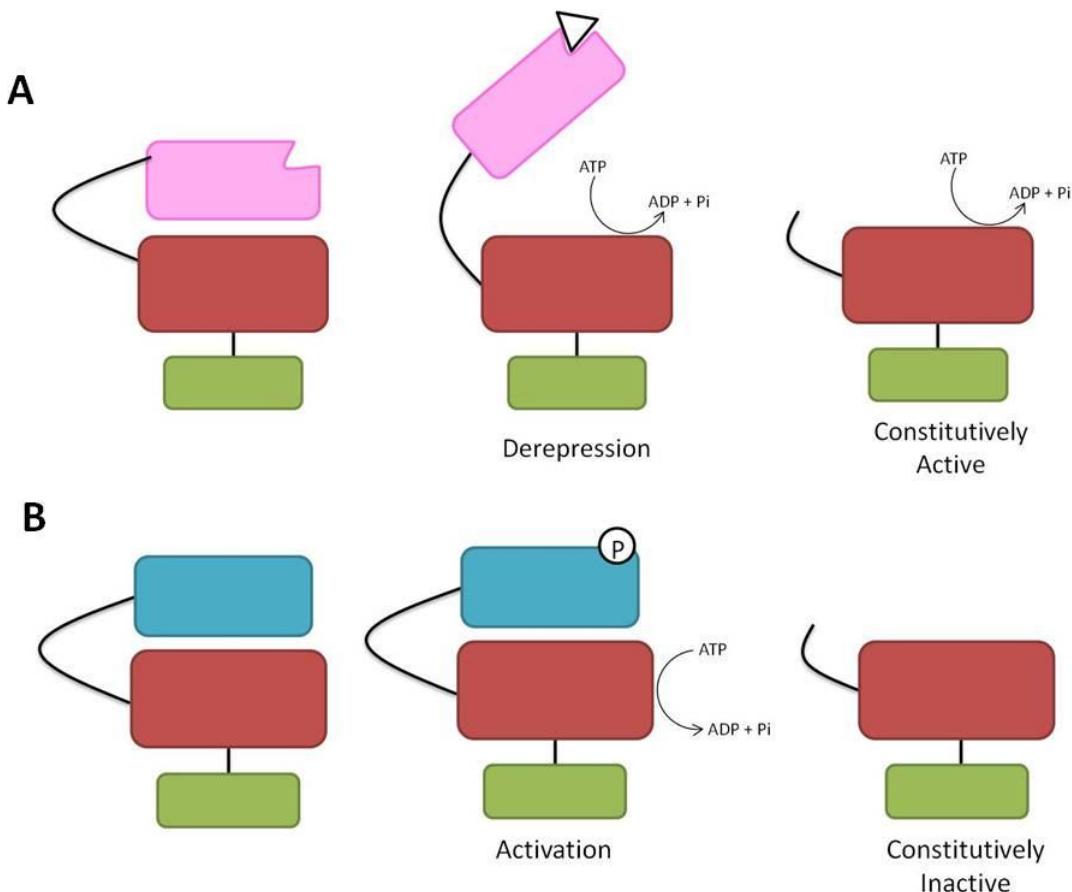
Figure 3.10 – Negative (A) and positive (B) control of AAA+ domain activity. In the more common mechanism of negative control, ligand binding (or phosphorylation) relieves the repression of the regulatory (R) domain on the central (C) domain which is intrinsically competent to hydrolyse ATP. The AAA+ domain is then able to carry out ATP hydrolysis. Accordingly, when the R domain is removed, the bEBP is active irrespective of the presence/absence of a signalling molecule or available kinase. In positive control, ligand binding or phosphorylation has a genuine stimulatory function. The phosphorylated or ligand bound form of the R domain activates the C domain which is not intrinsically competent to hydrolyse ATP. The AAA+ domain is then able to carry out ATP hydrolysis. Accordingly, when the R domain is removed, the bEBP is inactive irrespective of the presence/absence of a signalling molecule or available kinase.

(1)   Controlling AAA+ oligomerisation

As has been described, self-association of the AAA+ domains of the bEBP must occur in order to form the functional activator (Zhang *et al.* 2002; Rappas *et al.* 2007). Therefore the oligomeric determinants of the C domain represent an ideal target for the N-terminal regulatory domain in either a positive or negative mechanism of control. Structural studies of full-length and truncated forms of NtrC1 and DctD from *Aquifex aeolicus* and *Sinorhizobium meliloti* respectively indicate that the N-terminal R domain targets the oligomeric determinants of the AAA+ C domain in the mechanism of negative control (Figure 3.11B) (Meyer *et al.* 2001; Park *et al.* 2002; Lee *et al.* 2003; Doucleff *et al.* 2005a; Chen *et al.* 2008). The crystal structure of the NtrC1 protein composed of the regulatory domain joined to the central AAA+ domain by linker 1 (R-L$_1$-C) reveals a dimeric structure in which the arrangement of the subunits is incompatible with AAA+ ring assembly (Lee *et al.* 2003). Here, the unphosphorylated receiver domains form a homodimer that holds the AAA+ protomers in an inactive front-to-front configuration via interactions involving the coiled-coil linker between the R and C domains (linker 1). Structures of the activated regulatory domain indicate that phosphorylation disrupts the repressive interaction between the R and C domains, forming an alternative homodimer configuration and allowing reorientation of the AAA+ protomers into a front-to-back configuration (Doucleff *et al.* 2005a). This phosphorylation-dependent rearrangement allows self-association to take place to form an oligomer competent to hydrolyse ATP. In-line with this, a crystal structure of the NtrC1 C domain that is not subject to repression from the R domain shows a heptameric arrangement (Lee *et al.* 2003). It appears that the coiled-coil of linker 1 is critical to holding the central domains in an inhibitory configuration; its presence has become indicative of this type of regulation (Doucleff *et al.* 2005a). Indeed, in the ligand-binding XylR and DmpR proteins, mutational analysis has shown that the integrity of the linker between the regulatory and central domains is crucial for the repression of activity (Garmendia and de Lorenzo 2000; O'Neill *et al.* 2001). Despite employing the same mechanism of repression, the DNA-binding specificities of NtrC1 and DctD have evolved to target the promoters of genes linked to entirely different cell functions. The mechanism of regulation in NtrC also targets the oligomeric determinants of the AAA+ domain but is in stark contrast to that in NtrC1 and DctD (Figure 3.11A) (Doucleff *et al.* 2005a; De Carlo *et al.* 2006; Chen *et al.* 2008). A truncated form of the protein that lacks the R domain is constitutively inactive, indicating a genuine stimulatory rather than a derepressive role for phosphorylation. Activation of

oligomerisation occurs upon phosphorylation which exposes a hydrophobic patch on the R domain allowing it to bind to the N-terminal region of the central AAA+ domain. Recent X-ray solution scattering (SAXS/WAXS) and electron microscopy studies indicate that the R domain interacts with the C domain of an adjacent protomer on the outside edge of the AAA+ ring, promoting self-association and contributing to the stability of the resulting hexamer (De Carlo *et al.* 2006). Put simply, in the positive regulation of NtrC, phosphorylation of the regulatory domain *creates* a new interaction that leads to oligomerisation whereas in negative regulation, typified by NtrC1 and DctD, phosphorylation *releases* an interaction that leads to the formation of the functional oligomer (Figure 3.11). Interestingly, sequence analysis reveals a correlation between the mechanism of negative control in NtrC1 and DctD and the presence of a structured linker between R and C domains (linker 1) and an unstructured linker between C and D domains (linker 2). Conversely, the positively regulated NtrC contains an unstructured linker 1 and a structured linker 2. In both classes, the structured linker seems to play a significant role in stabilising the inactive dimer and the examination of the linker 1 and linker 2 sequences of other bEBPs may help identify whether self-association is subject to positive or negative control (Doucleff *et al.* 2005a). For example, mutation of the linker between the R and C domains in NtrC does not affect its activity (Wootton and Drummond 1989), in agreement with a function for the regulatory domain in positive rather than negative control. The bEBP NtrC4 from *Aquifex aeolicus* has a partially disrupted receiver–AAA+ domain interface and can assemble into active oligomers at high protein concentrations independent of phosphorylation, a process that does not occur with NtrC1 (Batchelor *et al.* 2008). The activated receiver domain has been shown to stabilize the hexameric form of NtrC4, thus functioning as an intermediate between the negative mechanism of NtrC1/DctD and positive mechanism of NtrC (Batchelor *et al.* 2008; Batchelor *et al.* 2009).

A Positive control (NtrC)

(1) Phosphorylation (2) Oligomerisation
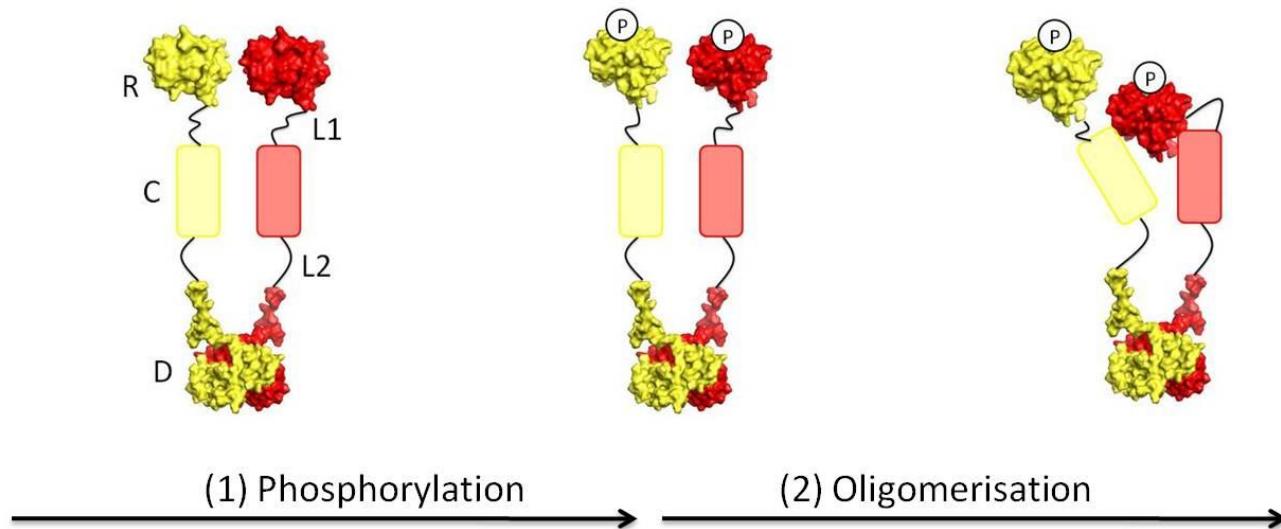
B Negative control (NtrC1/DctD)
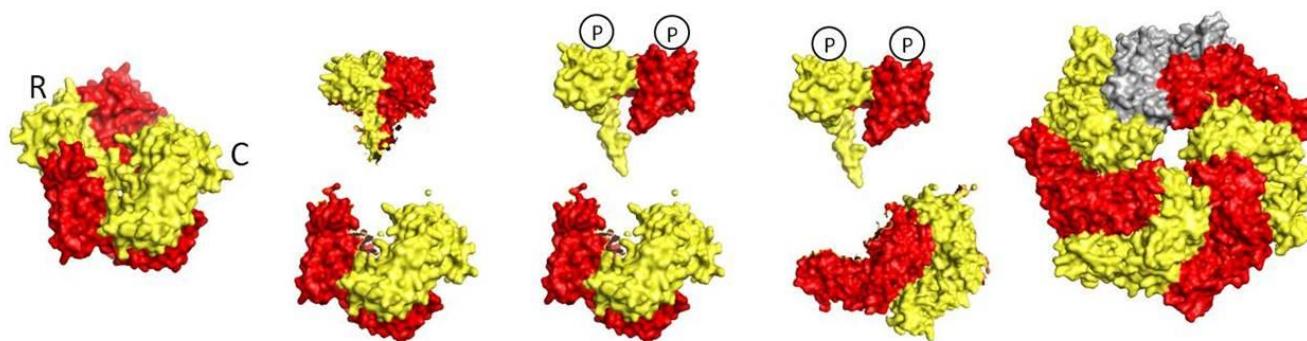
Figure 3.11 – Models of bEBP activation by phosphorylation through the promotion of oligomerisation by genuine stimulatory (**A**) **and derepressing** (**B**) **functions of the response regulator (RR) domain**. In activated NtrC the DNA-binding domain is hidden underneath the hexamer ring. For DctD and NtrC1, no information is available to define the position of DNA-binding domains. R = regulatory domain; L1 = linker 1; C = central domain; L2 = linker 2; D = DNA-bidning domain. Models were built using published structures: for NtrC fragments R (off-state PDB 1KRW, on-state 1KRX) and L2-D (PDB 1NTC); for NtrC1 fragment R (PDB 1ZY2), and fragments R-L1-C (PDB 1NY5) and L1-C (PDB 1NY6). Figure adapted from (Doucleff *et al.* 2005a).

(2) Controlling AAA+ ATPase activity

Ultimately the target of R domain-mediated regulation is the enzymatic activity of the bEBP. Where the R domain targets the oligomeric determinants, the effect is to promote or prevent formation of an oligomer that is capable of hydrolysing ATP. However, the regulatory domain of some bEBPs may specifically target the nucleotide hydrolysis machinery. This has been shown for the bEBP PspF which is regulated *in trans* through direct interaction between the activator and the negative regulator PspA (Figure 3.12) (Dworkin *et al.* 2000; Elderkin *et al.* 2002; Elderkin *et al.* 2005). Here oligomerisation is driven by the binding of ADP and ATP to the individual protomers (Joly *et al.* 2006); the 'DE' residues of the Walker B prevent nucleotide-independent hexamer formation (Joly *et al.* 2007). PspA has been shown to negatively regulate the ATPase activity of PspF through the formation of an interaction that is dependent on a surface exposed tryptophan residue (W56 of PspF) (Elderkin *et al.* 2002; Elderkin *et al.* 2005). Recently it has been shown that PspA-mediated inhibition of PspF ATPase activity is likely to involve repositioning of the conserved asparagine (N64 in PspF) involved in the sensing of the γ-phosphate during nucleotide hydrolysis. Substitutions of this asparagine in PspF do not prevent PspA-binding but ATPase activity is not significantly decreased as it is in the wild-type activator (Joly *et al.* 2008a). Consequently a model has been proposed to link the binding of PspA to the inhibition of ATP hydrolysis in PspF. Binding of PspA is detected via the W56 residue which relays this information to N64, via β-sheet 2. This leads to the repositioning of the N64 side chain, altering the distances between ATP, the conserved asparagine and the Walker B glutamate (E108 in PspF) (Joly *et al.* 2008a). These distances are thought to be critical for ATP hydrolysis and the coordination of resulting conformational changes in the AAA+ domain. Significantly it has been demonstrated that the inactive regulatory complex consists of approximately six PspA subunits and six PspF subunits (Joly *et al.* 2009). Therefore in contrast to the bEBPs NtrC1 and DctD, negative regulation of PspF activity is unlikely to target the oligomeric determinants (Lee *et al.* 2003; Doucleff *et al.* 2005a). In addition, PspA does not inhibit the interaction between PspF and σ[54], suggesting that negative regulation does not target the σ[54]-interaction surface of PspF. The PspA-PspF regulatory complex is instead expected to have an altered arrangement in the key ATPase determinants that form the catalytic site at the inter-protomer interfaces of the PspF hexamer. It may be that the inhibition of a pre-assembled PspF hexamer by PspA, allows the cell to rapidly respond to membrane damage (Joly *et al.* 2009).
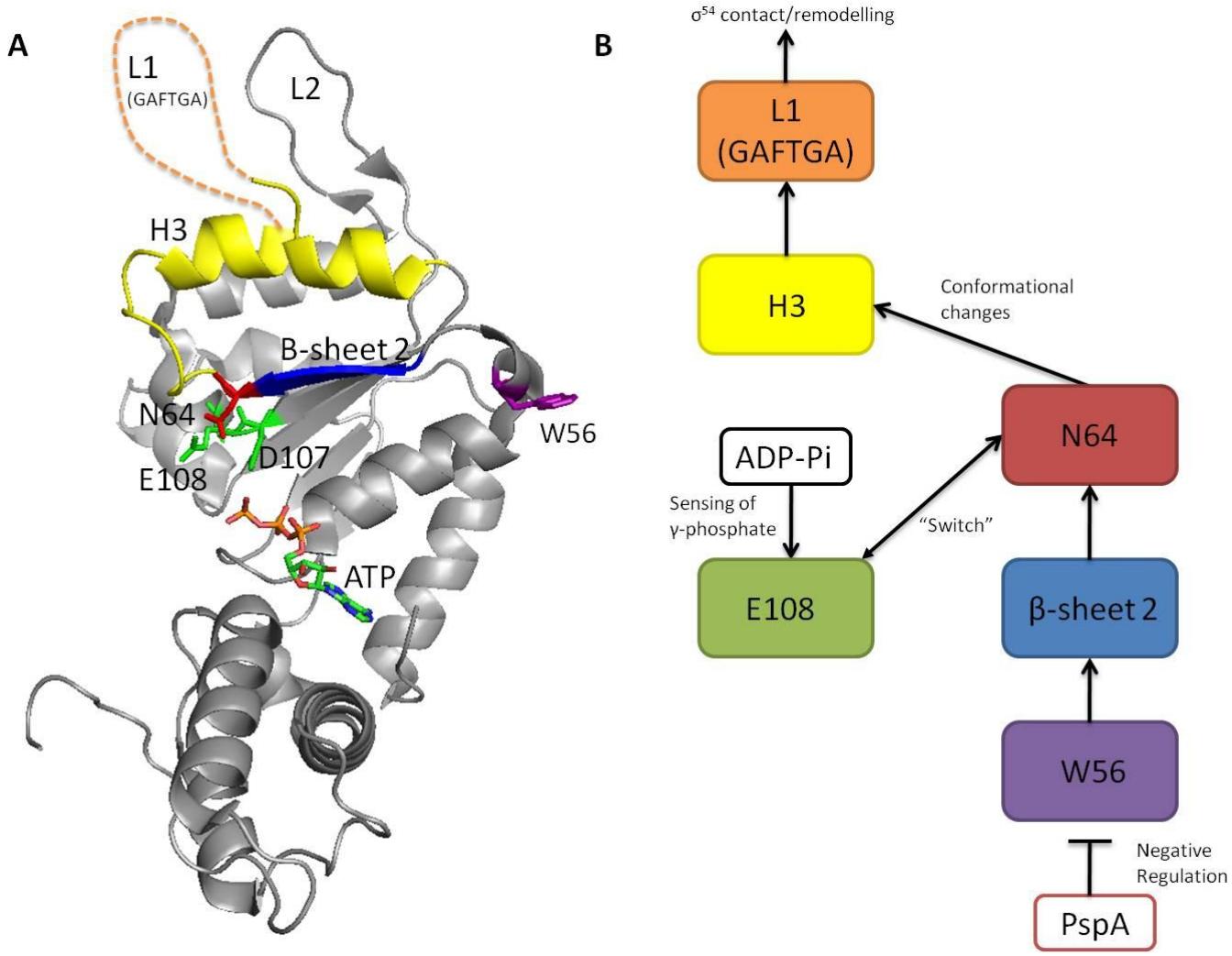
Figure 3.12 – Negative regulation of PspF AAA+ activity by PspA targets the nucleotide hydrolysis machinery via the W56 residue (Joly *et al*. 2008a). (**A**) Crystal structure of PspF$_{1-275}$ (PDB 2C96) in ATP-bound state showing the key residues involved. (**B**) Model of the signalling pathway coupling negative regulation to substrate remodelling. PspA directly interacts with PspF, an interaction that is detected via the surface exposed W56 (purple) residue of PspF. W56 relays this information to the conserved asparagine (N64, red) via β-sheet 2 (blue). This causes the repositioning of the Walker B glutamate (E108, green) to prevent ATP hydrolysis. Upon dissipation of the proton motive force (PMF), PspA inhibition is prevented (possibly facilitated by PspB and PspC) and ATP hydrolysis can occur, strengthening the σ^{54} interaction and leading to substrate remodelling. Removal of the γ-phosphate leads to a 90° rotation of the E108 side chain, breaking the interaction with N64. This change is translated to the GAFTGA-containing L1 (orange) via helix 3 (H3, yellow) and the loops compact back downwards.

3.4   Role of the carboxy-terminal domain

The C-terminal or DNA binding (D) domain contains a helix-turn-helix (HTH) motif that directs the binding of the bEBP to enhancer sites typically 80-150 bp upstream of the promoter (Xu and Hoover 2001; Studholme and Dixon 2003). FleQ, a regulator of flagellar biosynthesis from *Pseudomonas aeruginosa* is an atypical bEBP in that DNA binding can occur either upstream or downstream depending on the target promoter. On binding upstream from a distance, FleQ activates transcription as a typical bEBP via DNA looping but on binding downstream in the vicinity of the promoter, FleQ activates transcription via a novel mechanism that probably involves contacting the core RNAP (Jyot *et al.* 2002). In general, the role of the D domain in bEBPs can be considered to be two fold: (1) directing the binding of the activator to ensure a specific response; (2) facilitating the formation of, or stabilising the hexamer. The specificity of binding is maintained by well conserved enhancer binding sites (upstream activator sequences; UASs), bound by the second (recognition) helix of the HTH motif (Contreras and Drummond 1988). All sites exhibit a dyad symmetry and it is therefore unsurprising that the majority of bEBPs bind to DNA as dimers. This is supported by the crystal structure of ZraR and the NMR structure of NtrC that show dimerisation of the HTH motifs involving an α-helix, similar to that found in the FIS protein (Pelton *et al.* 1999; Sallai and Tucker 2005). All bEBPs bind to at least one enhancer site and as many as three have been identified upstream of the target promoter. NtrC dimers bind to two enhancer sites and recruit a third dimer from solution to form the functional hexamer upon phosphorlyation of the R domain (De Carlo *et al.* 2006). NorR is unusual in that it binds to three enhancer sites, each of which is essential for formation of an ATPase active-hexamer (Tucker *et al.* 2010a).

Oligomerisation has been shown to be DNA dependent in the bEBPs XylR (Perez-Martin and de Lorenzo 1996), NtrC (Rombel *et al.* 1998) and NorR (Tucker *et al.* 2010a). Where more than one UAS site is present, the binding of multiple bEBP dimers to enhancer DNA may lead to an increase in the local concentration of activator, thereby facilitating oligomerisation. However, at high concentrations some bEBPs have been shown to activate transcription without binding to enhancer DNA. Indeed C-terminally truncated forms of the activators PspF, NtrC, NifA and DctD have been shown to be active *in vivo* and *in vitro* (Morett *et al.* 1988; Huala and Ausubel 1989; Huala *et al.* 1992; Berger *et al.* 1994; North and Kustu 1997; Jovanovic *et al.* 1999). Intriguingly, some bEBPs such as *Chlamydia trachomatis* CtcC and *Helicobacter pylori* FlgR naturally lack the C-terminal

DNA-binding domain that is present in most other bEBPs (Beck *et al.* 2007). An N-terminally truncated version of FlgR (lacking the repressing RR domain) is competent to activate σ^{54}-dependent transcription from a promoter that naturally contains no upstream or downstream enhancer sites (Brahmachary *et al.* 2004). FlgR and CtcC (Koo and Stephens 2003) are the only activators of σ^{54}-dependent transcription in *H. pylori* and *C. trachomatis* respectively negating the need for enhancer-binding in these organisms. The energy savings gained by using such activators are likely to be small given the regulatory potential of using multiple DNA-binding bEBPs. Consistent with this *C. trachomatis* and *H. pylori* have only a limited biosynthetic capability (Brahmachary *et al.* 2004). In contrast FleT also lacks the C-terminal domain but is not the sole activator of σ^{54}-dependent transcription in *Rhodobacter sphaeroides*. Here, specificity is achieved through multiple σ^{54} paralogues that function at different sets of promoters (Poggio *et al.* 2002; Beck *et al.* 2007). For example, one of the paraolgues, RpoN1 functions with NifA to regulate the expression of the *nif* genes whilst RpoN2 is required for the transcription of the flagella genes.

Recent SAXS/WAXS structures and cryo-EM reconstructions of full-length, activated NtrC indicate a role for the DNA-binding domains in the stabilisation of the oligomer (Figure 3.13) (De Carlo *et al.* 2006). EM reconstructions of the bEBP bound to different nucleotides reveal significant changes in the position of the D domains during the ATP hydrolysis cycle (De Carlo *et al.* 2006). In the ADP.AlF$_x$-bound transition state in which σ^{54} contact is strengthened by interactions involving the GAFTGA motif, the DNA binding domains pack closely against the bEBP ring and are therefore likely to distort enhancer DNA. Upon phosphate release and disengagement of the GAFTGA-loop, the DNA binding domains appear to lose their tight association with the ATPase ring. Such conformational changes may stabilise the hexameric arrangement and/or facilitate the interaction between the bEBP and σ^{54} (De Carlo *et al.* 2006).
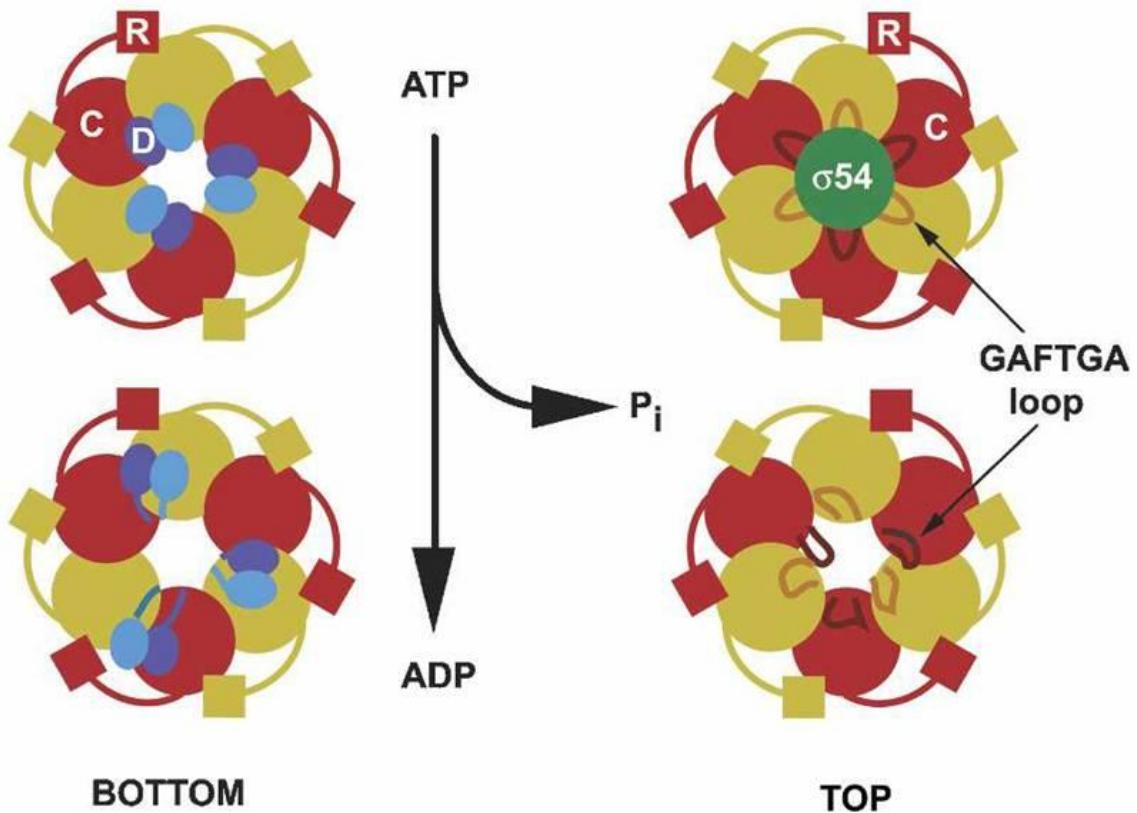
Figure 3.13 – Schematic representation of σ^{54}-remodelling by NtrC. Shows the conformational changes in the GAFTGA-loops of the central (C) domain (red/yellow circles) and the DNA binding (D) domain (blue/purple ovals) in activated NtrC. The regulatory (R) domains (red/yellow squares) are shown in the phosphorylated form that promotes hexamerisation. NtrC dimers are thought to bind to two UAS sequences and once activated, recruit a further dimer from solution in order to oligomerise. The D domains are located on the bottom of the bEBP ring whereas the GAFTGA loops contact σ^{54} (green) from the top. In the transition state of ATP hydrolysis, the interaction between the GAFTGA loops and σ^{54} is strengthened and the DNA binding domains form a tight association with the oligomeric ring. Upon phosphate release, the loops disengage from σ^{54} and the tight constraints upon enhancer DNA are relaxed. Figure from (De Carlo *et al.* 2006).

Chapter 4 – Introduction to NorR and the present work

As outlined in Chapters 1 and 3, the subject of this thesis, NorR, activates the expression of the anaerobic NO reductase (NOR), flavorubredoxin (NorV) as part of the response to nitrosative stress *E. coli* and is a bacterial Enhancer Binding Protein (bEBP) of the σ^{54}-dependent class. The structure and function of NorR will now be discussed in greater depth here.

4.1   NorR regulates *norV* expression in *E. coli* in response to NO

Formerly known as *ygaA*, the NorR protein was first identified as a putative regulator of unknown function in *E. coli* (Ramseier *et al.* 1994). It was subsequently shown to have 43% sequence identity with the NorR protein in *Ralstonia eutropha*, a denitrifying bacterium that contains a novel single component NO reductase (NOR) (Cramm *et al.* 1999; Pohlmann *et al.* 2000). The gene that encodes this protein is present in two copies; the *norB1* allele is present on the megaplasmid pHG1 and the *norB2* allele is located on the chromosome. Both produce functional proteins that can compensate for each other under physiological conditions (Cramm *et al.* 1997). The *norA1* and *norA2* genes are present immediately upstream of the *norB1* and *norB2* genes respectively and encode NorA proteins that share homology with the YtfE protein from *E. coli* (Justino *et al.* 2007; Todorovic *et al.* 2008). It was previously shown that denitrification in *R. eutropha* requires *rpoN* (Romermann and Friedrich 1985; Romermann *et al.* 1989) and a σ^{54} consensus promoter was identified upstream of the *norAB* cluster (Pohlmann *et al.* 2000). Furthermore the expression of both *norB* alleles was shown to require *rpoN* and *norR*. Indeed, in common with the majority of σ^{54}-dependent bEBPs, *R. eutopha* NorR is a tripartite protein with an amino-terminal regulatory (R) domain, a central (C) AAA+ domain and a carboxy-terminal DNA-binding (D) domain (Figure 4.1B). Pohlmann et al. constructed a *norA1-lacZ* fusion and demonstrated that in the presence of *norR*, the NO donor sodium nitroprusside (SNP) gave high levels of expression implying that NO is an efficient inducing agent. When the signalling domain of NorR was removed, the resulting truncated protein led to constitutive expression of *norA1*. Following demonstration that NorR in *R. eutropha* regulates *norAB*, it was shown that the *ygaA* gene product in *E. coli* regulates the expression of the flavorubredoxin NorV (Gardner *et al.* 2002; Gomes *et al.* 2002; Hutchings *et al.* 2002b). Expression of *norV* was demonstrated in the presence of nitrate, nitrite, NO gas and reactive nitrogen species (RNS) such as that produced by the NO donor SNP. This suggested that NorR activates the expression of *norV* in response to

NO (Figure 4.1A). On the basis of this regulation *ygaA* was renamed *norR*. As was the case in *R. eutropha* (Pohlmann *et al.* 2000), when the N-terminal signalling domain was removed, the truncated form of NorR induced expression both in the presence and absence of an NO source (Gardner *et al.* 2003).

4.2   Negative regulation of NorR activity

Many bEBPs have an additional N-terminal regulatory domain that stringently controls the activity of the AAA+ domain either positively or negatively in response to various environmental cues (Studholme and Dixon 2003). NorR contains an N-terminal regulatory GAF (cGMP-specific and –stimulated phosphodiesterases, *Anabaena* adenylate cylases and *Escherichia coli* FhlA) domain that has been predicted to bind NO. Since an N-terminally truncated form of NorR, lacking the regulatory GAF domain (NorRΔGAF), is competent to activate transcription in the absence of NO (Pohlmann *et al.* 2000; Gardner *et al.* 2003), NorR falls into a category of bEBPs in which the activity of the central AAA+ domain is negatively regulated by the N-terminal domain (Shingler 1996). Similar results have been obtained with XylR (Fernandez *et al.* 1995), DmpR (Shingler and Pavel 1995), NtrC1 and DctD (Lee *et al.* 2003; Doucleff *et al.* 2005a) in which the activity of the C domain is also negatively regulated by the R domain.

4.3   Mechanism of NO-sensing by NorR

Much of the recent research surrounding NorR has focussed on the mechanism of NO-sensing. NO responsive proteins are commonly "secondary" NO sensors whereby the principal function of the regulator is to sense another signal (Spiro 2007). Examples include SoxR, which principally responds to superoxide (Ding and Demple 2000), OxyR, which responds to hydrogen peroxide (Hausladen et al., 1996) and FNR, which responds to oxygen (Cruz-Ramos *et al.* 2002). NorR and the global regulator, NsrR, are the only known dedicated NO-sensors (Spiro 2007). Recent studies in NsrR support a model in which the repressor senses NO directly via a [2Fe-2S] cluster (Tucker *et al.* 2010b). A number of mechanisms for NO-sensing by NorR have been suggested (Gardner *et al.* 2003). Initially, it was proposed that NorR might find itself in the NtrC-class of activators that are part of two-component systems (Ninfa and Magasanik 1986; Weiss and Magasanik 1988). In this case, as a receiver (RR) domain, the N-terminal region of NorR would be phosphorylated by an NO-sensing histidine kinase (HK). In support of this hypothesis,
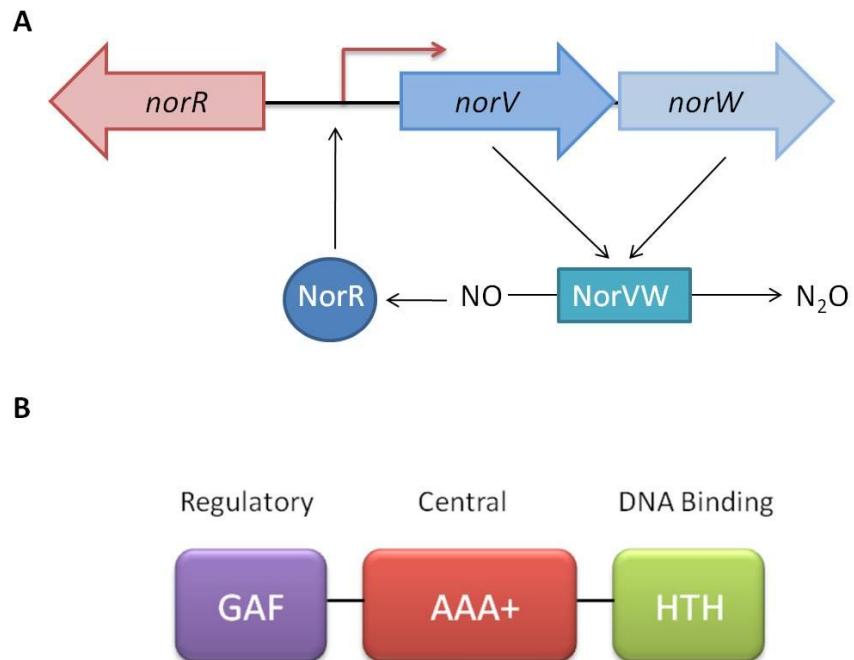
Figure 4.1 - NorR regulates *norV* expression in response to NO. (**A**) Schematic representation of the *norR-norVW* intergenic region. The flavorubredoxin, NorV and its redox partner, NorW together exhibit NO reductase (NOR) activity that allows *E. coli* to convert nitric oxide (NO) into nitrous oxide (N_2O). NorR is divergently transcribed upstream from the *norVW* genes and regulates their expression in response to NO. (**B**) Schematic showing the tripartite domain organisation of the bEBP NorR with an amino-terminal signalling and regulatory (GAF) domain (purple), a central nucleotide binding (AAA+) domain (red) and a carboxy-terminal DNA-binding (HTH) domain (green).

there are conserved aspartate residues in the N-terminal domain of NorR that could potentially be phosphorylated by a sensor kinase. However, the N-terminal signalling domain of *R. eutropha* NorR is devoid of phosphorylation signatures and the cognate histidine kinase is absent in the DNA sequences adjacent to the *norR* genes (Pohlmann *et al.* 2000). Another possibility suggested was that the NorR N-terminal domain could interact with a signal transduction protein in a manner reminiscent of the NifL/NifA and PspA-PspF systems (Elderkin *et al.* 2002; Gardner *et al.* 2003; Martinez-Argudo *et al.* 2004b; Martinez-Argudo *et al.* 2004c; Elderkin *et al.* 2005). Alternatively the NorR N-terminal domain might possibly sense NO directly, as previously demonstrated for the bEBP FhlA which binds formate (Hopper and Bock 1995). Indeed, a number of potential candidate residues that could form the NO-sensor in the N-terminal domain of NorR have been suggested (Gardner *et al.* 2003). The N-terminal domain contains a His^{111}-*X*-Cys^{113} site reminiscent of the Cys^{75}-*X*-Cys^{77} heme iron switch motif in the carbon monoxide sensing protein CooA of *Rhodospirillum rubrum* (Lanzilotta *et al.* 2000).

In order to investigate the mechanism of NO sensing in NorR, electron paramagnetic resonance (EPR) spectroscopy was carried out on whole cells of *E. coli* exposed to NO (D'Autreaux *et al.* 2005). A new EPR signal was observed in the $g = 4$ region only when the cells expressed NorR and were exposed to NO. This indicates that NorR contains a non-heme iron centre since similar spectra have been observed for several non-heme iron enzymes when exposed to NO (Arciero *et al.* 1983; Brown *et al.* 1995; Ray *et al.* 1999; Hauser *et al.* 2000; Clay *et al.* 2003). This characteristic EPR signal was observed in cells expressing the isolated GAF domain of NorR (GAF_{NorR}) but not in cells that expressed a form of the protein lacking the regulatory domain (NorRΔGAF), indicating that the non-heme iron centre is present within the N-terminal GAF domain. Purification and reconstitution of NorR and GAF_{NorR} with ferrous iron, gave identical NO-spectra to those observed with whole-cell EPR, confirming that the NorR GAF domain contains the non-heme iron centre. Overall, this work showed that the NorR GAF domain regulates the activity of the C domain in response to direct binding of NO at the non-heme iron centre to form a mononitrosyl complex (D'Autreaux *et al.* 2005). This is the first known example of a GAF domain using a transition metal as a mechanism of sensing and reveals a novel biological role for activation of a non-heme iron centre to form a high-spin $\{Fe(NO)\}^7$ ($S = 3/2$) complex. Subsequently, a model for the activation of transcription was proposed
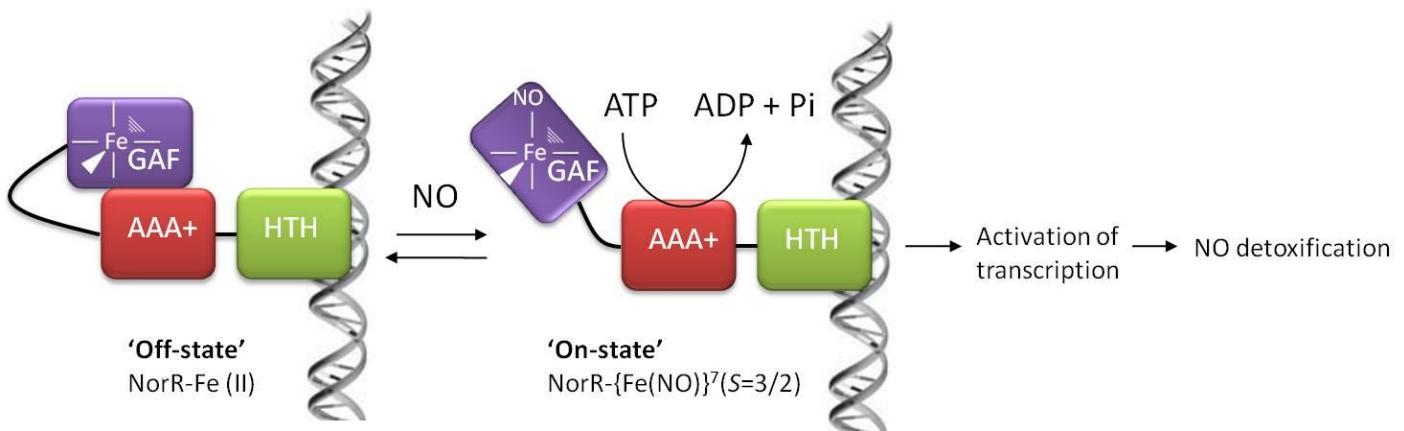
Figure 4.2 - Schematic of the proposed mechanism for transcriptional activation by NorR. For simplicity NorR is represented as a monomer and is shown in complex with DNA. NorR is a tripartite protein that consists of an N-terminal regulatory GAF domain (purple), a central ATPase-active domain (red) and a C-terminal DNA-binding domain (green) that contains a helix-turn-helix (HTH) motif. The binding of NO to the mononuclear non-heme iron centre in the GAF domain leads to the formation of a $\{Fe(NO)\}7$ (S = 3/2) mononitrosyl complex in which the GAF-mediated repression of the AAA+ domain is released. ATP hydrolysis can then occur, allowing expression of *norV* encoding flavorubredoxin. Adapted from (D'Autreaux *et al.* 2005).

(Figure 4.2). The central (AAA+) domain of NorR is intrinsically competent to hydrolyse ATP but in the absence of NO, the regulatory (GAF) domain represses the activity of the central domain to prevent the activation of transcription by NorR. In the presence of NO however, formation of the mononitrosyl $\{Fe(NO)\}^7$ (S =3/2) species triggers a conformational change that relieves the interdomain repression exerted by the GAF domain upon the AAA+ domain. NorR is then able to hydrolyse ATP leading to open complex formation and the expression of the *norVW* genes leading to NO detoxification by the flavorubredoxin.

4.4   The ligand environment of the non-heme centre in the NorR regulatory domain

The spectroscopic features of the NorR paramagnetic mononitrosyl-iron complex suggest that the iron centre has distorted octahedral symmetry and is coordinated by five or six ligands within the GAF domain (D'Autreaux *et al.* 2005). In order to study the coordination of the iron-centre in NorR, targeted mutagenesis was carried out at conserved residues within the regulatory domain (Tucker *et al.* 2007). As a result, five candidate ligands were proposed: D99, D131, C113, R75 and D96 (Figure 4.3). Variant forms of NorR containing substitutions at these positions gave proteins that were unable to bind iron or did not exhibit the characteristic g = 4 EPR signal after reconstitution *in vitro*. Therefore, these residues are likely to have a role in iron coordination. The identification of C113 as a candidate ligand is in accordance with a role in NO-sensing as previously suggested (Gardner *et al.* 2003). The C113, R75, and D131 residues are predicted to be part of a relatively rigid region of the GAF domain whereas D96 and D99 are predicted to be part of a random coil or loop structure. This raises the possibility that these latter residues may play a role in signalling the NO-response via a conformational change within this flexible region. Subsequently, a hexacoordinated model was proposed (Figure 4.3), based on the crystal structure of the GAF-B domain of 3', 5'-cyclic nucleotide phosphodiesterase (Martinez *et al.* 2002; Tucker *et al.* 2007). In the model, D99, D131, C113, R75 and D96 coordinate the Fe centre. Whilst D96 is likely to act as a bidentate ligand, it is also possible that a water molecule provides a sixth ligand. Futhermore, arginine is not an ideal ligand for transition metals but examples have been reported elsewhere such as for biotin synthase (Berkovitch *et al.* 2004). The predicted hexa-coodination of the iron centre suggests that one of the five predicted ligands would need to
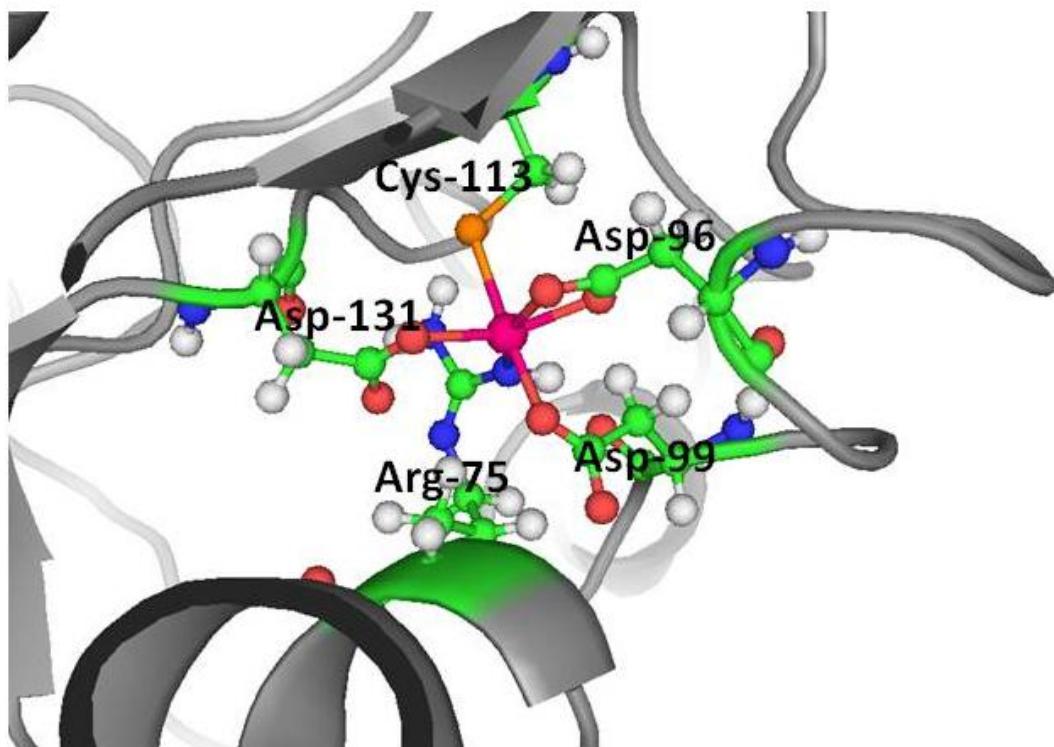
Figure 4.3 - Proposed model of the nitric oxide–sensing non-heme iron centre in the NorR regulatory protein. Structural model of the GAF domain of NorR based on the GAF-B domain of 3',5'-cyclic nucleotide phosphodiesterase PDB ID: 1MC0 (Martinez *et al.* 2002; Tucker *et al.* 2007) showing the iron centre (magenta) and proposed ligands C113, D96, D99, R75 and D131 (Ball and Stick: Carbon = green; Nitrogen = blue; Oxygen = red; Cysteine = orange; Hydrogen; grey) The R75 residue is the most likely to undergo ligand displacement upon NO binding.

be displaced in order to form the mononitrosyl-iron complex. R75 is the most likely candidate to relinquish a binding site for NO and may also stabilise the NO-bound form of the iron through hydrogen bonding (Tucker *et al.* 2007). Interestingly, the R81L and H111L substitutions do not affect the coordination at the non-heme iron centre but do give rise to an escape phenotype of NorR *in vivo*; the R81L and H111L variants are able to activate expression from a *norV-lacZ* promoter fusion both in the presence and absence of endogenously-produced NO (Tucker *et al.* 2007). Since these substitutions do not seem to affect NO-signalling, it is unlikely that the constitutive activity *in vivo* is due to changes in iron coordination or the conformation of the regulatory domain, to mimic the activated state. Rather, it is more likely that the R81L and H111L substitutions influence GAF-AAA+ interactions, suggesting a potential role for the R81 and H111 residues in the mechansism of interdomain repression.

4.5   Role of enhancer-DNA in NorR-dependent activation of transcription

In common with all bEBPs, NorR has been shown to bind enhancer sequences sites 80-150 bp upstream of the verified transcriptional start site (Tucker *et al.* 2004; Tucker *et al.* 2005). Gel retardation assays and DNase I footprinting experiments revealed the presence of three NorR binding sites that were confirmed using methylation protection experiments (Figure 4.4B). Results showed that the binding site closest to the *norV* promoter is protected from methylation at lower concentrations of the NorR protein than the second and third NorR binding sites. The quantified DNA binding data gives rise to a sigmoidal curve suggesting a cooperative binding mechanism (Tucker *et al.* 2004); NorR may have higher affinity for one site (NorR site 1), the occupation of which increases the affinity for the remaining sites (NorR sites 2 and 3). Comparison of the three *E. coli* NorR binding sites showed that sites 1 and 3 consist of the perfect inverted repeat GTCA-(N3)-TGAC, while site 2 consists of the imperfect repeat GTCA-(N3)-CGAC (Tucker *et al.* 2004). Alignment of the *norVW* promoter regions from the proteobacteria *Escherichia coli*, *Salmonella enteric*, *Salmonella typhimurium*, *Shigella flexneri*, *Erwinia carotovora* along with the chromosomal and megaplasmid copies of the *norAB* promoter from *Ralstonia eutropha* enabled the minimal consensus sequence GT(N7)AC to be established for NorR binding (Figure 4.4A). Bioinformatics did not subsequently reveal other predicted NorR binding sites upstream of σ^{54}-dependent promoters in *E. coli*, suggesting that the *norVW* transcriptional unit is the sole target for NorR (Tucker *et al.* 2004). Transcriptomics has revealed further potential targets of the NorR regulon including *ybiJ*
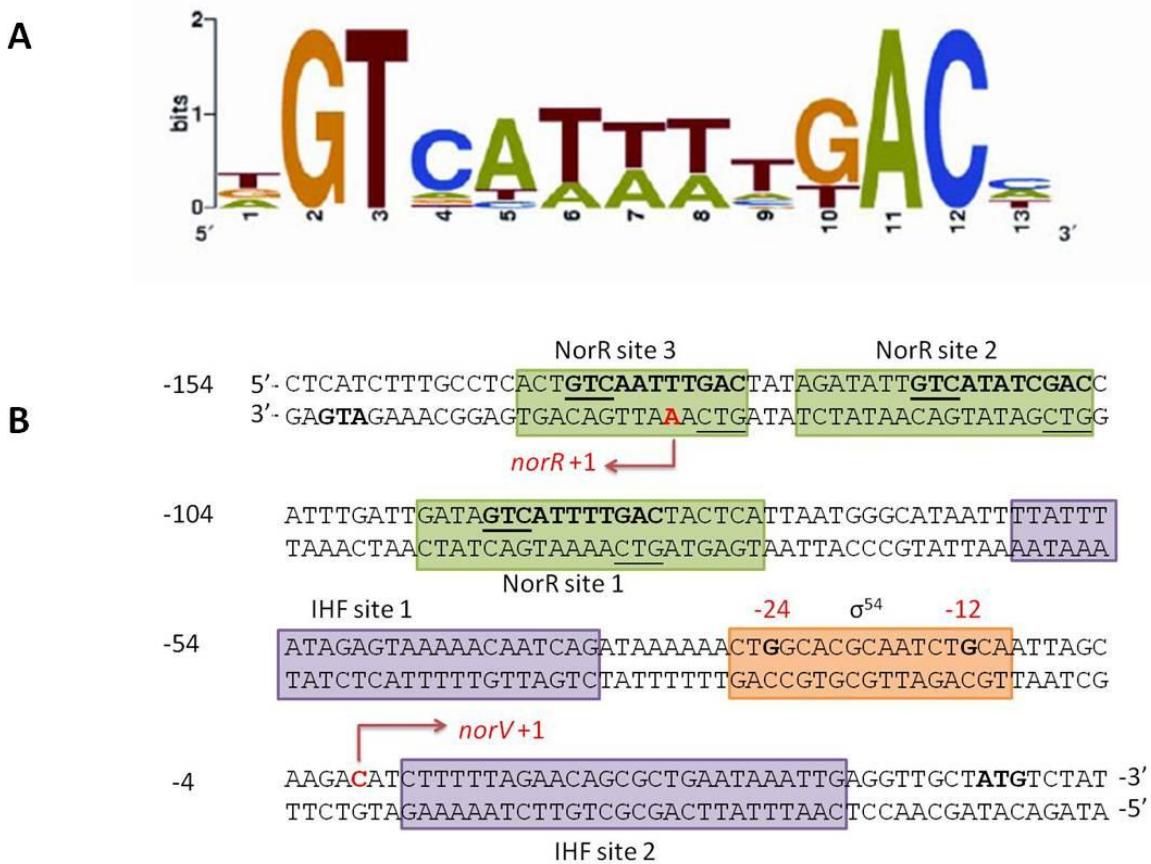
Figure 4.4 – Binding of NorR to conserved binding sites in the *norR-norVW* intergenic region. (**A**) The consensus NorR binding site based on sequences from the following organisms: *E. coli* (*norVW*); *Shigella flexneri* (*norVW*); *Pseudomonas aerugionosa* (*hmp*); *Salmonella Typhimurium* (*norVW*); *Salmonella enterica* (*norVW*); *Erwinia carotovora* (*norVW*); *Ralstonia eutropha* (*norAB*:megaplasmid); *Ralstonia eutropha* (*norAB*:chromosomal). Sequences were input into WebLogo (http://weblogo.berkeley.edu) as described in (Tucker *et al.* 2005). (**B**) The annotated *norR-norVW* intergenic region. The numbering on the left hand side is relative to the conserved -12 and -24 elements (indicated) of the predicted σ^{54} site (boxed in orange). The two regions protected by IHF in DNAase footprinting experiments are shaded in purple and labelled IHF site 1 and site 2 respectively. The three NorR binding regions (as determined by DNAse footprinting) are labelled NorR site 1, 2 and 3 respectively and are shaded in green. Bases protected from methylation or hypermethylated in methylation protection experiments are not indicated. The sequences corresponding to the NorR consensus binding site are in bold with inverted repeats indicated by underlining. Both the *norV* and *norR* transcript start sites are indicated in red whilst the start ATG start codons are in bold (Tucker *et al.* 2004).

118

(Mukhopadhyay *et al.* 2004) but neither NorR binding sites or a predicted σ^{54}-dependent promoter could be identified using bioinformatic analysis. This suggests that if NorR does target the *ybiJ* promoter, it does not do so directly. The identification of a minimal consensus sequence suggests that NorR activates σ^{54}-dependent transcription at promoters upstream of the *hmp* (*fhp*) gene that encodes a putative flavohemoglobin in *Pseudomonas putida*, *Pseudomonas aeruginosa* and *Vibrio cholerae*. In addition, potential NorR sites have been identified upsteam of a *norV*-like gene in *Vibrio vulnificus*. Overall, it seems that NorR may form part of a conserved mechanism among a diverse range of proteobacteria to control the expression of genes encoding enzymes that respond to NO. Indeed three NorR sites were later identified upstream of the *norAB* genes in *R. eutropha* (Busch *et al.* 2004).

A role for IHF in transcriptional activation of the *norVW* genes was confirmed by the identification of two IHF binding sites by DNA footprinting (Figure 4.4B) (Tucker *et al.* 2004). IHF site 1 is located between the promoter proximal NorR site and the σ^{54} promoter, consistent with a role in DNA looping. The location of the second site is intriguing since it is within the *norV* transcribed region. IHF binding sites have been previously identified in the coding sequence of other genes such as *csgD*. In this case, the IHF protein has been suggested to compete with another regulator OmpR in order to modulate the response to microaerophilic conditions (Gerstel *et al.* 2003). Indeed, as nucleoid-associated proteins, FIS and IHF has been shown to compete with other regulators in order to integrate additional signals and form more complex regulatory networks (Browning *et al.* 2000; Browning *et al.* 2004; Browning *et al.* 2005; Squire *et al.* 2009). However, since complete protection of this site was only achieved at very high concentrations of IHF, it is not thought to have a physiological role.

In order to assess the importance of each of the three NorR binding sites in *E. coli*, the enhancers were individually altered from the consensus GT-(N7)-AC to GG-(N7)-CC and introduced upstream of *norV-lacZ* promoter fusions on the *E. coli* chromosome. Disruption of any one of the three sites completely abolished the ability of NorR to activate transcription of *norVW in vivo* (Tucker *et al.* 2010a). Biochemical experiments have demonstrated that the ATPase activity of NorR is dependent not only on the presence of NO, but also on the enhancer DNA that contains the three NorR binding sites (D'Autreaux *et al.* 2005). In the absence of the regulatory GAF domain (NorRΔGAF), the requirement

for the NO-signal is relieved but enhancer DNA is still required to stimulate activity. When any of the three binding sites was individually altered from the consensus, the enhancer-dependent ATPase activity of NorRΔGAF was significantly diminished (Tucker *et al.* 2010a). Efficient open complex formation by NorR *in vitro* also required the three enhancer sites (Tucker *et al.* 2010a). The prerequisite for the three enhancer sites for transcriptional activation by NorR both *in vivo* and *in vitro* might reflect a requirement for three NorR dimers to assemble to form an inactive hexamer on the promoter DNA, given the dyad symmetry present at each of the sites. However, binding of NorR to a 21bp sequence encoding one of the enhancer sites (NorR site 1) stimulated both the ATPase activity and oligomerisation state to a certain extent, indicating that DNA binding *per se* promotes self-association and ATPase activity. Binding to a 66bp DNA fragment that contained all three enhancer sites stimulated ATPase activity and oligomerisation further, although not to the levels observed when a longer 266bp DNA fragment containing the intergenic region was used. This implies that the DNA flanking the enhancer sites has an important role in stabilising the NorR oligomer, possibly by wrapping around the hexamer. In agreement with this, electrophoretic mobility shift assays (EMSA) revealed a significant increase in the affinity and cooperativity of binding when the longer DNA fragment was present. Furthermore, negative-stain electron microscopy revealed the formation of protein-DNA complexes with the expected diameter of a NorR hexamer in the presence of the 266bp DNA, but not with the 66bp or 21bp fragments (Tucker *et al.* 2010a). Overall, this data supports a model in which three NorR dimers bind to the enhancer sites, inducing conformational changes that stimulate formation of a higher order oligomer, most probably a hexamer. The results suggest that this higher order NorR species is stabilised by extensive DNA interactions, possibly by wrapping around the hexamer to form a stable nucleoprotein complex.

4.6   NorR autoregulation

NorR is suggested to autoregulate its own expression due to the overlap of the *norR* transcriptional start site and the third putative NorR binding site (Figure 4.4B). Indeed, in the NorR protein of *R. eutropha*, analysis of a *norR1-lacZ* promoter fusion suggests that the transcription of the *norR1* gene (encoding the NorR1 megaplasmid copy) is negatively regulated by NorR1 (Pohlmann *et al.* 2000). In *E. coli,* the activity of a *norR-lacZ* fusion significantly increased in a *norR* mutant, indicating that the binding of NorR to the *norR-norVW* intergenic region prevents *norR* transcription (Hutchings *et al.* 2002b). In line with

this, DNA binding studies using probes with altered enhancer binding sites showed an *increase* in the activity of a *norR-lacZ* fusion (Tucker 2005). A model of steric hindrance has subsequently been proposed (Tucker 2005) in which NorR binds to the three enhancer sites, preventing binding of RNAP at the -35 and - 10 sites of the *norR* promoter (Figure 4.5). Formation of the higher oligomer might also facilitate distortion of the DNA to further prevent transcription of the *norR* gene. This would ensure that the low level of NorR required for *norV* expression is maintained whilst allowing transcription to occur if the level of NorR is sufficiently limited. Since NorRΔGAF readily binds to DNA *in vitro* and the expression of *norR in vivo* is not affected by NO (Hutchings *et al.* 2002b; Tucker *et al.* 2004), this mechanism of autoregulation is unlikely to require activation of NorR by NO. Similar mechanisms have been proposed for other members of the bEBP family such as the XylR regulator of the TOL plasmid in *Pseudomonas putida* (Bertoni *et al.* 1997).
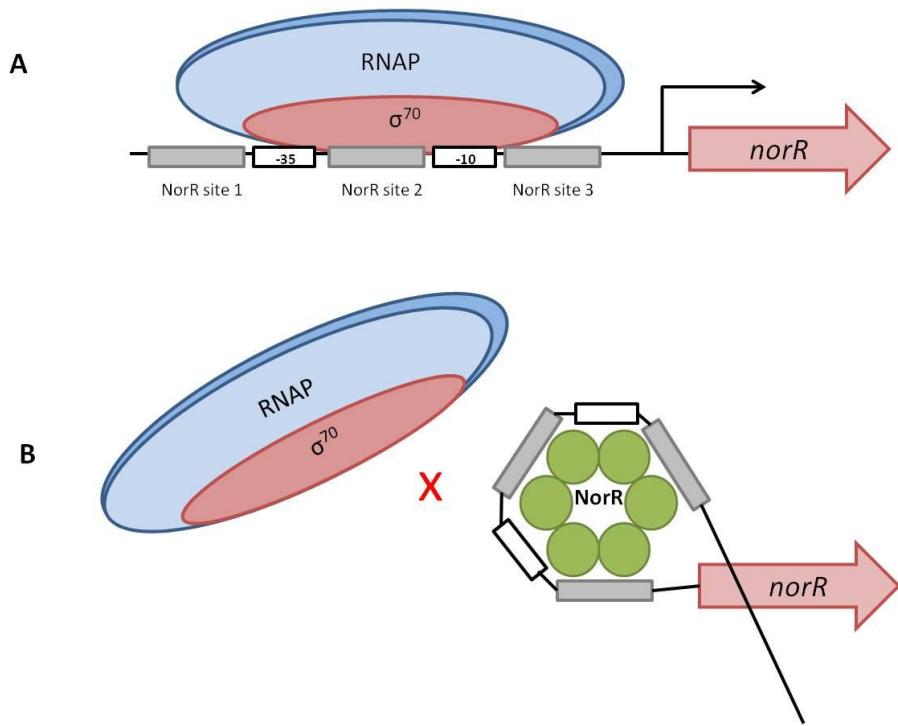
Figure 4.5 – Model of NorR dependent *norR* repression (Tucker 2005). The RNAP β and β' subunits are represented in blue and the σ^{70} subunit is represented as a red oval. The -10 and -35 promoter elements are indicated as white boxes and are labelled. The three NorR binding sites are represented as grey boxes and are also labelled. NorR is represented in hexameric form (green circles) and is labelled. (**A**) When the cellular concentration of NorR is low, the RNAP-σ^{70} holoenzyme is able to bind at the *norR* promoter and activate transcription. (**B**) As the level of NorR accumulates, it forms enhancer/UAS-bound oligomers that result in the wrapping of DNA around the complex. This nucleoprotein-like complex blocks the binding of σ^{70} to the -10 and -35 promoter elements, preventing *norR* expression.

4.7   Aims of this work

It is clear that interdomain repression is of central importance to the function of the bEBP NorR. Recent work has shown that NO is the signal that activates NorR via binding to the non-heme iron centre of the regulatory GAF domain (D'Autreaux *et al.* 2005) and the ligand environment in this domain has been extensively studied (Tucker *et al.* 2007). The importance of enhancer DNA has been shown in the formation of the active oligomer and its subsequent stabilisation, possibly by wrapping around the hexamer. Removal of the N-terminal domain produces a constitutively active form of NorR (Gardner *et al.* 2003), placing it in a sub-group of bEBPs that are regulated via interdomain repression (Shingler 1996). However, it remains unclear how the regulatory domain represses the activity of the central AAA+ domain in the mechanism of negative control in NorR. One possibility is that the regulatory domain targets the oligomeric determinants to prevent AAA+ activity in a manner reminiscent of the bEBPs NtrC1 and DctD (Lee *et al.* 2003; Doucleff *et al.* 2005a). In PspF, negative regulation by PspA is signalled to the nucleotide machinery to prevent ATP hydrolysis (Joly *et al.* 2008a). Alternatively, the NorR GAF domain may repress the activity of the AAA+ domain by another, unknown mechanism. In the absence of a crystal structure, it is not possible to discern the relative orientation of the regulatory and central domains but undoubtedly, an interface exists between them. Disruption of one or more of the points of contact by substitution of residues in either of the domains could therefore result in a constitutively active ("escape") phenotype. In this work it was decided to use random mutagenesis to search for mutations in the AAA+ domain that allow the NorR protein to escape GAF-mediated repression. Identification of substitutions in the AAA+ domain that cause a bypass of negative control and suppressor substitutions in the GAF domain that restore it, might help to characterise the interface of interdomain repression. Subsequent biochemical analysis of such NorR variants may reveal the mechanism by which the regulatory domain controls the activity of the catalytic AAA+ domain. It was hoped that structural studies in collaboration with the Zhang laboratory at Imperial College, London would support this approach.

Chapter 5 - Materials and Methods

5.1   Bacterial strains and plasmids

The bacterial strains and plasmids used in this work are listed in the appendix (section 12.1).

5.2   Buffers and solutions

5.2.1   Media

Liquid and solid media were prepared by dissolving the relevant amounts of reagents listed below in distilled water, followed by autoclaving at 121 ˚C and 15 PSI for 15 minutes. Solid media for plates was prepared by addition of bactoagar to liquid media prior to sterilisation. Agar plates were prepared by pouring approximately 20 ml of molten agar (containing required antibiotics and/or substrates) into each Petri dish and allowing it to set. Plates were then placed in a sterile laminar flow hood for 20 minutes to dry off excess liquid.

(i)   **Liquid media**

LB   1 % (w/v) Tryptone

                          0.5 % (w/v) Yeast extract

                          0.5 % (w/v) NaCl

TGYES   1 % (w/v) Tryptone

                          0.5 % (w/v) Yeast extract

                          1 % (w/v) NaCl

                          0.2 % (w/v) Glucose

MC Buffer   0.1M $MgSO_4$

                          5 mM $CaCl_2$

2 x YT   1.6 % (w/v) Tryptone

                          2.0 % (w/v) Yeast extract

                          0.5 % (w/v) NaCl

(ii)   **Solid media**

LB Plates	1 % (w/v) Tryptone
	0.5 % (w/v) Yeast extract
	0.5 % (w/v) NaCl
	1 % (w/v) agar
R Plates	1 % (w/v) Tryptone
	0.1 % (w/v) yeast extract
	0.8 % (w/v) NaCl
	1.2 % (w/v) agar
Soft Agar	1 % (w/v) Tryptone
	0.5 % (w/v) Yeast extract
	0.5 % (w/v) NaCl
	0.65 % (w/v) agar
	5 mM $CaCl_2$
	10 mM $MgSO_4$

5.2.3   Antibiotics and substrates

Antibiotics or and/or substrates were added to liquid or solid media when required at the following final concentrations:

Carbenicillin (Cb)	100 µg/ml
Kanamycin (Km)	50 µg/ml
Chloramphenicol (Cm)	35 µg/ml
Spectinomycin (Spc)	100 µg/ml
X-gal	40 µg/ml

5.2.4 Buffers for use with Polyacrylamide gels

5 x Loading Dye	0.25 % (w/v) bromophenol blue
	0.25 % (w/v) xylene cyanol ff
	50 % glycerol
Resolving buffer (4x)	1.5 M Tris-HCl (pH 8.8)
	0.4 % (w/v) SDS
Stacking buffer (4x)	0.5 M Tris-HCl (pH 6.8)
	0.4 % (w/v) SDS
Tris-Glycine SDS	192 mM Glycine
Running Buffer	25 mM Tris
	0.1 % (w/v) SDS
SDS-PAGE loading dye	63 mM Tris
	2 % (w/v) SDS
	10 % (v/v) Glycerol
	5 % (v/v) β-Mercaptoethanol
	0.001 % (w/v) Bromophenol blue
SDS-PAGE stain	41.5 % (v/v) Methanol
	16.5 % (v/v) Acetic acid
	0.1 % (w/v) Coomassie blue
SDS-PAGE destain	5 % Methanol
	10 % Acetic acid
Madams Buffer (5x)	25 mM Tris (pH 8.6)
	400 mM Glycine

2 x TAP	100 mM Tris-acetate (pH 7.9)
	200 mM potassium acetate
	16 mM magnesium acetate
	54 mM ammonium acetate
	2 mM dithiothretol
	7 % (w/v) PEG 6000
Acrilamide mix 80:1	40 % (w/v) acrylamide solution
	2 % (w/v) methylenebis*ac*rylamide solution
OPC loading dye	0.1 % (w/v) Xylene cyanol ff
	0.05 % (w/v) Bromophenol blue
	50 % (v/v) Glycerol
	2 mg heparin
Formamide loading dye	95 % Formamide
	20 mM EDTA pH 8.0
	0.1 % (w/v) Bromophenol blue
	0.1 % (w/v) Xylene cyanol ff

Buffers for western blotting

1 x TBS	10 mM Tris-HCl (pH 7.5)
	100 mM NaCl
TBST	20 mM Tris-HCl (pH 7.5)
	500 mM NaCl
	0.05 % (v/v) Triton-20
	0.2 % (v/v) Triton-X-100
Blocking buffer/	
Antibody Buffer	3 % BSA in 1 x TBS

5.2.5 Buffers for use with agarose gels

TBE buffer
135 mM Tris base
45 mM boric acid
2.5 mM Na_2EDTA

Loading dye
0.1 % (w/v) Xylene cyanol ff
0.05 % (w/v) Bromophenol blue
50 % (v/v) Glycerol

5.2.6 Buffers for β-galactosidase assay

Z-Buffer
0.06 M $Na_2HPO_4.7H_2O$
0.04 M $NaH_2PO_4.2H_2O$
0.01 M Kcl
0.001 M $MgSO_4.7H_2O$

Lysis Buffer
0.06 M $Na_2HPO_4.7H_2O$
0.04 M $NaH_2PO_4.2H_2O$
0.01 M Kcl
0.001 M $MgSO_4.7H_2O$
0.27 % (v/v) β-mercaptoethanol
0.005 % (w/v) SDS

O-Nitrophenyl β–D-Galactopuyranoside (ONPG)
4 mg/ml in Z-buffer

Stop Solution
1 M Na_2CO_3

5.2.7 Buffers for protein purification

Buffer A	100 mM Tris-HCl (pH 8.5)
	50 mM NaCl
	5 % Glycerol
Buffer B	100 mM Tris-HCl (pH 8.5)
	1 M NaCl
	5 % Glycerol
Buffer C	100 mM Tris-HCl (pH 8.5)
	200 mM NaCl
	8 mM DTT
	5 % Glycerol
Buffer D	100 mM Tris-HCl (pH 8.5)
	50 mM NaCl
	50 mM Imidazole
	5 % Glycerol
Buffer E	100 mM Tris-HCl (pH 8.5)
	50 mM NaCl
	500 mM Imidazole
	5 % Glycerol
NorR storage buffer	100 mM Tris-HCl (pH 8.5)
	80 % Glycerol

5.3 Microbiological methods

5.3.1 Preparation of competent *E. coli*

A 250 ml conical flask containing 100 ml of LB was inoculated with 500 µl of an overnight culture and grown at 37 $^{\circ}$C with shaking until the optical density at 650 nm (OD_{650}) was approximately 0.4. The cells were then harvested by centrifugation in 50 ml falcon tubes at 4500 RPM for 10 minutes at 4 °C. The cell pellets were gently resuspended in 12.5 ml of ice cold 100 mM $MgCl_2$. This cell suspension was then centrifuged at 4000 RPM for 10 minutes at 4 °C to harvest the washed cells. The cell pellet was then gently resuspended in 25 ml of ice cold $CaCl_2$ and incubated on ice for 20 minutes. Once again, the cells were harvested by centrifugation at 4000 RPM for 10 minutes at 4 °C. The cell pellet was then gently resuspended in 1 ml of ice cold 100 mM $CaCl_2$ and 20 % (v/v) glycerol. The competent cells were then stored in 200 µl aliquots at -80 °C until required.

5.3.2 Transformation of competent *E. coli* for cloning, complementation assays and overexpression.

50 µl of competent cells were added to approximately 0.5 µg plasmid DNA and incubated on ice for 30 minutes. The cells were then heat shocked at 42 °C for 90 seconds, after which they were incubated on ice for 2 minutes. 450 µl of sterile LB broth was added and the cells were incubated at 37 °C for at least 1 hour to allow them to recover. Typically 100 µl aliquots of the transformed cells were then spread onto separate LB agar plates containing the appropriate antibiotic(s). These plates were then incubated at 37 °C overnight.

5.3.4 Electroporation

Electroporation is a technique that is known to increase the transformation efficiency and is useful in forcing the uptake of ligated plasmids. A 1 ml volume of DH5α cells competent for electroporation was prepared by inoculating a 250 ml LB culture with 2.5 ml of overnight culture. This culture was divided into 50 ml volumes and placed at 37 °C until the OD_{600} was between 0.5-0.7. At this point cultures were incubated on ice for 15 mins. The culture was then centrifuged at 4 °C for 10 min at 4000 RPM. The supernatant was carefully discarded and the pellet gently resuspended in 200 ml of cold sterile MilliQ water. This centrifugation and resuspension step was repeated twice more before the pellet was resuspended in 50 ml cold MilliQ water. This was split into two 25 ml volumes before

a final centrifugation at 4 °C for 10 mins at 4000 RPM and resuspension of each pellet in 500 µl of cold MilliQ water.

200 µl of competent cells were added to 3 µl of sample (typically a ligation prepared by butanol precipitation) and a current applied using the Biorad gene pulser system (set to 2.5 Kv/400 Ohm/ 25 µF). 1 ml of LB was added to allow the cells to recover and 5 µl of recovered cells was plated out on LB-agar containing the appropriate antibiotic(s). Commonly, the remaining volume was used to inoculate a 5 ml overnight culture before plasmid purification the next day.

5.3.5   P1 Phage Transduction

P1 bacteriophage can be used to move sections of a bacterial genome from a donor strain to a recipient strain. Briefly, the phage infect the donor strain and randomly package up sections of the bacterial genome. The resulting lysate is then used to infect the recipient strain and the genomic DNA is inserted into the genome via homologous recombination. A selectable marker is then used to isolate a strain with the desired genomic change.

5.3.6   Preparation of P1 lysate

A 5 ml LB culture of the donor strain was set-up (with appropriate antibiotics) and placed shaking at 30 °C overnight. The next morning, this culture was added to 5 ml of MC buffer and left shaking at 37 °C for 30 minutes. Several dilutions of P1 phage were then prepared in LB ($1x10^{-1}$ to $1x10^{-3}$) and 0.1 ml of each phage solution was added to 0.15 ml of bacterial suspension. Solutions (including cell only control) were mixed and incubated at 15 °C without shaking. Each solution was mixed with 3 ml of soft agar in a glass universal (heated to 45 °C) and the contents of each tube poured onto a fresh, thick R plate. Once the R plates had hardened, they were placed at 37 °C for 5-7 hours and plates were periodically checked for partial lysis. To such plates, 2 ml of LB was added and left to stand for 10 minutes. The soft agar was then scrapped off using a disposable spatula and placed in an oakridge tube. 50 µl of chloroform was added per ml of volume, the contents mixed and left for 5 minutes. The tubes were centrifuged at 4 °C for 15 minutes at 12000 RPM. The supernatant was removed and to that, 50 µl of chloroform added per ml of volume and mixed. The resulting P1 lysate was stored in the dark at 4 °C.

5.3.7 P1 Transduction

A 5 ml LB culture of the donor strain was set-up (with appropriate antibiotics) and placed shaking at 30 ℃ overnight. The next morning, this culture was added to 5 ml of MC buffer and left shaking at 37 ℃ for 30 minutes. Three 2 ml volumes of this bacterial suspension were placed into 15 ml falcon tubes and centrifuged for 10 minutes at 8000 RPM. With the first tube, cells were resuspended in 0.2 ml TGYES (control). With the second tube, cells were resuspended in 0.1 ml TGYES and 0.1 ml P1 lysate added along with 10 μl of 0.1M $CaCl_2$. With the third tube, cells were resuspended in 0.1 ml TGYES and 0.1 ml of P1 phage added (non-lysate control). All tubes were incubated for 30 minutes at 37 ℃ without shaking. 5 ml of 1 M sodium citrate was added to each tube before centrifugation at 8000 RPM for 10 minutes. Cells were resupended in 5 ml LB and the centrifugation step repeated. Cells were then resuspended in 5 ml TGYES with 20 mM sodium citrate and left at 37 ℃ with shaking for 3 hours. Centrifugation was carried out at 8000 RPM for 10 minutes and cells then resuspended in 1.5 ml TGYES with 20 mM sodium citrate. Entire volumes were plated out onto LB plates and colonies selected for on the basis of antibiotic resistance. Successful transduction of the desired fragment was then confirmed by colony-PCR.

5.4   Mutagenesis

5.4.1   Site Directed Mutagenesis (Quickchange)

Quickchange is a commercially available mutant strand synthesis mutagenesis method produced by Stratagene. A methylated template plasmid is used to amplify mutated 'daughter' plasmids with DNA polymerase and oligonucleotides containing the desired point mutations. After thermal cycling, the reaction mixture is treated with *Dpn*I to digest only the methylated, non-mutated template DNA. The mutated, circular, nicked, double-stranded DNA is then transformed into ultracompetent cells which are able to repair the nicks in the mutated plasmid.

5.4.2   Site Directed Mutagenesis (Two-step PCR)

In this method three separate PCR reactions are carried out using two sets of oligonucleotide primers (Figure 5.1). External primers, outside of the region of interest are used to amplify the entire DNA fragment (appendix section 12.1.4). The mutagenic primers are complementary to each other and contain the desired point mutation(s) (appendix section 12.1.3). Initially two separate PCR reactions are carried out, using non-mutated DNA as a template. The first uses the forward external primer and the reverse mutagenic primer and the second uses the reverse external primer and the forward mutagenic primer. This produces two separate PCR products which include the desired mutation on opposite strands. Between 50-100 ng of each of these PCR products are used in a final reaction using only the external primers to amplify the entire region of interest producing the mutated PCR product. The Acuzyme system (Bioline) was used for all targeted changes using this protocol.
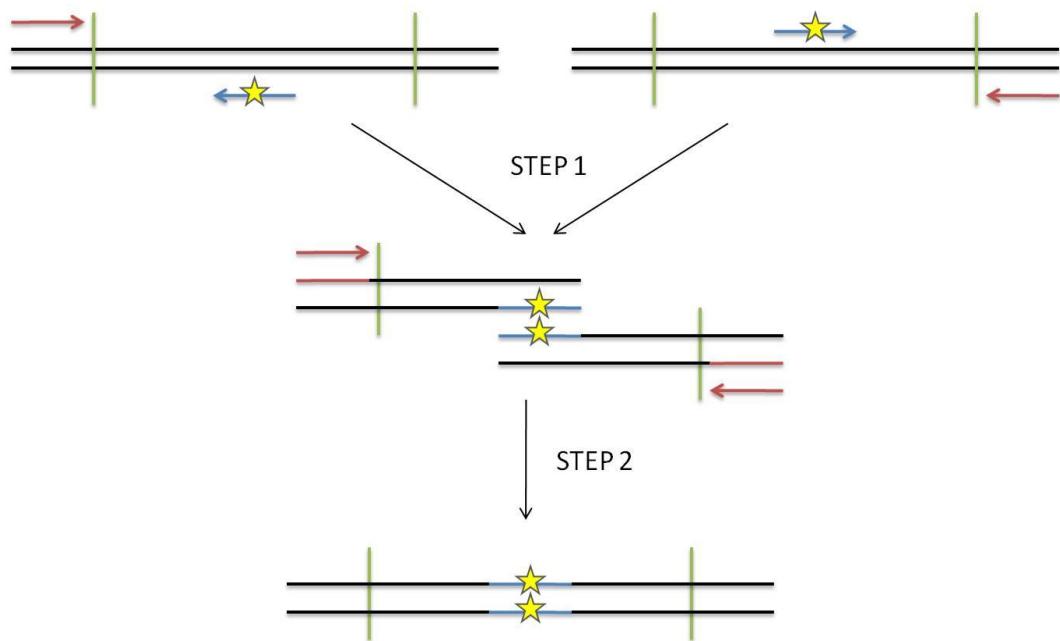
Figure 5.1 - Overview of the two-step PCR method of targeted mutagenesis
In the first step, two PCR reactions using opposite pairs of external primers (red arrows) and mutagenic primers (blue arrows) create two separate products that each include the desired mutation (yellow star). These products are mixed in the second step and the external primers used to complete the region of interest (between green lines). Restriction enzymes can then be employed to cut at appropriate sites outside this region, for easy cloning into a vector.

5.4.3 Error-prone PCR

Error-prone PCR is a technique based on the random incorporation of incorrect bases by a low-fidelity polymerase during a standard PCR reaction. Here, random PCR mutagenesis was carried out with *Taq* DNA polymerase under standard reaction conditions. To increase the frequency of incorporated mutations, a high cycle number was used (35 cycles). Reaction mixtures contained 75 ng of template DNA, 100 ng of each primer, 0.2 mM dNTPs, 1.5 mM $MgCl_2$ and 5 units of enzyme in a final volume of 50 ml. After the PCR reaction, reactions were purified and restriction digestion carried out to re-clone the randomly mutated region into a suitable vector for sequencing and *in vivo* characterisation.

5.5   Plasmids

5.5.1   Purification of plasmid DNA

Small scale preparations of plasmid DNA were carried out from 5 ml overnight cultures using the Qiaquick plasmid miniprep kit (Qiagen) as directed by the manufacturer. Large scale plasmid preparations were performed with Qiagen plasmid midiprep kits from 150 ml of overnight culture. This technique was typically used for preparation of low copy number plasmids and plasmid samples for DNA sequencing.

5.5.2   Plasmid Sequencing

Dye terminator sequencing reactions were carried out using the BigDye Terminator v3.1 kit (Applied Biosciences) and submitted to Genome Enterprise Ltd. as ready reactions for determination of the sequence by capillary electrophoresis and florescence detection. Reactions contained 100-200 ng/µl of plasmid DNA, 10 pmol/µl of primer, 1x BigDye v3.1 and 1x sequencing reaction buffer.

5.5.3   Butanol precipitation

This technique was commonly used to prepare ligated plasmids for electroporation. 1.2 ml of butan-1-ol (room temperature) was added to each sample which was mixed thoroughly before incubation at room temperature for 10 mins. Samples were then centrifuged at 13000 RPM for 4 mins and the supernatant carefully discarded. 1 ml of ice-cold 100 % ethanol was added and the sample vortexed briefly before incubation at -20 ℃ for 20 mins. Samples were again centrifuged at 13000 RPM, the supernatant carefully discarded and 1 ml of ice-cold 70 % ethanol added before vortexing briefly and incubating at -20 ℃ for 10 mins. A final centrifugation at 13000 RPM for 5 mins was carried out, the supernatant

discarded and the pellet vacuum-dried using a speed vac concentrator. The sample pellets were resuspended in 5 µl H$_2$0 and placed on ice until required.

5.6   DNA manipulation Methods

5.6.1   PCR purification

Following PCR, reactions were purified using the QIAquick PCR purification kit (Qiagen). This technique was also employed between two restriction digests to remove enzyme buffer, when a double digest could not be carried out. In addition, PCR purification was often employed following the restriction digestion of PCR products where the desired fragment was retained but unwanted products (<40mer) not. If possible, in order to achieve a high concentration of DNA, it was ensured that the 10 µg column capacity was reached during the PCR purification.

5.6.2   Restriction Digests

Up to 10 µg of DNA was typically digested in a 60 µl reaction, containing the appropriate buffer and enzyme units (as defined by the manufacturer). Bovine serum albumin (BSA) was added to the same volume as the buffer to aid the digestion which was carried out at the optimum temperature for the given enzyme (as defined by the manufacturer), typically 37 °C.

5.6.3   Agarose gel electrophoresis

Agarose gel electrophoresis was used for the size determination and purification of DNA. 1 % (w/v) of agarose was melted in 1 x TBE buffer using a microwave. Ethidium bromide was added to a final concentration of 1 µg/ml and poured into a gel mould. DNA samples were mixed with 5x loading dye alongside commercially available molecular weight markers and run out on the gel at 80 V for between 45 minutes and 1.5 hours, depending on DNA size. Since ethidium bromide intercalates between bases in DNA, visualisation of the different sized fragments was carried out using UV-light.

5.6.4   Gel extraction

DNA separated by electrophoresis was recovered and purified using the QIAquick gel extraction kit (Qiagen). DNA bands, visualised under UV light were removed from the surrounding agarose before purification of the DNA using the gel extraction kit. If possible, in order to achieve a high concentration of DNA, it was ensured that the 10 µg

column capacity was reached. This technique was typically used following restriction digestion to purify the digested fragments.

5.6.5   DNA dephosphorylation

During ligation, the DNA ligase enzyme will only splice together adjacent nucleotides if one nucleotide contains a 5' phosphate and the other contains a 3' hydroxyl group. Recircularisation of the vector can therefore be minimised by removing the 5' phosphate groups using shrimp alkaline phosphatase (SAP). A foreign DNA segment with 5' phosphate groups can then be ligated efficiently into dephosphorylated vector DNA. Samples were incubated with SAP (Roche, 1 unit/µg DNA) and 1 x SAP buffer for one hour at 37 °C and then for 20 minutes at 65°C to deactivate the phosphatase enzyme.

5.6.6   Ligation

T4 DNA ligase is employed to catalyse the formation of a phosphodiester bond between the 5' and 3' ends of DNA molecules, thus joining them together. Ligation reactions were mostly carried out at 16 °C overnight in the buffer supplied by the manufacturer (NEB). Typically a ratio of 5:1 of insert:vector was used with a minimum of 20 ng of dephosphorylated DNA vector in the reaction. Some ligation reactions were carried out using a Rapid DNA Ligation Kit as instructed by the manufacturer (Roche).

5.7   Construction of plasmids

5.7.1   Engineering constructs for determination of *in vivo* NorR activity

To determine the *in vivo* activity of NorR and its mutant derivatives, β-galactosidase assays were carried out in the MH1003 strain of *E.coli*, based on leaky overexpression from the pET21a plasmid. The pNorR plasmid contains the complete *norR* sequence with an *Nde* I site at the N-terminus (overlaps with the start codon ATG) and a *Bam* HI site following the stop codon at the C-terminus.

5.7.2   The pMJB1 plasmid

In order for targetted and random PCR-based mutagenesis techniques to be employed, suitable restriction sites were required to isolate either the GAF or AAA+ domain sequences after amplification before re-cloning into the *norR* sequence. However, there were no suitable restriction sites present that cut either side of the GAF and AAA+ domains and not elsewhere in the pNorR plasmid (pET21a). Therefore, silent mutations

were made either side of the AAA+ domain of NorR to create restriction sites without changing the *norR* coding sequence. Using the "WATCUT" online tool (http://watcut.uwaterloo.ca/watcut/watcut/template.php), the restriction sites for the enzymes *Mfe* I (*Mun* I) and *Sst* II (*Sac* II) were selected. These sites allowed for the isolation of the GAF domain sequence (1-495bp) using the *Nde* I and *Mfe* I sites and for the isolation of the AAA+ domain sequence (496-1341bp) using the *Mfe* I and *Sst* II sites. The location of the silent mutations in the *norR* coding sequence and the resulting restriction sites are shown in Figure 5.2.

5.7.3 Mutagenesis of the pMJB1 plasmid

(i) Random mutagenesis

Random mutagenesis of the AAA+ domain was carried out using the AAA+ Fwd and AAA+ Rev primers (appendix section 12.1.4) to amplify a 960bp region that includes the sequence encoding the AAA+ domain of NorR. The *Mfe* I and *Sst* II restriction enzymes could then be used to re-clone the mutagenic PCR product into the pMJB1 plasmid (cut with the same enzymes) to reconstitute the complete *norR* sequence. Random mutagenesis of the GAF domain was carried out in a similar fashion but instead used the T7long and GAF Rev primers (appendix section 12.1.4) to amplify a 593bp region that includes the GAF-domain encoding sequence. The *Nde* I and *Mfe* I restriction enzymes were then employed for re-cloning into the pMJB1 plasmid.

(ii) Targeted mutagenesis

Targeted mutagenesis using the 2-step PCR method was used to substitute residues in the GAF and AAA+ domains. For AAA+ domain substitutions, AAA+ Fwd and AAA+ Rev primers were used as external primers (appendix section 12.1.4) with the *Mfe* I and *Sst* II restriction enzymes used to re-clone the second step PCR product back into the pMJB1 plasmid. For GAF domain substitutions, T7long and GAF Rev primers (appendix section 12.1.4) were used as external primers with the *Nde* I and *Mfe* I restriction enzymes used to reconstitute the complete *norR* sequence in the pMJB1 plasmid.

```
AGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCT
AGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGAGTTTTTCCGTTGATGTGCTGGCGAAT
ATCGCCATCGAATTGCAGCGTGGGATTGGTCACCAGGATCGTTTTCAGCGCCTGATCACCACGCTACGT
CAGGTGCTGGAGTGCGATGCGTCTGCGTTGCTACGTTACGATTCGCGGCAGTTTATTCCGCTTGCCATC
GACGGTCTGGCAAAGGATGTACTCGGTAGACGCTTTGCGCTGGAAGGGCATCCACGGCTGGAAGCGATT
GCCCGCGCCGGGGATGTGGTGCGCTTTCCCGCAGACAGCGAATTGCCCGATCCCTATGACGGTTTGATT
CCTGGGCAGGAGAGTCTGAAGGTTCACGCCTGCGTTGGTCTGCCATTGTTTGCCGGGCAAAACCTGATC
GGCGCACTGACGCTCGACGGGATGCAGCCCGATCAGTTCGATGTTTTCAGCGACGAAGAGCTACGGCTG
ATTGCTGCGCTGGCGGCGGGAGCGTTAAGCAATGCGTTGCTGATTGAACAACTGGAAAGCCAGAATATG
CTGCCAGGCGATGCCACGCCGTTTGAAGCGGTGAAACAGACGCAGATGATTGGCTTGTCCCCTGGCATG
ACGCAACTGAAAAAAGAGATTGAGATTGTGGCGGCGTCCGATCTCAACGTCCTGATCAGCGGTGAGACT
GGAACCGGTAAGGAGCTGGTGGCGAAAGCGATTCATGAAGCCTCGCCACGGGCGGTGAATCCGCTGGTC
TATCTCAACTGTGCTGCACTGCCGGAAAGTGTGGCGGAAAGTGAGTTGTTCGGGCATGTGAAAGGAGCG
TTTACTGGCGCTATCAGTAATCGCAGCGGGAAGTTTGAAATGGCGGATAACGGCACGCTGTTTCTGGAT
GAGATCGGCGAGTTGTCGTTGGCATTGCAGGCCAAGCTGCTGAGGGTGTTGCAGTATGGCGATATTCAG
CGCGTTGGCGATGACCGTTGTTTGCGGGTCGATGTGCGCGTGCTGGCGGCGACTAACCGCGATTTACGC
GAAGAGGTGCTGGCAGGGCGATTCCGCGCCGATTTGTTTCATCGCCTGAGCGTGTTTCCACTTTCGGTG
CCGCCGCTGCGTGAGCGGGGCGATGATGTCATTCTGCTGGCGGGGTATTTCTGCGAGCAGTGTCGTTTG
CGGCAGGGGCTCTCCCGCGTGGTATTAAGTGCCGGAGCGCGAAATTTACTGCAACACTACAGTTTTCCG
GGAAACGTGCGCGAACTGGAACATGCTATTCATCGGGCGGTAGTTCTGGCGAGAGCCACCCGCAGCGGC
GATGAAGTGATTCTTGAGGCGCAACATTTTGCTTTTCCTGAGGTGACGTTGCCGACGCCAGAAGTGGCG
GCGGTGCCCGTTGTTAAGCAAAACCTGCGTGAAGCGACAGAAGCGTTCCAGCGTGAAACTATTCGTCAG
GCACTGGCACAAAATCATCACAACTGGGCTGCCTGCGCGCGGATGCTGGAAACCGACGTCGCCAACCTG
CATCGGCTGGCGAAACGTCTGGGATTGAAGGATTAAGGATCCGAATTCGAGCTCCGTCGACAAGCTTGC
GGCCGCACTCG
```

Figure 5.2 - Restriction site engineering in the pNorR plasmid.
Engineered restriction sites are highlighted in yellow with the base changes in red. The *Mfe* I/*Mun* I restriction site was created by making the C496T silent base substituion (CAACTG to CAATTG). The *Sst* II/*Sac* II restriction site was created by making the G1341C silent base substituion (CGGCGG to CCGCGG). The *Nde* I (CATATG) and *Bam* HI (GGATCC) sites are highlighted in light blue. The annealing sites for the forward (light green) and reverse (dark green) primers that flank the AAA+ domain coding sequence, used in the error-prone PCR are also shown (AAA+ Fwd and AAA+ Rev).

(iii)   **C-terminal truncation**

Mutagenesis was also carried out to truncate the C-terminal sequence of NorR to produce three ΔHTH constructs with different domain boundaries. Here, a single PCR reaction was performed using pMJB1 as a template and T7long as a forward primer (appendix section 12.1.4). Depending on the truncation required, reverse primers were designed to anneal to the *norR* sequence so that the last annealing base would be the point of truncation. Additionally, the reverse primers encoded a non-annealing stop codon, followed by a *Bam* HI site. Therefore, digestion of the PCR product using the *Nde* I and *Bam* HI enzymes created the C-terminally truncated *norR* sequence for re-cloning into pMJB1.

(iv)   **N-terminal truncation**

In previous work, the N-terminal sequence that encoded the first 170 amino acids of NorR was deleted from the pNorR plasmid to create the pNorRΔGAF plasmid. In this work, in order for the *in vivo* activity of N-terminally truncated NorR AAA+ variants to be assessed, the substitutions in the AAA+ domain were made using a targeted approach with pNorRΔGAF as a template in the PCR-based method. Since pNorRΔGAF does not contain the *Mfe* I or *Sst* II sites, the T7long and T7term external primers were used to amplify a region of the *norR* sequence containing the *Nde* I and *Bam* HI sites.

5.7.4   Overexpression

Since pMJB1 and its mutant derivatives are overexpression vectors based on pET21a, these plasmids were also used to overexpress and purify native NorR and its variants. However, in order to purify NorR with an N-terminal His-tag, the wild-type and mutant *norR* sequences were moved into the pETNdeM11 vector. This vector is based on the pETM11 vector (EMBL) and encodes an N-terminal, hexahistidine, TEV protease cleavable tag. In the pETNdeM11 vector, the *Nco* I site has been altered to an N*de* I site by silent mutation to allow easy cloning of the NorR sequence. The *Nde* I and *Bam* HI restriction enzymes were used to move the entire *norR* sequence from the pMJB1 plasmid into the multiple cloning site (MCS) of the pETNdeIM11 vector.

5.7.5   Mutagenesis of the pETNdeM11 plasmid

In order to assess the activity of NorR-His and its mutant derivatives *in vivo*, the antibiotic resistance-cassette used to maintain the *norR*-derivatives of the pETNdeM11 plasmid had to be changed. This was because the *E. coli* strain MH1003 is also maintained by the kanamycin resistance gene. Therefore the 2.0 Kb omega (Ω) cassette was removed from the 4.3 Kb pHP45Ω plasmid (Prentki and Krisch 1984) using the *Sma* I restriction enzyme (Figure 5.3). The 2.0 Kb insert was gel-purified away from the 2.3 Kb *Sma* I-digested pHP45 and then ligated into *Sma* I-digested pETNdeM11 that encoded the *norR* sequence of interest. This disrupted the kanamycin resistance gene whilst simultaneously inserting the sequence that encodes streptomycin (Sm)/ spectinomycin (Spc) resistance.

5.7.6   Mutagenesis of the pNPTprom plasmid

The pNPTprom plasmid is derived from the pUC19 cloning vector and encodes the 361bp fragment of the *norR-norVW* intergenic region between *Eco* RI and *Bam* HI sites. This DNA contains each of the 3 NorR-binding sites and was used in DNA-binding as well as Open Promoter Complex (OPC) EMSA assays. In order to study the functionality of NorR and its variants in the absence of the NorR binding sites, a 66bp sequence of the intergenic region was deleted. In order to remove the NorR-binding region, PCR-mediated deletion mutagenesis was employed (Lee *et al.* 2004). Two mutagenic primers were designed (appendix section 12.1.3). The forward mutagenic primer (pNPTprpm 66bp_DEL Fwd) anneals immediately downstream of the 66bp region to be deleted (Figure 5.4: shown in blue). At the 5' end there is a 9nt non-annealing sequence that is complementary to the region immediately upstream of the 66bp region (Figure 5.4: shown in green). The reverse mutagenic primer (pNPTprpm 66bp_DEL Rev) anneals immediately upstream of the 66bp region to be deleted. At the 5' end there is a 9nt non-annealing sequence that is complementary to the region immediately downstream of the 66bp region (Figure 5.4: shown in orange). In the first step, two PCR reactions were carried out; one with the M13 Fwd external primer and the reverse mutagenic primer and another with the M13 Rev primer and the forward mutagenic primer. This produced two partially complementary PCR products neither of which contained the 66bp sequence. In the second step, external primers were used to complete the 295bp intergenic region that is deleted for the NorR-binding region. Upstream (*Eco* RI) and downstream (*Bam* HI)
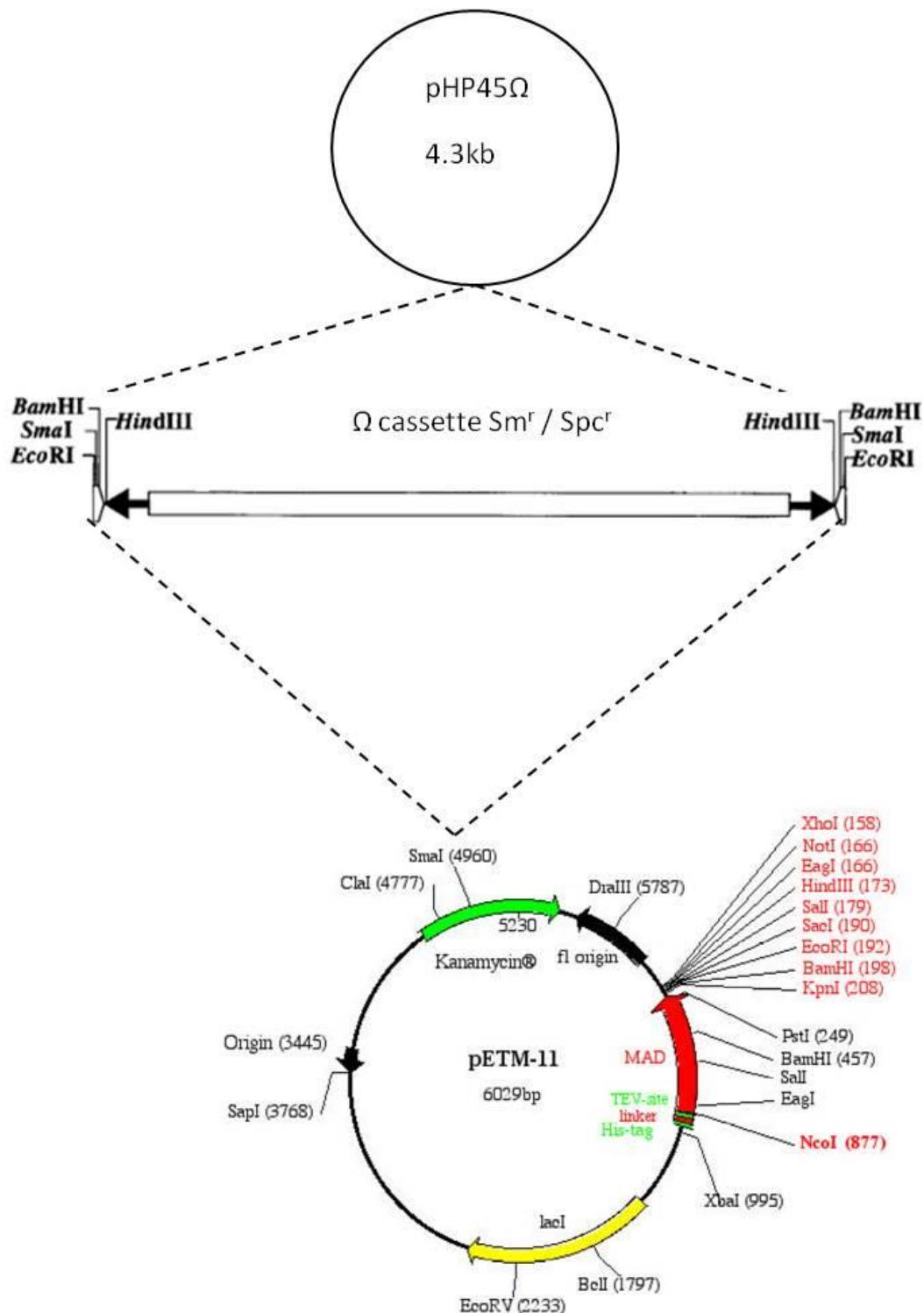
Figure 5.3 – Mutagenesis of the pETNdeM11 plasmid. To change the antibiotic resistance of pETNdeM11 from kanamycin (Km) to streptomycin (Sm)/ spectinomycin (Spc), the 2.0 kb omega (Ω) cassette was removed from the pHP45Ω plasmid (Prentki and Krisch 1984) using *Sma* I and then ligated into the *Sma* I site of the pETNdeM11 vector that already encoded the *norR* sequence of interest. The wild-type and mutant versions of *norR* could not be cloned into a pETNdeM11/Ω vector, since the Ω cassette contains multiple *Bam* HI sites. Figure adapted from pETM-11 vector map (http://www.embl-hamburg.de/~geerlof/webPP/vectordb/bact_vectors/).
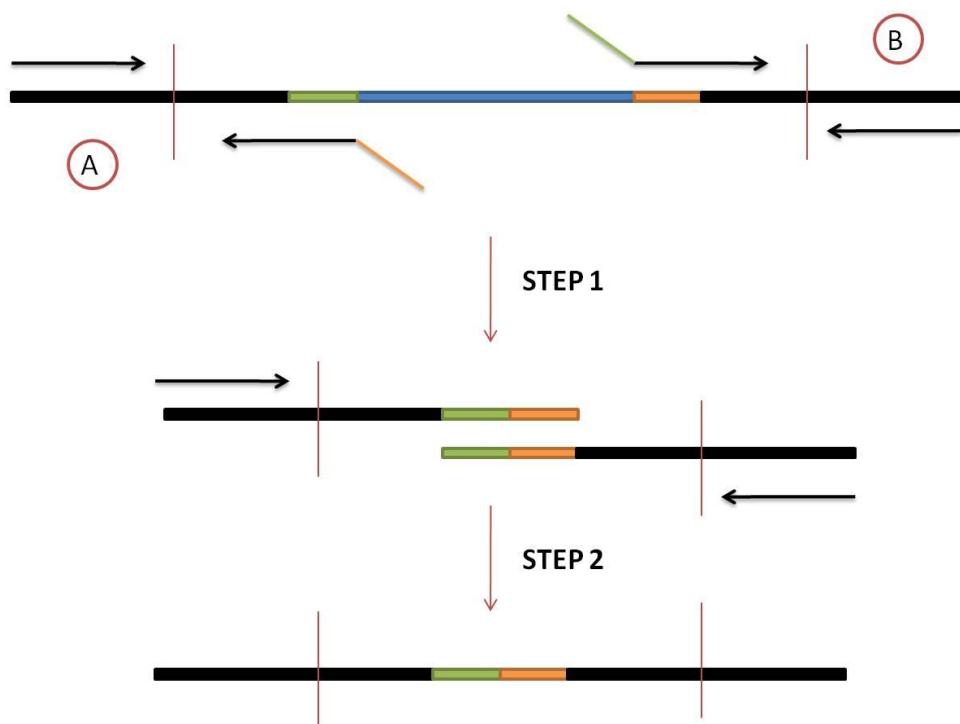
Figure 5.4 - PCR-mediated deletion mutagenesis (Lee *et al.* 2004). Two mutagenic primers are designed. The forward mutagenic primer anneals immediately downstream of the region to be deleted (blue). At the 5' end there is a 9nt non-annealing sequence that is complementary to the region immediately upstream of the region to be removed (shown in green). The reverse mutagenic primer anneals immediately upstream of the region to be deleted. At the 5' end there is a 9nt non-annealing sequence that is complementary to the region immediately downstream of the region to be removed (shown in orange). In the first step, two PCR reactions are carried out; one with the forward external primer and the reverse mutagenic primer (A) and another with the reverse external primer and the forward mutagenic primer (B). This produces two partially complementary PCR products neither of which contains the region to be deleted. In the second step, external primers are used to complete the product that now lacks the undesired sequence. Upstream and downstream restrictions sites can then be used to re-clone the new fragment into a suitable vector.

restriction sites were then used to re-clone the new fragment into the pNPTprom to give the pNPTprom2 plasmid (appendix section 12.1.2).

5.8   Protein methods

5.8.1   Overexpression of NorR

Expression of NorR was carried out using either the pET21a vector or the pETNdeM11 vector encoding native and N-terminal His-tagged NorR respectively. The appropriate plasmid was transformed into BL21(DE3) (appendix section 12.1.1) and colonies picked into 5 ml cultures that were left shaking at 37 ˚C during the day. 0.75 ml of this culture was used to inoculate a 50 ml overnight culture containing appropriate antibiotics which was then left shaking at 37 ˚C. The next morning, 10 ml of overnight culture was used to inoculate 1 l LB cultures containing antibiotic which were then left shaking at 37 ˚C until the OD_{600} was approximately 0.6. IPTG was then added to a final concentration of 1 mM and the culture left shaking at 37 ˚C for a further three hours. After 2-3 hours, cells were harvested by centrifuging at 5000 RPM for 10 minutes. Pellets were resuspended in 30 ml of breaking buffer in to which three mini EDTA-free protease inhibitor tablets were dissolved (Roche). The resuspension was stored at -80 ˚C until required. Both pre- and post- induction samples were removed to check for successful overexpression. 1 ml of culture was centrifuged at 14000 RPM for one minute. 900 μl was then discarded and to the remaining 100 μl, 100 μl 2x SDS-PAGE loading dye was added. Samples were then heated at 100 ˚C for five minutes before being loaded onto a 12.5 % polyacrylamide gel.

5.8.2   SDS polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a technique widely used to separate proteins by size. The resolving gel (12.5 % polyacrylamide) was prepared by mixing 12.5 ml 30 % acrylamide, 7.5 ml 4x resolving buffer, 9.55 ml of water and 0.3 ml 10 % SDS. 150 μl of 10 % ammonium persulfate (APS) and 10 μl of TEMED were added and the mixture poured to approximately 5 mm of the top of the gel mould. The stacking gel was prepared by mixing 2.5 ml of 4 x stacking buffer with 6.1 ml of distilled water and 1.33 ml of 30 % acrylamide solution. 50 μl 10 % APS and 5 μl TEMED was added, the mixture poured into the gel mould and the comb inserted until the gel had set.

An equal volume of 2x SDS-PAGE loading dye was added to the protein samples, which were then heated at 100 ˚C for 5 minutes before gel-loading. Samples were loaded using a

Hamilton syringe and run at approximately 150 V and 20 mA for 1 hour (ATTO corp.) in 1x Tris Glycine SDS buffer (Severn Biotech). Staining in approximately 20 ml of SDS-PAGE stain for 15 minutes, followed by destaining in 40 ml of SDS-PAGE destain overnight, enabled visualisation of the protein bands. Alternatively, Instant Blue (Expedeon) was used to enable rapid detection and visualisation of the protein bands.

5.8.3 Western Blotting

Western blotting is an analytical technique used to detect specific proteins and was employed to determine the stability of protein constructs when expressed in the β-galactosidase assay. Under denaturing conditions, it uses gel electrophoresis to separate proteins on the basis of size. The proteins are then transferred to a nitrocellulose membrane where they are detected by an antibody specific to the protein.

Before loading, the OD_{600} of cultures was measured and appropriate volumes taken to standardise them. Cells were then span down for 1 minute at 14000 RPM and the pellet resuspended in 100 μl LB and 100 μl 2x SDS-PAGE dye. Samples were boiled for 5 minutes and then loaded on a 12.5 % polyacrylamide SDS-PAGE gel. After electrophoresis, blotting was carried out using the Invitrogen XCell II Blot system. A piece of nitrocellulose membrane (pre-soaked in 1 x transfer buffer) was placed carefully onto the protein gel, in between two pieces of pre-soaked filter paper. Transfer was carried out at 30 V for 1 hour in 1 x transfer buffer, surrounded with dH_2O to ensure the assembly remained cool. After blotting, the membrane was transferred to a clean, square, plastic dish and two 10 minute washes with 1 x TBS carried out with gentle shaking at RT. Blocking of non-specific epitopes was carried out in 3 % BSA for 1 hour, shaking at RT or alternatively at 4 °C overnight, without shaking.

(i) **Preparation of primary antibody**

Anti-NorR antibody was generated against rabbits and the crude serum was aliquoted and stored at -20 °C. To improve the specificity of the antibody serum, it was cross-reacted with a cell extract of a strain that does not contain NorR (MH1003, appendix section 12.1.1). To prepare the cell extract, a 5 ml overnight MH1003 culture was centrifuged at 4000 RPM for 10 minutes and the pellet resuspended in 500 μl SP buffer. The cell resuspension was kept on ice and sonicated 5 times at an amplitude of 10 microns (15 seconds on and 15 seconds off). The lysed sample was then span down at 4000 RPM at 4

°C for 10 minutes. The resulting cell extract was placed on ice. The crude anti-NorR serum was mixed with the cell extract in a 1:1 ratio and incubated at 30 °C for 1 hour. It was then span down at 14000 RPM for 1 minute to remove precipitation and the supernatant (the cross-reacted primary antibody) stored at -20 °C until required.

(ii) **Washing and detection**

After blocking, the membrane was washed twice for 10 minutes in TBST. The cross-reacted NorR antibody was diluted 1 in 2500 in 3 % BSA, added to the membrane and then left shaking for 1 hour at RT. After treatment with primary antibody, the membrane was washed twice for 10 minutes in TBST and then for 10 minutes in 1 x TBS. An anti-rabbit IgG alkaline phosphatase secondary antibody (sigma) was diluted 1 in 5000 in 3 % BSA and added to the membrane. After shaking for 1 hour at RT, the membrane was washed four times for 10 minutes in TBST.

Secondary antibody detection was carried out using the SigmaFast system (Sigma). An BCIP/NBT (5-Bromo-4-chloro-3-indolyl phosphate/Nitro blue tetrazolium) tablet was crushed and dissolved in 10 ml of double distilled water and the solution added to the membrane. Once sufficient exposure had been achieved, the reaction was quenched using water and the membrane dried between two sheets of blotting paper.

5.8.4   Protein Purification

Cell resuspensions, produced from the overexpression of typically 3 litres of culture, were thawed in tepid water and once completely thawed, the cells were placed on ice. Cells were broken by French pressure disruption (1000 PSI) in two passes. The insoluble material was then removed by centrifugation in oakridge tubes at 15000 RPM for 30 minutes in the SS34 rotor (Sorvall Rc5B Plus centrifuge). The supernatant was then decanted into a clean falcon tube and placed on ice.

(i)   Purification of native NorR - Affinity Chromatography

For native protein preparations, NorR was purified using a 5 ml Hi-Trap Heparin HP column (Amersham Biosciences) equilibrated with buffer A using an Akta protein purifier. The clarified supernatant was loaded before NorR was eluted using buffer B. Samples for SDS-PAGE were taken to determine which fractions contained NorR.

(ii)   Purification of His-tagged NorR - Affinity Chromatography

This technique is a form of immobilised metal ion chromatography (IMAC) and is based on the ability of nickel to specifically interact with amino acids such as histidine. His-tagged proteins therefore bind strongly to the column, reducing impurities and are eluted using imidazole, which breaks the nickel - histidine interaction. For His-tagged protein preparations, two 1 ml Hi-Trap Ni-chelation columns (Amersham Biosciences) were connected in series and 3 ml of 100 mM $NiCl_2$ passed through the columns. The columns were washed with 2 ml of water and attached to the chromatograph (Akta protein purifier). The column was equilibrated using buffer D and the clarified supernatant loaded. NorR-His was eluted using buffer E. 1 ml fractions were collected and samples taken for SDS-PAGE, in order to determine the NorR-containing fractions.

(iii)   Purification of native and His-tagged NorR – Gel filtration

In both the native and His-tagged purification protocols, NorR shows a high tendency to precipitate after affinity chromatography, particularly at high concentrations. Therefore 2-4 ml of pooled NorR fractions was loaded onto a Superdex 200 16/60 column (Amersham Biosciences), pre-equilibrated in buffer C and the flow rate set at 1.0 ml/min. Depending on the oligomeric state, NorR eluted anything between 45 ml (the void volume) and 100 ml. NorR-containing fractions were again identified by SDS-PAGE. Prior to storage, protein

could be concentrated using an Amicon Ultra spin column (Millipore) with appropriate molecular weight cut-off. An equal volume of NorR storage buffer was added to selected fractions or concentrated samples. Aliquots were made and stored at -80 °C until required. In subsequent biochemical experiments, each aliquot was only used once to avoid repeated freeze – thaw which has been known to compromise NorR activity *in vitro*.

5.8.5 Bradford assay for protein concentration

This assay is widely used for protein concentration determination (Bradford 1976). Bradford reagent and BSA standards were supplied by Pierce and the assays were carried out according to the manufacturer's instructions. Rough estimates of protein concentration were carried out using the Nanodrop instrument.

5.8.6 ATPase activity assays

ATPase activity was measured using an assay in which production of ADP is coupled to the oxidation of NADH by lactate dehydrogenase and pyruvate kinase (Norby, 1988). The oxidation of NADH was monitored at 340 nm at 37°C using the PerkinElmer Lambda 35 UV/VIS Spectrometer. A master mix was prepared which contained 0.35 mM NADH, 3 mM phosphoenolpyruvate, 0.075 mg lactate dehydrogenase, 0.15 mg pyruvate kinase, 6 mM ATP and 4 mM $MgCl_2$ in 88.6 mM Tris-HCl (pH 8.5), 177.2 mM NaCl and 4.4 % glycerol. Reaction volumes were typically 0.5 ml with protein added to the master mix at the concentrations required. Reactions were carried out both in the absence and presence of 5 nM of a 266bp fragment of the *norR-norVW* intergenic region generated from the pNPTprom plasmid (Tucker *et al.*, 2004) using the NorRprom Fwd and NorRprom Rev primers (appendix section 12.1.3). ATPase activity was measured by observing the change in absorbance at 340 nm (ΔA). ΔA is calculated using the following equation: $\Delta A = $ (OD340 nm start - OD340 nm end) / (time end – time start). The total ATPase activity was then calculated using the equation: μmol ATP/min $= (\Delta A / 6220) \times 1 \times 10^6$.

5.8.7 Gel Retardation Assays (Electophoretic Mobility Shift Assay, EMSA)

EMSA is a common technique used to study protein-DNA interactions and can determine whether a protein is capable of binding a specific DNA sequence *in vitro*. The technique is often performed concurrently with footprinting or when studying transcription initiation.

(i)   **DNA binding assay**

(1)   Labelling of *norR-norV* intergenic region constructs

Gel shift assays were carried out to study the binding of NorR to either 361bp or 266bp fragments of the *norR-norVW* intergenic region *in vitro*. In order to study the binding to a 361bp fragment of the intergenic region, 5 μg of pNPTprom that contains the *norR-norVW* region blunt-end cloned into the SmaI site of pUC19 (Tucker *et al.* 2004) was digested using the *Eco* RI and *Bam HI* restriction enzymes. In addition, binding to a shorter 266bp product created by PCR using the NorRprom Fwd and Rev primers was studied. Measuring the binding to the *nifH* promoter derived from the pNH8 plasmid served as a negative control. Following restriction digestion or PCR, the sample was passed over a Qiagen PCR purification column to remove enzymes and buffers before Shrimp Alkaline Phosphotase (SAP) treatment (to remove the 5' phosphate groups). The intergenic region was double gel-purified using the Qiagen gel-extraction kit and finally eluted into a volume of 30 μl. To this, 2 μl of T4 polynucleotide kinase (Epicentre) was added along with 10 μl of 5x T4 polynucleotide kinase buffer, 2 μl [γ-32P] ATP (specific activity 10 μCi / μl) and distilled water to produce a final volume of 50 μl. This reaction was incubated at 37 ℃ for 15 minutes before inactivation of the enzyme at 65 ℃ for 15 minutes. 25 μl of mixture was then passed over a column containing sephadex G-50, span at 12000 RPM for three minutes to remove unincorporated nucleotide. The resulting labelled-DNA fragment was stored at -20 ℃ until required.

(2)   DNA binding and separation of retarded species

[32]P-labelled DNA fragments were diluted with distilled water to a volume to give roughly 30 counts per second. Diluted DNA was then added to 2x TAP buffer in a 1:5 ratio. 6 μl volumes of this mastermix were then added to 4 μl samples of purified NorR protein. Binding reactions were then incubated at 30 ℃ for 20 minutes before the addition of 3 μl OPC dye. Samples were loaded onto 4.2 % polyacrylamide gels made in MADAM's buffer and the gel was resolved at ~30 mA until the bromophenol blue dye front approached the bottom of the gel. Gels were then dried onto Whatman filter paper and exposed to film or quantified using a Fuji BAS 1000 Phosphorimager.

(ii)   **Open Promoter Complex assay**

Open complex assays were carried out using the 361bp fragment of the *norR-norVW* intergenic region, derived from the pNPTprom plasmid and ^{32}P-end labelled as in the gel

retardation assays. Open complex formation was assayed in TAP buffer as in the band-shift assays and contained 1 nM template DNA, 200 nM core RNA polymerase (Epicentre Biotechnologies), 200 nM σ^{54}-His, 130 nM IHF-His, 5 mM ATP (Pharmacia Biotech), 0.5 mM CTP (Pharmacia Biotech), 0.5 μg denatured salmon sperm DNA (Sigma Aldrich) and 0.15 mg BSA (NEB). The reaction components were pre-incubated for 10 min at 30 °C and reactions initiated by adding NorR/variants to a final concentration of 1 μM. After a further 20 min incubation at 30 °C, samples were mixed with 3 ml of dye mixture containing 50 % glycerol, 0.05 % bromophenol blue, 0.1 % xylene cyanol and 2 mg of heparin and the heparin resistant open complex resolved on a 4 % polyacrylamide gel as for the DNA-binding assays.

(iii)   Potassium permanganate footprinting of open complex
(1)   5' labelling of *norR-norV* intergenic region construct

Potassium permanganate targets single-stranded regions of DNA for cleavage and therefore can be used following open complex formation, using an EMSA approach to visualise the activation of transcription *in vitro* in a sequence specific manner. In this method, open complex formation and subsequent cleavage was carried out using a 266bp fragment of the *norR-norVW* intergenic region amplified from the pNPTprom plasmid using the norRprom Fwd and Rev primers (appendix section 12.1.3). The reverse primer was first 5'-end labelled with γ-^{32}P ATP using T4 polynucleotide kinase (Epicentre). PCR amplification from the pNPTprom template using the ^{32}P-labelled reverse primer and "cold" forward primer ensured that the 266bp product was single-end labelled. Following PCR, the DNA was purified using the Qiagen PCR purification kit to remove all buffers and unincorporated nucleotide. ^{32}P-labelled DNA was diluted to give 400-500 cps per reaction.

(2)   Preparation of G+A sequencing ladder

G+A sequencing markers were used to identify regions of enhanced cleavage following potassium permanganate cleavage of open complexes (Maxam and Gilbert 1977). 20 ng of 5' end labelled intergenic DNA was mixed with 2 μg of salmon sperm DNA and 1 μl of 4 % formic acid in a final volume of 20 μl and incubated at 37 °C for 25 minutes. 150 μl of 1 M piperidine was then added to the mixture which was incubated at 90 °C for 30 minutes. Samples were left on ice for 5 mins before addition of 1 ml butan-1-ol. After centrifugation, pellets were resuspended in 150 μl of 1 % SDS and a further 1 ml butan-1-

ol added. Two further precipitations/resupensions were then carried out using 0.5 ml butan-1-ol. The resulting pellets were vacuum dried and 25 µl of loading dye added (95 % Formamide, 20 mM EDTA pH 8.0, 0.1 % bromophenol Blue, 0.1 % Xylene Cyanol). 2.5 µl (200-300 cps) was loaded on each sequencing gel.

(3)   Potassium permanganate footprinting

Open complex reactions were carried out as described above. After the 20 min incubation step at 30 °C, 1 µl of 200 mM potassium permanganate was added to each reaction and samples incubated at RT for 4 mins. Reactions were stopped by adding 60 µl of stop solution (1.5 M β-mecaptoethanol, 0.3 M Na acetate, 0.1 mM EDTA, 2 mg/ml glycogen) and 750 µl ice-cold EtOH added before incubation on ice for 5 mins. Samples were centrifuged and the pellets resuspended in 300 µl Na acetate before the addition of 900 µl EtOH to each sample. Samples were again centrifuged and the pellets resuspended in 1 ml EtOH. The centrifugation step was repeated and the pellet vacuum dried before addition of 100 µl 1M piperidine and incubation at 95 °C for 30 mins. Samples were left on ice for 5 mins before addition of 1 ml butan-1-ol. After centrifugation, pellets were resuspended in 150 µl of 1 % SDS and a further 1 ml butan-1-ol added. Two further precipitations/resupensions were then carried out using 0.5 ml butan-1-ol. The resulting pellets were vacuum dried and 20 µl of loading dye added (95 % Formamide, 20mM EDTA pH 8.0, 0.1 % bromophenol Blue, 0.1 % Xylene Cyanol). Samples were denatured by heating at 90 °C for 3 mins and 10 µl loaded (200-250 cps) on a 6 % (w/v) polyacrylamide sequencing gel (acrylamide/bisacrylamide ratio, 19: 1) in 1x TBE, which had been pre-run until the temperature of the gel was 50 °C. Gels were run at 55 W and were dried and exposed to autoradiograph film or a phosphorimager screen.

5.9 Assaying NorR activity *in vivo*

5.9.1 Culture conditions

Transcriptional activation by NorR *in vivo* was measured by introducing wild-type and mutant plasmids into MH1003, a *norR*::*cat* derivative of *E. coli* strain MC1000 with a *lacZ* reporter fusion to the *norVW* promoter inserted at the phage λ attachment site (appendix section 12.1.1) (D'Autreaux *et al.* 2005). Cultures were grown with shaking in 50 ml of LB medium at 37 °C until the OD_{650} reached 0.2-0.3, at which point glucose was added to the culture to a final concentration of 1 %. Cultures were then split into 8   ml Bijou bottles and were grown anaerobically overnight at 37 °C with or without potassium nitrite (4 mM). Under these conditions, NorR is activated by the NO that is generated endogenously by nitrite reduction in *E. coli* (D'Autreaux *et al.* 2005). The next morning, β-galactosidase assays were carried out.

5.9.2 β-galactosidase assay

β-galactosidase is able to hydrolyse o-nitrophenyl-β-D-galactopyraniside (ONPG), releasing o-nitrophenol which has a yellow colour detectable at 420 nm. Here, ONPG was used to monitor gene expression from the *norV-lacZ* promoter fusion that is activated by the *norR* transgene. Initially the OD_{600} of the overnight cultures was measured and recorded. 30 μl of each culture was added to 970 μl of Lysis buffer in appropriately marked glass tubes. 20 μl of chloroform was then added to each tube. Tubes were vortexed for 10 seconds and incubated at 28 ˚C for five minutes. Each reaction was initiated by the addition of 200 μl of ONPG, at which point a timer was started. When a yellow colour had developed, the assay was quenched using 500 μl of 1 M sodium carbonate. The time course of the reaction was then recorded. 300 μl of the quenched samples was transferred to a 96-well microtitre plate and the OD_{420} and OD_{550} measured and recorded using a BioTek plate reader. In each case, the β-galactosidase activity (Miller units) was calculated using the equation: $(1000*OD_{420} - (1.75*OD_{550})) / (t*0.03*OD_{600})$; where t is the time in minutes.

Chapter 6 - Error-prone PCR mutagenesis of the AAA+ domain in NorR

As a member of the AAA+ family of proteins, NorR uses the energy derived from ATP hydrolysis to activate transcription. However, the regulatory (GAF) domain ensures that the activity of the AAA+ domain is repressed in the absence of the signalling molecule, NO (D'Autreaux *et al.* 2005). Since this regulation is absent in a variant form of NorR that lacks the N-terminal domain (Pohlmann *et al.* 2000; Gardner *et al.* 2003), direct contact(s) between residues of the GAF and AAA+ domains are likely to mediate this mechanism of interdomain repression. To explore the interface between the N-terminal regulatory (GAF) domain and the central ATPase (AAA+) domain of NorR, PCR mutagenesis of the coding sequence of the AAA+ domain was employed to create mutations that disrupt the mechanism of interdomain repression (Figure 6.3A). Error-prone mutagenesis exploits the random incorporation of incorrect nucleotides by a low-fidelity polymerase during a standard PCR reaction. In order to employ this mutagenic technique to the sequence encoding only the central domain of NorR, specific primers that amplify the AAA+ domain sequence were used. In addition, suitable restriction sites were required to isolate the AAA+ domain sequence after amplification and to enable re-cloning back into the *norR* sequence. However, there were no suitable restriction sites either side of AAA+ domain sequence that were absent in the remainder of the *norR* sequence or the the pET21a vector. Therefore, two silent mutations were engineered either side of the sequence encoding the AAA+ domain to create restriction sites that did not change the *norR* coding sequence. A base change C496T created the *Mfe* I/ *Mun* I restriction site and a further base change G1341C created the *Sst* II/ *Sac* II restriciton site at the 5' and 3' ends of the *norR* AAA+ sequence respectively. The resulting plasmid was named pMJB1 and the *Mfe I – Sst II* fragment formed the template for the random PCR mutagenesis.

6.1   Confirmation of wild-type activity of mutant plasmid pMJB1

Although the engineered mutations were silent in terms of the coding sequence, it was important to establish that they did not alter the expression of the NorR protein as a result of codon changes. Transformation of the pMJB1 and pNorR plasmids into the MH1003 *E. coli* strain (appendix section 12.1.1) enabled comparison of the expression of NorR proteins from sequences with and without the engineered codon changes respectively. MH1003 is a *norR*::*cat* derivative of the *E. coli* strain MC1000 (appendix section 12.1.1) with a *lacZ* reporter fusion to the *norVW* promoter inserted at the phage λ attachment site (D'Autreaux *et al.* 2005). Transformation of this strain, which lacks a chromosomal copy

of the *norR* gene, with a plasmid encoding NorR is the basis of the β-galactosidase that relies upon "leaky" expression from the pET21a vector. In the presence of NO, wild-type NorR activates the expression from the *norV-lacZ* promoter, and the accumulation of β-galactosidase indicates the level of expression and transcriptional activity of NorR. The subsequent measurement of β-galactosidase activity confirmed that creation of the restriction sites does not alter the activity of NorR with respect to activation of transcription at the *norV* promoter (Figure 6.1).

6.2   Searching for escape mutants in the AAA+ domain

The overall strategy for identifying mutations in the AAA+ domain which allow the protein to escape GAF-mediated repression of activity is shown in Figure 6.2. 35 cycles of a standard PCR reaction were conducted using the AAA+ Fwd and Rev primers (appendix section 12.1.4) and the low-fidelity Taq polymerase. Several, separate reactions were set-up to maximise the number of novel base variations within the AAA+ sequence of *norR*. Multiple PCR reactions were pooled and then purified. The amplified sequence encoding the AAA+ region of NorR was subsequently digested with the *Mfe* I and *Sst* II restriction enzymes that cut at the engineered restriction sites. The resulting population of *Mfe* I – *Sst* II fragments was cloned into pMJB1, similarly treated to remove the *Mfe* I – *Sst* II fragment, to reconstitute the complete *norR* sequence. This created a population of plasmids with randomly incorporated base substitutions within the AAA+ sequence of *norR*. Electroporation of the plasmid population into the *E. coli* strain DH5α was conducted and plasmid purification carried out before transformation of the sample into the *E. coli* strain MH1003 (appendix section 12.1.1).

6.3   Identification of constitutive mutants

Identification of constitutive mutants was based around the ability of the variant forms of NorR encoded by the mutant plasmid population to activate expression from the *norV-lacZ* promoter fusion of MH1003 in the absence of an NO-source. Transformants were plated out on LB agar containing suitable antibiotics and the X-gal substrate (40 μg/ml). In the absence of an NO-source, the regulatory (GAF) domain represses the activity of the central (AAA+) domain of NorR (D'Autreaux *et al.* 2005). Therefore the wild-type protein is largely unable to activate transcription and significant production of β-galactosidase does not occur. In some plasmids random PCR mutagenesis of the AAA+ sequence might lead
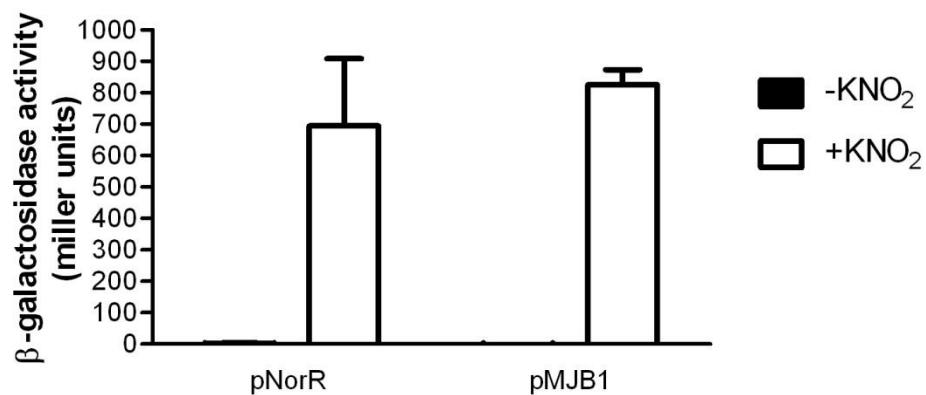
Figure 6.1 – Comparison of the *in vivo* activity of wild-type NorR and a derivative expressed from a sequence containing two silent mutations that introduced the *Mfe* I and *Sst* II restriction sites. Transcriptional activation was assessed using the β-galactosidase (*norV-lacZ* reporter) assay after transformation of the MH1003 complementation strain with the pMJB1 and pNorR plasmids that contain *norR* sequences with and without the engineered codon changes respectively. 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.
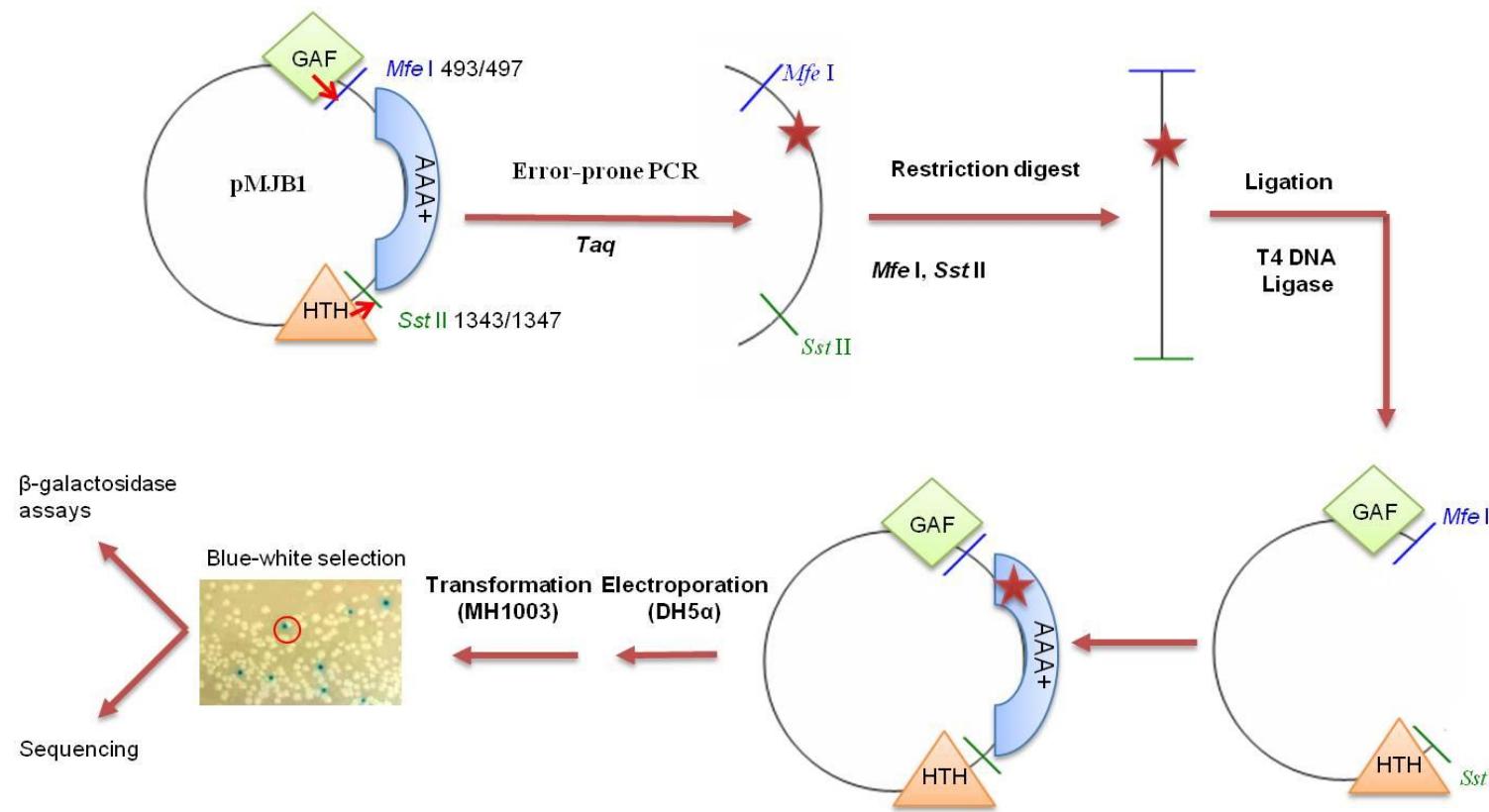
Figure 6.2 - Summary of the strategy used to identify constitutive mutants of the AAA+ domain in NorR. pMJB1 plasmid encoding the three NorR domains is shown; N-terminal regulatory domain (GAF) = Green; central ATPase domain (AAA+) = blue; C-terminal DNA binding domain (HTH) = orange. Since no suitable restriction sites were present, silent mutations were made by SDM to create the *Mfe* I (493/497) and *Sst* II (1343/1347) restriction sites either side of the AAA+ domain. Error-prone PCR mutagenesis was then carried out using the AAA+ Fwd and AAA+ Rev primers that anneal outside of the restriction sites, either side of the central domain. A low fidelity Taq polymerase was used to ensure that mutation events occurred during the amplification (indicated by the red star). Restriction digestion using the *Mfe* I and *Sst* II enzymes removed the "mutated" AAA+ sequence and electrophoresis followed by gel extraction enabled purification of the PCR insert. Non-treated pMJB1 was digested simultaneously and the larger "empty" vector isolated in the same way. Ligations were then set-up to clone the "mutated" AAA+ sequence back into *norR*. Electroporation of the ligated mutant plasmid sample into DH5α was conducted and plasmid purification carried out before transformation of the sample into the *E. coli* strain MH1003. Constitutive mutants were identified by a blue-white screen based on the ability of the *norR* gene to produce a protein that can activate expression of a *norV-lacZ* fusion. In the absence of NO, constitutive mutants activate expression of β-galactosidase that cleaves the X-gal substrate to produce a blue product. Mutant colonies were picked for sequencing and determination of their activity by the β-galactosidase assay.

to the expression of NorR variants that escape GAF-mediated repression of activity. Such mutant derivatives will therefore be able to activate transcription from the *norV-lacZ* fusion and β-galactosidase will cleave the X-gal substrate to produce a blue colour. Blue colonies that represented the expression of a constitutively active form of NorR were observed at a frequency of less than 5 % (Figure 6.3B). Within this population, the majority of colonies were pale blue in colour, suggesting a moderate NorR activity, compared to a small number of dark blue colonies that indicated a higher expression of β-galactosidase. All dark blue colonies as well as a large number of pale blue colonies identified in the screen were streaked out to ensure homogeneity before purification of the variant plasmids. Overall, three separate rounds of random PCR mutagenesis were carried since amplification commonly resulted in multiple copies of the *Mfe* I – *Sst* II fragment with the same base change. This facilitated the identification of the less frequent, dark blue colonies that represented variants of NorR with particularly strong phenotypes. In total 100 colonies identified in the screening were picked for plasmid purification and sequencing.

6.4   Escape mutants of the AAA+ domain of NorR

After screening, sequenced plasmids from blue colonies were transformed back into the *E. coli* strain, MH1003, for analysis of the *in vivo* phenotype (Figure 6.4). This strategy produced variant versions of NorR that exhibited activity in cultures grown in the absence of endogenous NO, generated by the presence of potassium nitrite. Plasmids isolated from pale blue colonies in the screen gave rise to NorR variants with only a small β-galactosidase activity in the absence of endogenously generated NO, typically only 10-fold higher compared to the wild-type protein (Figure 6.4A). Furthermore, in the presence of an NO-source, these variants exhibited two to five-fold lower activities compared to wild-type NorR. Since, the K226E, E249K, K274R, Y305C, Y305N, V323A, H346Y and S349N NorR variants still require NO to activate transcription *in vivo*, these substitutions do not appear to significantly disrupt GAF-mediated repression of AAA+ activity. In contrast, plasmids isolated from dark blue colonies in the screen led to the expression of NorR variants with significant β-galactosidase activity in the absence of potassium nitrite (Figure 6.4B). In some cases (L256F, E276G, G266D and S292L) activity in the absence of the NO-signal was similar to that exhibited by a truncated version of NorR lacking the GAF domain (NorRΔGAF). This phenotype suggests that repression by the GAF domain has been completely bypassed, resulting in loss of regulation upon the AAA+ domain. In other
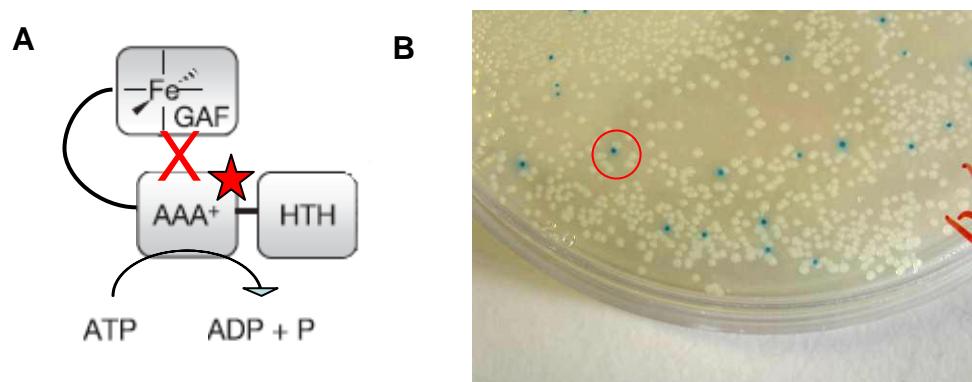
Figure 6.3 - NorR constitutive AAA+ mutants (**A**) Schematic of constitutively active NorR mutants. A substitution in the AAA+ domain is anticipated to interrupt the GAF-AAA+ interface, preventing interdomain repression and activating the protein in the absence of NO. (**B**) Transformation of mutated pMJB1 into MH1003 (in the absence of NO) yields blue colonies when the X-gal substrate is present that indicate an escape phenotype.
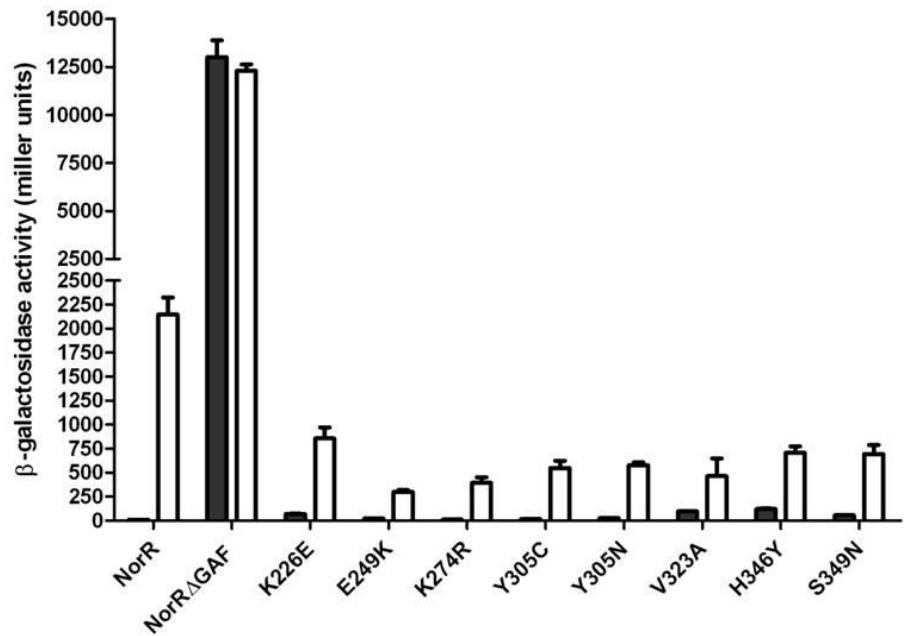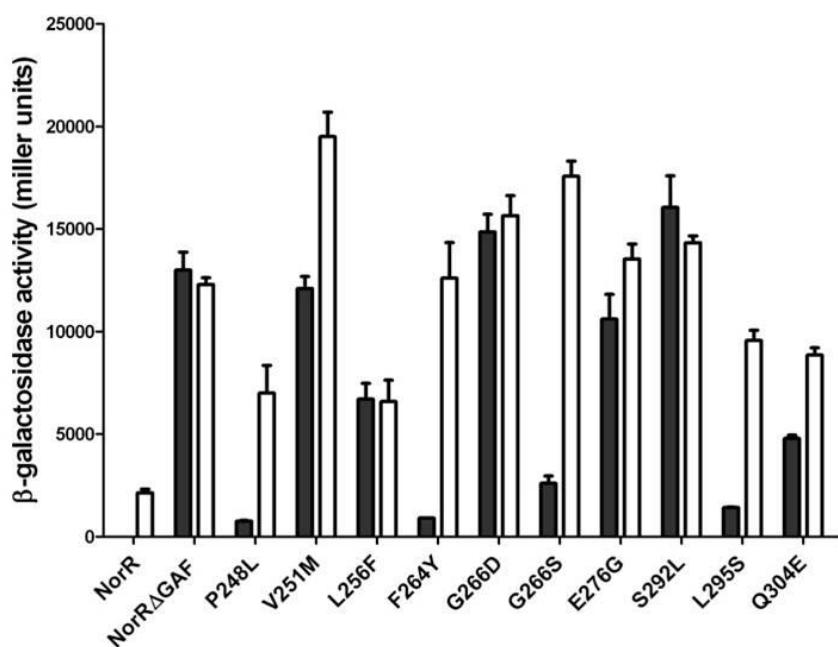
Figure 6.4 - **Transcriptional activation by NorR AAA+ domain variants *in vivo* as measured by the *norV-lacZ* reporter assay**. (**A**) *in vivo* activities of NorR variants expressed from plasmids that gave rise to a pale blue colony-phenotype in the screen. (**B**) *in vivo* activities of NorR variants expressed from plasmids that gave rise to a dark blue colony-phenotype in the screen. 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.

cases (P248L, V251M, F264Y, G266S, L295S and Q304E) some repression in the absence of NO was evident, indicative of a partial bypass phenotype. In order to predict the location of residues identified in the random mutagenesis, a structural model of the NorR AAA+ domain of NorR was built (by R.A Dixon) based on the crystal structure of the NtrC1 monomer (1NY5 Chain A) (Lee *et al.* 2003). Using this NtrC1 structure as a basis for the structural model had the advantage that the surface-exposed loops were well-defined unlike many other available bEBP structures. The model predicted that the majority of the substitutions are located in either helix 3 (H3), helix 4 (H4) or loop 1 (L1) of the AAA+ domain of NorR (Figure 6.5). These are the structural features in the central domain that undergo nucleotide-dependent conformational changes during the ATPase cycle to promote engagement with σ^{54}. In a structural model of the the NorR oligomer (built by R.A Dixon) based on the ATP-bound forms of the NtrC1 heptamer (PDB ID: 3MOE) (Chen *et al.* 2010), the majority of residues identified are clearly located at the upper surface of the ring and towards the centre of the pore within the region of σ^{54}-interaction (Figure 6.5C). Therefore, this region may represent a target of the GAF domain in the mechanism of interdomain repression in NorR.

6.4.1   P248L, V251M, S292L, L295S and L256F substitutions

The P248 and V251 residues are predicted to be located at the base of helix 3a in the model of the NorR AAA+ domain based on the NtrC1 structure (Figure 6.5) (PDB ID: 1NY5 Chain A) (Lee *et al.* 2003). Based on their relative positions, they are predicted to form an H-bonding interaction (Figure 6.6B). The V251M mutation gave rise to significant activity in the absence of an NO-source with β-galactosidase activity similar to that of NorRΔGAF (Figure 6.4B). However, in the presence of endogenously-generated NO, the activity of the V251M variant increased 1.5 to 2-fold suggesting that although this substitution leads to significant "escape", AAA+ activity is still partially repressed by the GAF domain. It is possible that substitution of the valine residue to methionine at position 251 is enough to disrupt the 248-251 interaction and significantly reduce the repression exerted by the GAF domain upon the AAA+ domain. However, unlike NorR-V251M, the P248L variant only exhibited low β-galactosidase activity in the absence of the NO-signal which was strongly induced in the presence of an NO-source (Figure 6.4B). In this case, a weaker escape phenotype might be explained by the ability of a leucine residue at the same position to maintain the potential H-bond with V251. Alternatively, disruption of the H-bonding interaction may not be responsible for the phenotypes observed *in vivo*.
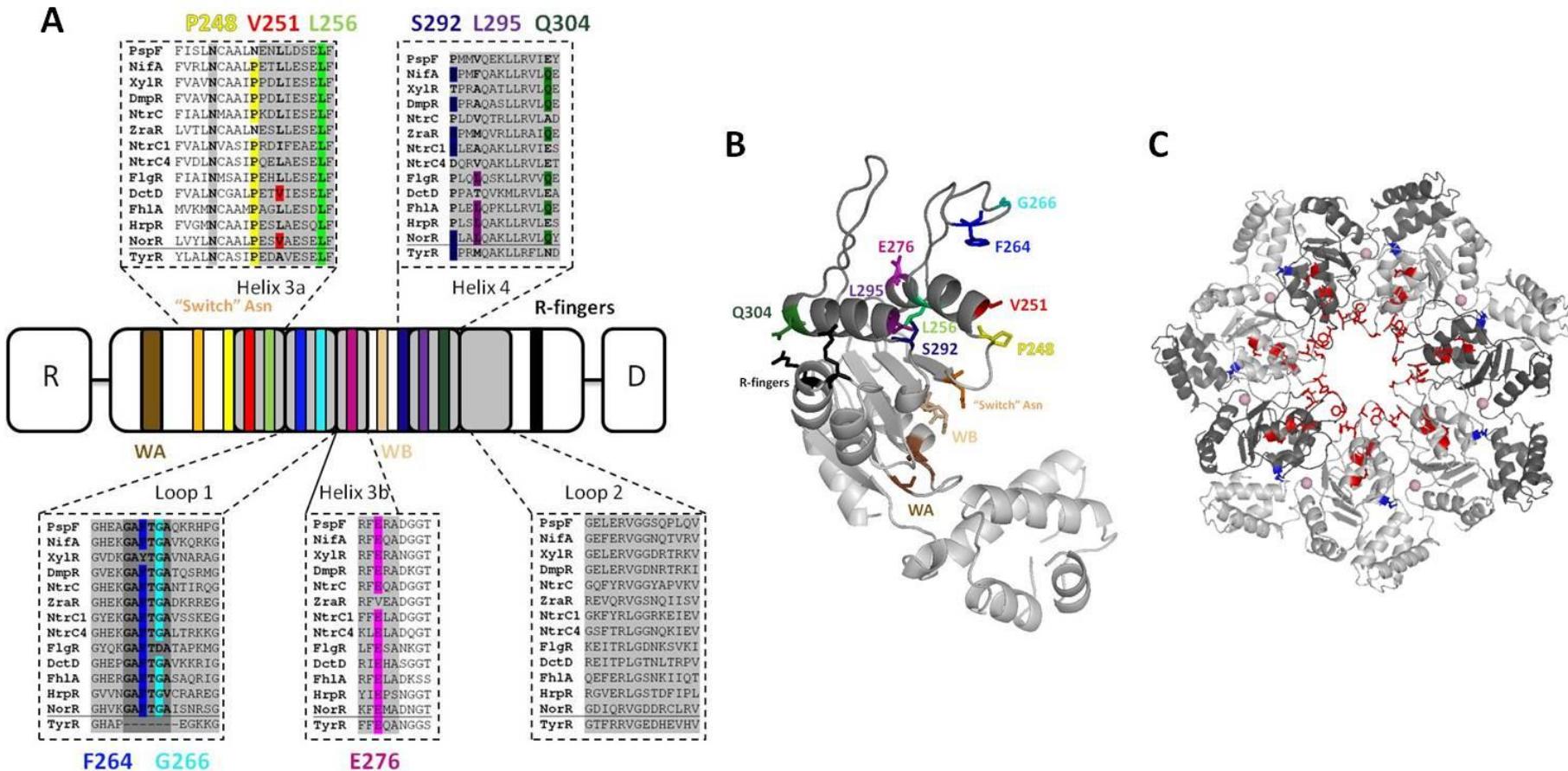
Figure 6.5 - (**A**) **Domain map and sequence alignment of bEBP AAA+ domains**. The locations of the residues identified in the random mutagenesis are indicated relative to the Helix 3a, Loop 1, Helix 3b, Helix 4 and Loop 2 motifs (shaded in grey). Residues that are identical to the NorR sequence (underlined) in the bEBP alignment are shaded. The locations of the Walker A, "Switch" Asn, Walker B and R-finger motifs are also indicated by colour coding. Alignments were conducted using ClustalW (www.ebi.ac.uk/clustalw/) using the sequences from UniProtKB/Swiss-Prot (http://www.expasy.ch/): PspF (*E. coli*), NifA (*A. vinelandii*), XylR (*P. putida*), DmpR (*Pseudomonas* sp.), NtrC (*E. coli*), ZraR (*E. coli*), NtrC1 (*A. aeolicus*), NtrC4 (*A. aeolicus*), FlgR (*H. pylori*), DctD (*S. meliloti*), FhlA (*E. coli*), HrpR (*P. syringae*), NorR (*E. coli*), TyrR (*E. coli*). R = regulatory domain; D = DNA binding domain. (**B**) **Structural model of the AAA+ domain of NorR based on the NtrC1 structure** (Lee *et al.* 2003) (1NY5 Chain A). The residues identified in the mutagenesis and key conserved motifs are labelled and shaded, correlating to the domain map in (**A**). The F264 and G266 residues form part of the GAFTGA motif that contacts σ[54]. (**C**) **Structural model of the NorR oligomer based on the ATP bound NtrC1 heptamer** (Chen *et al.* 2010) (PDB ID: 3MOE). Alternative monomers are indicated by dark and light shading. For clarity the ATP molecules are not shown but the catalytic magnesium ions are shown as light pink spheres. The majority of residues identified are clearly located in the region of the NorR oligomer that undergoes conformational change upon ATP hydrolysis i.e. at the upper surface and towards the centre of the ring (shaded in red). The exception is the Q304 residue (shaded in blue) which is not expected to have a role in modulating the conformation of the σ[54]-interaction surface.

Likewise, the S292 and E249 residues are also predicted to interact via an H-bond (Figure 6.6C). The activity of the S292L variant *in vivo* was very similar to that of the truncated version of wild-type NorR (NorRΔGAF), indicating that the mutation allows the central domain to fully escape repression (Figure 6.4B). Potentially, the serine to leucine substitution prevents this H-bond contact from forming between the two residues, thereby disrupting the interface between the GAF and AAA+ domains. Error-prone PCR also identified the E249K variant of NorR. This substitution would also be expected to disrupt a potential H-bonding interaction between the 249 and 292 residues but the E249K variant only has a mildly constitutive phenotype (Figure 6.4A). Possibly, the lysine residue maintains the interface between GAF and AAA+ domains by forming a polar interaction with another nearby residue. Alternatively, the disruption of this potential H-bond may not be responsible for the *in vivo* phenotype of the S292L variant. Indeed, the S292 residue is also predicted to form an H-bonding interaction with L295, so disruption of this potential interaction may be responsible for the escape phenotype of S292L. The L295S and L256F mutants also showed strong activity in the absence of NO and according to the structural model the L295 residue (H4) may form a hydrophobic interaction with L256 in the opposite helix (H3) (Figure 6.6D), although the model calculates this distance as 4.0 Å. The L256F variant appeared to fully escape repression by the GAF domain unlike the L295S protein which remained partially inducible (Figure 6.4B). The disruption of this potential hydrophobic interaction by the L to S change at position 295 or by the L to F change at position 256, may offer an explanation for the constitutive phenotypes observed *in vivo*.

6.4.2   The E276G substitution

The E276 residue is predicted to be located at the tip of helix 3b in the structural model (Figure 6.5). The E276G variant showed a fully constitutive phenotype (Figure 6.4B), indicating complete escape from GAF-mediated repression of AAA+ activity. E276 is predicted to form a salt bridge with the R310 residue (Figure 6.6E) and the disruption of this interaction might therefore explain the escape phenotype of the E276G variant. In order to further investigate the escape of repression in this variant, targeted mutagenesis was carried out at positions 276 and 310 (Figure 6.7). If the loss of an E276-R310 polar interaction was responsible for the constitutive activity, then exchange of the negatively charged glutamate for a positively charged residue would be expected to result in a similar phenotype. Similarly, exchange of the positively charged arginine for a negatively charged
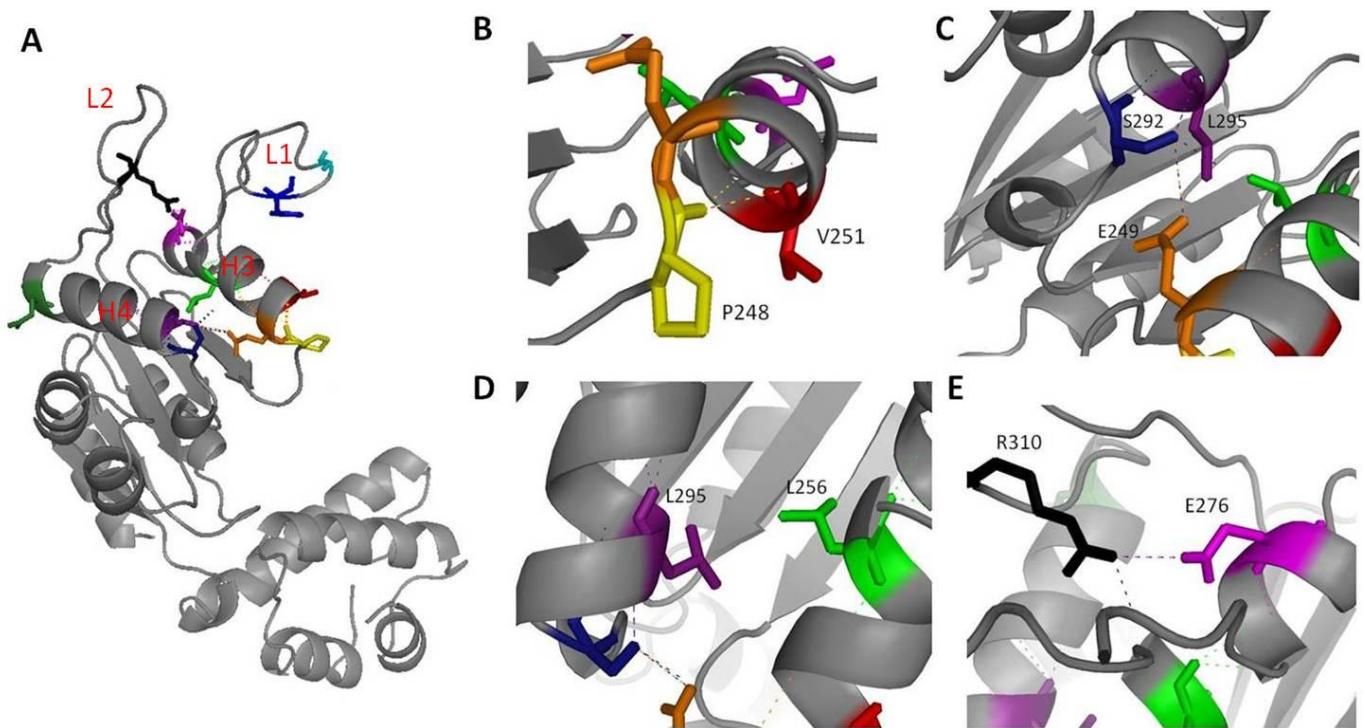
Figure 6.6 - Model of the expected interactions between residues in the AAA+ domain of NorR, based on the structure of NtrC1 (Lee *et al.*, 2003) (1NY5 Chain A; NtrC1 inactive dimer) (**A**) Overall fold of the AAA+ domain based upon a single monomer of the NtrC1 (PDB ID: 1NY5 Chain A of inactive dimer) (Lee *et al.* 2003). Helix 3 (H3), Helix 4 (H4), Loop 1 (L1) and Loop 2 (L2) are labelled. (**B**) Predicted H-bond between the P248 and V251 residues of Helix 3. (**C**) Predicted H-bond between the E249 (Helix 3) and S292 (Helix 4) residues. (**D**) Predicted contact between the L295 (Helix 4) and L256 (Helix 3) residues. (**E**) Predicted salt-bridge between the R310 (Loop 2) and E276G (Helix 3) residues. The residue side chains are coloured as in Figure 6.5 with the additional residues R310 and E249 coloured in black and orange respectively.

residue would also be expected to produce a constitutive phenotype. However, lysine, arginine, and histidine substitutions at position 276 all gave phenotypes similar to wild-type NorR. Additionally, an alanine substitution gave a wild-type response, suggesting that this residue is not of crucial importance in the mechanism of interdomain repression. Mutagenesis of the R310 residue (the predicted interaction partner of E276) to either glutamate or alanine resulted in a null phenotype, indicating that this residue is of crucial importance in NorR (Figure 6.7). It appears then, that the breaking of a potential E276-R310 polar interaction is not the explanation for the escape phenotype and that the glycine change at position 276 enables escape from GAF-mediated repression by some other mechanism.

6.4.3 The Q304E substitution

The only substitution identified in the random mutagenesis screen that gave rise to a significant escape phenotype but is predicted to be located outside the region of nucleotide-induced conformational change was Q304. The equivalent residue in NtrC1 is most probably involved in inter AAA+ domain subunit interactions. β-galactosidase assays showed that the Q304E variant had a partially constitutive phenotype (Figure 6.4B). The mutant protein showed significant activity in the absence of the NO-signal but an NO-source was required to enable full activation. The Q304E variant of NorR is the subject of Chapter 9.

6.4.4 Mutations in the GAFTGA motif of NorR give rise to constitutive activity

Error-prone PCR identified three substitutions within the highly conserved GAFTGA motif which forms part of a loop on the surface of the AAA+ domain that makes contact with σ^{54} during the ATP-hydrolysis cycle (Bordes *et al.* 2003). Structural studies have shown that the GAFTGA loop (also known as loop 1), assisted by loop 2 is in an extended conformation in the ATP-bound transition state and is therefore well positioned for contact with region I of σ^{54} (Rappas *et al.* 2006; Bose *et al.* 2008a). Upon phosphate release, the GAFTGA loops collapse down towards the surface of the AAA+ domain, allowing for relocation of σ^{54} and open complex formation (Figure 3.4D). Therefore the GAFTGA motif is critical in coupling the ATP-dependent conformational changes that occur in the AAA+ domain to the isomerisation of holoenzyme.
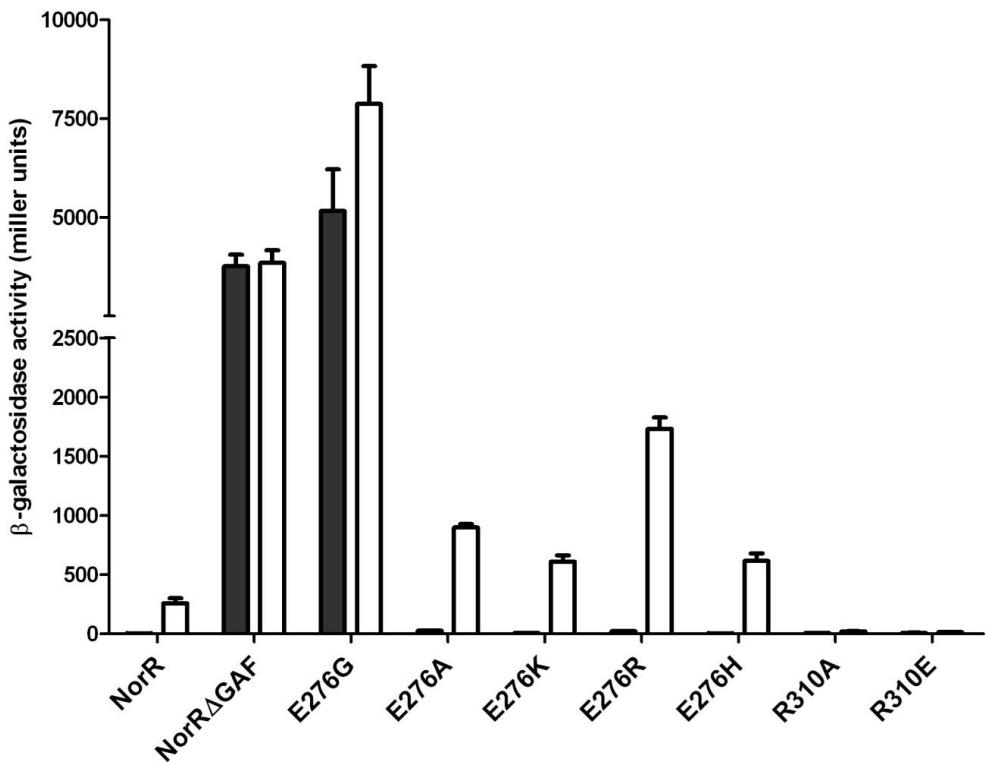
Figure 6.7 - Transcriptional activation by variants of the E276 and R310 residues of NorR
***in vivo* as measured by the *norV-lacZ* reporter assay**. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking the GAF domain (residues 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.

β-galactosidase assays revealed that the F264Y substitution enabled partial escape from interdomain repression (Figure 6.4B). This phenylalanine is part of the key GAFTGA motif found in all bEBPs and has recently been implicated in the remodelling of the Eσ^{54} closed complex through an interaction with the -12 DNA fork junction structure (Zhang *et al.* 2009). Two substitutions at the second glycine of the GAFTGA motif (position 266) were identified in the random mutagenesis of the NorR AAA+ domain. The most notable of these was the G266D mutation, which allowed full escape from the GAF-mediated repression of NorR activity (Figure 6.4B). This might seem surprising given that this loop is required to contact σ^{54} to drive open complex formation (Buck *et al.* 2006) and that substitutions at G266 are likely to influence the conformational flexibility of this loop. In order to determine which amino acids at the G266 position give rise to constitutive activity, the glycine residue was substituted for each of the other 19 natural amino acids (Figure 6.8A). Like the GAFTGA variant G266D, the G266N variant of NorR was fully active and furthermore escaped GAF-mediated repression *in vivo*. In addition to the aspartate and asparagine substitutions that gave rise to constitutive activity; glutamine, serine, cysteine, and methionine all gave activity in the absence of an NO-source. Whereas asparagine and aspartate changes gave fully constitutive phenotypes, the other changes were still partially subject to regulation by the GAF domain. Surprisingly, the glutamate substitution did not produce a functional NorR protein. The remaining amino acid changes all resulted in non-functional proteins and Western blotting confirmed that this was not due to instability (Figure 6.8B).
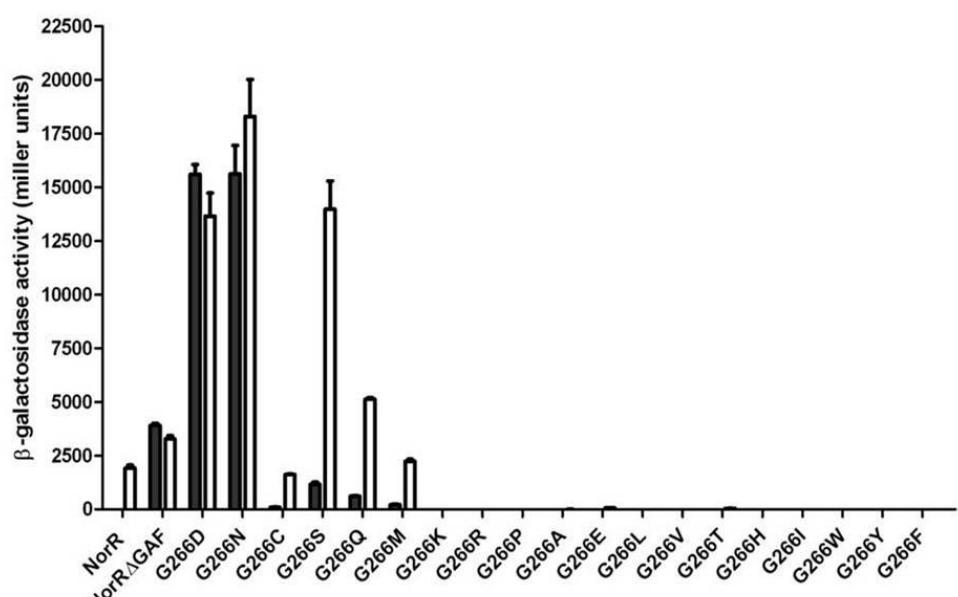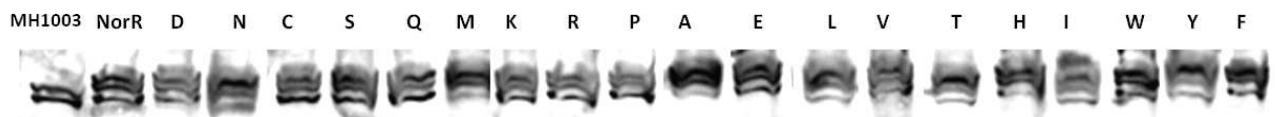
Figure 6.8 – Mutagenesis at position 266 of NorR (**A**) Transcriptional activation by mutants of the G266 residue of NorR *in vivo* as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated 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. (**B**) Western blot analysis indicating the stability of NorR G266 variants *in vivo* when cultures are grown in the absence of potassium nitrite. "NorR" refers to the wild-type protein. Lane 1 shows the result of Western blot analysis against the *E. coli* strain MH1003 only (i.e. without expression of NorR). Here, the uppermost band, correlating to NorR is absent.

6.5   Testing the requirement for the GAF domain in the GAFTGA variant G266D

β-galactosidase assays showed that the G266D mutant-version of NorR is fully constitutive, indicating a complete escape from GAF-mediated repression of AAA+ activity (Figure 6.4B). Although this GAFTGA-variant is fully constitutive, the N-terminal domain could still be required for the full-escape phenotype observed. To investigate this, the G266D and G266N substitutions were introduced into a construct that lacks the first 170 residues (pNorRΔGAF). Surprisingly, the resulting G266DΔGAF and G266NΔGAF variants exhibited a much lower activity *in vivo* than the full-length G266D and G266N variants respectively (Figure 6.9). Unlike G266, the Q304 residue is not predicted to have a role in modulating the conformation of the σ^{54}-interaction surface and the Q304E variant remains partially subject to GAF-mediated repression. When the Q304E substitution was introduced into a NorRΔGAF construct, the NorR variant was able to fully escape interdomain repression (Figure 6.9). In order to confirm a genuine requirement for the N-terminal GAF domain in the NorR variant G266D, Western blot analysis was carried out to compare the stability of the full-length and truncated constructs. Unfortunately, the NorR primary antibody was unable to detect the version of NorR that lacks the GAF domain, presumably since part of the N-terminal domain forms the epitope for the antibody. Since ΔGAF and Q304EΔGAF constructs exhibited a strong constitutive phenotype in the assay, it is unlikely that the G266DΔGAF variants are unstable although this cannot be entirely ruled out.

6.6   Testing the requirement for the GAF domain in other AAA+ variants of NorR

Next, it was important to investigate whether the potential requirement for the GAF domain is a feature of all of the AAA+ variants identified here. As mentioned previously, the Q304E variant remained partially subject to repression and as expected, the N-terminally truncated form was able to fully escape regulation by the GAF domain (Figure 6.9). Likewise, those variants that remained partially inducible *in vivo* (P248L, V251M, L295S and F264Y) became fully constitutive when the GAF domain was absent (Figure 6.10). However, as was the case for the G266D and G266N variants, other AAA+ variants (P248L, V251M, L256F, E276G, and F264Y) showed reductions in activity in the ΔGAF form. Since it was not possible to test the stability of such constructs, reduced stabilities in variants that lack the first 170 residues cannot be ruled out. Overall, only the activities of the Q304E, L295S and S292L variants were not reduced in the absence of the N-terminal
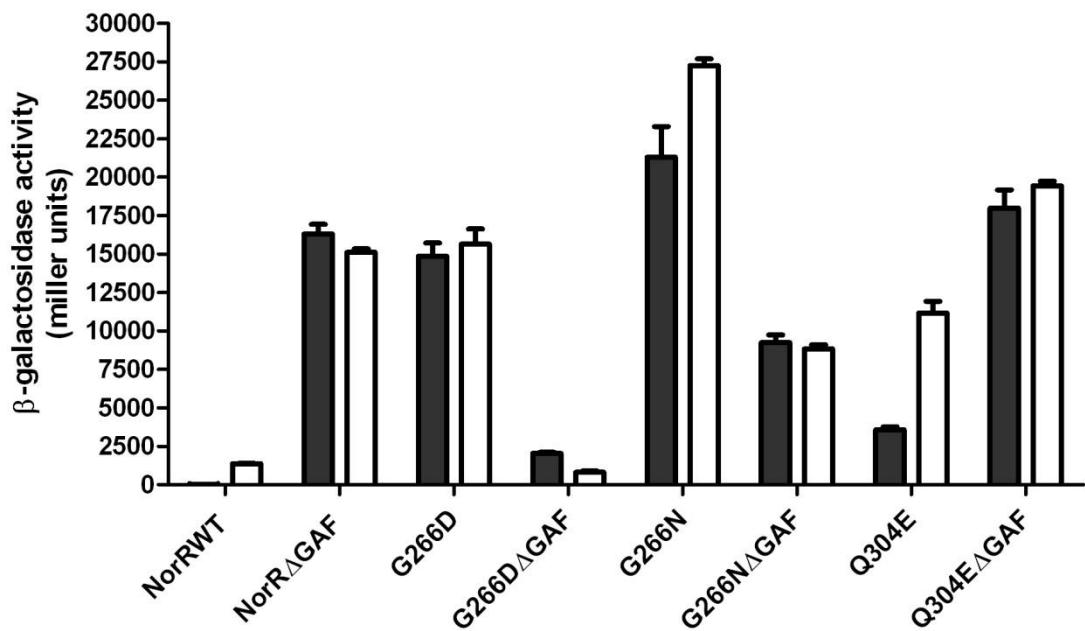
Figure 6.9 - Transcriptional activation *in vivo* by the G266D, G266N and Q304E variants in full-length and truncated forms (Δ1-170) as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated 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.
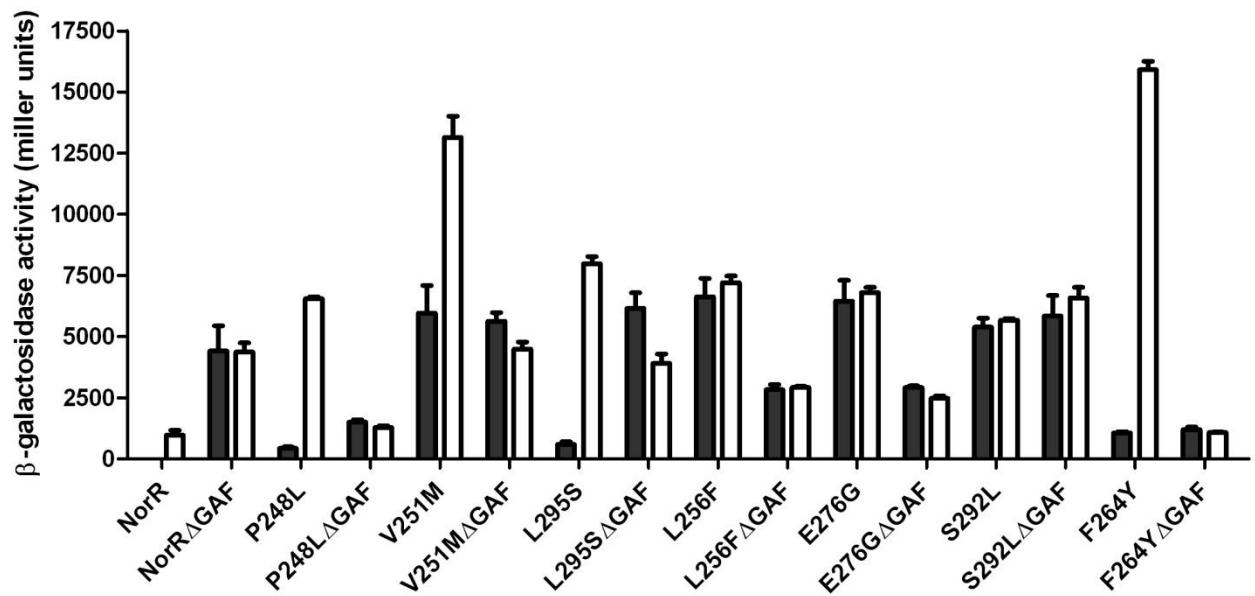
Figure 6.10 - Transcriptional activation *in vivo* by AAA+ variants in full-length and truncated forms (Δ1-170) as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking the GAF 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.

domain and interestingly these are the only substitutions identified in helix 4 (H4) of the AAA+ domain.

6.7   Testing the requirement for the GAF domain in His-tagged AAA+ variants

Since the NorR primary antibody was unable to detect the stability of wild-type and mutant versions of the protein that lack the GAF domain, it was decided to make tagged-versions of the full-length and truncated constructs and test the stabilities of expressed proteins *in vivo* using a primary antibody raised to detect the presence of the tag. In order to do this, the sequences that encode wild-type and mutant versions of NorR and NorRΔGAF were cloned into the pETNdeM11 vector (appendix section 12.1.2). The resulting constructs expressed proteins with a TEV cleavable, hexahistidine tag at the N-terminus and were later employed in protein purification (Chapters 7 and 8). The *E. coli* strain used to measure β-galactosidase activity expressed from the *norV* promoter, MH1003 (*norR*::cat *norV-lacZ*), is maintained by the chloramphenicol and kanamycin resistance genes and hence there was no oppotunity to select for the presence of the pETNdeM11 vector (which is also maintained by kanamycin) when expressed in the MH1003 strain. Therefore a strategy was devised to insert the omega (Ω) cassette, which encodes streptomycin/spectinomycin resistance, from the pHP45Ω plasmid (Prentki and Krisch 1984) into the *norR*-derivatives of the pETNdeM11 plasmid at the *Sma I* site (Chapter 5, Figure 5.3). The resulting His-tagged full-length and truncated constructs were then assayed for NorR activity in the *E. coli* strain MH1003. In contrast to the non-tagged constructs, there was no reduction in NorR activity for the G266DΔGAF-His and G266NΔGAF-His proteins compared to their full-length counterparts (Figure 6.11). It appears therefore that the decrease in activity observed for the non-tagged constructs *is* due to decreased stability. Presumably, the presence of the hexahistidine tag at the N-terminus of the truncated constructs is able to stabilise the variant proteins.
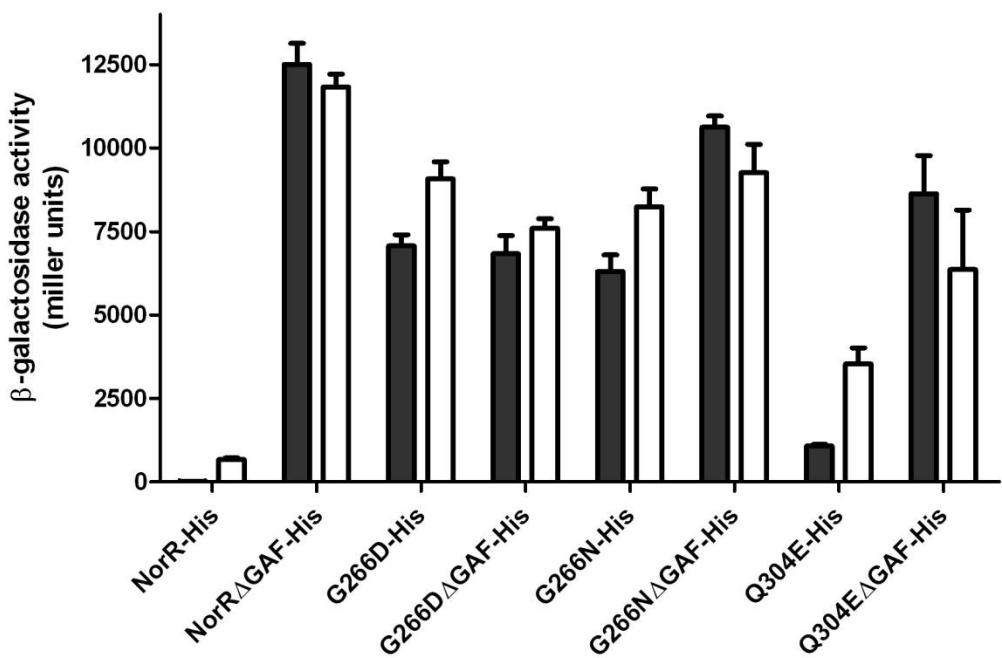
Figure 6.11 - Transcriptional activation *in vivo* by the G266D, G266N and Q304E variants in full-length and truncated forms (Δ1-170) with N-terminal hexahistidine tags, as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "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.

6.8 Influence of GAF domain substitutions on the G266D phenotype

In order to further investigate the role of the N-terminal (GAF) domain in the GAFTGA variant G266D, mutagenesis was carried out to substitute residues in the regulatory domain of the NorRG266D construct. In previous work, substitution of conserved residues prior to spectroscopic characterisation led to the identification of five candidate ligands to the non-heme iron centre (Tucker *et al.* 2007). The C113, D96 and D131 side chains are predicted ligands to the iron centre since the C113G, C113S, D96A and D131A variant proteins are unable to bind iron or do not exhibit the characteristic $g = 4$ Electron Paramagnetic Resonance (EPR) signal after reconstitution *in vitro* (Tucker *et al.* 2007). The EPR and UV-visible spectra of the D99A variant protein show that this substitution alters the structure of the Fe(NO) complex, making it a good candidate as an iron ligand (Tucker *et al.* 2007). Although arginine generally does not make a good ligand for transition metals, R75 has emerged as a fifth ligand in the model of the hexa-coordinated iron centre. Despite a normal complement of iron, EPR of the R75K variant is unable to detect the signal characteristic of the mononitrosyl complex in whole cells or after reconstitution of the iron *in vitro* (Tucker *et al.* 2007). The R75K, D96A, D99A, C113G, C113S or D99A substitutions were combined with the G266D change in the AAA+ domain of NorR to examine their influence on the escape phenotype of the G266D mutant protein. Results showed that substitutions at each of the five candidate ligands resulted in a reduction of NorR activity. Such reduction in activity could conceivably arise by affecting coordination at the iron centre, disrupting transmission of the NO-signal from the ferrous iron or by causing structural perturbations in the GAF domain. The former is more likely when the reduction in iron binding of the variants or the inability to form the mononitrosyl complex, observed *in vitro*, is considered (Tucker *et al.* 2007). As well as a reduction in activity, a number of the GAF-substitutions enabled escape from the GAF-mediated regulation of AAA+ activity e.g. D99A, D131A. Here, constitutive phenotypes could arise either as a consequence of disruption of the GAF-AAA+ interface, a conformational change in the N-terminal domain similar to that of the activated form, or a change in the coordination environment of the iron centre to mimic the NO-activated state. Considering the predicted role of these residues as ligands to the iron centre, the latter appears the most likely. Irrespective of the phenotype caused by the single GAF-substitutions, β-galactosidase assays revealed that these changes did not cause a reduction in the ability of G266D to activate transcription (Figure 6.12). This confirms that the NO-sensing function of the
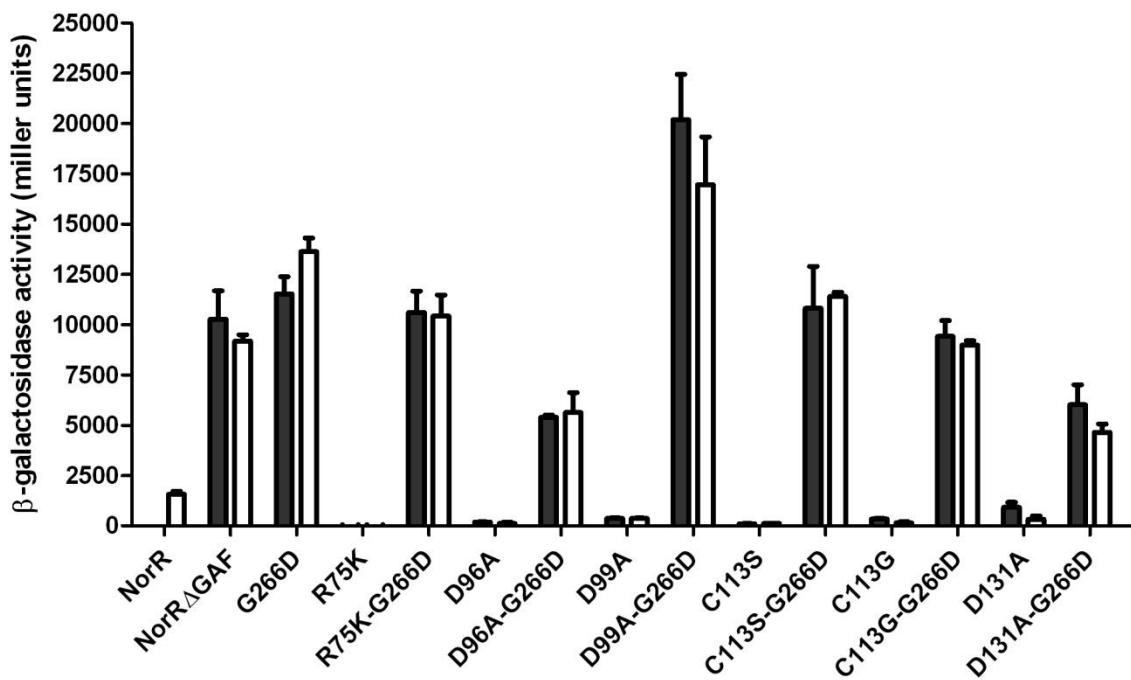
Figure 6.12 – Activities of G266D variants *in vivo* as measured by the *norV-lacZ* reporter assay when additional substitutions are made in the GAF domain to substitute residues predicted to form ligands to the non-heme iron centre (Tucker *et al.* 2007). Substitutions are indicated on the x axis. "NorR" 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.

GAF domain does not contribute to the constitutive phenotype of the G266D mutant observed *in vivo*.

Unlike substitutions at the predicted iron ligands of the non-heme iron centre of NorR, changes at the R81 and H111 residues do not appear to disrupt the iron centre since Electron Paramagnetic Resonance (EPR) reveals an Fe(NO) complex with identical spectral properties to the wild-type protein (Tucker *et al.* 2007). In addition, *in vivo,* the R81L, H111L and H111Y variants responded to the addition of nitrite suggesting that NO-induced signalling functions normally in these variants (Figure 6.13A). Therefore these residues are not predicted to have a role in coordinating the ferrous iron. Although not highly conserved in NorR proteins, the R81 residue is predicted to be surface exposed and a charged residue here would be well situated to form interaction(s) with residues in the AAA+ domain. The H111 residue is predicted to be found closer to the coordination site and may have a role in transmitting the "on" signal from the NO-bound iron to affect interactions between the GAF and AAA+ domains. The predicted location of R81 and H111 raises the possibility that these residues contribute to the mechanism of interdomain repression in NorR. In order to assess the effect of the R81L, H111L and H111Y substitutions on the ability of the G266D variant to activate transcription *in vivo*, these changes were made in the GAF domain of NorRG266D to create double mutants. β-galactosidase assays showed that the R81L substitution (but not the H111Y or H111L changes) in the GAF domain resulted in a significant decrease in the activity of the G266D mutant protein and Western blotting has shown that this is not due to a decrease in stability (Figure 6.13B). Therefore, it appears that substitution(s) in the GAF domain at R81 can partially restore the interface of interdomain repression in the G266D protein. Since R81 is predicted to be surface exposed and well placed for AAA+ domain contact, this suppression event suggests that the R81 residue may have a significant role in maintaining GAF-AAA+ interactions. The role of R81 in interdomain repression will be studied further in the Chapter 7.
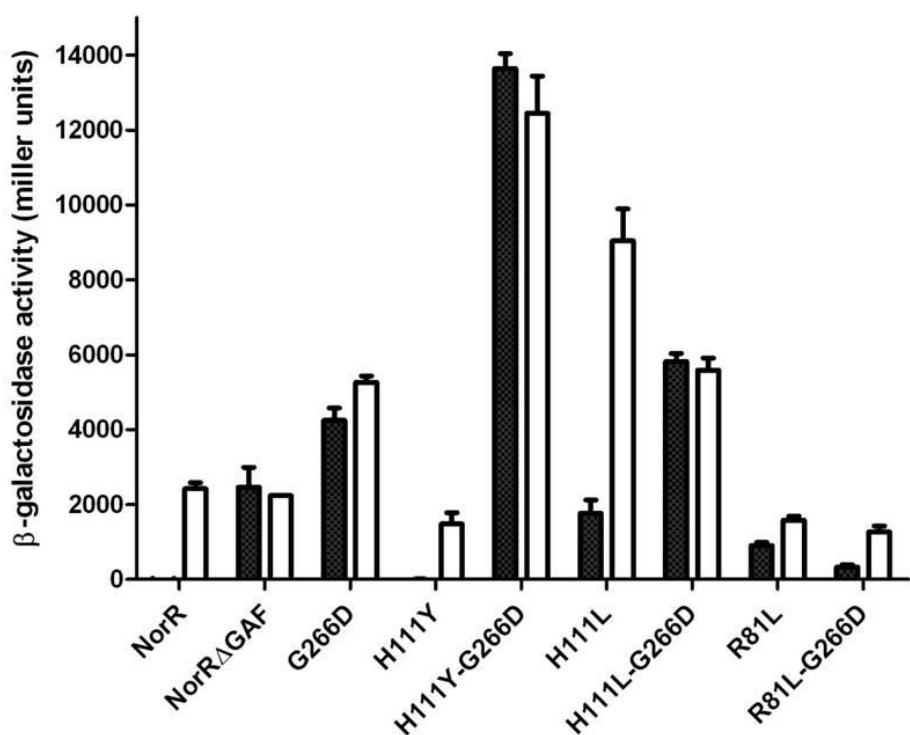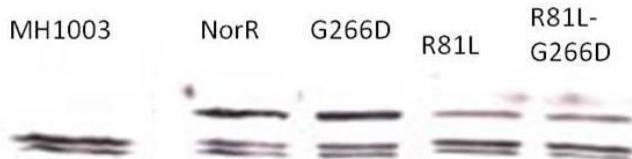
Figure 6.13 - Activities of G266D variants *in vivo* as measured by the *norV-lacZ* reporter assay when additional substitutions are made at the predicted surface-exposed R81 and H111 residues of the GAF domain (Tucker *et al.* 2007). Substitutions are indicated on the x axis. "NorR" refers to the wild-type 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. (**B**) Western blot analysis indicating the stability of NorR variants *in vivo* when cultures are grown in the absence of potassium nitrite. "NorR" refers to the wild-type protein. "MH1003" refers to the *E. coli* strain only. The uppermost of three bands that is not detected in the MH1003 strain correlates to NorR and its variants in the Western analysis.

6.9   Discussion

The lack of NO-responsive regulation in truncated forms of NorR that lack the GAF domain (D'Autreaux *et al.* 2005), clearly places NorR in the class of bEBPs that are negatively regulated. Here, a random mutagenic approach has been used to identify mutations in the central (AAA+) domain of NorR that allow the protein to escape repression by the N-terminal regulatory domain. The *in vivo* screening approach was based on the cleavage of the X-gal substrate by β-galactosidase, expressed from a *norV-lacZ* fusion and led to the identification of two types of NorR variant. Variants with weak escape phenotypes correlating to pale-blue colonies in the screen were more frequent but showed minimal transcriptional activity in the absence of an NO-source (Figure 6.4A). Therefore the K226, E249, K274, Y305, V323 and H346 residues are unlikely to have a significant role in the mechanism of interdomain repression. In contrast, variants with strong escape phenotypes correlating to dark-blue colonies in the screen were less common but showed significant constitutive activity in the β-galactosidase assay (Figure 6.4B). Substitutions of the AAA+ domain that result in such a phenotype indicate that the repression of the regulatory (GAF) domain upon the central (AAA+) domain has been bypassed. These include the changes L256F, E276G, G266D and S292L that gave rise to fully constitutive phenotypes and therefore appear to completely bypass the GAF-mediated repression mechanism. Other substitutions (P248L, V251M, F264Y, G266S, L295S and Q304E) were still responsive to NO, indicating a partial bypass phenotype. Overall three rounds of random PCR mutagenesis were sufficient to generate a plethora of constructs encoding mutant-versions of NorR. When coupled to an effective screening strategy, this is a powerful approach to identify variants of a particular phenotype. However, error-prone PCR is limited by the codon changes that can be made when any one (or more) base change occurs at a given position in the DNA sequence. Therefore it is possible that alternative substitutions at residues not deemed here to have a significant role in repression, may lead to stronger escape phenotypes. Significantly, the majority of substitutions identified here that gave rise to strong escape-phenotypes are all found in helix 3, helix 4 or loop 1. This region of the AAA+ domain undergoes significant conformational change throughout the ATP hydrolysis cycle to promote contact with σ^{54} (Rappas *et al.* 2006). Therefore, one explanation for the escape phenotypes of such variants is that substitutions prevent the GAF-mediated repression of this region thereby enabling σ^{54}-interaction without the need for NO-activation of NorR activity. Unlike the other substitutions identified in the AAA+ domain that lead to strong escape phenotypes, the

Q304E change is at a residue not expected to have a role of modulating the conformation of the σ^{54}-interaction surface. Predicted to be located at the tip of helix 4, changes at this residue are predicted to influence inter-protomer interactions, based on structures of related bEBPs. The Q304E variant and its mechanism of escape from GAF-mediated repression is the subject of Chapter 9.

It is noteworthy that the E276G substitution is at a residue predicted to play a major part in the conformational changes that couple the energy derived from ATP hydrolysis to σ^{54}-contact. A glycine substitution would be expected to change the nature of helix 3b, of which E276 is predicted to be a part. In PspF, the equivalent residue (E97) makes a salt-bridge with the key R131 residue (R310 in NorR) and the NtrC1 structure reveals a potential salt-bridge that forms between the R310 and E276 residues (Figure 6.6E). As part of the nucleotide-driven conformational change proposed in PspF (Figure 3.4D), PspF-R131 breaks this contact to establish a salt bridge with the PspF-E81 residue that leads to the rotation of helix 3 and return to the ADP bound state, when σ^{54} is released (Rappas *et al.* 2006). Therefore, the loss of the NorR E276-R310 interaction may lock the L1 loop in a conformation that is resistant to interdomain repression by the GAF domain. However, targeted mutagenesis revealed no correlation between the loss of the potential E276-R310 interaction and the *in vivo* phenotype, suggesting that the E276G variant escapes the GAF-mediated repression mechanism by some other, unknown mechanism.

Of particular interest is the identification of escape variants located in the highly conserved GAFTGA motif. This critical motif couples ATP-dependent conformational changes in the AAA+ domain to isomerisation of the Eσ^{54}-complex (Bordes *et al.* 2003). It is remarkable that substitutions in the surface exposed GAFTGA loop are able to prevent negative regulation by the GAF domain but still retain the ability to interact with σ^{54} and activate open complex formation. In the majority of bEBPs, substitutions in the GAFTGA motif cause a severe defect on the ability of the protein to activate transcription (Table 3.2). This includes substitutions at position 264 (GA<u>F</u>TGA) (Wang *et al.* 1997; Gonzalez *et al.* 1998; Li *et al.* 1999; Wikstrom *et al.* 2001; Bordes *et al.* 2003; Zhang *et al.* 2009). This work has identified the F264Y variant of NorR that was competent to activate transcription *in vivo* but required the addition of an NO-source to become fully active (Figure 6.4B). It may be that other substitutions at this position cause more significant disruption of the GAF-AAA+ interface but that the F264Y change is the only one that can be tolerated for NorR

activity. Work in NifA and PspF has shown that only substitution of the phenylalanine for a tyrosine residue permits transcriptional activation (Gonzalez *et al.* 1998; Zhang *et al.* 2009), indicating that the aromatic ring at this position is important for bEBP activity. In agreement with the observation that a F264Y substitution in NorR does not prevent transcriptional activation *in vivo*, ~ 7% of the 291 bEBPs available on the NCBI database have a naturally occurring tyrosine at this position. In PspF, the equivalent F85Y protein is the only variant of this position competent to interact with σ^{54} and form the ADP.AlF$_x$ "trapped" complex. However, in contrast to the F264Y variant in NorR, the F85Y variant of PspF did not activate transcription *in vivo* (Zhang 2010), although it was competent for low-level transcriptional activation *in vitro* (Zhang *et al.* 2009). It has been suggested that the presence of a tyrosine residue at this position of the motif in some bEBPs may ensure the lower level of σ^{54}-dependent-transcription that is required at certain promoters. In agreement with this proposal, a naturally occurring tyrosine is present in the bEBP HrpS that together with HrpR activates transcription from the *hrpL* promoter in *P. syringiae* (Hutcheson *et al.* 2001). When the tyrosine is substituted for a phenylalanine to produce the consensus "GAFTGA" motif, there is a 50 % increase in transcription from the *hrpL* promoter (Zhang *et al.* 2009). In the case of NorR, *in vivo* data suggest that the presence of a tyrosine at position 264 does not decrease the level of transcription and it may be that co-variance elsewhere in the bEBP, σ factor or promoter sequence is required to ensure a lower level of transcription.

Few studies have been carried out to explore the role of the second glycine in the GAFTGA motif. Targeted mutagenesis at the G266 position of NorR revealed that thirteen of nineteen possible amino acid changes resulted in a variant protein that was unable to activate transcription irrespective of whether an NO-source is present (Figure 6.8A). The non-functional nature of most substitutions at this position is not unexpected, given the importance of the GAFTGA motif and its high conservation in bEBPs. In agreement with the *in vivo* phenotype of the NorR variant G266K, the equivalent substitution in NtrC (G219K) results in the expression of a protein that fails to activate transcription. However, this variant had improved DNA binding activity and the ATPase activity was 50% of the activated wild-type (North *et al.* 1996), suggesting that the lysine substitution may result in a null phenotype by preventing interaction with σ^{54}. Overall, positively charged or aromatic residues are apparently not tolerated at this position in NorR, which may reflect a requirement for the σ^{54}-interaction. In contrast, the cysteine, serine, glutamine and

methionine substitutions at position 266 of NorR resulted in proteins that were competent to activate transcription (Figure 6.4A). These variants showed varying levels of activity in the absence of an NO-source but in all cases were still subject to the GAF-mediated repression of full NorR activity. The G266C variant of NorR exhibited a phenotype similar to that of the wild-type, indicating that the presence of a cysteine residue at the tip of loop 1 does not prevent interaction with σ^{54}, nor normal regulation of AAA+ activity. The G219C variant of NtrC is competent to form open complexes *in vitro* but intriguingly can only do so in the absence of enhancer DNA (Yan and Kustu 1999). This defect may be explained by changes in the relative juxtaposition of the DNA binding and ATPase domains observed during the ATPase cycle (De Carlo *et al.* 2006).

Here, random PCR and targeted mutagenesis has identified the G266D and G266N substitutions respectively that both gave rise to fully constitutive phenotypes, indicating a complete escape from the GAF-mediated repression of AAA+ activity (Figure 6.8A). Indeed, further substitutions designed to disrupt the detection of NO at the non-heme iron centre of the GAF domain had no effect on the activity of the G266D variant (Figure 6.12). Therefore, the NO-sensing function of the regulatory domain does not contribute to the constitutive phenotype observed *in vivo*. In order to confirm that the G266D and G266N mutant-versions or NorR are "true" escape variants, the aspartate or asparagine substitutions were introduced into the NorRΔGAF construct (Figure 6.9). It appears that removal of the N-terminal domain caused a decrease in activity in all AAA+ variants identified here except for those with substitutions in helix 4 (Figure 6.10). Since the ΔGAF and Q304EΔGAF proteins exhibited strong constitutive phenotypes, it seems unlikely that similar constructs with alternative substitutions should be unstable, although Western blotting was unable to detect any NorR construct that lacks the first 170 amino acids. Interestingly, β-galactosidase assays showed that the complete bypass phenotypes of the G266D and G266N are not affected in the absence of the GAF domain when an additional TEV-cleavable, hexahistidine tag is present at the N-terminus (Figure 6.11). It is possible that the N-terminal His-tag stabilises such proteins and therefore the GAF domain probably does not contribute to the activity of the GAFTGA variants G266D and G266N, suggesting that they are "true" escape variants.

In line with the identification of active G266D and G266N variants of NorR, a number of bEBPs exist with naturally occurring aspartate or asparagine residues at position 266

(Figure 6.14). These include two theoretical proteins that are each predicted to contain a GAF domain (Swiss-Prot Q7M8U4 from *Wolinella succinogenes* and Swiss-Prot Q7UK84 from *Rhodopirellula baltica*). A protein (Swiss-Prot Q8EN58) from *Oceanobacillus iheyensis* is predicted to contain a PAS domain whilst a number of other theoretical proteins are predicted Response Regulators (RRs) of two-component systems and contain a REC (cheY-homologous receiver domain) domain. Notable amongst this group is the bEBP GAFTDA-containing FlgR which is phosphorylated by FlgS, activating it as a bEBP for the transcription of genes required for flagella biosynthesis (Spohn and Scarlato 1999; Brahmachary *et al.* 2004). *Helicobacter pylori* and *Helicobacter hepaticus* FlgR both contain the GAFTDA motif but only FlgR from *H. pylori* also lacks the C-terminal DNA-binding domain that is present in most other bEBPs. This suggests that the loss of the DNA binding domain in *H. pylori* occurred after the divergence of the *Helicobacter* (Brahmachary *et al.* 2004) and since both contain GAFTDA motifs, the presence of an aspartate at position 266 is not linked to the enhancer-independent function of *H. pylori* FlgR. As is the case for FlgR, the other bEBP-like proteins with altered GAFTGA motifs are presumably still subject to normal regulation of AAA+ activity. In contrast to the G266D and G266N NorR variants, such proteins therefore would not be predicted to exhibit escape phenotypes.

Overall, the identification of the fully constitutive GAFTGA variants gives rise to the hypothesis in which regulation of NorR activity is mediated by the GAF domain which prevents interaction of the GAFTGA motif with σ^{54} in the absence of NO. Reduction in the constitutive activity of the G266D variant by the additional R81L change (Figure 6.13A) may suggest that this phenotype can be suppressed to restore the NO-dependent activity of NorR. The means by which the GAFTGA variant G266D escapes repression and the role of the R81 residue in interdomain repression will be the subject of Chapter 7.

Figure 6.14 – AAA+ proteins with altered coding-sequences in the highly conserved GAFTGA motif of bEBPs. (**A**) Alignments were conducted using ClustalW (www.ebi.ac.uk/clustalw/) using the sequences (*accession numbers in red*) from UniProtKB/Swiss-Prot (http://www.expasy.ch/): P37013/NorR (*E. coli* K12); Q8A7J5, Q8A9F9, Q8A7Q5 (*B. thetaiotaomicron*); Q7MW89, Q7MXY0 (*P. gingivalis*); Q7MAA8, Q7M8U4 (*W. succinogenes*); Q7UK84, Q7UJ61 (*Rhodopirellula baltica*); O25408 (*H. pylori*); Q7VFP0 (*H. hepaticus*); Q8EN58 (*O. iheyensis*). The sequence of the GAFTGA motif is boxed in green. (**B**) The predicted domain structures of the unusual-"GAFTGA" containing AAA+ proteins according to the SMART database (http://smart.embl-heidelberg.de/). The accession number, organism, recommended protein name and GAFTGA motif conservation are listed.

182

Chapter 7 - Investigating the escape mechanism of the GAFTGA variant G266D

7.1 Introduction

In the previous chapter, error-prone mutagenesis identified substitutions in the AAA+ domain that allow NorR to bypass repression by the GAF domain. The majority of these substitutions are located in the vicinity of the surface exposed loops that couple ATP-dependent conformational changes in the AAA+ domain to isomerisation of the $E\sigma^{54}$-complex (Bordes *et al.* 2003). Mutations were also identified in the GAFTGA motif that directly contacts σ^{54}. Together, this suggests that in NorR, negative regulation may target the σ^{54}-interaction surface. In other members of the bEBP family, negative regulation has been shown to target oligomerisation (Lee *et al.* 2003; Doucleff *et al.* 2005a). Since AAA+ domain self-association is required to form an activator that is competent to hydrolyse ATP (Zhang *et al.* 2002; Rappas *et al.* 2007), the oligomeric determinants are an ideal target of the regulatory domain in either a positive or negative mechanism of control. The bEBPs NtrC1 and DctD are subject to negative control of AAA+ activity. In these cases, the unphosphorylated receiver domains form a homodimer that maintains the AAA+ domains in an inhibitory front-to-front configuration. Upon phosphorylation, this homodimer undergoes rearrangement, allowing reorientation of the AAA+ protomers into the front-to-back configuration, stimulating the formation of the active oligomer (Figure 3.11). In contrast, in the bEBP PspF, the activity of the central domain is controlled *in trans* by the PspA protein which targets the ATPase hydrolysis machinery (Joly *et al.* 2008a). PspA has been shown to negatively regulate PspF activity in this way via an interaction that is dependent upon a surface-exposed tryptophan residue (W56 of PspF) (Elderkin *et al.* 2002; Elderkin *et al.* 2005). In order to further study the mechanism by which the GAFTGA variants of NorR escape GAF-mediated repression, *in vitro* studies were performed to identify any changes in enhancer DNA-binding, oligomerisation and ATP hydrolysis. This work was carried out to determine whether or not NorR employs a novel mechanism for negatively regulating bEBPs.

7.2 Purification of Δ1-170 forms of NorR

In order to study the biochemical properties of the GAFTGA variants, the mutant versions of NorR were purified. Variants were first purified in full-length form but initial results suggested that substitutions in the AAA+ domain severely reduced the percentage of soluble protein, even when purified via an N-terminal histidine tag (Chapter 8). Therefore, N-terminally truncated versions of NorR that lack the first 170 amino acids were

engineered into the pETNdeM-11 vector. This is a vector based on pETM-11 (appendix section 12.1.2) with a silent mutation to convert the *Nco* I site into an *Nde* I site to allow for easy cloning of the *norR* sequence. The resulting overexpression vectors encoded wild-type and variant forms of NorRΔGAF with additional N-terminal, TEV cleavable, hexahistidine tags. This enabled the protein constructs to be purified to a relatively high concentration using nickel affinity chromatography (Figure 7.1A and B). Typically, wild-type and mutant derivatives of NorRΔGAF-His eluted in the range of 100 mM to 200 mM imidazole. Under these conditions, the protein had a propensity to precipitate. Therefore, 4-5 ml of the eluted fractions were quickly loaded onto a Superdex 200 16/60 column (Amersham Biosciences) which separated NorR from any remaining impurities as well as removing the imidazole from the sample. NorRΔGAF-His and variants thereof typically eluted in the range of 70-80 ml, corresponding to a molecular weight in the monomer-dimer range (Figure 7.1C and D).

7.3   The G266D mutation does not affect enhancer binding of NorR *in vitro*

Since the oligomerisation state and hence the activity of the AAA+ domain of bEBPs is often controlled by the regulatory domain, it was important to question whether the NorR GAFTGA substitutions 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.* 2010a), it was first investigated whether the GAFTGA mutations influence DNA binding. Electrophoretic Mobility Shift Assays (EMSA) were employed to measure the binding of purified wild-type and GAFTGA variants of the truncated (Δ1-170) form of NorR to a 361 bp DNA fragment containing the three enhancer sites upstream of the *norV* promoter. Results showed that the affinity of NorRΔGAF-His for enhancer DNA was not significantly influenced by the presence of the G266D and G266N substitutions (Figure 7.2). Dissociation constants (Kd) were calculated as 100 nM in each case. Therefore, the GAFTGA substitutions do not bypass the GAF-mediated repression of the AAA+ domain by altering the affinity of binding to enhancer DNA.
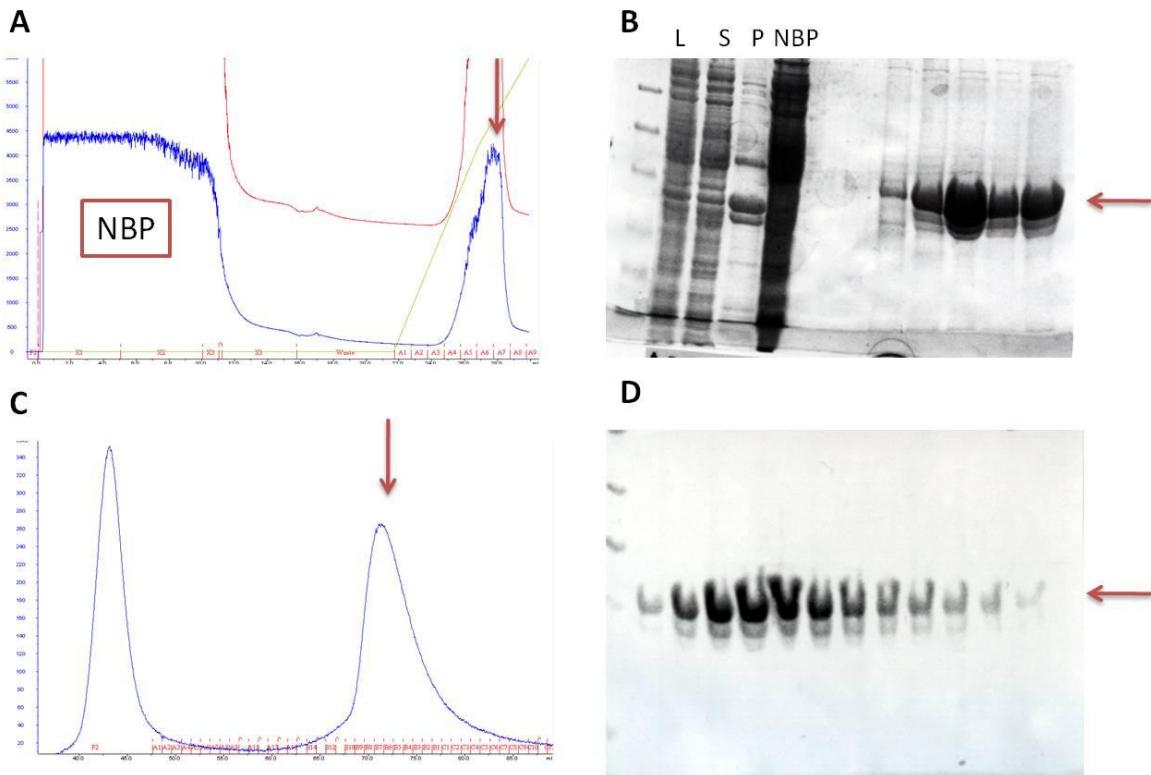
Figure 7.1 – Purification of NorRΔGAF-His by affinity chromatography and gel filtration. (**A**) Nickel affinity chromatography showing non-binding pool (NBP) and protein eluted using an increasing concentration of imidazole. The NorRΔGAF-His bound to the Ni^{2+} column and eluted in the range of 100 mM-200 mM imidazole. (**B**) SDS-PAGE gel of bound protein eluted by increasing imidazole concentrations. L = lysate, S = supernatant, P = pellet, NBP = Non Binding Pool. (**C**) Gel filtration of selected NorR-containing affinity fractions using the 124 ml superdex 200 16/60 column. NorRΔGAF-His eluted at 70-80 ml. The peak around the void volume (43 ml) did not contain any protein (**D**) SDS-PAGE gel of the eluted protein from gel filtration. In each case the presence of the NorRΔGAF-His protein (39.72 kDa monomer) is indicated by a red arrow. The Q304E and G266D variants exhibited similar purification profiles when expressed in a form that lacked the first 170 amino acids.
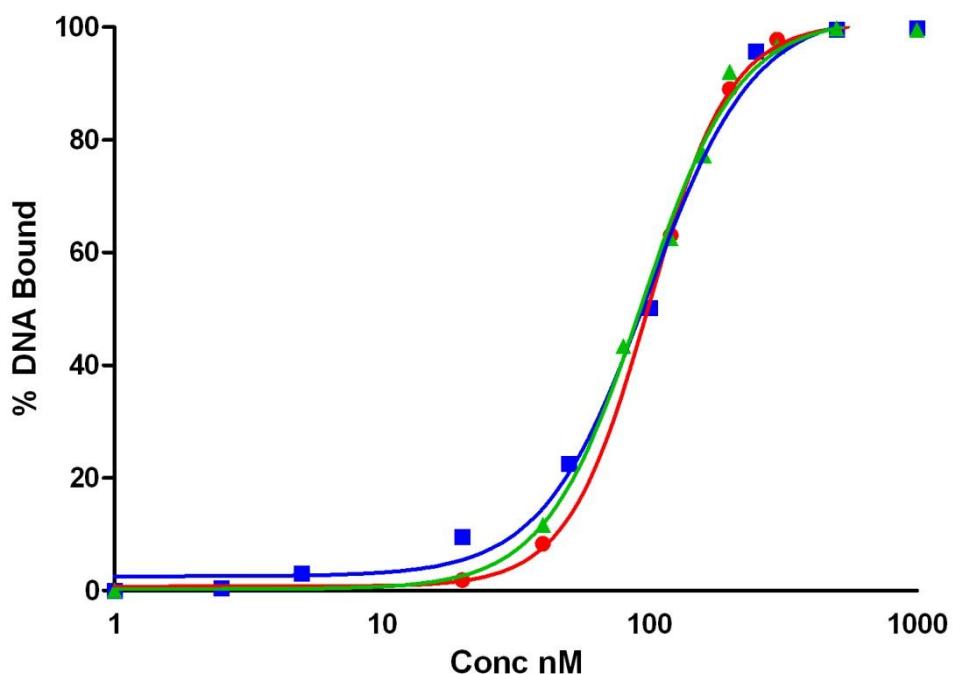
Figure 7.2 - Enhancer binding activity of the G266DΔGAF-His (blue squares) **and G266NΔGAF-His** (green triangles) **variants compared to NorRΔGAF-His** (red circles) **as determined by EMSA**. The percentage of fully shifted DNA was quantified using a Fujix BAS 1000 phosphoimager. The G266 substitutions do not significantly affect the affinity of NorR for the 361bp fragment of the *norR-norVW* intergenic region that contains the 3 NorR binding sites.

7.4   The G266D mutation does not affect oligomerisation of NorR *in vitro*

NorR is known to bind to three binding (or enhancer) sites within the *norR-norVW* intergenic region, each of which is essential for the ability of the bEBP to hydrolyse ATP and form open complexes (Tucker *et al.* 2004; Tucker *et al.* 2005; Tucker *et al.* 2010a). Furthermore, it appears that DNA flanking these sites promotes the formation of the NorR oligomer (Tucker *et al.* 2010a). To determine the effect of the G266D substitution on enhancer dependent NorR oligomer formation, purified protein was sent to two different collaborating research groups. Analytical gel filtration experiments in the absence and presence of a 266bp DNA fragment containing the three enhancer sites were performed by Tamaswati Ghosh of Prof. Xiaodong Zhang's group at Imperial College, London. Based on reference elution volumes obtained with different protein standards, unbound G266DΔGAF-His eluted as an apparent monomer/dimer species (Figure 7.3A). In agreement with this, Electrospray-Mass Spectrometry (ES-MS) experiments performed by Ahyoung Park of Prof. Carol Robinson's group at the University of Oxford, indicated that in the absence of DNA, the G266D and G266N variants of NorRΔGAF are in equilibrium between monomeric and dimeric states (data not shown). Gel filtration experiments showed that the presence of the 266bp DNA fragment shifted the eluted protein peak towards a higher molecular mass species (Figure 7.3A) indicating formation and stabilization of a higher order nucleoprotein complex. These elution profiles are similar to those reported recently for wild-type NorRΔGAF (Tucker *et al.* 2010a). Analysis of the purified protein-DNA complex using negatively-stained electron microscopy, allowed visualisation of higher order ring-shaped particles with dimensions of 125 Å in diameter (Figure 7.3B) consistent with a hexameric ring observed for NorRΔGAF in complex with the 266bp DNA fragment (Tucker *et al.* 2010a). No oligomeric particles were seen in the electron micrographs for protein alone (Figure 7.3C). We conclude from these studies that the G266D mutation does not apparently influence the oligomeric assembly of the AAA+ domain or the requirement for enhancer sites to stabilise the formation of a higher order oligomer.
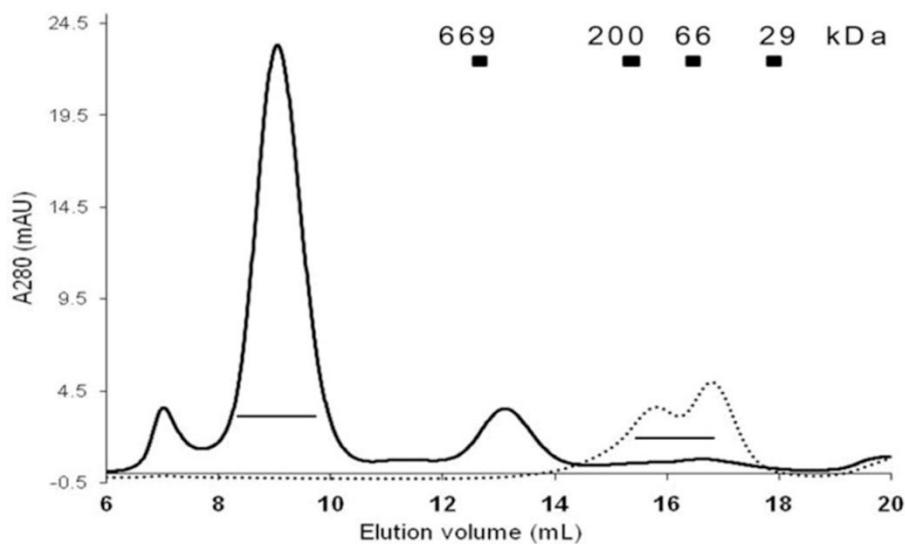
Figure 7.3 - Enhancer-dependent higher order oligomeric assembly of the G266DΔGAF-His variant. (**A**). Gel filtration chromatography of 9 μM G266DΔGAF-His variant in the absence (*dotted line*) and presence (*solid line*) of 0.75 μM 266bp dsDNA (molar ratio of 12:1 monomer : DNA), containing all three enhancer sites, performed at 4 °C using a Superose 6 column (24 ml). The presence of DNA stabilises a higher order oligomeric form of G266DΔGAF-His. The lines below the elution peaks represent the fractions analyzed by negative-stain electron microscopy. Corresponding molecular weight of standard globular proteins are indicated at their elution volume. (**B,C**) Negative-stain EM studies. Shown are raw micrographs of G266DΔGAF-His alone (**C**) and in complex with 266 bp DNA (**B**), scale bar 100 nm. Ring-shaped oligomeric particles were only observed in the presence of DNA. **Experiments were conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

7.5   G266 bypass variants show enhancer-dependent ATPase activity *in vitro*.

In bEBPs the ATP hydrolysis site is configured through interactions between adjacent AAA+ protomers in the hexameric ring (Schumacher *et al.* 2008). Previously, it has been established that enhancer DNA is required for ATP hydrolysis by NorR and that the three binding sites upstream of the *norV* promoter are necessary for activation of ATPase activity, consistent with the requirement for DNA for formation of a functional higher order oligomer (Tucker *et al.* 2010a). Low levels of ATP hydrolysis were observed in the absence of enhancer DNA (Figure 7.4, closed bars/closed squares). Consistent with previous studies with a non-his-tagged form of NorRΔGAF (Tucker *et al.* 2010a), ATPase activity was strongly stimulated by the presence of promoter DNA. Under these conditions ATP hydrolysis by NorRΔGAF-His increased as a sigmoidal response to increasing protein concentration indicative of positive cooperativity, with a lower rate of increase exhibited at concentrations above 250 nM (Figure 7.4A, open bars/open squares). The absence of increased activity at higher protein concentrations may reflect saturation of the enhancer sites consistent with the observed DNA binding constant (100 nM as reported above, Figure 7.2). In order to confirm that the activity observed is due to turnover by NorR and not a contaminating protein, an additional alanine substitution was made at the D286 residue which forms part of the Walker B motif. The DExx residues of this motif are highly conserved in AAA+ proteins and are essential for catalytic activity. Structural studies of PspF$_{1\text{-}275}$ bound to ATP and AMPPNP show that D107 (equivalent to D286 in NorR) is closely positioned to the γ-phosphate and therefore well placed to catalyse ATP hydrolysis. Accordingly, a D107A variant of PspF was completely unable to hydrolyse ATP *in vitro* (Tucker *et al.* 2010a). In line with this, the ΔGAF-His protein with the additional D286A substitution is largely inactive (Figure 7.5). Therefore the turnover measured here is due to the catalytic activity of NorRΔGAF-His and not a contaminating protein.
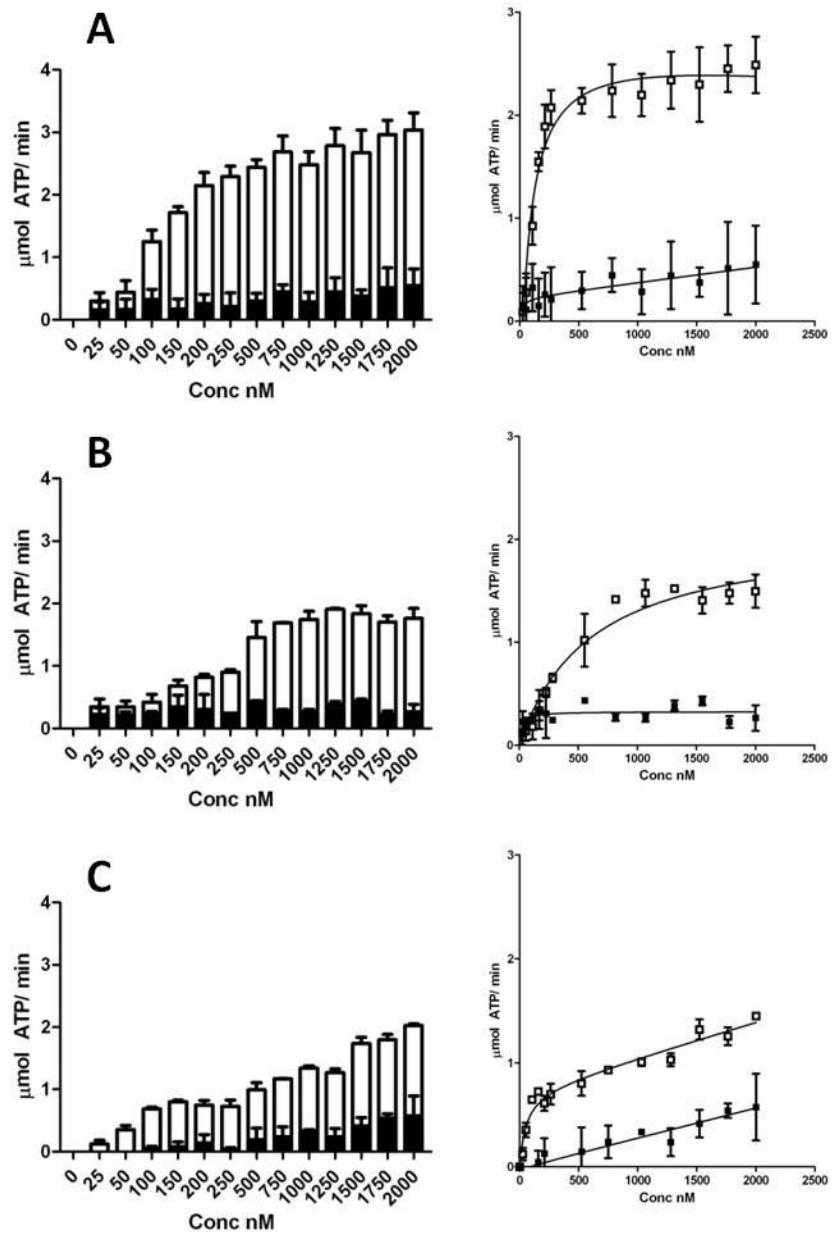
Figure 7.4 - ATPase activity of the NorRΔGAF-His (A), G266DΔGAF-His (B) and G266NΔGAF-His (C) variants in response to protein concentration and the presence of enhancer DNA. Each data set is shown in (stacked) bar chart and graphical form. Non-linear regression was carried out using GraphPad Prism software. Assays were conducted either in the absence (closed bars or closed squares) or presence (open bars or open squares) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.
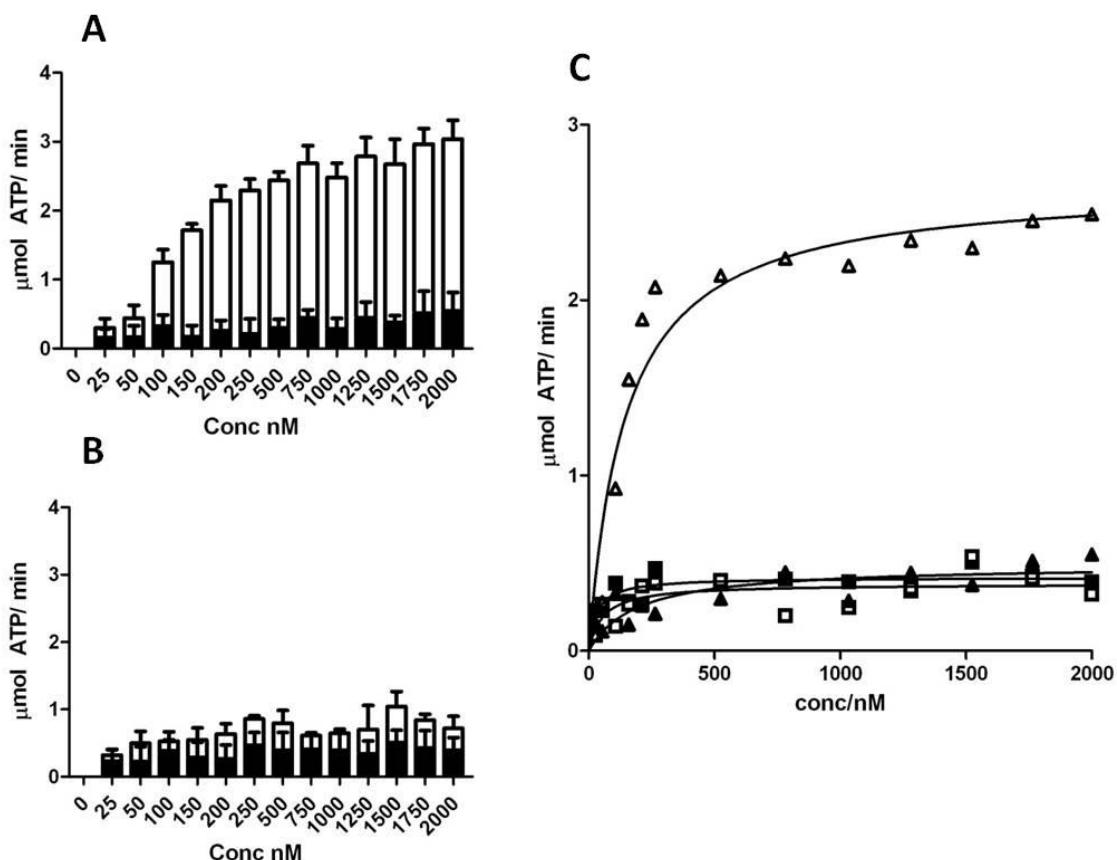
Figure 7.5 - ATPase activity of the NorRΔGAF-His (A) and D286AΔGAF-His (B) variants in response to protein concentration and the presence of enhancer DNA. Each data set is shown as a (stacked) bar chart. (**C**) Graph showing the ATPase activity of NorRΔGAF-His (triangles) and D286AΔGAF-His (squares). For clarity error-bars are not included. Non-linear regression was carried out using GraphPad Prism software. All assays were conducted either in the absence (closed bars or closed shapes) or presence (open bars or open shapes) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

Since the GAFTGA motif relays nucleotide-dependent interactions at the ATP hydrolysis site to enable contact with σ^{54}, the influence of the G266 substitutions upon ATPase activity was next assessed. Results showed that like NorRΔGAF-His, the G266DΔGAF-His and G266NΔGAF-His proteins exhibited enhancer-dependent ATPase activity. In the absence of the 266bp DNA, required for the formation of a stable oligomer, only low levels of ATP hydrolysis were observed (Figure 7.4B and C, closed bars/closed squares). However, in the presence of DNA that contained the three enhancer sites, ATPase activity was strongly stimulated (Figure 7.4B and, open bars/open squares). Furthermore, when the additional D286A substitution was made in a G266DΔGAF construct, the expressed Walker B-mutant derivative was inactive with respect to ATP turnover (Figure 7.6). This indicated that the ATPase activity observed is the result of turnover by the G266D variant and not a contaminating protein. Although ATP hydrolysis by the G266D and G266N variants was also stimulated by the enhancer sites, the response to protein concentration was less cooperative than observed with NorRΔGAF-His and activities were lower than those of the wild-type protein even at a relatively high protein concentration (2 µM) (Figure 7.4, compare A, B and C).

7.6   Testing the requirement for ATPase activity in the NorR variant G266D

Structural studies of PspF have shown that the energy from ATP hydrolysis is used to drive a series of conformational changes in the AAA+ domain that lead to relocation of the surface exposed loops including the GAFTGA-containing loop 1 (Rappas *et al.* 2006). Earlier it was shown that the GAFTGA variants G266D and G266N hydrolyse ATP in an enhancer-dependent manner when expressed in a form of NorR that lacks the N-terminal GAF domain. However, it is possible that a substitution in the L1 loop might promote a conformation of the AAA+ domain that is able to engage with σ^{54}, without the need for ATP hydrolysis. In order to confirm that the GAFTGA variant G266D is still dependent upon the ATPase activity of NorR, the additional D286A change was made to substitute the catalytically-essential Walker B aspartate side chain. Data *in vitro* showed that this residue is essential for ATP turnover by the bEBP since the Walker B-mutant derivative of G266DΔGAF-His was unable to turnover ATP *in vitro* (Figure 7.6). β-galactosidase assays showed that the D286A substitution abolished the activity of the NorR and NorRΔGAF proteins *in vivo* (Figure 7.7). These null phenotypes are presumably due to the inability of the proteins to turnover ATP, required to drive the remodelling of the closed
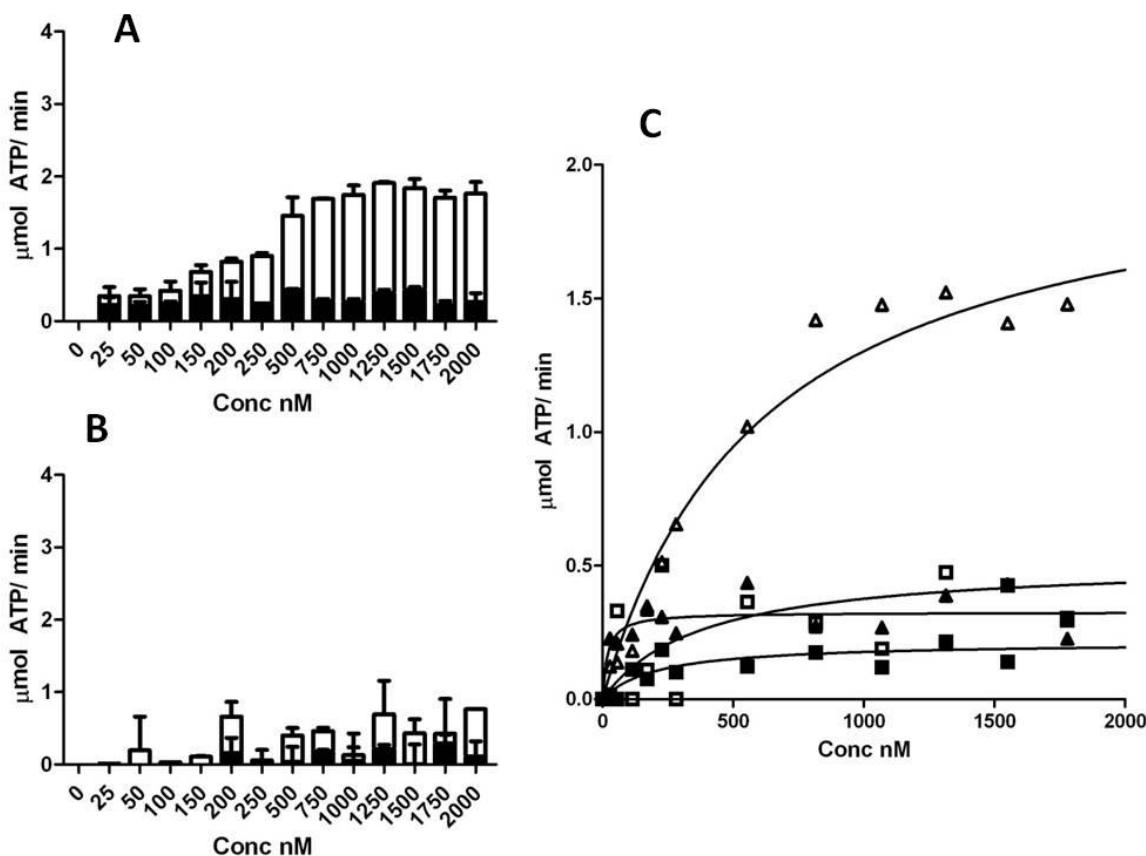
Figure 7.6 - ATPase activity of the G266DΔGAF-His (A) and G266D-D286AΔGAF-His (B) variants in response to protein concentration and the presence of enhancer DNA. Each data set is shown in (stacked) bar chart form. (**C**) Graph showing the ATPase activity of G266ΔGAF-His (triangles) and D286AΔGAF-His (squares). For clarity error-bars are not included. Non-linear regression was carried out using GraphPad Prism software. All assays were conducted either in the absence (closed bars or closed shapes) or presence (open bars or open shapes) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

complex. Likewise, the additional D286A substitution prevented the G266D variant from being able to activate transcription at the *norV* promoter (Figure 7.7). Therefore, the G266D substitution does not negate the requirement for ATPase activity and it can be concluded that the G266D variant of NorR does not escape the GAF-mediated repression mechanism by modelling the post-ATP hydrolysis conformation of the AAA+ domain.

7.7   Negative regulation in NorR does not directly target the ATP hydrolysis machinery

Previously it was shown that the escape mutants G266D and G266N of the GAFTGA motif exhibit enhancer-dependent ATPase activity in a manner similar to the wild-type form of NorRΔGAF-His (Figure 7.4). This would suggest that the GAFTGA variants do not escape the repression mechanism by significantly altering the ATPase behaviour of the AAA+ domain. Furthermore, the ability of the G266D variant to hydrolyse ATP has been shown to be essential to activate transcription, suggesting escape from repression does not involve modelling the conformation of the post-ATP hydrolysis state (Figure 7.7). It is clear that repression by the GAF domain results in an ATPase-inactive protein, since full-length NorR does not exhibit ATPase activity in the absence of NO (D'Autreaux *et al.* 2005). Next, we wanted to provide evidence that the N-terminal GAF domain does not directly target the ATPase hydrolysis machinery in the mechanism of interdomain repression.

In PspF, negative regulation occurs *in trans* by interaction with the PspA protein at the W56 residue of PspF. Structural studies of PspF bound to different nucleotides identified N64 as the key residue that couples ATP hydrolysis to conformational changes in the AAA+ domain of bEBPs (Rappas *et al.* 2006). Although N64 variants were able to bind PspA, the ATPase activity of the bEBP was not reduced (Joly *et al.* 2008a). This suggests that in PspF, negative regulation by PspA at the W56 residue is signalled to the nucleotide machinery, including N64 to prevent ATPase hydrolysis. To explore the role of this conserved asparagine in the regulation of NorR, substitutions were made at the equivalent residue (N243) and NorR activity assessed in the absence and presence of an NO-source in the *E. coli* strain MH1003. In contrast to PspF, N243 variants of NorR were still subject to repression. The N243A and N243S variants were unable to activate transcription in the absence of an NO-source but became active once NO-dependent derepression had occurred
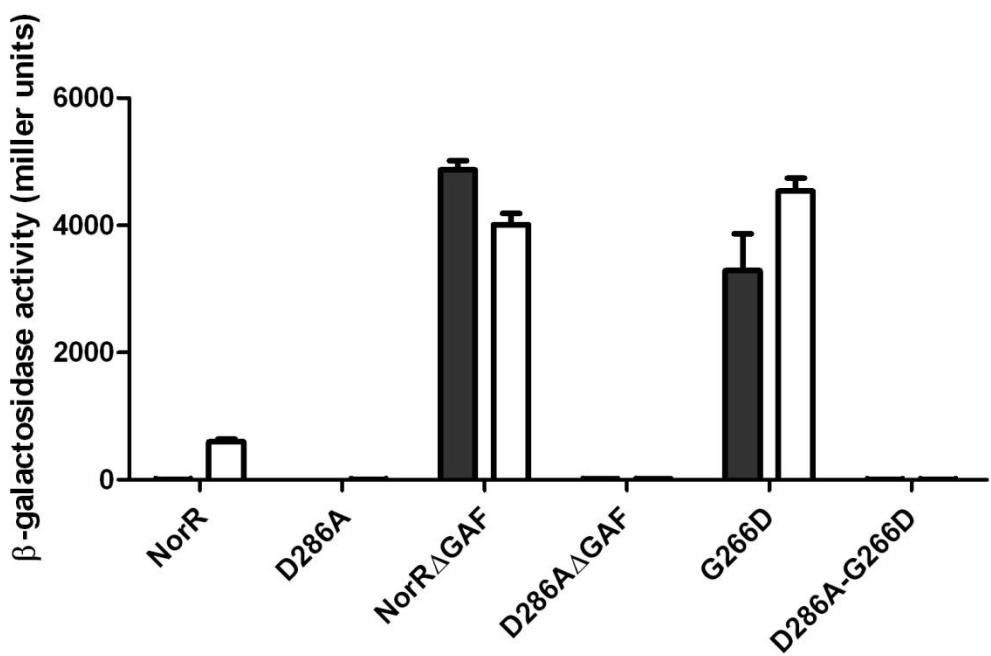
Figure 7.7 - Activities of NorR and the NorR G266D variant *in vivo* when the additional D286A substitution is made at the Walker B motif in the AAA+ domain, as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" 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.
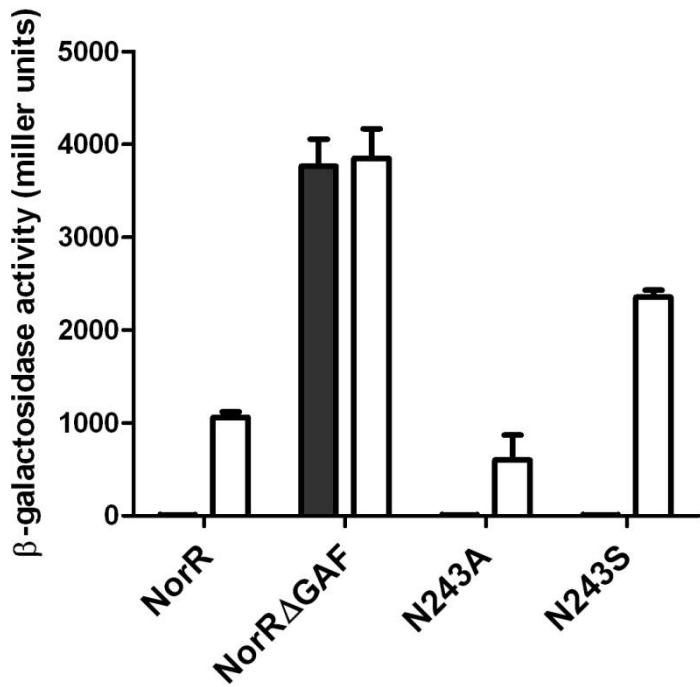
Figure 7.8 - Transcriptional activation by NorR *in vivo*, when substitutions are made at position 243 of the AAA+ domain, as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking 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.

(Figure 7.8). This would suggest that regulation does not occur directly through the nucleotide machinery as it does in PspF but instead occurs via a different mechanism.

7.8   The GAFTGA variants can activate open complex formation *in vitro*

To further test the functionality of the G266 variants *in vitro*, assays were conducted to measure their ability to catalyse the conversion of the σ^{54}-RNA polymerase closed complexes to open promoter complexes. Although NorR-DNA complexes exhibit heparin resistance, open promoter complexes can be visualised as heparin resistant super-shifted species on non-denaturing gels (D'Autreaux *et al.* 2005). In the presence of all the components required for open complex formation, the G266D and G266N variants were competent to form the super shifted species, as in the case of NorRΔGAF (Figure 7.9A compare lanes 3, 5, 7 and 9). Open complex formation was ATP-dependent as expected (Figure 7.9A lanes 2, 4, 6 and 8). In order to probe the nature of the open complexes formed, footprinting of complexes was carried out using potassium permanganate, which targets cleavage to single stranded DNA regions, hence providing sequence-specific information. In all cases enhanced cleavage was observed corresponding to T residues located between -11 to +1 at the *norV* promoter, consistent with the expected footprint (Figure 7.9B). Notably, the band intensity observed with the G266 variants was decreased in comparison with NorRΔGAF or NorRΔGAF-His (Figure 7.9B compare lanes 3,4,5 and 6), a feature that was not clearly visible when the proportion of open complex formed was compared (Figure 7.9A compare lanes 3,5,7 and 9). The reduction in the intensity of the footprint could reflect the slight reduction in total ATP turnover observed for the GAFTGA variants at 1.5 μM when compared to the wild-type (Figure 7.4, compare A, B and C). Despite this, overall the results confirm that the G266 variants are competent to interact with σ^{54} and can activate transcription *in vitro*.

Figure 7.9 - Open promoter complex formation by AAA+ variants. (**A**) Heparin resistant complexes formed by NorRΔGAF, NorRΔGAF-His, G266DΔGAF-His and G266NΔGAFHis, on the 361bp DNA fragment carrying the *norR-norVW* intergenic region. In all cases the final NorR concentration was 1500 nM. Reactions contained no NorR (lane 1), NorRΔGAF (lanes 2 and 3), NorRΔGAF-His (lanes 4 and 5), G266DΔGAF-His (lanes 6 and 7), and G266NΔGAF-His (lanes 8 and 9). Reactions loaded in lanes 1, 3, 5, 7 and 9 contained ATP (final concentration 5 mM which was absent in lanes 2, 4, 6 and 8. Arrows indicate the position of free DNA, NorR bound DNA and the open promoter complexes. (**B**) Potassium permanganate footprinting of the 266bp *norR-norVW* promoter fragment after open complex formation initiated by NorR. Lane 1 is a G+A ladder. Lane 2 is a control without activator present. Lanes 3, 4, 5, and 6 show footprinting after initiation of open complexes in the presence of 1 μM (final concentration) ΔGAF, ΔGAF-His, G266DΔGAF-His and G266NΔGAF-His respectively. The arrow marks the *norVW* transcriptional start and the positions of the enhanced cleavage at T bases are indicated.

7.9 Evidence for direct intramolecular interaction between the GAF domain and the σ^{54}- interaction surface

From the biochemical results presented thus far it seems likely that the GAFTGA mutations do not bypass intramolecular repression solely on the basis of changes in DNA binding, ATP hydrolysis or oligomerisation state. To gain more insight into the nature of the interactions between the GAF and AAA+ domains, a genetic suppression strategy was employed. Previously, mutagenesis of conserved residues in the GAF domain identified the R81L change that when engineered in a construct encoding the GAFTGA variant G266D, suppresses the constitutive phenotype (Chapter 6). To further investigate the role of this residue in the regulation of AAA+ activity, the arginine residue was substituted for each of the other 19 natural amino acids (Figure 7.10A). *In vivo* assays for transcriptional activation by NorR showed that the R81 residue is critical for the negative regulation of the AAA+ domain by the GAF domain. When the side chain was absent, the R81A variant showed significant activity in the absence of an NO-source. Hydrophobic changes (including R81L) resulted in significant constitutive activity whereas negatively charged residues and serine substitutions not only prevented negative control but also stimulated NorR activity beyond wild-type levels. R81D, R81N, R81Q and R81E gave rise to two to three-fold more activity than NorRΔGAF in the *E. coli* strain MH1003. The only substitution that exhibited the normal NO-response was the R81K change, providing further evidence that the R81 residue contributes to the repression mechanism through the formation of a polar contact between the GAF and AAA+ domains.

Since the R81 residue appears to be critical for inter-domain repression, it was decided to investigate whether R81 is required for positioning the GAF domain in the vicinity of the GAFTGA motif. As discussed in Chapter 6, the R81L substitution was able to suppress the constitutive activity of the GAFTGA variant G266D so that the GAF-mediated repression of AAA+ activity in the off-state was almost completely restored (Figure 6.13A). Western blotting analysis showed that the reduction in activity is not due to a decrease in the stability of this double mutant (Figure 6.13B). Interestingly, the R81L mutation had a similar effect on other constitutively active variants located in the key region of the AAA+ domain that is predicted to undergo conformational change upon ATP hydrolysis (Figure 7.11A). As well as reducing the activities of the GAFTGA variants G266D, G266N, G266S and F264Y, the R81L substitution was also able to suppress the "escape"
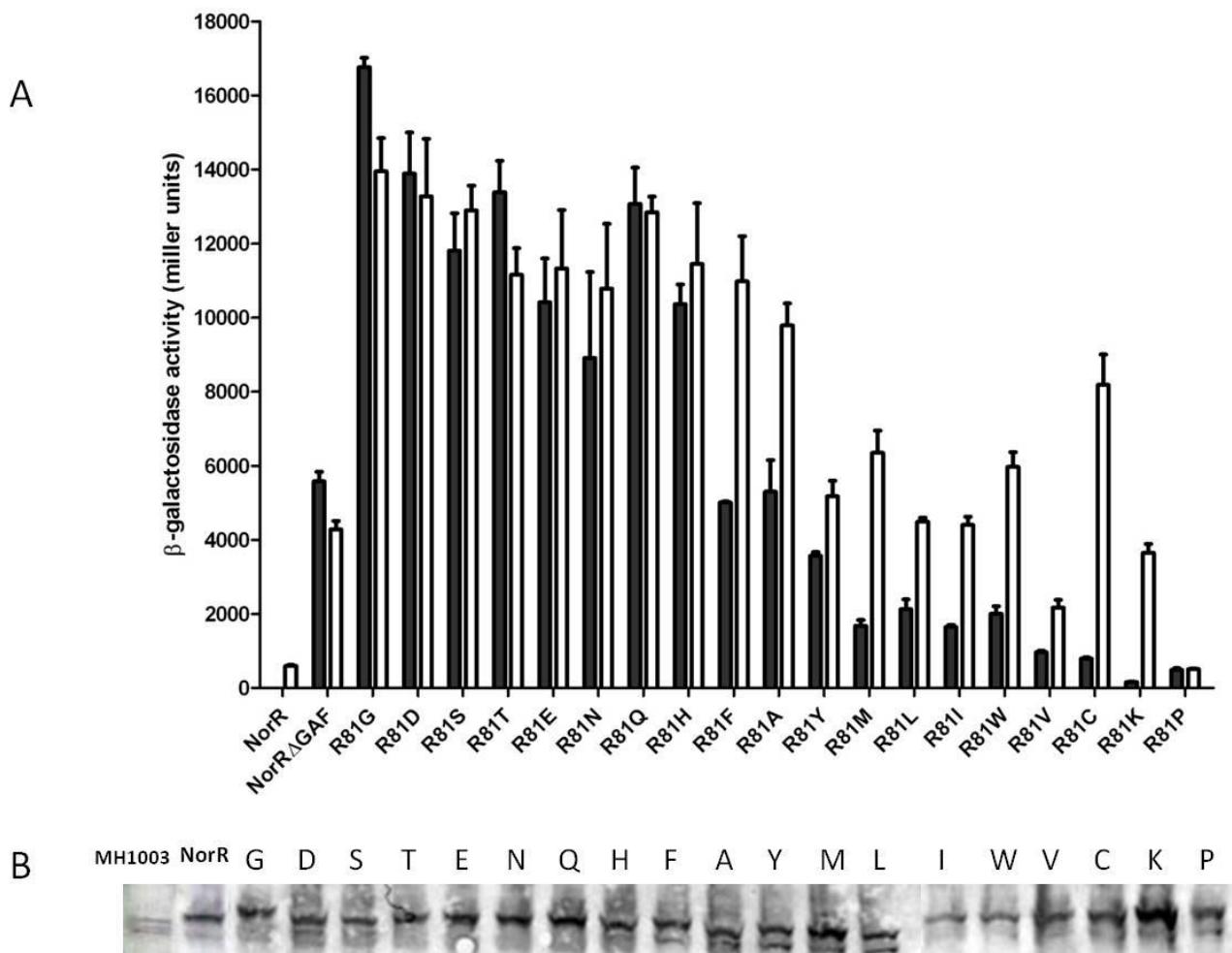
Figure 7.10 - (A) Activities of R81 variants *in vivo* as determined by the *norV-lacZ* reporter assay. "NorR" represents the wild-type protein and "NorRΔGAF", the N-terminally truncated (Δ1-170) protein. Cultures were grown either in the absence (black bars) or presence (white bars) of 4 mM potassium nitrite, which induces endogenous NO production. Error-bars show the standard error of the three replicates carried out for each condition. **(B) Western blot analysis indicating the stability of NorR variants *in vivo* when cultures are grown in the absence of potassium nitrite.** "NorR" refers to the wild-type protein. "MH1003" refers to the *E. coli* strain only. The uppermost band that is not detected in the MH1003 strain correlates to NorR and its variants in the Western analysis.
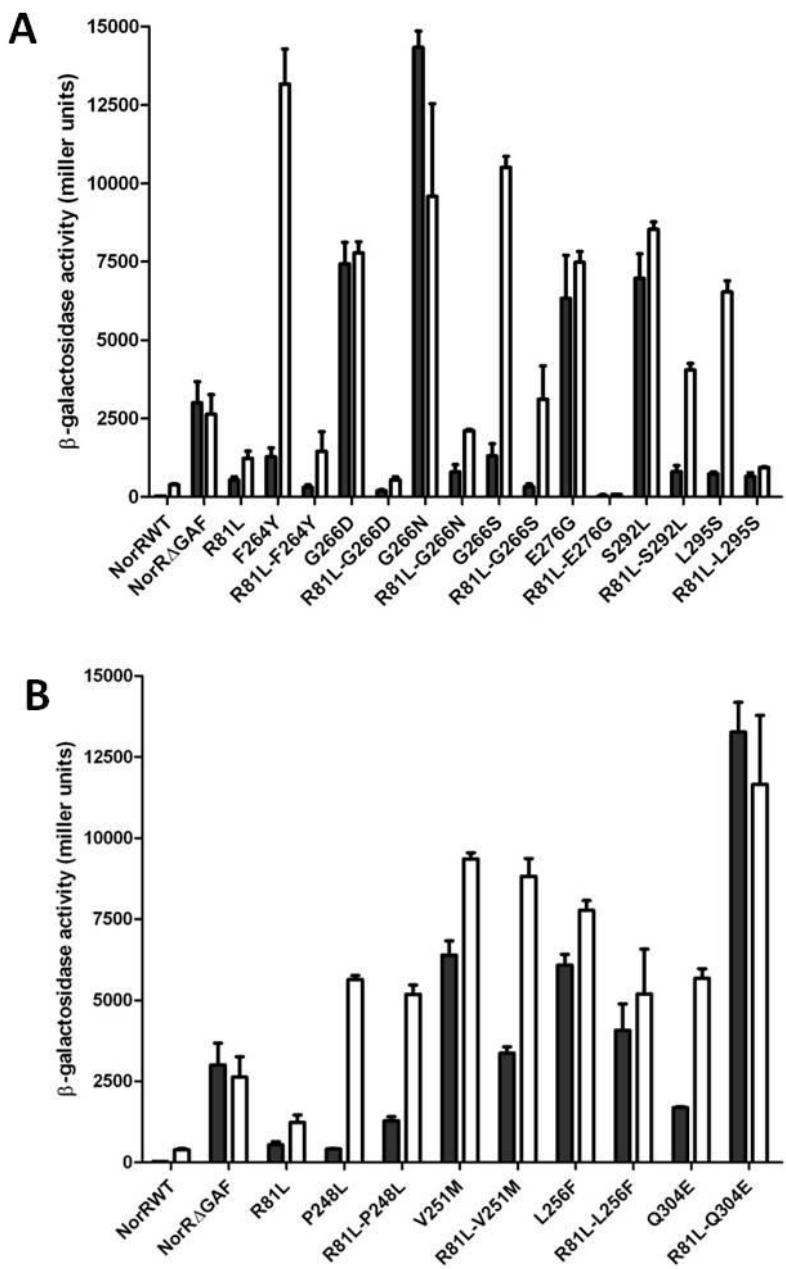
Figure 7.11 - Influence of the R81 substitution on the activity of AAA+ domain variants. (**A**) NorR AAA+ variants that are effectively suppressed by the R81L substitution in the GAF domain. (**B**) NorR AAA+ variants that are not significantly suppressed by the R81L substitution in the GAF domain. "NorR" is the wild-type protein. "NorRΔGAF" is the N-terminally truncated (Δ1-170) protein. Activities *in vivo* (as determined by the *norV-lacZ* reporter assay) were measured 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.

phenotypes of the E276G, S296L and L295S variants *in vivo*. Most noticeable was the suppression of the E276G mutant phenotype. The E276G substitution allows complete escape from GAF-mediated repression but when the additional R81L substitution was present in the regulatory domain, the protein was almost entirely inactive, even in the induced state. This suggests that when both substitutions are present in the two domains that make up the interface of interdomain repression, the NorR protein is unable to make the NO-dependent transition from the off to the on-state. However, a number of other escape variants located in the key region of conformational change in the AAA+ domain were not effectively suppressed (Figure 7.11B). The P248L, V251M and L256F constitutive phenotypes were not reduced when the R81L mutation was additionally present in the regulatory domain. These substitutions are predicted to cluster in the lower region of Helix 3 (H3), furthest away from the GAFTGA-containing loop 1 (L1) and loop 2 (L2) (Figure 6.5). The Q304 residue is predicted to be at the base of helix 4 in the AAA+ domain of NorR and is not expected to have a role in coordinating movements in the GAFTGA loop upon transition to the "on" state. In accordance with this, the Q304E mutation was not suppressed by the R81L substitution. Instead, when combined with Q304E, the R81L substitution enabled complete escape from inter-domain repression (Figure 7.11B).

Next, the specificity of the suppression of the G266D-variant phenotype was determined by making different substitutions at the R81 position. Only hydrophobic changes including R81L, V, I and F were able to suppress the escape phenotype of G266D (Figure 7.13). It is possible that such changes introduce a new hydrophobic contact that helps restore interactions between the GAF and AAA+ domains. Moreover other substitutions such as the charge change R81D have no effect on the constitutive activity of the G266D NorR variant (Figure 7.12). Overall, the constitutive activity of the G266D variant and the specific suppression of this phenotype by hydrophobic changes at the R81 position suggest that the GAF domain may target the GAFTGA motif to prevent σ^{54} contact in the absence of the NO signal. Furthermore, the R81 residue is critical in maintaining the repression mechanism in NorR and may form essential contacts at the interface of interdomain repression.
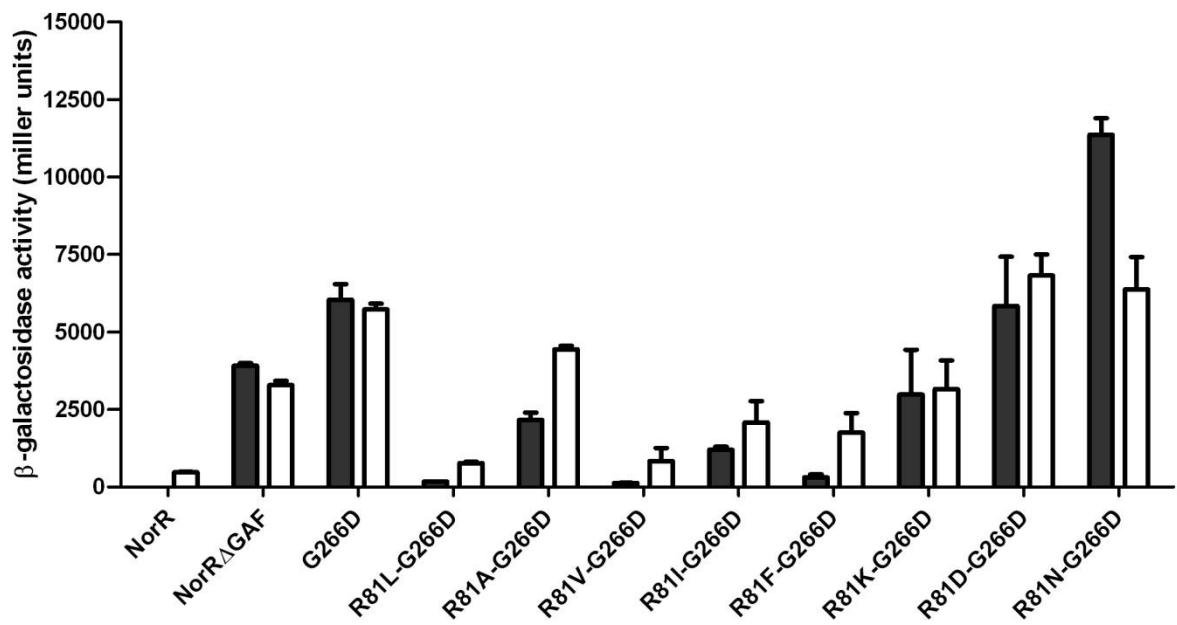
Figure 7.12 - Suppression of the constitutive phenotype of the NorRG266D variant by hydrophobic substitutions made at the R81 residue in the GAF domain. "NorR" is the wild-type protein. "NorRΔGAF" is the N-terminally truncated (Δ1-170) protein. Activities *in vivo* (as determined by the *norV-lacZ* reporter assay) were measured 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.

7.10 Discussion

As discussed in Chapter 6, error-prone mutagenesis of the NorR AAA+ domain identified the GAFTGA variants G266D and G266N that fully escape the mechanism of GAF-mediated repression *in vivo* (Figure 6.8A). Here, biochemical analysis of the GAFTGA variants has been conducted to investigate the mechanism by which these mutant-versions of NorR are able to elude the repression of AAA+ activity. Since self-association of bEBP AAA+ domains is required to form the functional ATPase (Zhang *et al.* 2002; Rappas *et al.* 2007), the oligomeric determinants of the central domain represent an ideal target for the regulatory (GAF) domain in the absence of NO. However, purification of the GAFTGA variants G266D and G266N via an N-terminally His-tagged form of NorR that lacks the first 170 amino acids suggested that the substitutions caused no major change in the oligomeric state. As was the case for wild-type NorRΔGAF, the G266 variants eluted at a volume corresponding to a monomeric/dimeric molecular weight in gel filtration during protein purification (Figure 7.1). This is in full-agreement with ES-MS studies, performed in collaboration, which reveal that *in vitro* samples of the NorRΔGAF, G266DΔGAF and G266NΔGAF proteins, in the absence of DNA, consisted of a monomeric and a dimeric population. Furthermore, gel-filtration followed by Cryo-EM of eluted fractions, performed in collaboration, revealed that the G266D substitution did not alter the enhancer-dependent oligomerisation of NorR, at least in the context of the NorRΔGAF protein (Figure 7.3). Importantly, higher order oligomers were only observed in the presence of DNA that contains the three NorR binding sites. These particles were the expected size of a NorR hexamer, as has been observed for the wild-type protein (Tucker *et al.* 2010a) suggesting that the G266D variant does not exhibit any major changes in oligomerisation. Therefore the explanation that the GAFTGA substitutions bypass the mechanism of interdomain repression by locking the AAA+ domain in a constitutive hexameric oligomerisation state seems unlikely. In the response regulators NtrC1 and DctD, the N-terminal regulatory domain represses the activity of the AAA+ domain by preventing self-association (Figure 3.11B). A structured coiled-coil linker between the regulatory and central domain stabilises an inactive dimeric conformation of the AAA+ domain. Phosphorylation enables the rearrangement of dimers from a front-to-front into a front-to-back configuration, allowing oligomerisation to take place (Lee *et al.* 2003; Doucleff *et al.* 2005a). In agreement with data here suggesting the mechanism of repression in NorR does not target oligomeric determinants, the linker region between the GAF and AAA+ domains of NorR is not predicted to form a coiled-coil helix, a structural

feature that is also absent in negatively regulated NtrC4 and positively regulated NtrC (Batchelor *et al.* 2008).

In agreement with the ability of the GAFTGA variants to self-associate in an enhancer-dependent manner, here the G266D and G266N mutant-versions of NorR also hydrolysed ATP in the presence of a 266bp section of the *norR-norVW* intergenic DNA (Figure 7.4). The slight reduction in total ATP turnover and decrease in cooperatively in response to protein concentration cannot be explained by a decrease in the affinity of binding to enhancer DNA since the GAFTGA variants and wild-type NorR all bound to a 361bp section of the *norR-norVW* intergenic with a Kd of ~ 100 nM (Figure 7.2). Since the enhancer DNA is likely to be fully saturated with protein at concentrations above 300 nM, the G266 substitutions may alter the stability of the nucleoprotein complexes, perhaps by influencing protomer interactions that impact upon the ATP hydrolysis site. When the ability of the GAFTGA variants to activate transcription *in vivo* as well as to turnover ATP i*n vitro* is considered, it is not surprising that they were also able to form open complexes *in vitro* (Figure 7.9A). In accordance with the ATPase data, potassium permanganate footrpinting revealed a slight decrease in the intensity of the expected footprint upstream of the transcriptional start site (Figure 7.9B). Because the NorRG266D and NorRG266N proteins are still capable of hydrolysing ATP in a manner strictly dependent upon the presence of enhancer DNA, it is unlikely that such variants exhibit escape phenotypes *in vivo* by preventing repression of ATP hydrolysis. In the case of PspF, which does not contain an amino-terminal regulatory domain, the activity of the AAA+ domain is negatively controlled in response to the PspA protein (Figure 3.12). Repression does not seem to be achieved by controlling oligomerisation of the AAA+ domains, but rather by inhibition of ATP hydrolysis (Joly *et al.*, 2009). Inhibition is mediated by the interaction of PspA with a surface exposed tryptophan residue (W56) on PspF, which is likely to communicate with the ATP hydrolysis site. The N64 residue in PspF was shown to play an important role in the glutamate "switch" that functions to translate nucleotide hydrolysis the remodelling of the surface-exposed L1 and L2 loops (Rappas *et al.* 2006). Importantly, N64 variants are still able to bind PspA, but their ATPase activity is no longer inhibited (Joly *et al.* 2008a) suggesting that negative regulation by PspA at the W56 residue is directly signalled to the nucleotide machinery via N64 to prevent ATPase hydrolysis by PspF. In contrast, the N243A and N243S variants in NorR were still subject to regulation by the N-terminal GAF domain *in vivo* (Figure 7.8), suggesting that the ATP hydrolysis

site is not the direct target of the GAF domain in the mechanism of negative control in NorR.

The data here suggests that the mechanism of interdomain repression in NorR does not regulate the activity of the AAA+ domain through the control of self-association or via the direct targeting of the ATP hydrolysis machinery. This leads to an alternative hypothesis in which the N-terminal GAF domain prevents interaction between the AAA+ domain of the activator and σ^{54} in the absence of NO. In agreement with this model, the majority of substitutions identified in the random PCR mutagenesis of the AAA+ domain (Chapter 6) are predicted to be located in a region of the AAA+ domain that undergoes conformational change to couple ATP hydrolysis to the relocation of the surface-exposed L1 and L2 loops. The role of the σ^{54}-interaction surface in negative regulation by the GAF domain is further supported by genetic suppression data. The R81 residue in the GAF domain appears to play a crucial role in the repression mechanism since an alanine substitution at this position lead to significant constitutive NorR activity whilst the R81D and R81E charge changes gave rise fully constitutive phenotypes (Figure 7.10). Hydrophobic substitutions, particularly leucine, restored repression only when combined with specific bypass mutations in the AAA+ domain, including those in the GAFTGA loop (Figure 7.11A). Previous structural modelling of the GAF domain suggested that the R81 residue is surface exposed (Tucker *et al.* 2007) and therefore well placed to make contact with the AAA+ domain. In the model, it is located at the opposite end of an α-helix to the R75 residue (Figure 7.13), which is postulated to be a ligand to the hexa-coordinated iron and is the most suitable candidate to be displaced upon NO binding (Tucker *et al.* 2007). Therefore it is possible that formation of the mononitrosyl iron complex would displace the R75 ligand causing a conformational change in the helix that repositions R81. Interactions between the R81 residue and residue(s) in the AAA+ domain may thus facilitate the switch from the "off" to the "on" state.

In conclusion, NorR represents another mechanism of negative regulation in which the N-terminal regulatory domain targets the σ^{54}-interacting region of the AAA+ domain that includes the GAFTGA motif. In the case of those "rare" bEBPs that contain a naturally occurring aspartate or asparagine residue at the second glycine of this motif (Figure 6.14), it is anticipated that the AAA+ domain of such proteins is still subject to regulation by the N-terminal domain. Indeed, the FlgR protein from *H. pylori* contains a "GAFTDA" motif

but still requires phosphorylation by FlgS in order to activate transcription (Brahmachary *et al.* 2004). Therefore these bEBPs are not likely to be regulated by targeting the σ^{54}-interaction surface as may be the case in NorR; instead control may be at the point of oligomerisation or ATP hydrolysis.
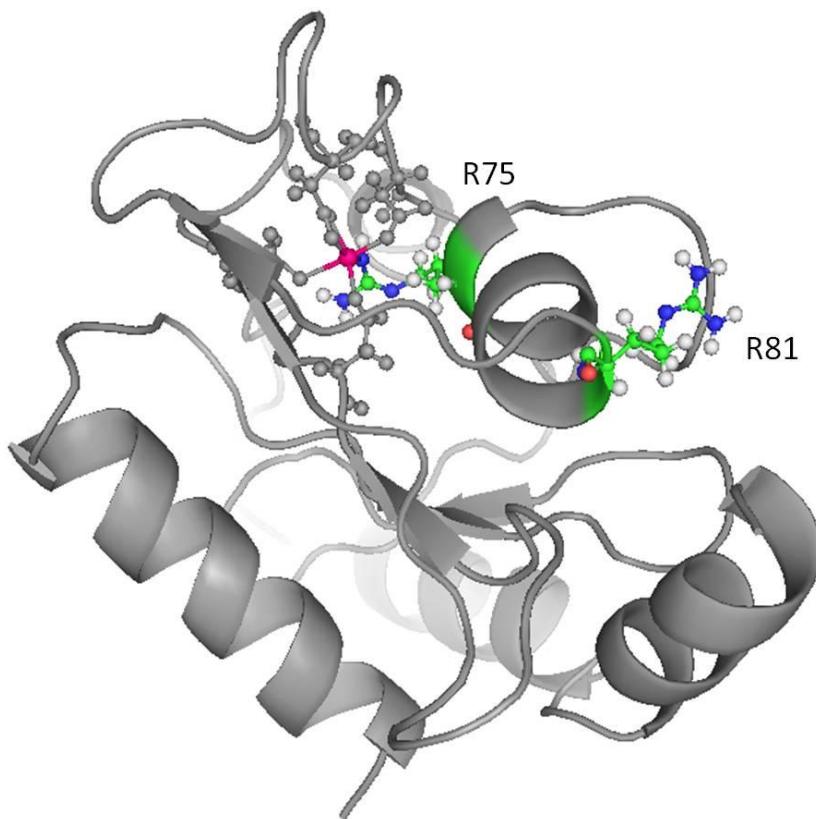
Figure 7.13 - The role of the R81 residue in the mechanism of interdomain repression in NorR. Structural model of the GAF domain of NorR based on the GAF-B domain of 3',5'-cyclic nucleotide phosphodiesterase PDB ID: 1MC0 (Martinez *et al.* 2002; Tucker *et al.* 2007) showing the iron centre and proposed ligands, relative to the R81 residue. The iron centre is indicated (magenta) as well as the proposed ligands C113, D96, D99, and D131 (in grey ball and stick). The fifth ligand, R75 and the R81 residue are highlighted (ball and stick: Carbon = green; Nitrogen = blue; Oxygen = red; Cysteine = orange; Hydrogen; grey). The R75 residue is the most likely to undergo ligand displacement upon NO binding. The structural model predicts that R81 is surface exposed and at the opposite end of an α-helix that also contains the R75 ligand. Therefore, NO-binding and subsequent displacement of R75 may lead to conformational changes that involve re-positioning of the R81-containing α-helix. This may cause disruption of GAF-AAA+ interactions (that could involve R81), leading to the activation of NorR.

Chapter 8 - *In vitro* studies of the full-length GAFTGA-variant G266D

8.1   Introduction

In Chapter 6, the GAFTGA substitution G266D was identified that enabled the NorR protein to bypass GAF-mediated repression *in vivo*. *In vitro* studies of this unique variant together with genetic suppression data *in vivo*, support the hypothesis that the N-terminal GAF domain targets the σ^{54}-interaction surface in the mechanism of interdomain repression (Chapter 7). Such work employed a version of the NorR G266D variant purified via an N-terminal His-tag in a construct that lacked the sequence encoding the first 170 amino acids (G266DΔGAF-His). This gave a high concentration of pure protein (100-150 µM) that was amenable to biochemical studies (Figure 7.1). In contrast, the full-length form of the G266D variant was significantly less soluble than G266DΔGAF even when purified via a hexa-histidine tag.

Recently, the effect of different domain combinations on the overall oligomeric state of bEBPs has been investigated in NtrC4 from *Aquifex aeolicus* (Batchelor *et al.* 2009). Electrospray mass spectrometry (ES-MS) experiments showed that full-length NtrC4 forms hexamers but that the isolated AAA+ domain forms heptamers. Likewise structural studies of the AAA+ protein p97 have revealed rings with 6-fold symmetry for the intact protein (Zhang *et al.* 2000) but with 7-fold symmetry for a truncated form of the protein (Davies *et al.* 2008). Taken together, these studies stress the importance of examining the oligomeric state of a full-length, intact AAA+ protein or bEBP. In the previous chapter, comparisons between the wild-type and GAFTGA-variant forms of NorR were made in a construct that lacked the N-terminal regulatory domain. Therefore it was decided to assess the role of the GAF domain in the GAFTGA variant by characterising the full-length G266D protein *in vitro*. In order to do this, further purification trials were carried out to maximise the concentration and purity of protein available for biochemical studies of the full-length G266D-His protein.

8.2 Purification of full-length (1-504) GAFTGA variants

Initially, the full-length form of the G266D variant protein was overexpressed and purified without a tag using heparin affinity chromatography. However, unlike wild-type NorR which has been previously purified using this method (D'Autreaux *et al.* 2005; Tucker 2005), the GAFTGA variant was largely insoluble when overexpressed in its non-tagged form (Figure 8.1). Therefore, the *norR* sequences encoding the GAFTGA substitutions G266D and G266N were moved into the pETNdeM11 vector to enable purification via an N-terminal, TEV-protease cleavable, hexahistidine tag. To increase the solubility of the GAFTGA-variant proteins further, overexpression was carried out at lower temperature. This is thought to help the large amounts of protein being overexpressed in the cell to fold correctly, increasing solubility. *E. coli*, BL21(DE3) cells expressing the G266D-His or G266N-His proteins were grown to an OD_{600} of 0.6 and then immediately cooled on ice before the addition of IPTG to a final concentration of 0.5 mM. Cells were then left shaking overnight for 12-16 hours at 7 °C. The next day, cells were harvested and resuspended in breaking buffer containing EDTA-free protease inhibitors (Roche); pellets were stored at -80 °C until required.

Purification of the His-tagged full-length GAFTGA-variant proteins was more successful than the non-tagged versions (Figure 8.2). Overexpression produced high levels of protein and although 90 % of the expressed protein was insoluble, a significant amount was able to bind to the nickel column. The GAFTGA variants G266D and G266N eluted in the range of 100 mM – 200 mM imidazole as is typical for NorR and its mutant derivatives. In contrast to other NorR variants, the G266 mutant proteins were not prone to precipitation under these conditions. It is likely that the precipitation usually observed in the presence of imidazole is a concentration-dependent phenomenon and the lower concentration of the full-length G266 variants when purified means that precipitation does not occur so readily. In order to remove the imidazole and separate contaminants from the NorR-containing fractions, gel filtration was carried out using a Superdex 200 16/60 column (Amersham Biosciences). Significantly, the G266 variants eluted over a broad range of volumes around the void volume (43 ml). This suggests the formation of a large protein complex. Indeed, gel filtration standards (BioRad) revealed that the eluted peak centred on a volume corresponding to ~1020 kDa or ~17 NorR monomers (data not shown). In order to minimise aggregation, the salt concentration was increased (200 mM) to produce a
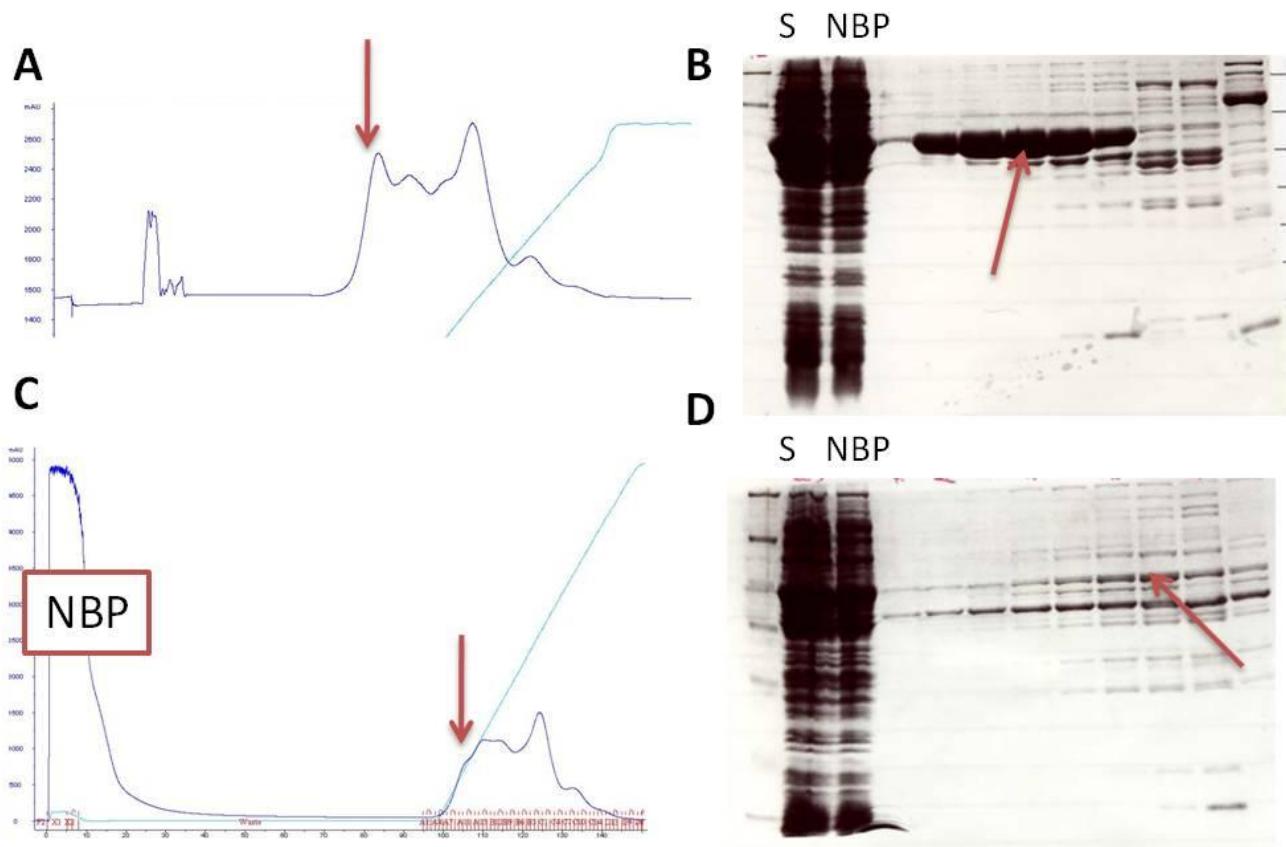
Figure 8.1 – Purification of full-length NorR and G266D by affinity chromatography. (**A**) Heparin affinity chromatography for NorR and G266D (**C**) showing non-binding pool (NBP; not shown for NorR) and protein eluted using an increasing concentration of NaCl. (**B**) SDS-PAGE gel of bound NorR and G266D (**D**) protein eluted by increasing NaCl concentrations. S = supernatant, NBP = Non Binding Pool. The presence of the protein (55.25 kDa monomer) is indicated by a red arrow. Both the wild-type NorR and variant G266D proteins eluted in the range of 200 mM-500 mM NaCl but the G266D protein was significantly less soluble. NorR-containing fractions were subsequently loaded onto a 124 ml superdex 200 16/60 column for gel filtration. Wild-type NorR eluted in the range of 65-75 ml but pure fractions of the G266D variant were not obtained (data not shown)
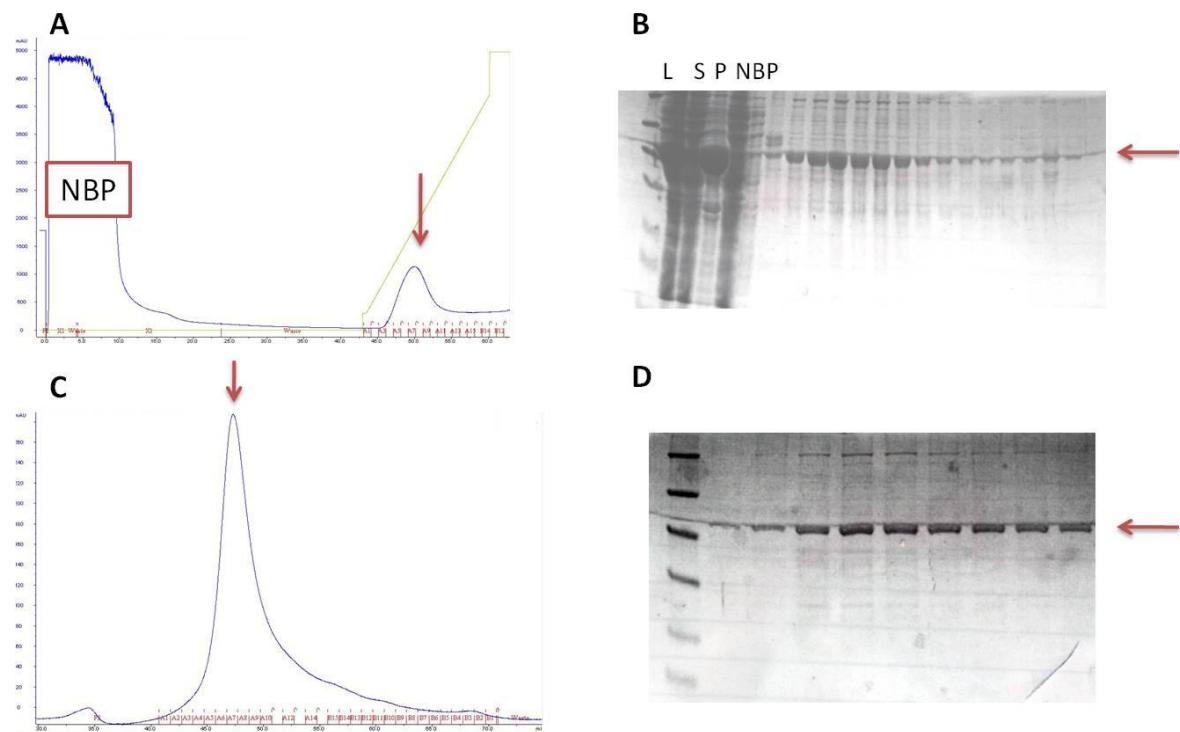
Figure 8.2– Purification of NorRG266D-His by affinity chromatography and gel filtration. (**A**) Nickel affinity chromatography showing non-binding pool (NBP) and protein eluted using an increasing concentration of imidazole. G266D-His bound to the Ni^{2+} column and eluted in the range of 100 mM-200 mM imidazole. (**B**) SDS-PAGE gel of bound protein eluted by increasing imidazole concentrations. L = lysate, S = supernatant, P = pellet, NBP = Non Binding Pool. (**C**) Gel filtration of selected NorR-containing affinity fractions using the 124 ml superdex 200 16/60 column. G266D-His eluted over a broad range of fractions around the void volume (43 ml). (**D**) SDS-PAGE gel of the eluted protein from gel filtration. Purified protein was at a low concentration (1-3 µM) and contained a small proportion of impurities. In each case the presence of the G266D-His protein (58.5 kDa monomer) is indicated by a red arrow.

"salting-in" effect, in which protein solubility is increased slightly. Here ions from the salt associate with the surface of the protein, increasing the concentration of "free" water. The thiol-containing reducing agent, dithiothreitol (DTT), was also added to the gel filtration buffer (8 mM) in an attempt to separate multimers in what was anticipated to be a large oligomeric assembly. However, neither the presence of salt nor DTT was effective in preventing this larger "complex" from forming. The resulting G266D and G266N-containing fractions were of a low concentration (1-3 μM) and contained a small proportion of impurities. If higher concentrations were required, fractions could be concentrated using Amicon Ultra (Millipore) centrifugal units, span at a lower than recommended speed to prevent protein precipitation.

8.3   The full-length G266D variant forms enhancer-independent, higher order oligomers *in vitro*

To determine the effect of the G266D substitution on the oligomeric state of NorR in the context of the full-length protein, analytical gel filtration experiments using the G266D-His protein were carried out by Tamaswati Ghosh from Prof. Xiaodong Zhang's group at Imperial College, London. In agreement with size-exclusion chromatography, performed as part of the purification process, G266D-His eluted at a volume corresponding to a high molecular weight. Using the superdex 200 16/60 preparative column (Amersham Biosciences), the protein eluted over a broad range close to the void volume (Figure 8.2). The superpose 6 analytical column (Amersham Biosciences) confirmed that the protein is likely to form a complex with a high molecular weight but gave better separation of this species from the void volume. At 3 μM, the major peak was at 10.5 ml, corresponding to a complex larger than the 669 kDa molecular weight marker, thyroglobulin. Although an accurate molecular weight cannot be measured using this technique, based on the monomer (58.5 kDa), this volume corresponds to a complex greater than 11 monomers in size. In order to assess the impact of enhancer binding upon the oligomerisation state of the full-length variant protein, nucleoprotein complexes were prepared by our collaborators at Imperial College, London, using purified G266D-His and a 266bp fragment of DNA carrying the *norR-norVW* intergenic containing the 3 NorR binding sites. Analytical gel filtration revealed that the G266D-His protein in complex with the 266bp DNA gave a major peak at 9.5 ml. This volume is similar to that of the nucleocomplex formed between the G266DΔGAF-His protein and intergenic DNA (Figure 7.3A). To more accurately
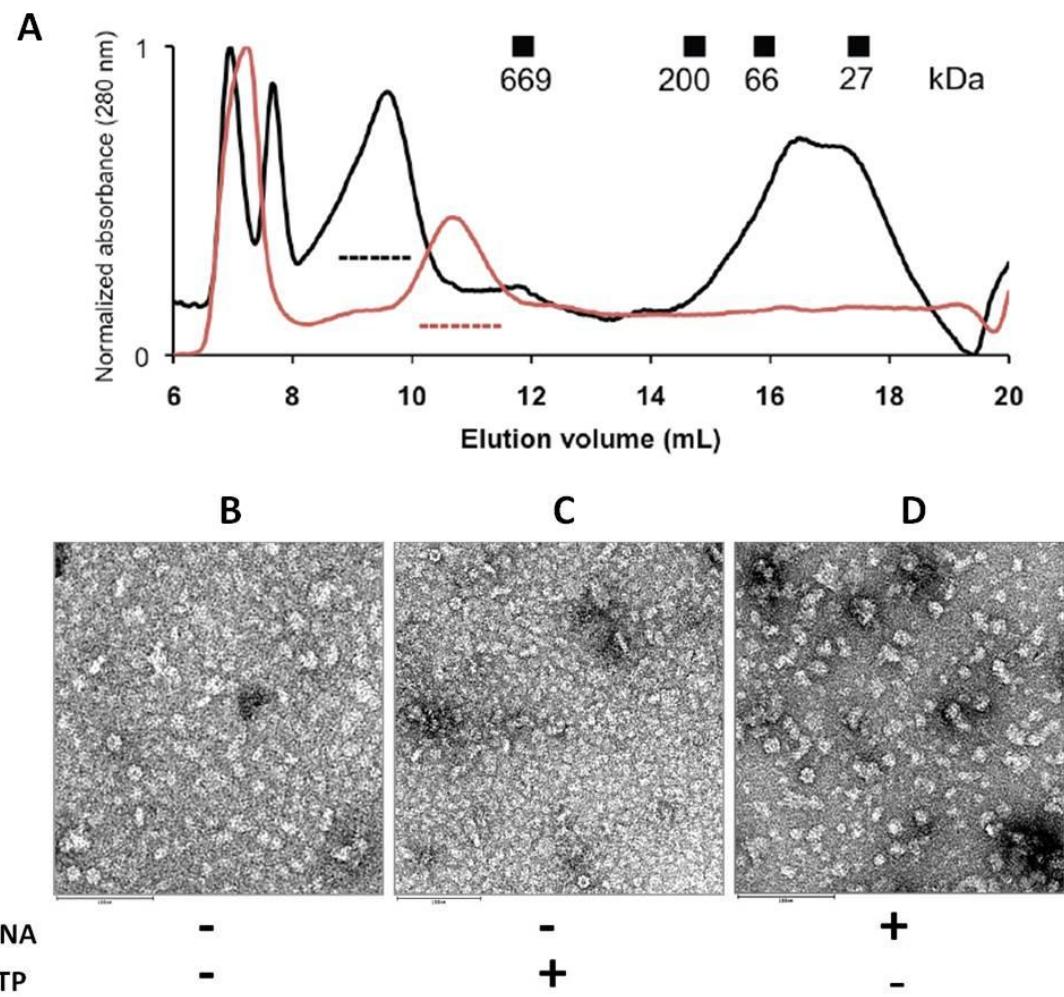
Figure 8.3 - Enhancer independent higher order oligomerisation of G266D-His. (**A**) 3 µM NorR-G266D alone (*red line*) or in complex with the 266 bp DNA fragment at a molar ratio of 12:1 monomer:DNA molar ratio (*black line*) analysed by gel filtration at 4 °C using a Superose 6 column (24 ml). The protein eluted as a higher order oligomer at 10.5 ml and the nucleoprotein complex elutes at 9.5 ml (dashed lines). (**B**), Negative-Stain Electron Microscopy (EM) of the G266D-His (full-length) protein in the absence of the 266bp fragment of the *norR-norVW* intergenic region and ATP, in the presence of ATP (**C**) and in the presence of DNA (**D**). Ring-shaped oligomeric particles were observed under all conditions. Scale bar 100 nm. **Experiments were conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

examine the oligomeric state of the full-length G266D-His protein in the presence and absence of enhancer DNA, fractions collected from analytical gel-filtration of G266D-His alone and in complex with 266bp DNA (Figure 8.3A, dashed lines) were subjected to negative-stain electron microscopy. Analysis of the protein and nucleoprotein complexes by negative-stain EM revealed the formation of large oligomeric particles irrespective of the presence of enhancer DNA (Figure 8.3B). This is in stark contrast to the G266DΔGAF-His protein which was previously shown to oligomerise in an enhancer-dependent manner (Figure 7.3), as is observed for NorRΔGAF (Tucker *et al.* 2010a). To confirm that the GAFTGA-variant is not required to bind to enhancer DNA in order to form this higher oligomeric assembly, a truncated protein was purified lacking the C-terminal DNA-binding domain (Δ442-504) was expressed and purified. Gel filtration data indicated that this version of the protein still eluted over a broad range of volumes close to the void volume of the Superdex 200 16/60 column (Amersham Biosciences) (data not shown). Therefore, the G266D-His variant does not appear to need to bind to enhancer DNA in order to form the high molecular weight complex observed by gel filtration and Cyro-EM *in vitro*. Overall the results demonstrate that the GAFTGA variant of NorR is able to oligomerise in the absence of enhancer DNA but only when the regulatory GAF domain is present. The fact that both the wild-type and G266D variant forms of ΔGAF-His oligomerise in an enhancer-dependent manner suggests that the enhancer-independent oligomerisation of full-length G266D-His is due to the substitution at position 266 and not simply due to the presence of the regulatory GAF-domain. To confirm this, full-length NorR-His was purified and subjected to analytical gel filtration and negative-stain EM analysis. Compared to the G266D variant, wild-type, full-length NorR-His was significantly more soluble and gel filtration gave pure protein at high concentration (10-15 µM) (Figure 8.4). Significantly, the protein did not elute as part of the void volume (45 ml) of the superdex 200 16/60 preparative column (Amersham Biosciences), as was the case for the G266D-His protein. Instead, the protein eluted at 60-70 ml, roughly corresponding to the volume of a NorR trimer. This is in agreement with previous size-exclusion chromatography studies using non-tagged NorR (Tucker 2005). Furthermore, cross-linking studies revealed the presence of a species roughly three times the size of a NorR monomer (Justino *et al.* 2005a). Although, this trimeric species may represent an intermediate of higher oligomeric assembly, it is not expected to have a biological role given that bEBPs are known to form larger rings. After protein purification, analytical gel filtration was performed by our
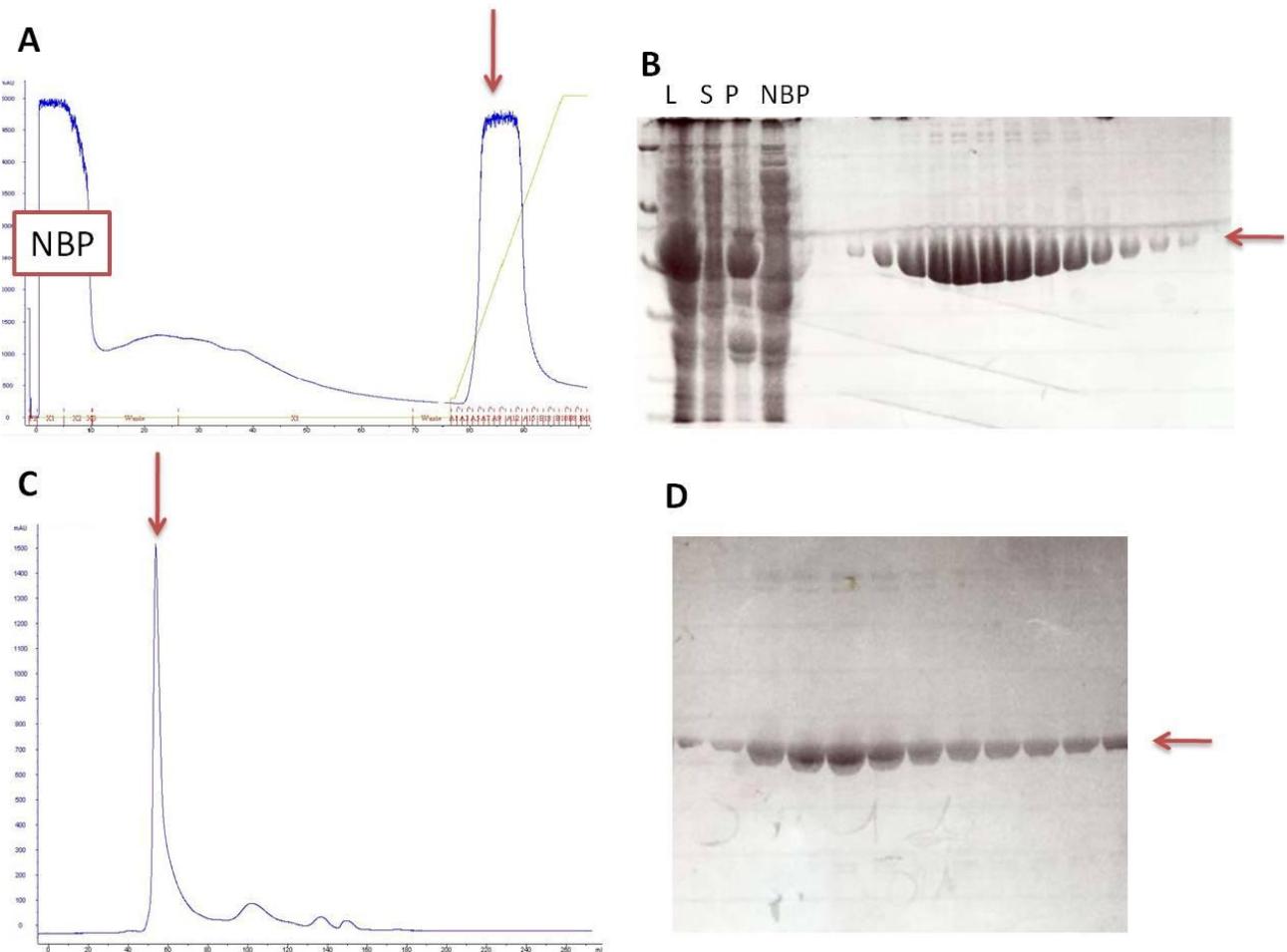
Figure 8.4 – Purification of NorR-His by affinity chromatography and gel filtration. (**A**) Nickel affinity chromatography showing non-binding pool (NBP) and protein eluted using an increasing concentration of imidazole. NorR-His bound to the Ni^{2+} column and eluted in the range of 100 mM-200 mM imidazole. (**B**) SDS-PAGE gel of bound protein eluted by increasing imidazole concentrations. L = lysate, S = supernatant, P = pellet, NBP = Non Binding Pool. (**C**) Gel filtration of selected NorR-containing affinity fractions using the 124 ml superdex 200 16/60 column. NorR-His eluted at 50-55 ml. (**D**) SDS-PAGE gel of the eluted protein from gel filtration. In each case the presence of the NorR-His protein (58.44 kDa monomer) is indicated by a red arrow.

collaborators at Imperial College, as was carried out for the variant protein G266D-His. Results indicate that the wild-type, full-length protein forms oligomers in a manner that is strictly dependent upon the presence of enhancer DNA. Analytical gel filtration using the superpose 6 column (Amersham Biosciences) revealed the formation of a nucleoprotein complex that elutes at 9.5 ml (Figure 8.5A), as is seen for the G266DΔGAF-His (Figure 7.3) and G266D-His (Figure 8.3) proteins. Subsequently, negative-stain EM analysis revealed the presence of oligomeric ring-like particles only in the presence of a 266bp section of the *norR-norVW* intergenic region (Figure 8.5B). The requirement for enhancer DNA for oligomerisation of wild-type NorR is in agreement with recent work indicating that the *norR-norVW* intergenic DNA acts as a key ligand for the activation of NorR as a transcription factor (Tucker *et al.* 2010a). The binding of DNA may lead to conformational changes, translated to the AAA+ domain via the C-terminal domains, to promote oligomerisation. In the case of full-length G266D, it appears that inter-protomer interactions that are dependent on the regulatory domains can take place without the requirement for DNA-induced conformational changes. However, in the absence of higher-resolution structural information, it is not known whether the G266D substitution stimulates the same inter-protomer contacts that are DNA-dependent in wild-type NorR.

8.4   Testing the requirement for enhancer binding in the GAFTGA-variant G266D *in vivo*

Negative-stain electron microscopy of the full-length GAFTGA variant has revealed that the G266D substitution negates the requirement for enhancer DNA for oligomerisation, although it is not clear whether such oligomeric rings represent a functional and physiologically relevant form of the protein. Next, it was decided to investigate the requirement for enhancer binding on the ability of the G266D variant to activate transcription *in vivo*. Since Cryo-EM indicates oligomerisation can occur in the absence of DNA, it was decided to measure the ability of the G266D protein to activate transcription *in vivo* when unable to bind to the *norR-norVW* intergenic region. Therefore, C-terminal truncations were made to remove the helix-turn-helix (HTH) motif (Δ444-504, Δ442-504, and Δ436-504) in the NorRΔGAF and NorRG266D proteins. Results showed that both NorRΔGAF and NorRG266D were unable to activate transcription at the *norV* promoter when their ability to bind to enhancer DNA was removed (Figure 8.6A). However, Western blotting showed that one of the truncations resulted in an unstable protein, both in the context of wild-type NorR and G266D (Figure 8.6B). The inability of the stable
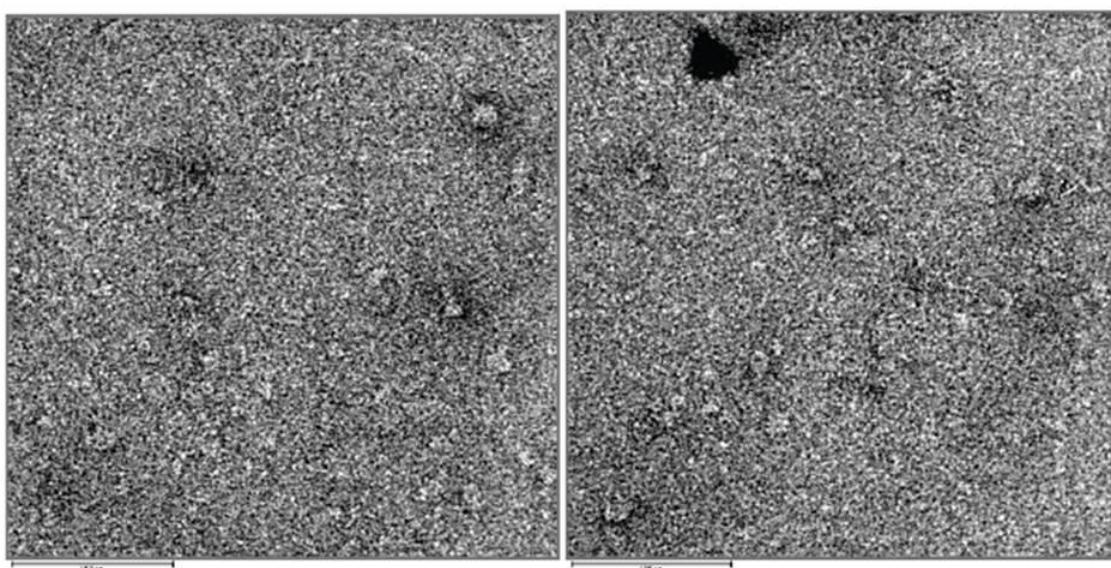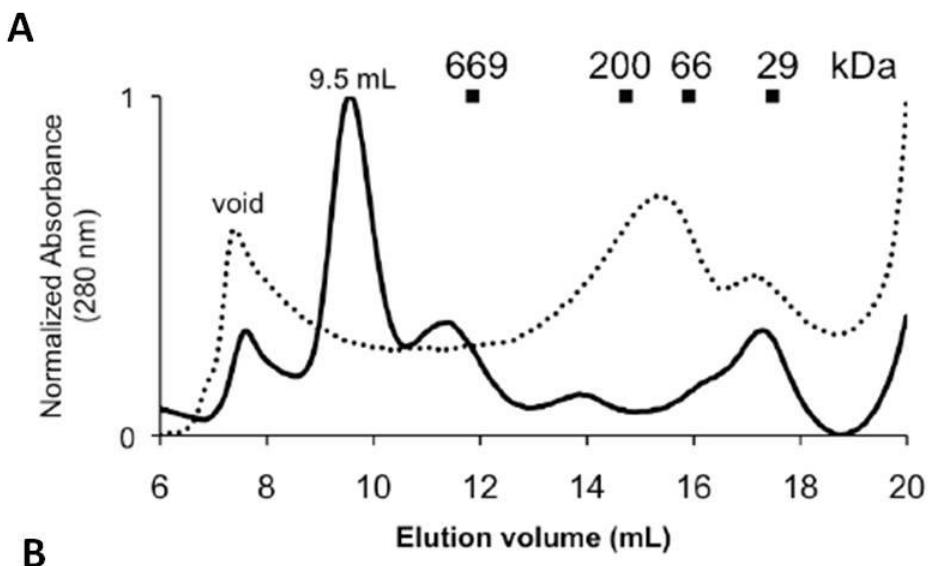
Figure 8.5 - Enhancer-dependent higher order oligomeric assembly of full length NorR-His. (**A**) **Gel filtration chromatography** of 3 μM His-tagged NorR in the absence (*dotted line*) and presence (*solid line*) of 0.4 μM 266bp dsDNA (molar ratio of 12:1 monomer : DNA), containing all three enhancer sites, performed at 4 °C using a Superose 6 column (24 ml). The His-tagged wild type nucleoprotein complex eluted at 9.5 ml. Corresponding molecular weights of standard globular proteins are indicated relative to their elution volumes. (**B**) **Negative-stain Electron Microscopy**. Shown are raw micrographs of NorR-His in complex with 266bp DNA fragment. Scale bar 100 nm. Ring-shaped higher order oligomeric particles were observed only in the presence of DNA. Particles of other sizes were also observed, which may represent intermediate assembly-states of the protein. **Experiments were conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

218
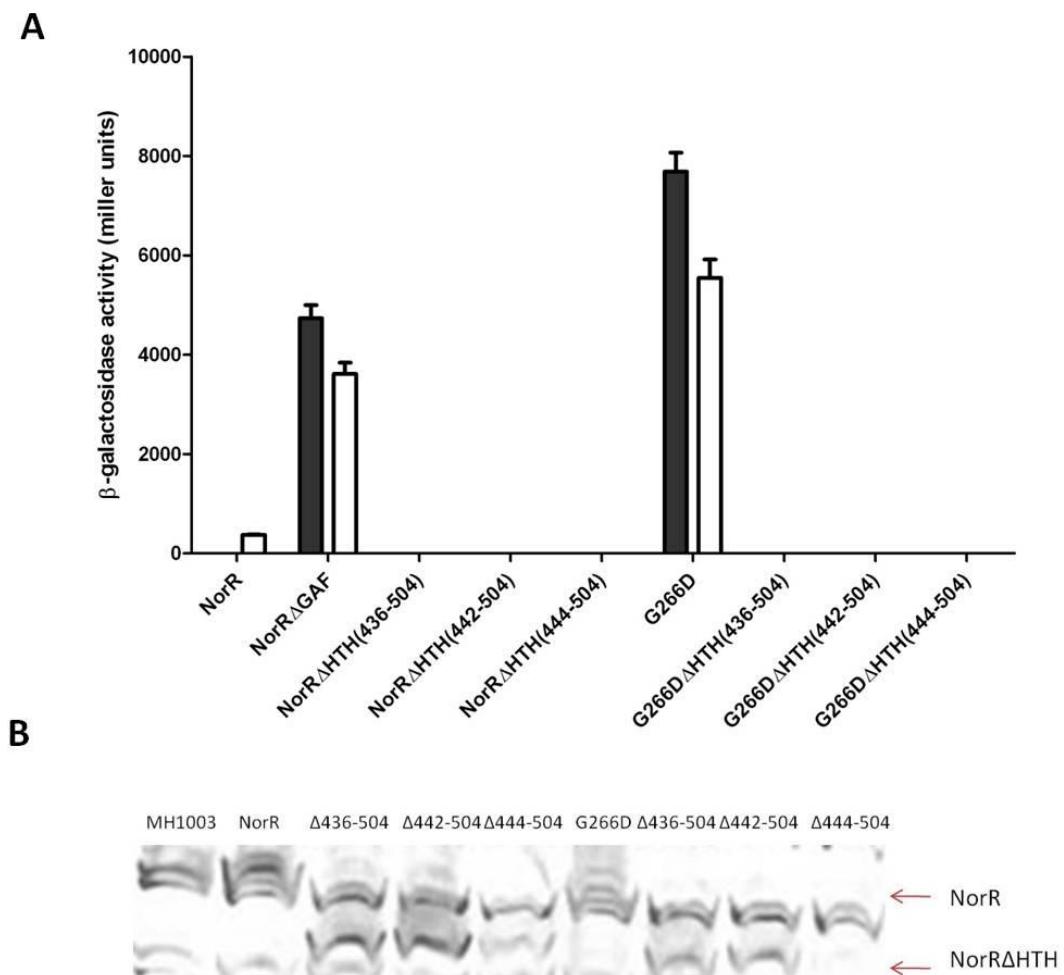
Figure 8.6 – (A) Activities of NorR and the G266D variant when C-terminal truncations are made in the *norR* sequence. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein, "NorRΔGAF" refers to the N-truncated form lacking the GAF domain (residues 1-170) and "NorRΔHTH" refers to the C-truncated form lacking the helix-turn-helix (HTH) motif (residues 436-504, 442-504 or 444-504). 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. **(B) Western blot analysis indicating the stability of NorR variants *in vivo* when cultures are grown in the absence of potassium nitrite.** The locations of the bands corresponding to the full-length and ΔHTH constructs are indicated by red arrows. "MH1003" refers to the *E. coli* strain only. The uppermost band that is not detected in the MH1003 strain correlates to full-length NorR and its variants. Results showed that the Δ444-504 constructs were unstable *in vivo*.

truncations to activate transcription indicates that NorR and the GAFTGA-variant G266D can only function as transcriptional activators when bound to enhancer DNA. Recent work has suggested that DNA-induced conformational changes are essential for the activation of NorR (Tucker *et al.* 2010a). Therefore, although the G266D protein can oligomerise in the absence of DNA, binding to enhancer DNA is likely to be required to activate it as a transcription factor. Alternatively, the ability of a NorR hexamer to activate transcription may not depend on changes that occur upon enhancer binding. The binding of NorR to DNA *in vivo* may simply serve to anchor the activator and orientate it relative to the holoenzyme bound at the promoter. In order to assess the overall requirement for enhancer binding whilst allowing *in cis* activation of the holoenzyme at the promoter, the ability of NorR to activate transcription was assessed in *E. coli* strains with individually altered enhancer sites. When the consensus sequence of binding site 1 (S1), site 2 (S2) or site 3 (S3) was altered from GT-(N7)-AC to GG-(N7)-CC, the NorRΔGAF protein was able to bind to the other two sites but was unable to hydrolyse ATP and as a result was largely unable to activate open complex formation (Tucker *et al.* 2010a). In line with this, the NorR and NorRΔGAF proteins were mostly inactive *in vivo,* in strains where either one of the three enhancer sites was altered (Figure 8.7). This emphasises the importance of the 3 NorR binding sites for the formation of the active oligomer. Interestingly, the GAFTGA-variant G266D was significantly more able to activate transcription when one of the 3 sites had been altered (Figure 8.7). When NorR site 3 (S3) was altered to GG-(N7)-CC, the G266D protein still showed 50 % activity compared to its activity when all three sites were present in their consensus forms. Overall this suggests that the GAFTGA variant is less dependent than wild-type NorR upon enhancer binding for the formation of an active oligomer. Since, each of the three binding sites is required to induce oligomerisation (Tucker *et al.* 2010a), this data is in agreement with gel filtration and electron microscopy data which together suggest that the G266D substitution stimulates oligomerisation in full-length NorR.

8.5   3D-reconstruction of the full-length G266D-His protein in the absence of enhancer DNA

In order to understand the molecular architecture of the enhancer-independent G266D-His oligomers, visualised in negative-stain EM *in vitro*, particle images were selected by Tamaswati Ghosh of Prof. Xiaodong Zhang's group at Imperial College, London and single particle reconstruction carried out to generate a 3D structure. Initially 3701 particles
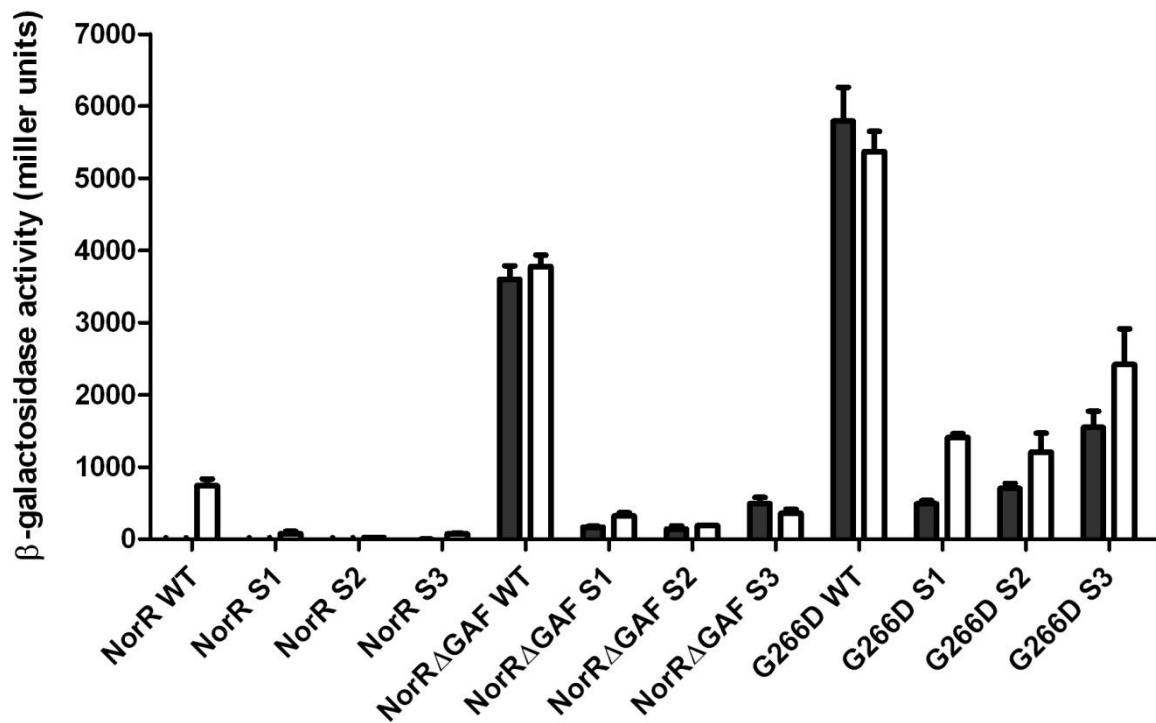
Figure 8.7 – *In vivo* transcriptional activation by GAFTGA variant G266D in the absence of NorR binding site 1, 2 or 3. NorR constucts were transformed into strains of *E. coli* with either three wild-type (WT) NorR binding sites (GT-(N7)-AC) or with one of three NorR binding sites (S1, S2, S3) altered to GG-(N7)-CC. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking the GAF domain (Δ1-170). 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.

were picked prior to the generation of class averages (Figure 8.8A). The most striking observation after single particle analysis was the seven-fold symmetry of the protein rings with seven clear regions of EM-density around the central cavity (Figure 8.8B). This seven-fold symmetry was used to generate the initial 3D before iterative refinement of the structure. The final reconstruction of G266D-His contained 1600 particles in 110 class averages, with the resolution estimated to be 24 Å. Importantly, there was a good correlation between the class averages and the reprojections, providing confidence in the 3D model (Figure 8.8 B and C). Looking at the 3D reconstruction of G266D-His, the EM-density appears to be composed of three major layers with an expanded central pore (Figure 8.9). The major, central density consists of a clear heptameric ring 165 Å wide with the central cavity 63 Å in diameter. Seven distinct EM-densities extend downwards from the heptameric ring in a claw-like shape that forms the bottom layer of the 3D-reconstruction. The third, upper-layer of the model is composed of seven further areas of EM-density that protrude upwards from the central ring. From the side, the three-layered structure is 77 Å in height.

The formation of heptamers by bEBPs has long been a matter for debate with several proteins of the AAA+ class showing heptameric arrangements in either their crystal or Cryo-EM structures (Miyata *et al.* 2000; Lee *et al.* 2003; Akoev *et al.* 2004). In the case of the bacterial protein-disaggregating chaperone, ClpB, heptamers are able to form in the absence of nucleotide but in the presence of ATP or ADP, hexamerisation occurs (Kim *et al.* 2000; Akoev *et al.* 2004). Therefore, in order to determine whether the formation of G266D hexamers also requires nucleotide, nucleoprotein complexes were prepared in the presence of ATP. Results show the formation of higher order oligomers with seven-fold symmetry, indicating that the presence of nucleotide is unable to prevent heptamerisation in the full-length GAFTGA variant G266D (Figure 8.3C).

8.5.1   An atomic model of the G266D-His protein

In order to examine the predicted interactions between the seven protomers of the heptameric ring, collaborators performed manual fitting of crystal structures into the 3D-reconstruction (Figure 8.10). Initially a model of the monomeric NorR AAA+ domain was built based on the crystal structure of the ADP-bound AAA+ domain of NtrC1 from *Aquifex aeolicus* (PDB ID: 1NY6). This AAA+ model was then fitted into the central EM
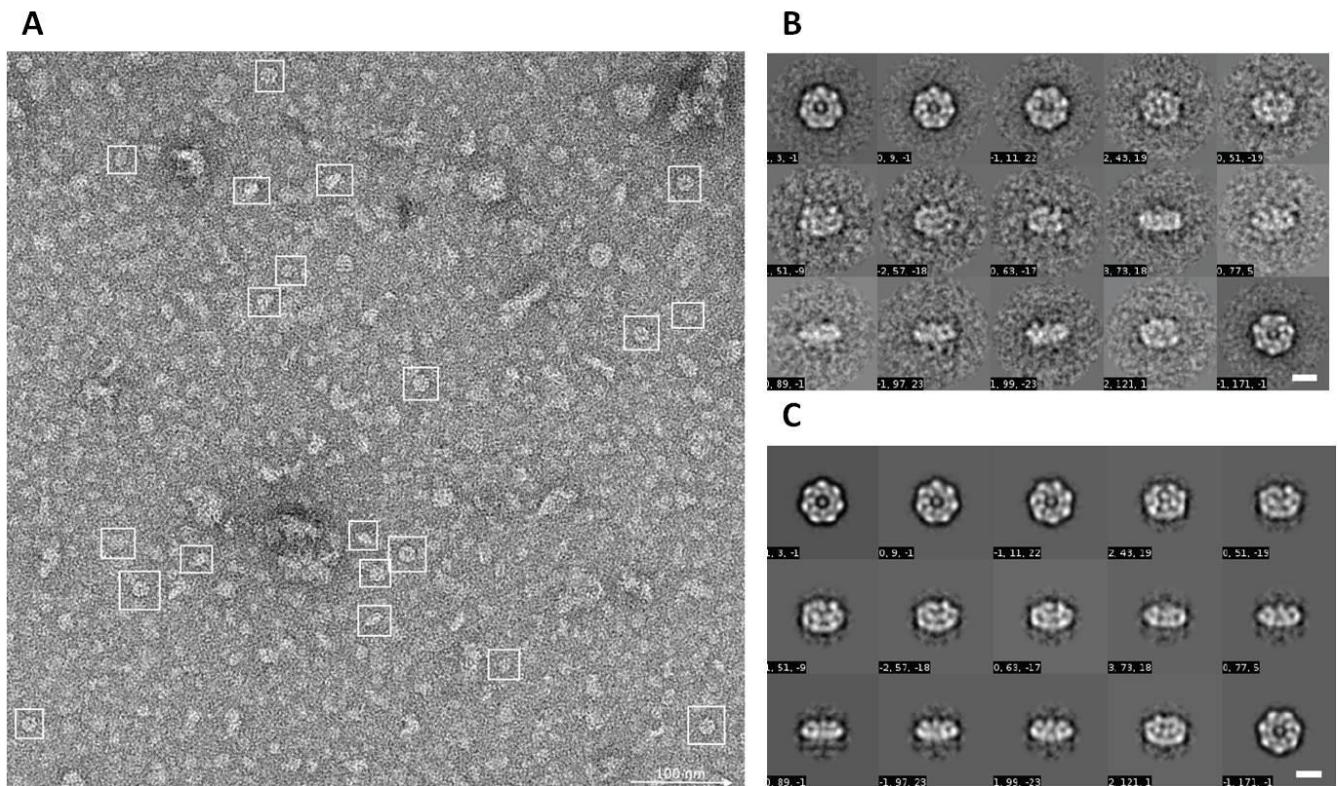
Figure 8.8 - Negative stain EM analysis of G266D-His. (**A**) Raw micrograph collected on the CM200 electron microscope at a magnification of 50000x. Some of the selected particles are boxed. The scale bar is 1000 Å. (**B**) Selection of the best class averages (~ 10 particles/class) and (**C**) their corresponding reprojections from the initial 3D model. **Experiments were conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**
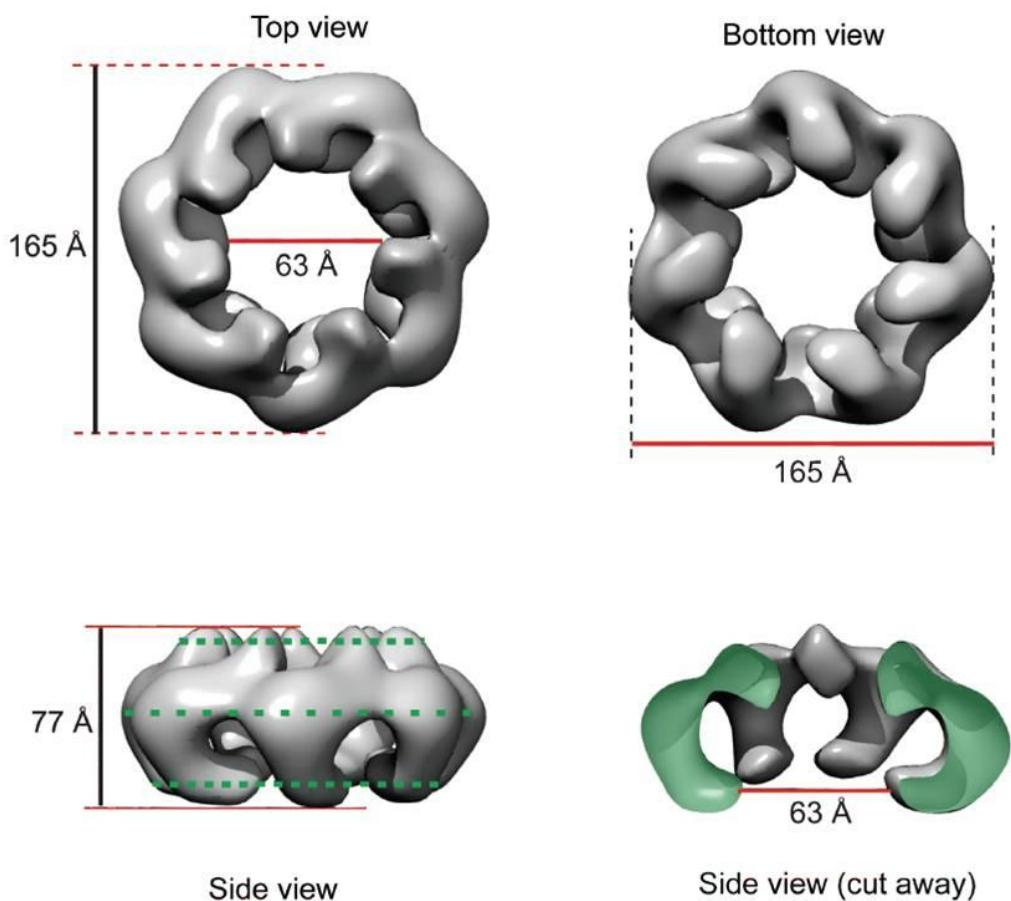
Figure 8.9 - 3D reconstruction of full length NorR-G266D at 26 Å resolution. A surface representation including 100 % of the expected volume of the EM density obtained from the 3D reconstruction of the G266D complex is shown in different orientations. The structure consists of three layers, when viewed from the side (dashed green lines). The main body of the molecule (central heptameric ring) has a diameter of 165 Å and a central opening of 63 Å. Small density lobes are found at the ends of each of the seven claw-shaped densities that extend down from the central ring; no connections appear between the individual subunits (bottom and side 'cut-away' views). A side view has been cut open to reveal the expanded central cavity and the channel spanning the entire length of the molecule. **Analysis was conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**
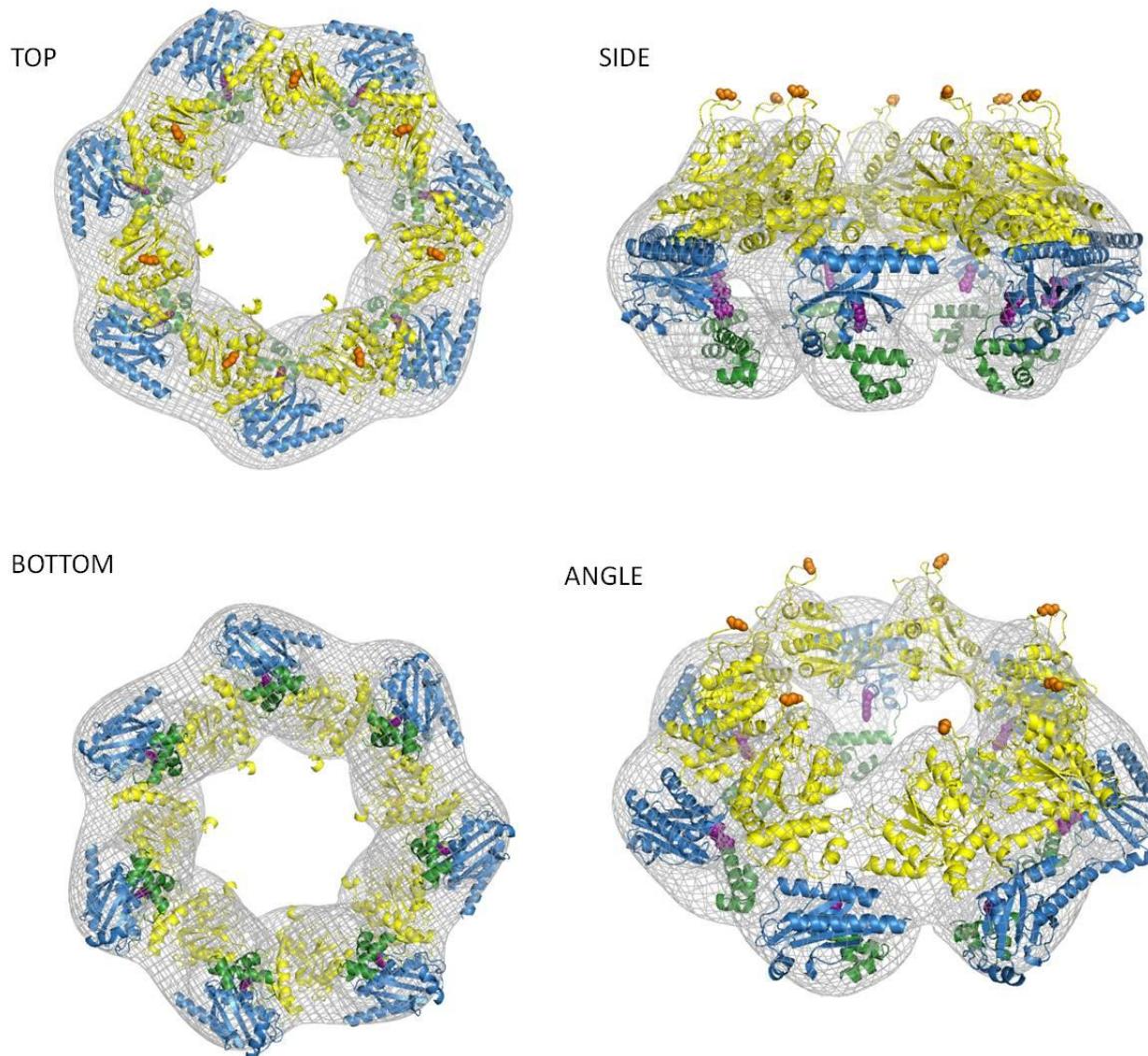
Figure 8.10 - Assignment of the NorR domains in the 3D reconstruction of NorRG266D-His. Fitting of the atomic models of the individual NorR domains into the 3D density map (grey mesh), shown in the top view (*top-left*), bottom view (*bottom-left*), side view (*top-left*) and titled side view (*bottom-right*) orientations. Seven AAA+ domains (yellow) were fitted into the central ring density, with individual GAF domains (blue) positioned at the outer rim of the ring. The DNA binding domains (shown in green) were fitted into the density lobes found at the bottom of the central ring. Also shown are the positions of the G266 (orange spheres) and R81 (purple spheres) residues, both of which are thought to play key roles in maintaining the mechanism of repression. **Analysis was conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

density so that the N-terminal α/β subdomain containing helix 3 and helix 4, occupied the upper-layer of density that extends upwards in the 3D-reconstruction. This ensured that the L1 and L2 loops were surface-exposed in the model as is observed in the crystal structures of the ZraR hexamer and the NtrC1 heptamer (Lee *et al.* 2003; Sallai and Tucker 2005). However, in order to fit neatly within the EM-density, the C-terminal helix of the monomeric NorR AAA+ model was rotated approximately 90° such that it faced downwards towards the bottom of the central ring. Following modelling of the NorR AAA+ domain, homology modelling based on the crystal structure of the Acinetobacter phosphoenolpyruvate-protein phosphotransferase GAF domain (PDB ID: 3CI6) was used by collaborators to generate a model of the NorR GAF domain. In the 3D-reconstruction of G266D-His, additional EM-density is present at the outside edge of the ring formed by the AAA+ domains. Therefore the model of the monomeric NorR GAF domain was fitted into these peripheral regions. This arrangement is similar to that observed in the EM model of full-length activated NtrC (De Carlo *et al.* 2006). Together the AAA+ and GAF domains of the seven protomers account for the central EM density that forms the heptameric ring. Finally, a model of the NorR C-terminal DNA-binding domain was generated based on the crystal structure of the DNA-binding domain from ZraR. Secondary structure prediction revealed that the NorR C-terminal domain is likely to contain a three-helix bundle, something observed in the crystal structure of the ZraR D domain (Sallai and Tucker 2005). The first helix (α1) is implicated in the dimerisation of the C-terminal domains whilst the second and third (α2 and α3) helices together form the helix-turn-helix motif. The ZraR DNA-binding domain was manually fitted into the EM-density below the central body of the heptameric ring to complete the atomic model for the full-length GAFTGA variant.

The proposed atomic model for G266D-His is shown in Figure 8.10. Examination of the model revealed that the NorR AAA+ domains are arranged in a front-to-back configuration in agreement with other bEBPs including PspF, NtrC1 and ZraR (Lee *et al.* 2003; Rappas *et al.* 2005; Sallai and Tucker 2005; De Carlo *et al.* 2006). Importantly, the N-terminal regulatory (GAF) domains are positioned on the periphery of the ring, away from L1 and L2 loops which are surface exposed and well placed for σ^{54}-interaction at the top of the heptamer. In this state, the GAF domains appear to interact with the AAA+ domain of the same protomer but there is no inter-protomer association facilitated by the regulatory

domains. Furthermore, based on the model no inter-protomer interactions are predicted to form between GAF domains. However, it is possible that the regulatory domains contribute to the stability of the oligomeric assembly by restricting the relative conformations of the AAA+ α-helical and α/β subdomains. This may explain why enhancer-independent oligomerisation was only observed for the full-length GAFTGA variant and not for G266DΔGAF-His. The arrangement of the GAF domains on the outer-edge of the AAA+ ring may account for the large 165 Å diameter that is observed in the reconstruction. In contrast, the wild-type and G266D variant forms of NorRΔGAF-His are 125 Å in diameter. This is consistent with hexameric rings formed by PspF that lacks a regulatory domain and by ZraR, which was crystallised in its absence. The NtrC1 heptameric AAA+ ring, crystallised in the absence of the receiver domain also adopts a similar diameter of approximately 124 Å (Lee *et al.* 2003). Finally, fitting of the ZraR C-terminal, DNA-binding domain within the density lobes beneath the central plane of the ring suggests that binding to enhancer DNA would take place on the opposite surface of the protein from the upper, σ^{54}-interaction surface. The DNA-binding domains do not occupy the full EM-density below the central plane and this might reflect the independent motion of the C-terminal domains relative to the AAA+ domain that may result from the flexible (>22 amino acids) linker between the two domains. In contrast to the EM model of NtrC (De Carlo *et al.* 2006), the DNA-binding domains are not in close enough proximity to dimerise. However, it has been suggested that this may be a result of the seven-fold symmetry imposed during the refinement of the 3D-reconstruction (Ghosh 2010).

8.5.2   The 3D-reconstruction predicts that the G266D-His protein is ATPase inactive *in vitro*

One of the most striking features of the NorRG266D-His reconstruction is the significantly larger diameter of the central pore (~63 Å) compared to the AAA+ ring of the crystal structure of NtrC1 (~20 Å) (Figure 8.11, compare A and B). Since both structures consist of seven protomers, this is likely to result in differences in the separation and organisation of individual domains. In order to assess the implications of these differences on the function of the ATPase, the AAA+ inter-protomer interactions in the NorR 3D-model were compared to those present in the crystal structure of NtrC1[C]. In particular the position of key residues required for the ATPase activity of bEBPs was examined at the interface
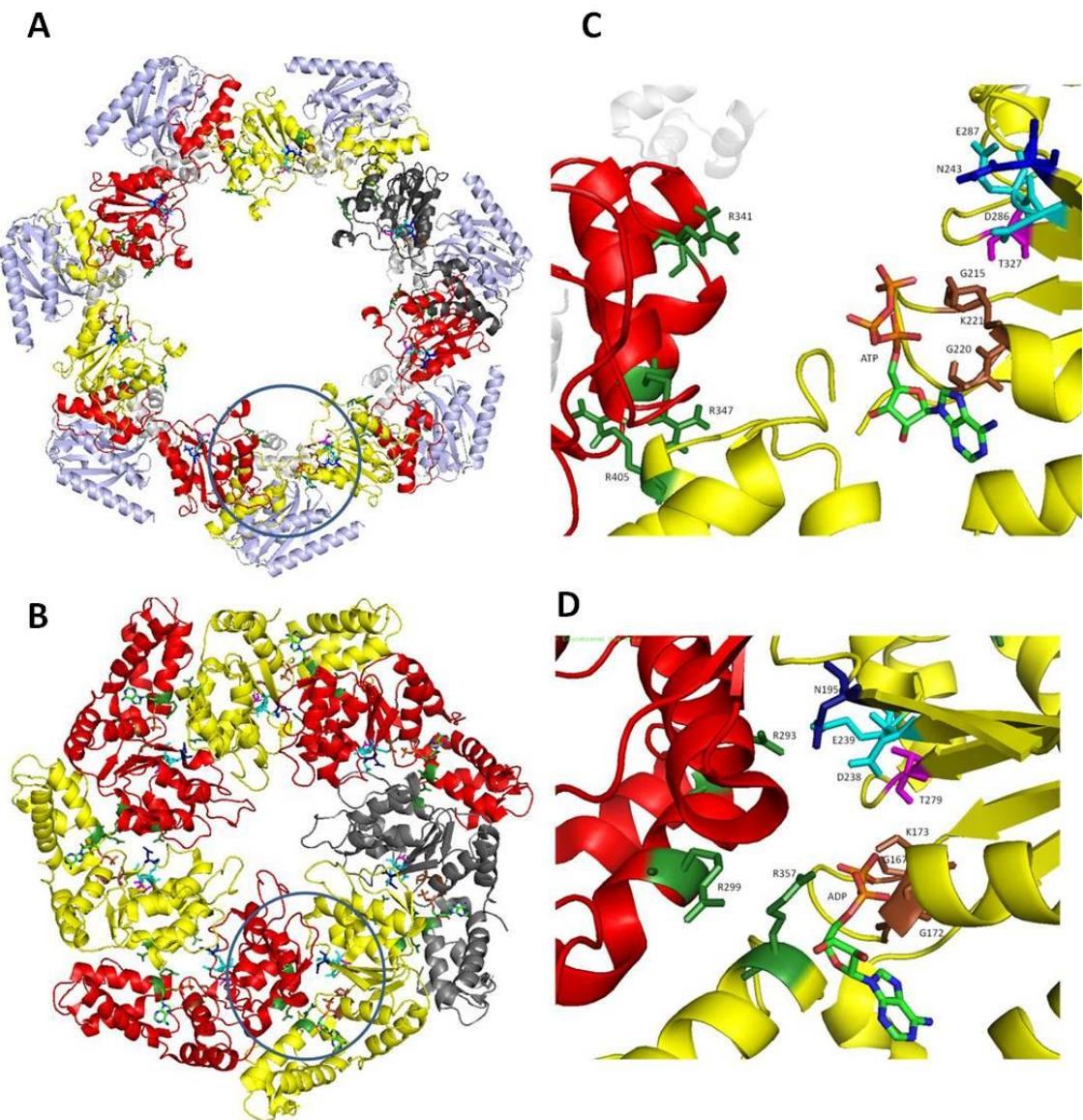
Figure 8.11 – Comparison of ATP hydrolysis sites in the NorRG266D-His heptamer model and the crystal structure of the heptameric NtrC1. (**A**) Top view of the G266D variant model showing the heptameric ring assembly. Consecutive AAA+ domains are shown in alternating red and yellow colours with the seventh monomer in dark grey. GAF domains are in light blue and DNA binding domains in light grey. (**B**) The crystal structure of the AAA+ domain of NtrC1 (PDB 1NY6). Consecutive AAA+ domains are shown in alternating red and yellow colours with the seventh monomer in dark grey. Both the regulatory domains and DNA binding domains are absent in the crystal structure. An example of the interprotomer interface at which the ATP hydrolysis site is found is shown in both models (blue circles). (**C**) Close-up of the ATPase active site located between adjacent AAA+ subunits in the atomic model of NorRG266D-His and in the crystal structure of NtrC1 (**D**). Conserved residues implicated in ATP binding and hydrolysis, inter-subunit catalysis and relaying nucleotide states to the surface exposed L1/L2 loops are indicated: Walker A residues G220/G172, K221/K173, G215/G167 (brown), Walker B residues D286/D238 and E287/E239 (cyan), "switch" asparagine N243/N195 (dark blue), Sensor I T327/T279 (magenta), sensor II residue R405/R357 (green), and *trans*-acting putative R-fingers R341/R393 and R347/R299 (green). An ADP molecule is present in the active site of NtrC1 (**C**) whilst a molecule of ATP has been positioned in the NorR model (**D**). **Analysis was conducted as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

between two AAA+ protomers that forms the site for hydrolysis (Schumacher *et al.* 2008). The NtrC1 AAA+ heptamer was crystallised in its ADP-bound form and an ATP molecule was also positioned in the NorR model in the cleft between two subdomains, such that the phosphate backbone wrapped around the P-loop (residues 214-222). The Walker A motif forms a P-loop with the consensus GxxxxGK[T/S] and contacts the phosphates of ATP (Saraste *et al.* 1990). It has been shown to be essential for hydrolysis in a number of bEBPs including PspF and NtrC (Rombel *et al.* 1999; Schumacher *et al.* 2004). The position of the conserved lysine of the Walker A motif in NtrC1 (K173) is in line with this role. However, in the 3D-reconstruction, the NorR side chain (K221) faces away from the phosphates of ATP. The Sensor II arginine has been implicated in ATP hydrolysis via the coordination of the γ-phosphate (Schumacher *et al.* 2006) and the position of this residue in NtrC1 (R357) correlates with this. However, in the G266D-His reconstruction, the R405 residue points away from the nucleotide. The open conformation at the ATP hydrolysis site also appears to alter the position the highly conserved Walker B "DE" residues. The aspartate side chain is suggested to have a role in Mg^{2+} coordination as well as the activation of a water molecule for the nucleophilic attack of the γ-phosphate. However in the NorR model, the D286 and E287 side chains are not well placed to carry out these functions unlike the D238 and E239 residues in NtrC1. This may be due to the inability of the R-finger residue (R341) to H-bond to the Walker B aspartate (D238), a contact observed in the NtrC1 structure (R293-D238) that helps to position the Walker B glutamate. Overall, the more open conformation of the G266D-His 3D-reconstruction, compared to the NtrC1 crystal structure suggests that key ATPase determinants are not suitably placed for efficient turnover of nucleotide. Next, it was decided to test this hypothesis by carrying out ATPase activity assays *in vitro*.

8.6 The G266D-His protein exhibits contaminating ATPase activity *in vitro*.

The position of key determinants in the 3D-model of G266D-His suggests that this protein is unable to hydrolyse ATP. Therefore, it was important to determine whether the heptameric G266D-variant protein maintained the interactions that result in a functional ATPase *in vitro*. Non-activated wild-type NorR has been shown to be inactive for ATP hydrolysis irrespective of whether DNA containing the *norR-norVW* intergenic region is present (D'Autreaux *et al.* 2005). A truncated variant lacking the regulatory GAF domain, (NorRΔGAF) showed a low level of ATP hydrolysis in the absence of enhancer DNA
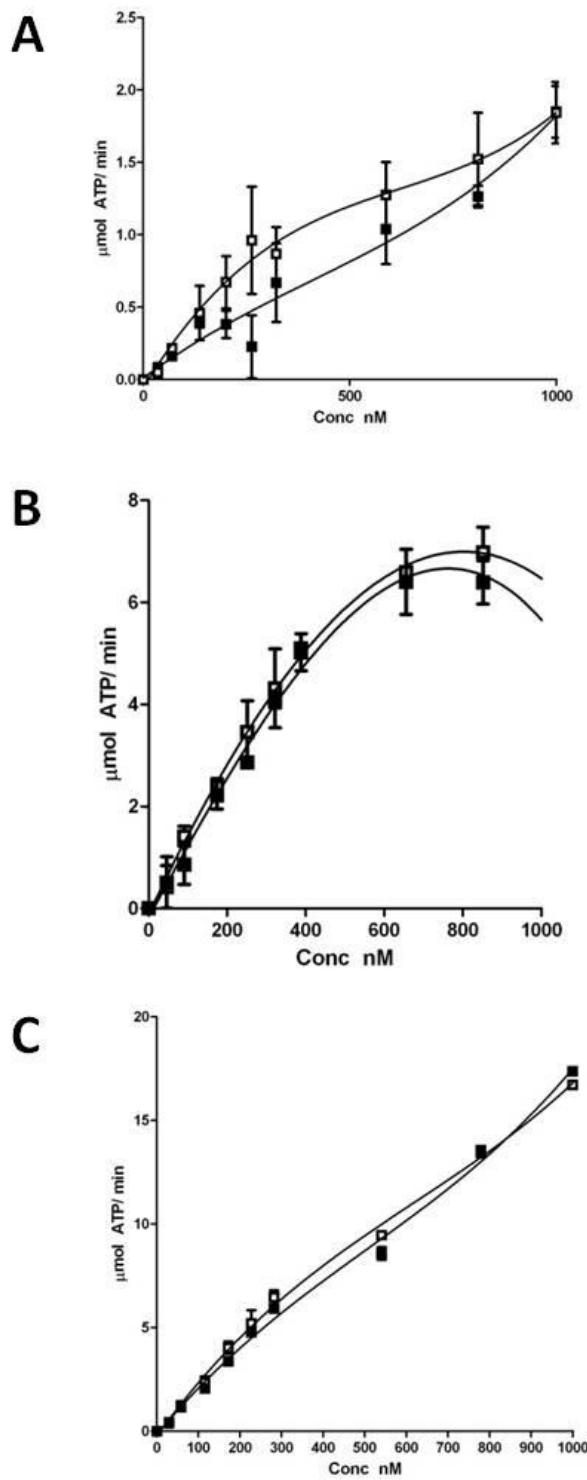
Figure 8.12 - ATPase activity of the G266D-His (A), G266N-His (B) and G266D-D286A-His (C) variants in response to protein concentration and the presence of enhancer DNA. Non-linear regression was carried out using GraphPad Prism software. Assays were conducted either in the absence (closed bars or closed squares) or presence (open bars or open squares) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

(Figure 7.4A). The presence of DNA presumably stimulates oligomerisation with the appropriate interactions occurring between residues of adjacent protomers to form the functional hexamer. This is reflected in the significant increase in ATPase activity that was observed as enhancer DNA was added (Figure 7.4A). Initial results revealed that the G266D and G266N variants both exhibited enhancer-independent ATPase activity. A high level of ATPase activity was observed that was not dependent on the presence of a 266bp DNA fragment containing the three NorR enhancer sites of the *norR-norVW* intergenic region (Figure 8.12). This suggests that the large oligomers visualised in the Cryo-EM with a central pore of ~63 Å in diameter *are* competent to hydrolyse ATP. To confirm this result, an additional substitution was made at the essential Walker B motif (D286A). This mutation renders bEBPs inactive for ATPase activity as observed for NorRΔGAF (Figure 7.5) and G266DΔGAF (Figure 7.6). The D286A-G266D double mutant was expressed prior to purification of the protein that eluted from the gel filtration column at a volume corresponding to ~960kDa or ~16 NorR monomers (data not shown). This volume is similar to that observed for the single-mutant G266D (~1020kDa or ~17 monomers). Importantly, no reduction in ATPase activity was observed when the D286A substitution was additionally present, suggesting that the G266D substitution is not responsible for the ATP turnover previously observed. Since the G266D protein eluted over a large range of volumes around the void volume of the gel filtration column, it may be that it co-purified with one or more contaminating proteins. Changes in the protocol for the purification of the GAFTGA variants were not successful in removing contaminating ATPases (data not shown). Overall, *in vitro* assays were unable to determine whether or not the heptameric form of NorR is functional for ATP hydrolysis. However, when the position of the key catalytic residues in the 3D-reconstruction is considered, it can be concluded that it is most unlikely that the G266D-His protein can turnover ATP *in vitro*.

8.7   Full-length G266D does not activate open complex formation *in vitro*

Since it was not possible to discern whether the full-length G266D protein was able to hydrolyse ATP *in vitro*, it was decided to test the functionality further by conducting open promoter complex (OPC) assays. The presence of a contaminating ATPase would not be expected to interfere with the OPC assay since other proteins should be unable to activate transcription from the *norV* promoter within the physiological range of activator concentrations. Whilst the assay conditions enabled NorRΔGAF-His to form open complexes (Figure 8.13A), it was not possible to determine whether the G266D variant
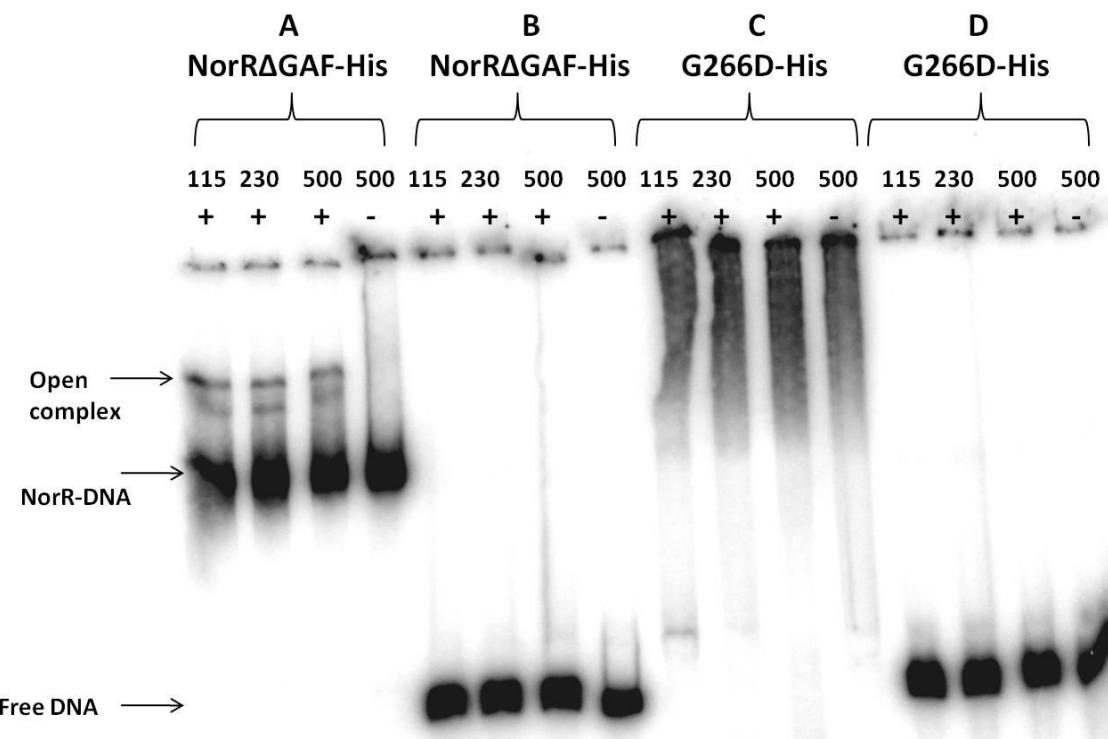
Figure 8.13 – Analysis of Open promoter complex formation by the GAFTGA variant G266D. Heparin resistant complexes formed by NorRΔGAF-His (A) and G266D-His (C) on a 361bp DNA fragment carrying the *norR-norVW* intergenic region. Complex formation by NorRΔGAF-His (B) and G266D-His (D) on the 295bp DNA fragment that does not contain the NorR-binding region, is also shown. The presence or absence of ATP is indicated above each lane by the -/+ sign and the NorR concentrations are indicated in nM. Arrows indicate the position of free DNA, NorR bound DNA and the open promoter complexes.

was functional. Assays indicated that the G266D variant bound effectively to a 361bp section of the *norR-norVW* intergenic region but that the NorR-DNA complex was unable to migrate into the polyacrylamide gel to form a distinct DNA-bound species (Figure 8.13C). This is in line with gel filtration experiments that indicated that this variant forms a large mulitimeric assembly. Studies *in vivo* showed that the G266D variant was able to activate transcription when either one of the three NorR binding sites was altered from the consensus sequence (Figure 8.7). Although the transcriptional activity was reduced, the GAFTGA variant appeared to be less reliant than NorRΔGAF on the three NorR binding sites to facilitate the formation of a functional oligomer. Because of this, and the fact that Cryo-EM data showed enhancer-independent oligomerisation (Figure 8.3B), it is possible that the G266D variant might be able to activate transcription in the absence of all three NorR binding sites in the intergenic region. Therefore, PCR-based deletion mutagenesis was carried out to remove a 66bp sequence of the *norR-norVW* intergenic region (section 5.7.6). The resulting 295bp fragment was used in the same way as the 361bp fragment to detect the formation of open complexes. Neither the NorRΔGAF nor G266D variant proteins were able to activate transcription from a 295bp fragment that lacks the three NorR binding sites (Figure 8.13B and 8.13D). It appears that for both the wild-type and G266D-variant, NorR requires the ability to bind to enhancer DNA in order to activate transcription. This is in agreement with the inability of NorR and the GAFTGA-variant to activate transcription in the absence of the C-terminal DNA-binding domain (Figure 8.6). These data emphasise the importance of enhancer-DNA in the activation of NorR as a transcription factor, as has been previously demonstrated (Tucker *et al*. 2010a).

Since the large oligomeric assembly that forms when the G266D variant is purified in full-length form prevented identification of the open promoter complexes in the standard OPC assay, we decided to try to detect the presence of open complex using potassium permanganate footprinting. In contrast to the standard open promoter complex assay, this technique does not analyse the migration of protein-DNA complexes but rather detects transcriptional activation by the cleavage of single stranded DNA regions, including the region immediately upstream of the start-site. For the non-tagged and tagged versions of NorRΔGAF, we observed enhanced cleavage corresponding to T residues located between -11 to +1 in the *norV* promoter, consistent with the expected footprint (Figure 8.14, lanes 3 and 4). However, no enhancement of cleavage was observed when G266D-His was used as an activator in the reaction (Figure 8.14 lane 5). Taken together, these data suggest that the
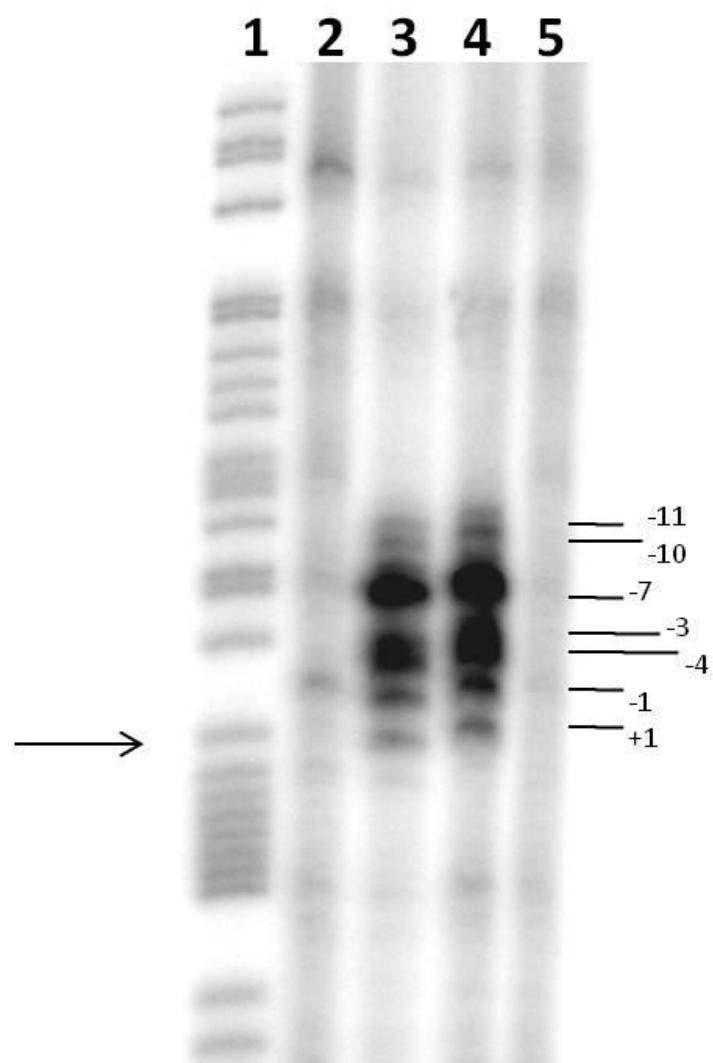
Figure 8.14 - Potassium permanganate footprinting of the 266bp *norR-norVW* promoter fragment after open complex formation initiated by NorR. Lane 1 is a G+A ladder. Lane 2 is a control without activator present. Lanes 3, 4, and 5 show footprinting after initiation of open complexes in the presence of 1 µM (final concentration) ΔGAF, ΔGAF-His and G266D-His respectively. The arrow marks the *norVW* transcriptional start and the positions of the enhanced cleavage at T bases are indicated.

full-length G266D "heptameric" species previously observed in Cryo-EM analysis is non-functional with respect to open promoter complex formation.

8.8   Discussion

In the previous chapter, the N-terminal regulatory domain of NorR was proposed to target the GAFTGA-containing loop 1 (L1) in order to prevent access of the activator to σ^{54} in the absence of NO. Biochemical analysis of the G266D variant protein in the context of GAF-truncated NorR suggested that neither the oligomeric determinants nor the ATP hydrolysis site serve as targets in the mechanism of interdomain repression, as has been demonstrated in related bEBPs. However, a number of recent studies have highlighted the importance of examining the activity and structure of such proteins in their full-length forms. Despite difficulties with the low-solubility of full-length G266D-His, low-levels of protein (1-3 μM) were eventually purified. The first clue indicating that the presence of the regulatory domain altered the properties of the GAFTGA variant was provided by size-exclusion chromatography. The truncated form of the protein (G266DΔGAF-His) eluted at a volume corresponding to a monomer or dimer (70-80 ml on the Superdex 200 16/60 preparative column). In contrast, full-length G266D-His eluted over a broad range of volumes close to the void (43 ml) (Figure 8.2). Based on the elution of gel filtration standards, this peak centred on a volume corresponding to ~1020 kDa or ~17 monomers. The analytical Superose 6 column provides better separation of larger protein complexes and gel filtration carried out by collaborators at Imperial College indicated that the G266D-His complex was greater in size than the 669 kDa molecular weight marker, thyroglobulin (Figure 8.3A). This suggests that *in vitro*, the full-length GAFTGA variant forms a complex with greater than 11 monomers. Structural studies of the related proteins p97 and MCM have identified double-AAA+ rings (Rouiller *et al.* 2002; Beuron *et al.* 2003; Costa *et al.* 2006b; Davies *et al.* 2008) and G266D-His may form similar complexes. Since the protein does not elute at a discrete volume during gel filtration, it is likely that a range of oligomeric particles assemble.

Significantly, cryo-electron microscopy has revealed that the G266D-His protein is competent to form oligomeric rings in the absence of the three NorR binding sites (Figure 8.3B). Since full-length, wild-type NorR-His does not form such complexes in the absence of DNA (Figure 8.5), it can be concluded that the G266D substitution causes this dramatic change in behaviour rather than the presence of the regulatory GAF-domain *per se*. The

question of why the G266D substitution stimulates enhancer-independent oligomerisation of NorR only when the GAF domain is present remains unanswered. The 3D-model of G266D-His suggests that there are no inter-protomer interactions that involve the regulatory domain (Figure 8.10) although it is possible that the GAF domains help stabilise the higher oligomer by restricting the relative conformations of the α-helical and α/β subdomains of the AAA+ domain. In contrast, the GAFTGA variant was shown to exhibit a complete bypass phenotype *in vivo* (Figure 6.4B), suggesting that the GAF domain does not contribute to the activity of the G266D protein. This is supported by the observation that when the regulatory domain was absent altogether in a hexa-histidine-tagged construct expressed from the pETNdeM-11 vector (G266DΔGAF-His), there was no reduction in the *in vivo* activity of the escape mutant (Figure 6.11). However, the GAF domain may contribute to the activity of the non-tagged protein *in vivo,* since G266DΔGAF showed a reduction in activity (Figure 6.9).

Previous data supports a model in which three NorR dimers bind to the enhancer sites, inducing conformational changes that lead to the formation of a hexamer (Tucker *et al.* 2010a). Each of the three NorR binding sites has been shown to be essential for oligomerisation, ATP hydrolysis and open complex formation. Here, it has been shown that the GAFTGA variant G266D must still bind to enhancer DNA in order to activate transcription. A truncated form of the protein that lacks the C-terminal DNA binding domain was inactive *in vivo* (Figure 8.6). Furthermore *in vitro,* the full-length protein was unable to form open promoter complexes when the upstream DNA did not include the enhancer sites (Figure 8.13). Interestingly, the G266D variant was partially active *in vivo* when either site 1 (S1), site 2 (S2) or site 3 (S3) was altered from the consensus (Figure 8.7). Indeed when NorR site 3 (S3) was altered from GT-(N7)-AC to GG-(N7)-CC, the activity of the G266D variant showed only a two-fold reduction compared to the activity when all three sites were present in their consensus forms. This phenotype may be explained by an increase in the cooperativity of DNA-binding at the remaining intact sites, induced by the G266D substitution. The 3D-reconstruction suggests that the regulatory GAF domains are located close enough to the DNA-binding domains to mediate such an effect (Figure 8.10). The G266D substitution may also have an indirect effect on DNA binding through the stimulation of GAF-AAA+ interactions that increase cooperative binding at the enhancer sites. An alternative explanation is that the greater propensity of full-length G266D to oligomerise (as demonstrated by the formation of DNA-independent

heptamers) enables the protein to form active oligomers *in vivo* in the presence of two rather than three enhancer sites.

Particularly important is the observation that the G266D-His oligomeric particles are larger than the expected size of a NorR hexamer (125 Å) and correspond to a ring with 7-fold symmetry (165 Å). Heptameric structures have also been observed in the bEBPs NtrC1 (Lee *et al.* 2003; Chen *et al.* 2010) and NtrC4 (Batchelor *et al.* 2009) as well as in the AAA+ proteins RuvB (Miyata *et al.* 2000); MCM (Yu *et al.* 2002; Costa *et al.* 2006b; Costa *et al.* 2006a), ClpB (Kim *et al.* 2000; Akoev *et al.* 2004); HslU (Rohrwild *et al.* 1997); Lon (Stahlberg *et al.* 1999); magnesium chelatase (Reid *et al.* 2003) and p97 (Davies *et al.* 2008). Since the physiological relevance of such structures has been a matter for debate it was important to verify whether the G266D-His protein that forms heptamers *in vitro* was functional with respect to ATP hydrolysis and open complex formation. Although the presence of contaminating ATPases meant that the ability of the GAFTGA variant to turnover ATP could not be determined (Figure 8.12 A and C), given the inability to form open complex (Figure 8.13) it might also be expected that the protein is inactive with respect to ATP hydrolysis. In agreement with this hypothesis, the 3D-reconstruction revealed that the key residues involved in hydrolysis are too distant relative to the expected position of the nucleotide for efficient hydrolysis (Figure 8.11).

Although in some AAA+ proteins, functional roles for two higher oligomeric forms have been proposed (Table 3) e.g. RuvB, MCM, ClpB (Kim *et al.* 2000; Miyata *et al.* 2000; Yu *et al.* 2002; Akoev *et al.* 2004; Costa *et al.* 2006b; Costa *et al.* 2006a), it is becoming increasingly clear that the active form of the protein is most commonly hexameric. The most common reason for the formation of heptameric rings is the truncation of proteins that lead to non-native stoichiometries. Therefore, if possible, the oligomeric state of an *intact* AAA+ protein should be assessed. Interestingly, in the case of the NorR variant G266D, heptameric rings form when the protein is intact and truncation of the N-terminal domain results in the formation of the 6-membered ring. However, it is probable that the existence of the two configurations is not due to the presence or absence of the GAF domain *per se* since both the full-length and N-truncated forms of wild-type NorR form hexamers. In other cases, heptameric rings can often be converted into rings with 6-fold symmetry upon binding of a key ligand or cofactor. The bacterial protein ClpB forms heptamers in the absence of nucleotide but rearranges to form hexamers when ATP or

ADP binds (Kim *et al.* 2000; Akoev *et al.* 2004). However, the addition of ATP was not sufficient to prevent the formation of G266D-His heptamers (Figure 8.3C). In the case of RuvB from *Thermus thermophilus* and the MCM protein from *Methanothermobacter thermoautotrophicus,* heptamers form in the absence of DNA but hexamerisation occurs when DNA is present (Kim *et al.* 2000; Miyata *et al.* 2000; Yu *et al.* 2002; Akoev *et al.* 2004; Costa *et al.* 2006b; Costa *et al.* 2006a). However, in the current analysis, the addition of a DNA fragment carrying the three enhancer sites to the G266D-His protein did not facilitate the conversion of the 165 Å diameter oligomers into smaller 125 Å hexamers (Figure 8.3D). However, OPC assays have confirmed that the full-length variant is able to bind to a DNA fragment containing the three enhancer sites in a heparin-resistant manner. It is possible that the 7-membered ring represents a configuration of NorR, stimulated by the G266D substitution that is capable of binding to enhancer DNA but unable to undergo the DNA-induced conformational changes required to form the NorR hexamer.

The non-functional nature of the G266D variant *in vitro* contradicts *in vivo* data that indicates that the G266D substitution enables full escape from the GAF-mediated repression mechanism (Figure 6.8A). This would suggest that the ability of the GAFTGA variant to form heptamers *in vitro* does not contribute to the proteins ability to activate transcription *in vivo*. It is likely that the conditions inside the cell are suited to the formation of the active hexamer rather than the inactive heptamer. The assembly of the higher order oligomer from NorR dimers using DNA as a scaffold may promote the formation of the hexamer over the heptamer in this case. It would be interesting to test therefore whether the G266D heptamer can be denatured and re-folded in the presence of enhancer DNA to encourage the formation of an active NorR hexamer.

In the previous chapter it was shown that the N-terminal, regulatory domain of NorR is likely to target the σ^{54}-interaction surface in the mechanism of interdomain repression. In this model, the GAF domain "clamps-down" over a region of the AAA+ domain that includes the GAFTGA-containing loop 1, required for σ^{54}-interaction. The G266D substitution leads to a complete bypass phenotype *in vivo* (Figure 6.8A), suggesting that the GAF domain adopts a conformation close to its position in the on-state. This proposed conformation can be examined in the 3D-reconstruction of the G266D-His protein (Figure 8.10). The most striking feature of the 3D-model is the large size of the heptameric ring. The reconstruction is 165 Å in diameter in contrast to the predicted diameter of 125 Å for

the hexamers formed by the wild-type and G266D variant of NorRΔGAF-His (Figure 8.9). This size may be explained in part by the presence of the regulatory GAF domain which appears to occupy the EM-density on the periphery of the AAA+ ring, away from the σ^{54}- interaction surface (Figure 8.10). Based on the model of interdomain repression proposed in Chapter 7 and the position of the GAF domain in the reconstruction, it is possible that a "swing-out" movement occurs from the top to the outer edge of the AAA+ ring (Figure 8.15). The random coiled-coil linker present between the GAF and AAA+ domains is of sufficient length (29 residues) to enable such a transition. Furthermore, in the proposed off-state conformation, the critical R81 residue would be in a position to interact with the α/β subdomain. Structural analysis of the wild-type protein in the absence of NO would be helpful in order to identify a more exact position of the regulatory domain in the off-state.
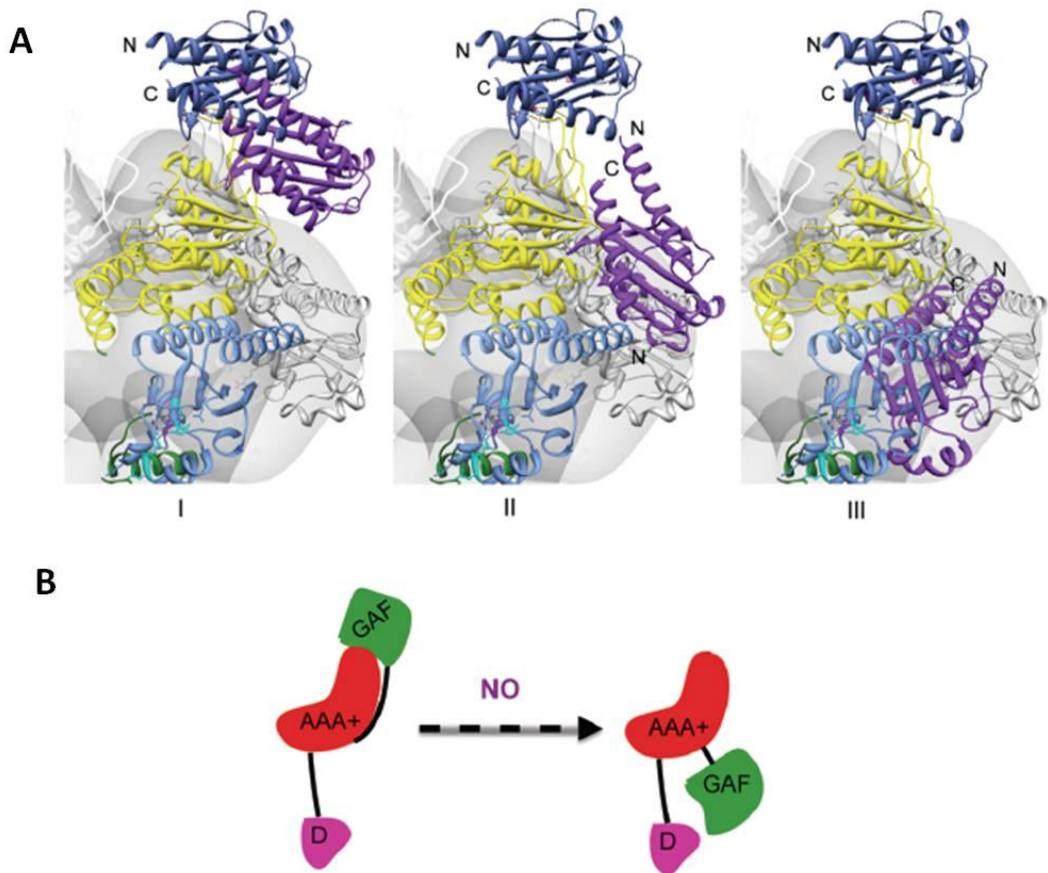
Figure 8.15 – Proposed model of GAF-domain relocation upon the release of repression. (**A**) Shows the "swing-out" movement (monomer in purple, positions I-III) of the GAF domain depicting the potential 'off-state' (dark blue) and 'on-state' (light blue) conformations adopted by the NO-sensing domain. The central AAA+ domain is also shown (monomer in yellow) (**B**) Schematic to show the potential relocation of the GAF domain as NorR becomes activated upon binding to NO. **Modelling was conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

Chapter 9 - Investigating the escape mechanism of the Q304E variant

9.1 Introduction

Of the ten substitutions identified in the screen that gave rise to a significant escape phenotype, only the Q304E change was at a residue predicted to be located outside the region of nucleotide-induced conformational change (Figure 6.5). Since the Q304E variant of NorR is still subject to NO-dependent regulation of AAA+ activity, it might be inferred that the Q304E substitution does not directly disrupt the GAF-AAA+ interface. To investigate the mechanism by which the Q304E variant of NorR escapes GAF-mediated repression of AAA+ activity, the protein was purified prior to structural and biochemical studies.

9.2 Targeted mutagenesis at position 304

Initially, the escape-mechanism of the Q304E variant was studied by making a number of other substitutions at this position (Figure 9.1). β-galactosidase assays showed that removal of the amino acid side chain (i.e. alanine substitution) does not affect the ability of NorR to activate transcription in response to NO, suggesting that the Q304 residue does not have a role in maintaining GAF-mediated repression of AAA+ activity. This is unsurprising, since the Q304 residue is not predicted to be located within the region of nucleotide-induced conformational change (Figure 6.5 B and C) that is proposed to be the target of the GAF domain in the mechanism of negative control in NorR. The Q304D variant of NorR gives a similar level of activity to the Q304E variant, with a partial escape from repression (Figure 9.1). In contrast, the Q304N and Q304R variants have null phenotypes and are unable to activate transcription irrespective of the presence of nitrite in the β-galactosidase assay, although stability was not tested. The requirement of a polar carboxyl group in the side chain of the 304 residue for escape is reflected in the ability of the glutamate and aspartate, but not the glutamine (wild-type) or asparagine changes to produce constitutive activity. Interestingly, the non-polar proline substitution also gives significant activity in the absence of an NO-source (Figure 9.1). Such a residue at this position would cause significant distortion in and around helix 4 (H4) and therefore may lead to "escape" by an indirect mechanism.
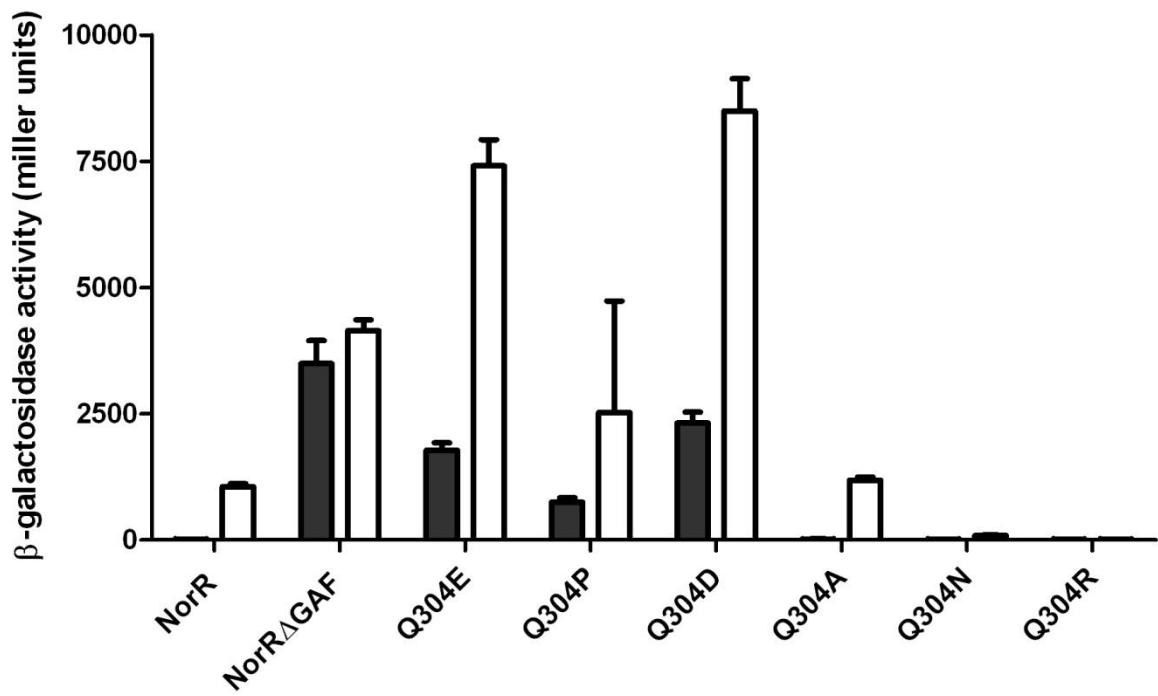
Figure 9.1 - Transcriptional activation by mutants of the Q304 residue of NorR *in vivo* as measured by the *norV-lacZ* reporter assay. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking the GAF domain (residues 1-170). 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.

9.3   Influence of GAF domain substitutions on the Q304E phenotype

Since β-galactosidase assays demonstrate that the Q304E substitution produces a partial escape phenotype (Figure 9.1), it was of interest to determine what role the GAF domain plays in the regulation of NorR activity. In order to confirm that the NO-sensing function of the GAF domain contributes to the phenotype of the Q304E variant, targeted substitutions were made at residues known to disrupt the non-heme iron centre in the GAF domain (Tucker *et al.* 2007). The Y98L, R75K, D99A, H111Y and C113S variants gave rise to a null phenotype *in vivo* and the mutant proteins were unable to activate transcription in the absence or presence of an NO-source. In Chapter 6 it was shown that when such additional substitutions were made in the fully constitutive GAFTGA variant G266D, there was no reduction in NorR activity *in vivo*. In the partial-escape variant Q304E, the apparent NO-induced activity in the on-state suggests a requirement for an intact iron-centre. Therefore, it might be expected that such substitutions would result in a reduction of activity in the on-state to the level observed in the off-state. Surprisingly however, these additional substitutions instead led to an *increase* in the activity of the Q304E variant protein (Figure 9.2). In the case of the R75K-, D99A- and H111Y-Q304E variants, the additional GAF-domain substitutions gave rise to a fully-escaped phenotype. The Y98L-Q304E and C113S-Q304E variants were still subject to partial regulation by the GAF domain but there was a significant increase in the activity of the double mutant compared to the Q304E-variant. This suggests that the non-heme iron-centre may play a role in maintaining the interface of interdomain repression in the Q304E variant of NorR. Furthermore, the differences in the requirement for NO-activation between the GAFTGA variant G266D and the Q304E variant further suggests that these mutant-versions of NorR escape repression by entirely different mechanisms.
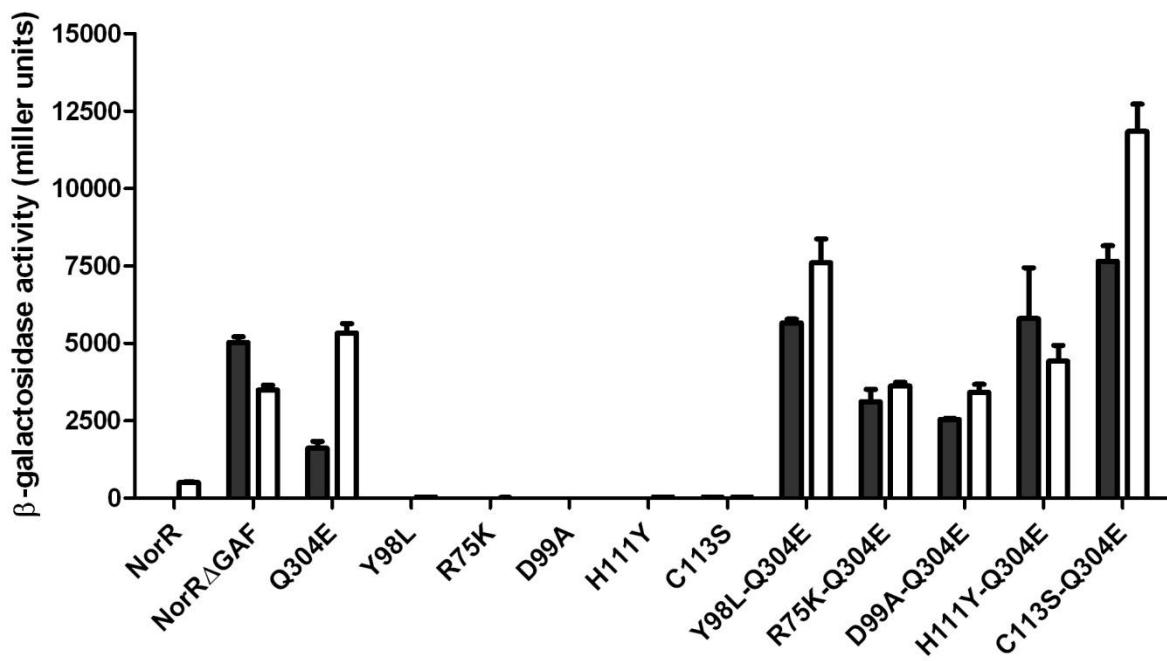
Figure 9.2 - Activities of Q304E variants *in vivo* as measured by the *norV-lacZ* reporter assay when additional substitutions are made in the GAF domain to disrupt the non-heme iron centre (Tucker *et al.* 2007). Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein and "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.

9.4   *In vitro* studies of the partial escape variant Q304E

9.4.1   Purification of full-length NorRQ304E

In order to further study the mechanism of escape in the Q304E variant, the protein was biochemically characterised. Initially, the full-length form of the Q304E variant protein was overexpressed and purified without a tag using heparin affinity chromatography. However, like the GAFTGA variant G266D, the Q304E mutant protein was largely insoluble compared to wild-type NorR when overexpressed in its non-tagged form (Figure 9.3A and B). Therefore, the protein was overexpressed and purified in the pETNdeM11 vector which encodes an N-terminal, TEV-cleavable, hexahistidine tag (Figure 9.3C and D). The Q304E-His protein was much more soluble than the previously characterised G266D-His protein. As is typical for NorR and its variants, the protein eluted in the range of 100 mM -200 mM imidazole and under these conditions, the Q304E protein showed a propensity to precipitate. Therefore, Q304E-containing fractions were loaded immediately onto a Superdex 200 16/60 column (Amersham Biosciences) to remove the imidazole as well as any impurities (Figure 9.3 E and F). NorRQ304E-His had an elution volume of around 65 ml, corresponding to a molecular weight in the dimer-trimer range, as was observed for wild-type NorR-His. Concentration using an Amicon Ultra spin column (Milipore) gave on average 10-15 μM of pure protein.

9.4.2   Purification of Q304EΔGAF

In order to investigate the effect of the Q304E substitution on NorR activity in the absence of the N-terminal GAF domain, a Q304E variant that additionally lacks the first 170 residues was purified via an N-terminal, TEV-cleavable, hexahistidine tag. As is the case for NorRΔGAF-His and the GAFTGA-mutant derivatives, Q304EΔGAF-His is highly soluble when overexpressed. Nickel affinity chromatography resulted in pure protein that elutes in the range of 100 mM to 200 mM imidazole. As was the case for all NorR constructs, under these conditions there was a propensity for the protein to precipitate at high concentrations. Gel filtration was employed to remove the imidazole and Q304EΔGAF-His eluted in the range of 70-80 ml as was observed for wild-type NorRΔGAF-His, corresponding to a molecular weight in the monomer-dimer range. Concentration using an Amicon Ultra spin column (Milipore) typically gave 100-150 μM of pure protein.
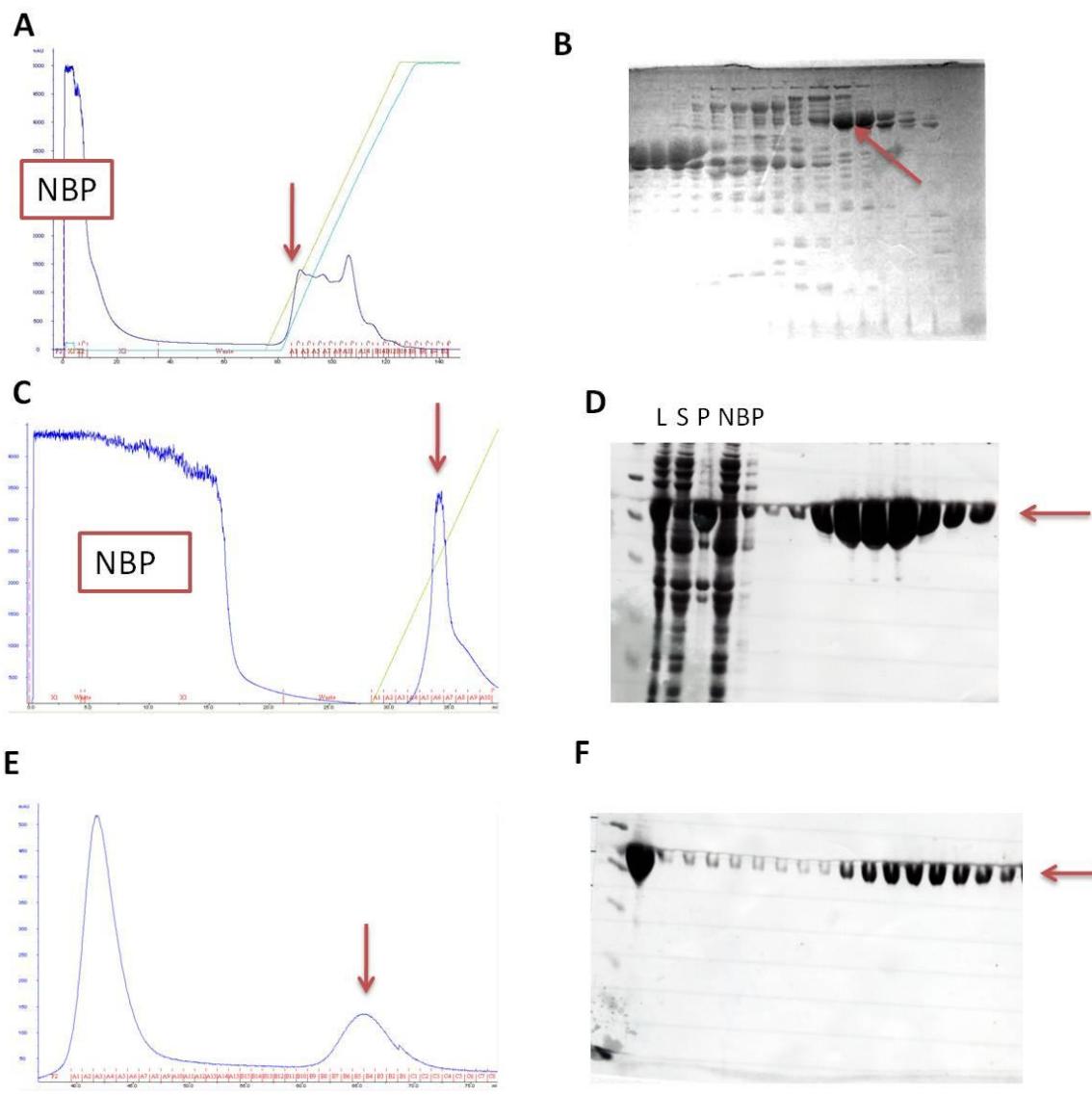
Figure 9.3 – Purification of full-length Q304E (**A**) Heparin affinity chromatography for non-tagged Q304E showing non-binding pool (NBP; not shown for NorR) and protein eluted using an increasing concentration of NaCl. (**B**) SDS-PAGE gel of bound protein. The presence of the protein (55.25 kDa monomer) is indicated by a red arrow. Both the wild-type NorR (Figure 70 A and B) and variant Q304E proteins eluted in the range of 200 mM-500 mM NaCl but the Q304E protein was significantly less soluble. NorR-containing fractions were subsequently loaded onto a 124 ml superdex 200 16/60 column for gel filtration. Wild-type NorR eluted in the range of 65-75 ml but pure fractions of the Q304E variant were not obtained (data not shown). (**C**) Purification of Q304E-His by affinity chromatography. Nickel affinity chromatography showing non-binding pool (NBP) and protein eluted using an increasing concentration of imidazole. The Q304E-His protein was significantly more soluble than the non-tagged form and bound to the Ni^{2+} column , eluting in the range of 100 mM-200 mM imidazole. (**D**) SDS-PAGE gel of bound protein eluted by increasing imidazole concentrations. L = lysate, S = supernatant, P = pellet, NBP = Non Binding Pool. (**E**) Gel filtration of selected NorR-containing affinity fractions using the 124 ml superdex 200 16/60 column. Q304E-His eluted at 60-70 ml. The peak around the void volume (43 ml) did not contain any protein (**D**) SDS-PAGE gel of the eluted protein from gel filtration. In each case the presence of the Q304E-His protein (58.44 kDa monomer) is indicated by a red arrow.

9.4.3 The Q304E mutation does not affect enhancer binding of NorR *in vitro*

Transcription factors of the bEBP family are commonly regulated through the control of their oligomeric states. Binding of NorR to the three enhancer sites in the *norR-norVW* intergenic region has been shown to be essential for the formation of stable oligomers capable of hydrolysing ATP. Furthermore enhancer DNA appears to be a key ligand in the activation of NorR as a transcription factor (Tucker *et al.* 2010a). Therefore, it was first investigated whether the Q304E mutation influences the binding of NorR to enhancer DNA. Electrophoretic Mobility Shift Assays (EMSA) were employed to measure the binding of the purified wild-type and Q304E variants in full length and truncated ($\Delta 1$-170) forms to two different fragments of the *norR-norVW* intergenic region. The 266bp and 361bp fragments both contain the three NorR binding sites. Recent work has demonstrated that the NorRΔGAF protein has a greater affinity for the longer 266bp fragment than a minimal 66bp fragment spanning the NorR binding region only (Tucker *et al.* 2010a). Therefore the DNA that flanks the NorR binding sites has been implicated in stabilising the NorR-DNA complex, possibly via DNA-wrapping. In order to investigate the extent to which flanking DNA stimulates enhancer-binding, the longer 361bp fragment was also employed in the EMSA experiments. The affinity of NorR and NorRΔGAF for the intergenic fragments was not significantly influenced by the presence of the Q304E substitution (Figure 9.4). For NorR-His and Q304E-His, dissociation constants (Kd) were estimated as 1-2 nM for the 361bp fragment (Figure 9.4A, closed symbols) and roughly 5 nM for the 266bp fragment (Figure 9.4B, closed symbols). For NorRΔGAF-His and Q304EΔGAF-His, dissociation constants (Kd) were estimated as 75-100 nM for the 361bp fragment (Figure 9.4A, open symbols) and 150-200 nM for the 266bp fragment (Figure 9.4B, open symbols). Therefore, it can be concluded that the Q304E substitution does not partially bypass the GAF-mediated repression of the AAA+ domain by altering the affinity of binding to enhancer DNA. However, the results do confirm that NorR and its variants have an increased affinity for the longer 361bp fragment compared to the 266bp fragment of the intergenic region, in line with the hypothesis that flanking DNA stabilises the hexamer. NorR is likely to bind to the intergenic region *in vivo* with a very high affinity and this may contribute to the rapid response of NorR to NO-induced stress. Intriguingly, the EMSA experiments also reveal that the full-length forms of wild-type NorR and its variants have a greater affinity for enhancer DNA than truncated forms that lack the GAF domain ($\Delta 1$-170).
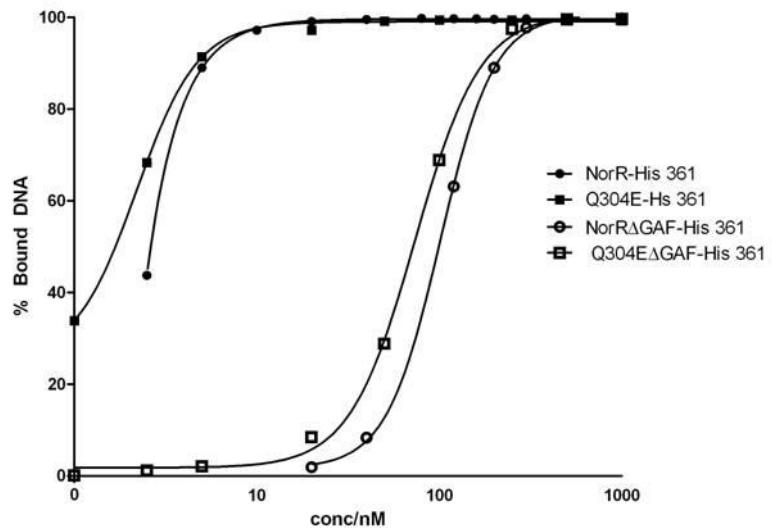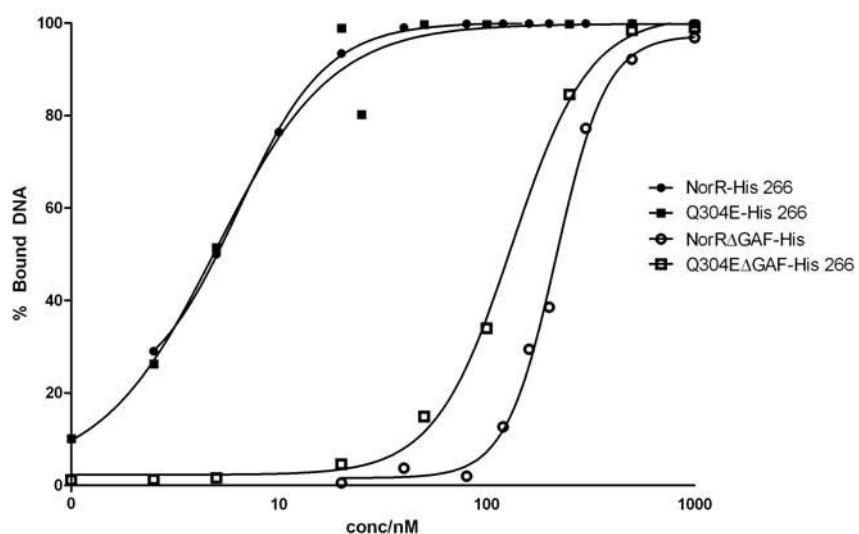
Figure 9.4 - Enhancer binding activity of the Q304E-His (closed squares) **and Q304EΔGAF-His** (open squares) **variants compared to NorR-His** (closed circles) **and NorRΔGAF-His** (open circles) **as determined by EMSA**. (**A**) Binding of NorR to a 361bp fragment of the *norR-norVW* intergenic region. (**B**) Binding of NorR to a 266bp fragment of the *norR-norVW* intergenic region. The percentage of fully shifted DNA was quantified using a Fujix BAS 1000 phosphoimager. The Q304E substitution does not significantly affect the affinity of NorR for either the 361bp or 266bp fragment of the *norR-norVW* intergenic region that contains the 3 NorR binding sites.

9.4.4 The GAF domain contributes to the DNA-binding activity of NorR

The DNA-binding assays in Figure 9.4 show that the Q304E substitution does not significantly affect the affinity of enhancer-binding in either the full-length NorR protein or a variant that lacks the first 170 amino acids (NorRΔGAF), both of which have hexahistidine tags at their N-terminus. Surprisingly, the full-length form of the tagged protein bound to both 361bp and 266bp *norR-norVW* intergenic DNA with a significantly greater affinity than the truncated variant. This suggests that the N-terminal GAF domain contributes to the strength of DNA-binding in NorR. An alternative explanation is that the hexahistidne tag that includes a linker region with a TEV-cleavage site is responsible for the differences in affinity. The tag may not influence DNA binding when placed at the N-terminus of the GAF domain in full-length constructs but may have a negative effect on DNA binding when placed at the N-terminus of the AAA+ domain in the ΔGAF constructs. In order to assess the role of the GAF domain in DNA binding in NorR, EMSA assays were conducted to compare the binding of tagged and non-tagged NorR to both the 266bp and 361bp fragments. The affinity of binding was significantly reduced in the absence of the regulatory domain irrespective of whether the protein is tagged or non-tagged (Figure 9.5). This suggests that the regulatory domain in NorR does contribute to the affinity of DNA-binding, at least *in vitro*. In contrast, *in vitro* ATPase and OPC assays as well as *in vivo* β-galactosidase assays have indicated that the absence of the GAF domain only removes the requirement for induction by NO rather than reducing the ability to activate transcription *per se*.

Although the GAF-domain contributes to DNA-binding in this assay for both NorR and NorR-His, there is a clear difference in the *size* of the reduction in affinity between tagged and non-tagged proteins. For example, binding to the shorter 266bp fragment (Figure 9.5B) had a dissociation constant (Kd) of 8.58 for non-tagged NorR and 30.1 for non-tagged NorRΔGAF, a reduction in affinity of 3.5-fold (red arrow). For the tagged proteins NorR-His and NorRΔGAF-His, dissociation constants (Kd) were 6.36 and 215.9 respectively, a reduction in affinity of approximately 35-fold (blue arrow). Hence, although the N-terminal His-tag had little effect on the affinity of the full-length protein for DNA, its presence caused a ten-fold greater reduction in the affinity when the GAF domain is absent. Similar effects were observed when the binding to the longer 361bp fragment was studied. This data is in contrast to β-galactosidase assays which showed no significant reduction in the *in vivo* activity of the NorRΔGAF protein when the N-terminal tag is
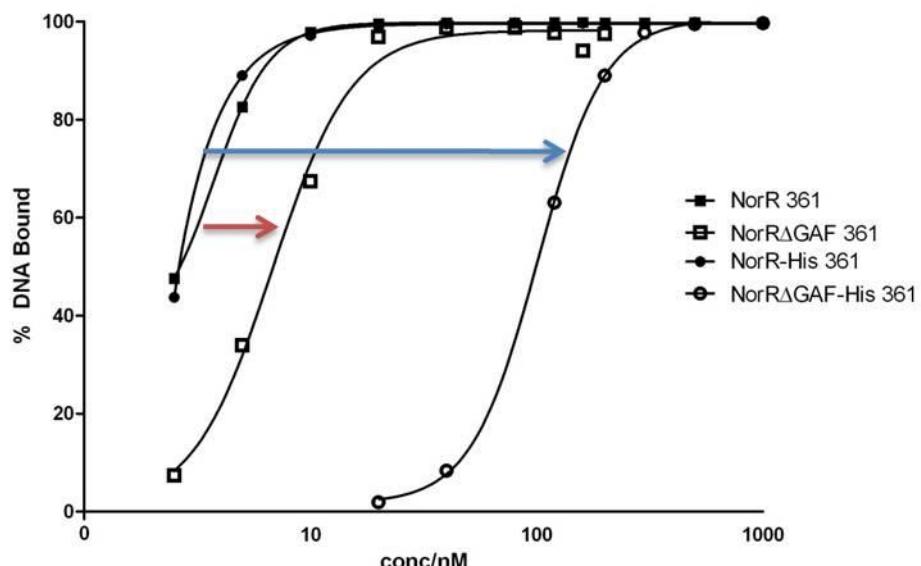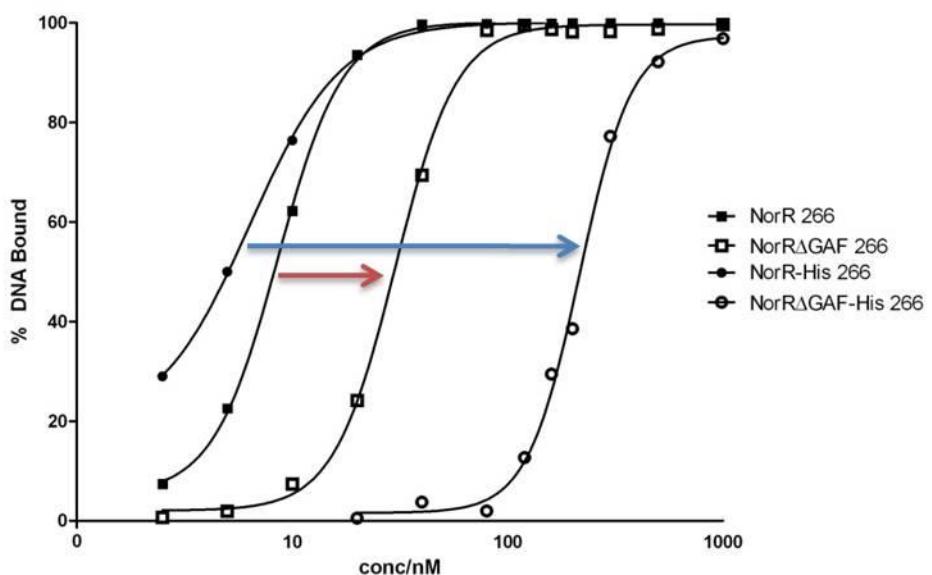
Figure 9.5 - Enhancer binding activity of non-tagged NorR (closed squares) **and non-tagged NorRΔGAF** (open squares) **compared to NorR-His** (closed circles) **and NorRΔGAF-His** (open circles) **as determined by EMSA**. (**A**) Binding of NorR to a 361bp fragment of the *norR-norVW* intergenic region. (**B**) Binding of NorR to a 266bp fragment of the *norR-norVW* intergenic region. The percentage of fully shifted DNA was quantified using a Fujix BAS 1000 phosphoimager. The absence of the N-terminal GAF domain significantly reduced the affinity of both the non-tagged NorR (red arrows) and NorR-His proteins (blue arrows) for either the 361bp or 266bp fragments of the *norR-norVW* intergenic region that contained the 3 NorR binding sites. However the presence of the N-terminal tag appeared to enhance this effect.

present (Compare Figure 6.9 and Figure 6.11). This is somewhat surprising since based on the dissociation constants calculated here, a ~7-fold higher protein concentration is required to give 50 % occupancy for NorRΔGAF-His compared to NorRΔGAF. It may be that *in vivo*, sufficient activator is present to activate transcription despite changes in the affinity of enhancer-binding.

In order to investigate the effect of the Q304E substitution on the oligomeric state of NorR, size-exclusion chromatography was performed. Results showed that the Q304E substitution did not alter the elution volume relative to wild-type NorR in either the full-length or N-terminally truncated forms. Unbound Q304E-His eluted at around 65 ml using the Superdex 200 16/60 column (Figure 9.3E) and the Q304EΔGAF-His protein eluted in the range of 70-80 ml (for typical profile see Figure 7.1C); corresponding to molecular weights in the dimer/trimer and monomer/dimer range respectively. In agreement with this, Electrospray-Mass Spectrometry (ES-MS) experiments performed by Ahyoung Park of Prof. Carol Robinson's group at the University of Oxford indicated that in the absence of DNA, the Q304E variant of NorRΔGAF-His is in equilibrium between monomeric and dimeric states (data not shown). Therefore, the oligomeric state of the Q304E variant appeared to be unaffected, at least under the conditions used for the purification of the protein. Next, it was important to determine whether the Q304E variant was able to oligomerise in the presence of intergenic DNA that contains the three NorR binding sites. Analytical gel filtration experiments were conducted in the presence of a 266bp DNA fragment that includes the *norR-norVW* intergenic region by T. Ghosh of Prof. Xiaodong Zhang's group at Imperial College, London. In agreement with size-exclusion data from the protein purification, unbound Q304E and Q304EΔGAF variants eluted at similar volumes previously observed for NorR and NorRΔGAF. The presence of 1 mM ATP did not seem to alter the oligomeric state in these variants. In both cases, in the presence of the 266bp DNA fragment, the protein peak was shifted to elute at ~9 ml, indicating the formation of a higher order nucleoprotein complex. This is similar to the elution volumes observed for wild-type NorR-His (Figure 8.5A), G266D-His (Figure 8.3A) and the N-truncated proteins NorRΔGAF-His (Figure 6B) (Tucker *et al.* 2010a), and G266DΔGAF-His (Figure 7.3A) in the presence of DNA. In order to determine the exact oligomeric state of the Q304E variants, individual unbound and DNA-bound fractions from gel filtration (Figure 9.6A) were analysed by negative-stain electron microscopy (EM). Results showed that the full-length Q304E-His protein oligomerised in an enhancer-dependent manner to

form hexameric-like particles 146 Å in diameter (Figure 9.6C). In the absence of 266bp DNA, only smaller particles were observed (Figure 9.6B). This is in contrast to the full-length GAFTGA mutant G266D-His which oligomerised in an enhancer-independent manner to form particles 165 Å in diameter and with 7-fold symmetry (Figure 8.9). However, when a higher protein: DNA ratio (24:1 based on the monomer) was employed, larger oligomeric assemblies were visualised (Figure 9.7A). 2800 particles were selected and class averages showed heptameric particles 160 Å in diameter, similar to the dimensions of the G266D-His oligomer. In the G266D variant, the heptamers visualised were formed in the absence of enhancer DNA and it is not known here whether or not the Q304E-His heptamers are bound to the 266bp fragment. Cryo-EM analysis of the AAA+ proteins p97 and MCM identified double-rings (Rouiller *et al.* 2002; Beuron *et al.* 2003; Costa *et al.* 2006b; Costa *et al.* 2006a; Davies *et al.* 2008) and 2D-side-views of the Q304E-His heptamer typically revealed a two-tiered (~14mer) oligomer (Figure 9.7B). However, smaller 140 Å particles were also observed and the dataset showed significant heterogeneity. Consequently further single particle reconstruction was discontinued. In line with what has been observed for the N-terminally truncated proteins NorRΔGAF (Tucker *et al.* 2010a) and G266DΔGAF-His (Figure 7.3), the Q304EΔGAF-His protein formed hexameric-like particles but only in the presence of DNA. The presence or absence of ATP did not appear to affect oligomerisation in this case (Figure 9.8B and 9.8C). Taken together, the gel filtration, Cryo-EM and ES-MS data suggest that the Q304E substitution does not significantly affect the oligomerisation of NorR. Hexamerisation of the NorR variant is strictly dependent on the presence of enhancer DNA, although at high protein concentrations, a small proportion of heptamers appear to form for the full-length protein. However, it is possible that this substitution influences the oligomerisation state in a manner that cannot be visualised under the conditions employed here.
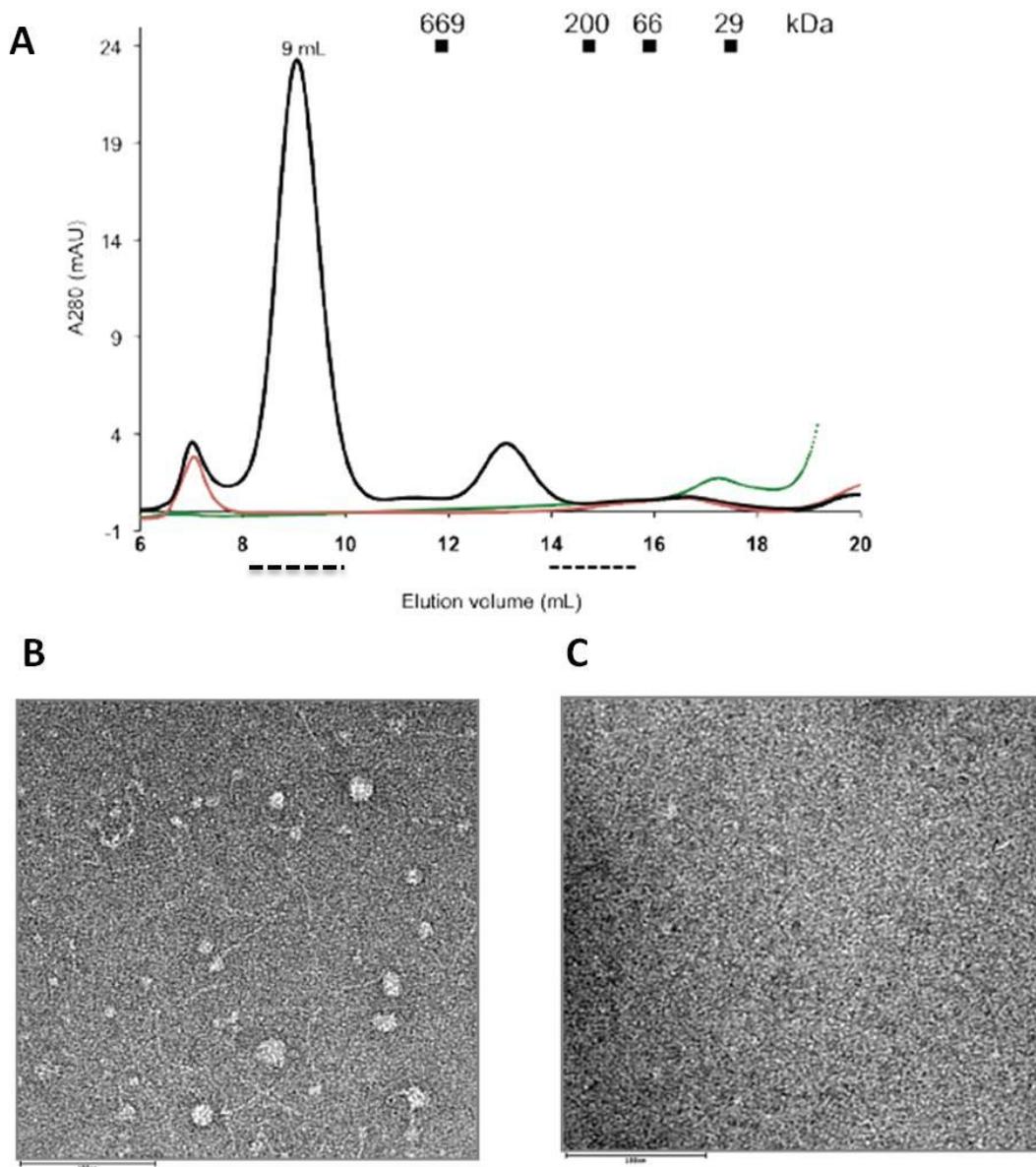
Figure 9.6 – Enhancer-dependent higher order oligomeric assembly of Q304E-His mutant. (**A**) Gel filtration chromatography of 6 μM Q304E-His variant alone (*red line*), in the presence of 1 mM ATP (*green line*) and the presence of 0.2 μM 266bp dsDNA (molar ratio of 12:1 monomer:DNA), containing all three enhancer sites (*black line*), performed at 4 °C using a Superose 6 column (24 ml). The presence of DNA stabilises a higher order oligomeric form of full length Q304E variant. Complex peak eluted at 9 ml and the fractions were visually analyzed by negative stain electron microscopy. The dotted lines below the elution peaks (9 ml and 15 ml) represent the fractions analyzed by negative-stain electron microscopy. Corresponding molecular weights of standard globular proteins are indicated at their elution volumes. (**B,C**) Negative-stain EM studies. Shown are raw micrographs of Q304E-His alone (**C**) and in complex with 266bp DNA (**B**). The 266bp DNA can be visualised as short fibres in the micrographs, scale bar 100 nm. At the 12:1 ratio, ring-shaped oligomeric particles were only observed in the presence of DNA. **Experiments were conducted by Tamaswati Ghosh as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

253
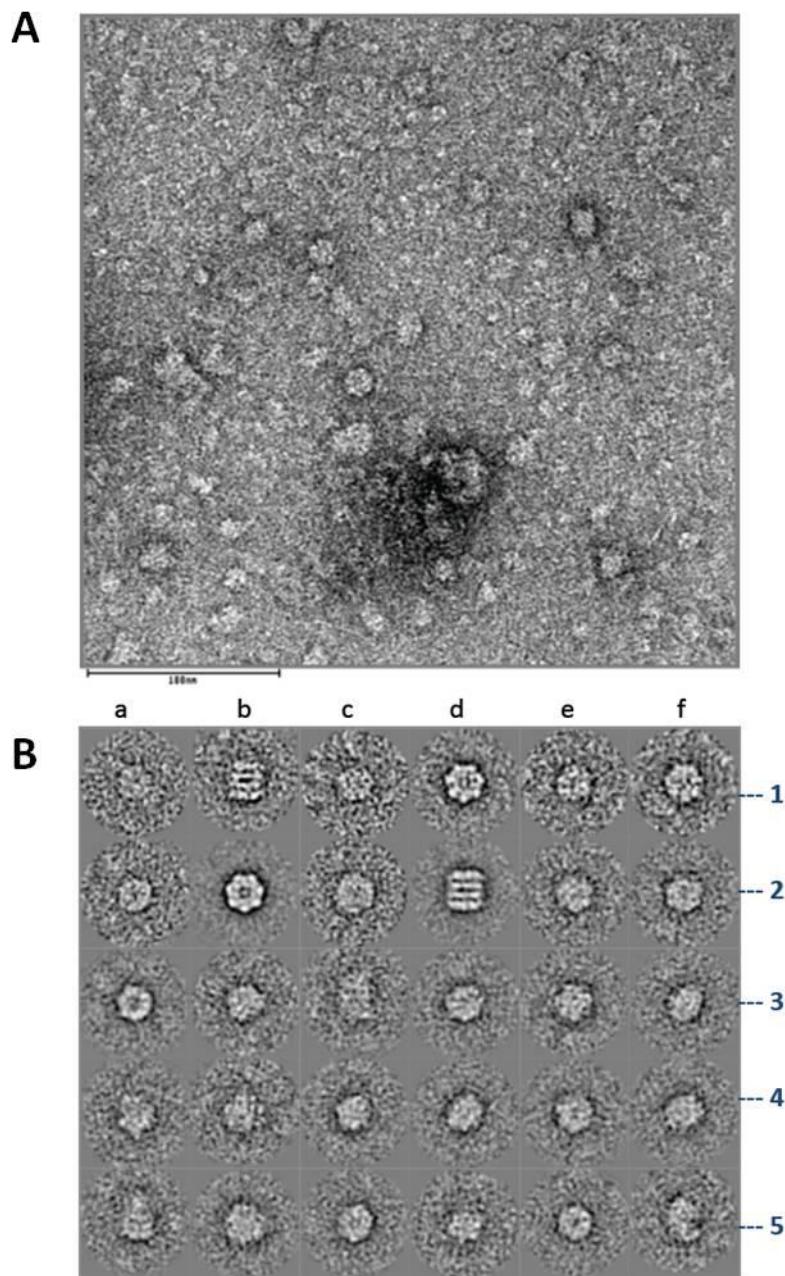
Figure 9.7 - Q304E variant can assemble into heptamer rings. (**A**) Raw micrographs of Q304E-His in complex with the 266bp DNA fragment (protein: DNA ratio of 24:1 based on monomer). Scale bar is 1000 Å. (**B**) Selection of class averages generated. The rows (1-5) and columns (a-f) of the figure panel are labelled. Locations 1.d, 1.f and 2.b are typical top views of a NorR heptamer. 1.b and 2.d represent side views of a double heptamer (14-mer). **Experiments were conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

	A	B	C
DNA	−	−	**+**
ATP	−	**+**	−
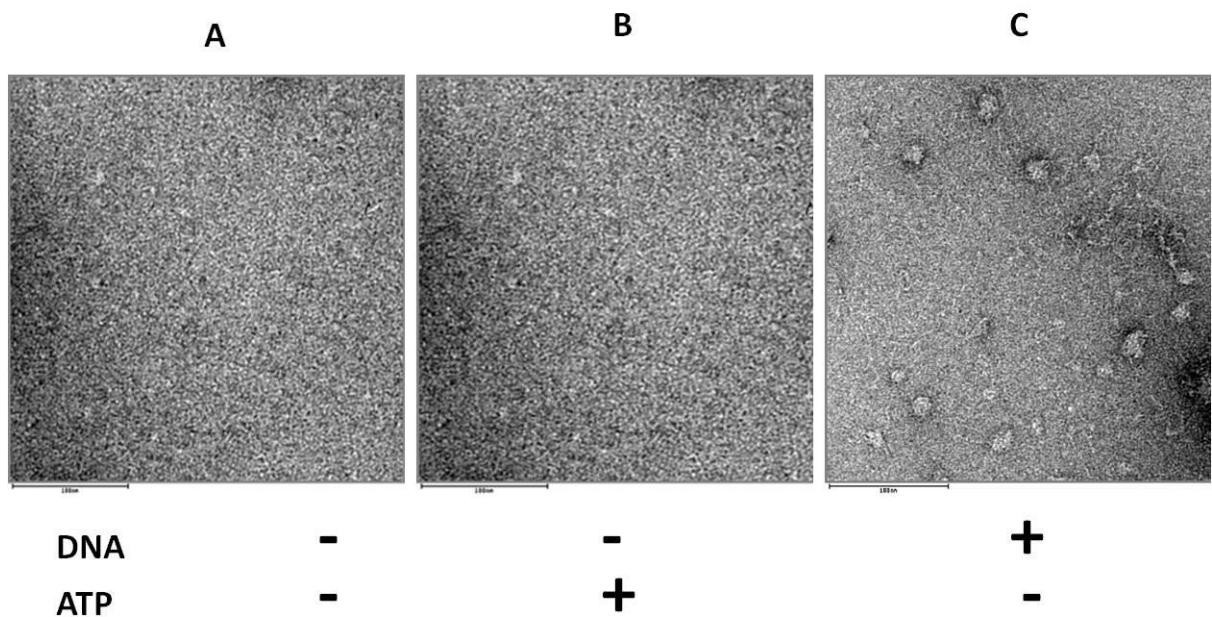
Figure 9.8 - Negative-Stain Electron Microscopy of the Q304EΔGAF-His (Δ1-170) protein in the absence of the 266bp fragment of the *norR-norVW* intergenic region and ATP (**A**), in the presence of ATP (**B**) and in the presence of DNA (**C**). Ring-shaped oligomeric particles were observed in the presence of DNA. Scale bar 100 nm. Protein: DNA ratio of 12:1 based on monomer. **Experiments were conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

9.5   3D-reconstruction of the full-length Q304E-His protein in the presence of enhancer DNA

Cryo-EM analysis of unbound and DNA-bound protein has revealed the enhancer-dependent hexamerisation of the full-length Q304E variant. Previous single particle analysis of the G266D-His protein led to the generation of a 3D-reconstruction of the heptamer in the absence of DNA (Chapter 8). However, this form of the protein was shown to be non-functional and is therefore probably not physiologically relevant. As a partial-escape variant, 3D-reconstruction of the Q304E-His protein might allow the DNA-bound hexameric structure of the "on-state" to be examined. Therefore, negatively-stained EM images of the Q304E-His protein bound to 266bp DNA were collected by collaborators at Imperial College, London before single particle reconstruction (Figure 9.9A).

An initial reconstruction was generated with no symmetry constraints. However, iterative refinement of the model did not give rise to consistent side or intermediate views. Therefore, to determine the shape of the ring, six-fold symmetry was imposed and the model refined via iterative cycles until the reprojections showed strong correlation to the class-averages used in the reconstruction. Results showed the appearance of EM-density at the bottom of the central pore with connecting density to both the bottom and the top faces of the ring. This arrangement is very different to that observed for the 3D-reconstruction of the G266D-His heptamer. In order to confirm that this density at the centre of the ring was not due to an artefact of staining or due to the six-fold symmetry constraints, this 3D model was refined with no symmetry constraints. After iterative refinements, one of the two faces displayed three-fold symmetry, correlating with the binding of three NorR dimers to enhancer DNA. Consequently, three-fold symmetry was imposed to improve the reconstruction given the relatively small dataset (Figure 9.9B and C). The final 3D reconstruction of Q304E-DNA complex obtained by applying three-fold symmetry included 1900 particles in 60 class averages with a resolution of 24 Å. Picking more particles to increase the size of the data-set would increase the signal to noise ratio and the overall resolution achieved. The structure is composed of two stacked rings with the upper ring structure consistent with an oligomer of six NorR subunits, 155 Å in diameter (Figure 9.10A). The central pore is 45 Å in diameter and appears blocked, although this could be due to the symmetry-constraints imposed during the reconstruction process. When viewed from the side, the top-face of the ring structure appears dome-like and is 54 Å in height. In
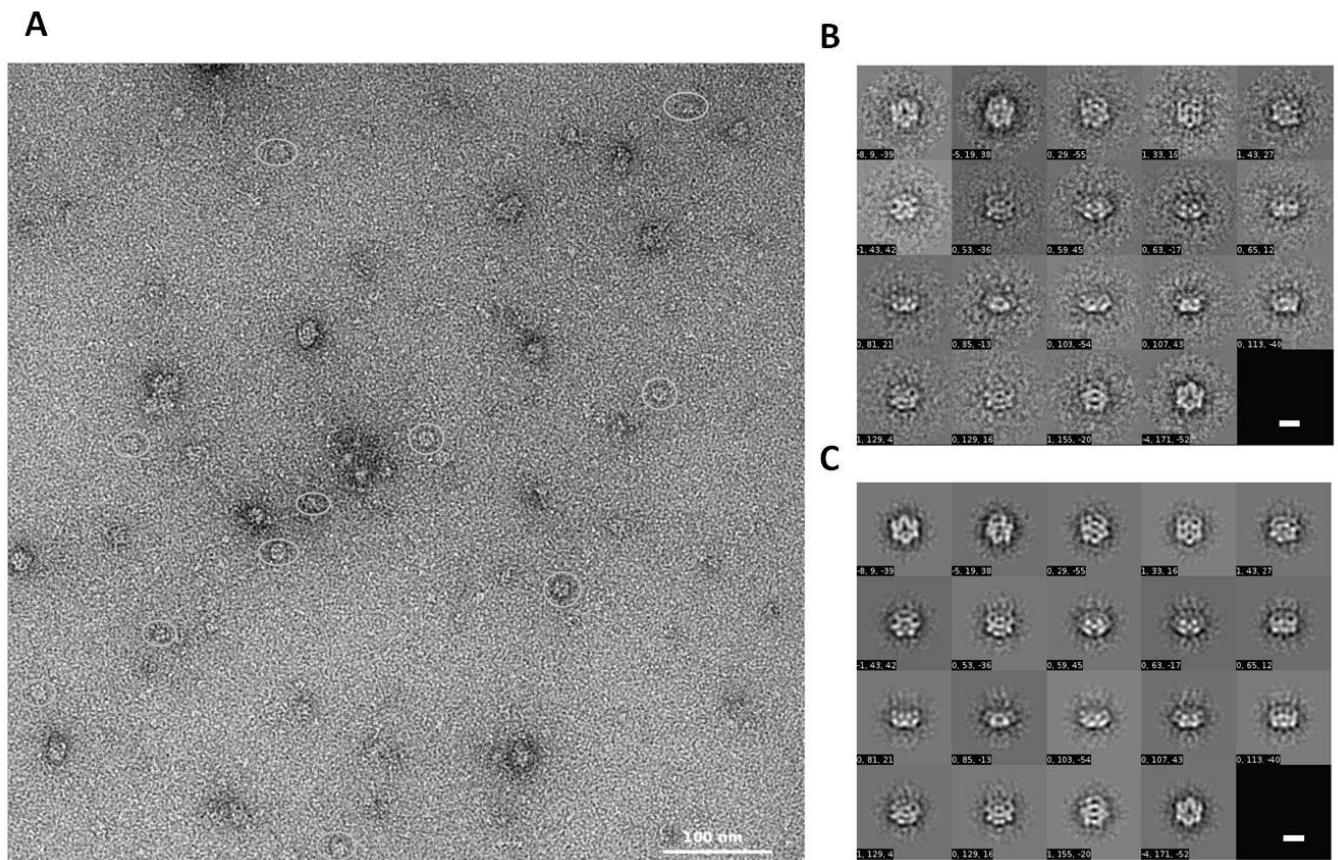
Figure 9.9- Negative stain EM analysis of Q304E-His. (**A**) Raw micrograph collected on the CM200 electron microscope at a magnification of 50000x. Some of the selected particles are circled. The scale bar is 1000 Å. (**B**) Selection of the best class averages (~ 10 particles/class) and (**C**) the corresponding reprojections from the final six-fold symmetric 3D model. **Experiments were conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**
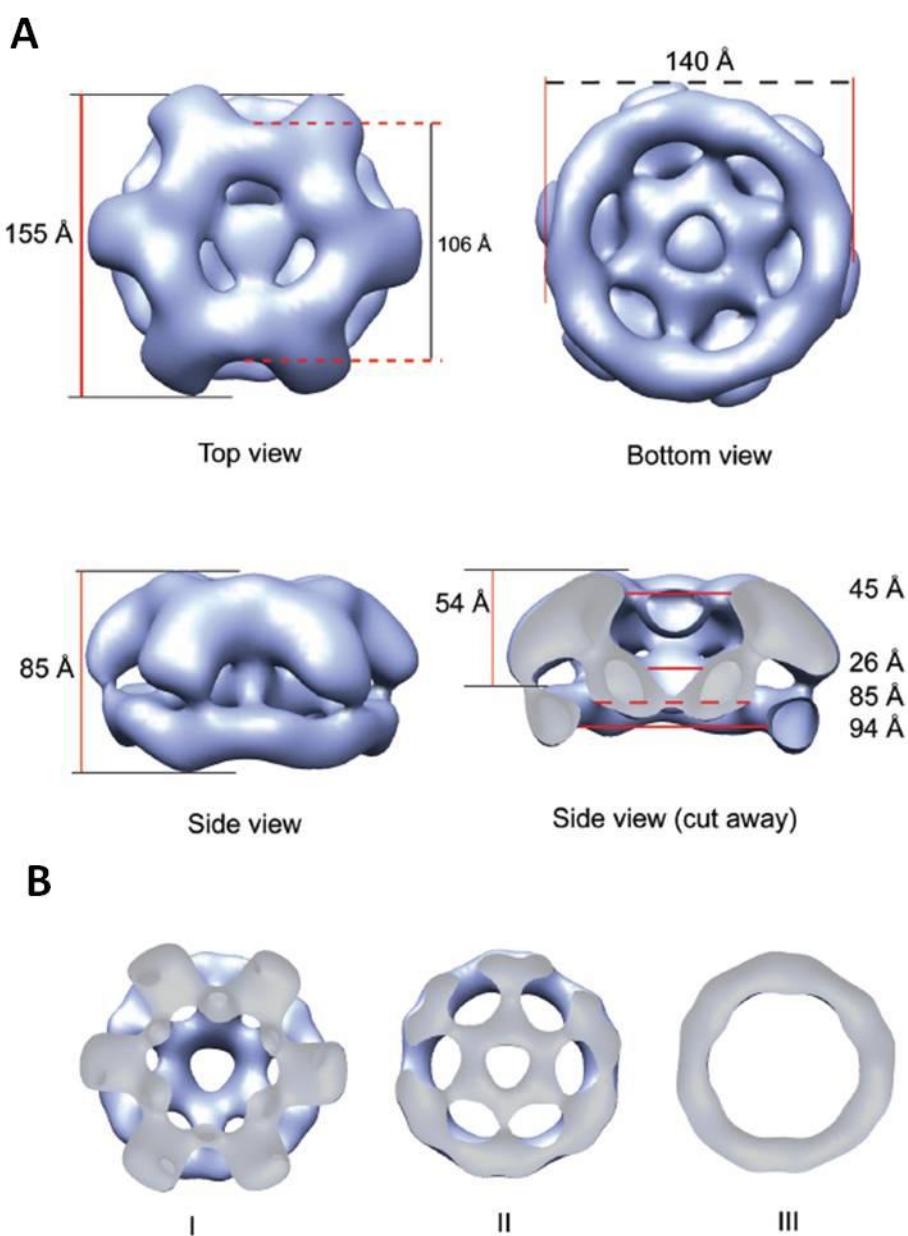
Figure 9.10 - 3D-reconstruction of Q304E-His bound to the 266bp DNA fragment containing all three enhancer sites. (**A**) A surface representation from the 3D reconstruction refined with three-fold symmetry constraints is shown in different orientations. The overall dimensions for the complex are given. A side-view has been cut-open to reveal the central chamber spanning the entire length of the molecule. (**B**) 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 densities, connected and surrounding the symmetry axis, are found in the cavity just below the hexameric ring with clear connections to the top (*A, side views*) and bottom (*II*) rings in the structure. **Analysis was conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

contrast, the lower ring shows no clear hexameric symmetry and is instead rounded in shape. It has a diameter of 140 Å with a wide (94 Å) central opening. Such dimensions are in line with the structures of other full-length, hexameric AAA+ proteins, including ADP-bound p97 (Beuron *et al.* 2003) and DNA-bound MCM (Costa *et al.* 2006b; Costa *et al.* 2006a). However, unlike these structures the 3D-reconstruction of Q304E-His does not have a hollow central channel. Six regions of EM-density are found just below the upper, hexameric ring and above the lower, circular ring (Figure 9.10B). This small, ring-like density has a diameter of 85 Å and has a central pore of 26 Å. Clear EM-density connects this ring to both the upper and lower rings in the 3D-model. In order to evaluate whether this density is an artefact due to the imposed symmetry constraints, the reconstruction was compared to that of the only other published Cryo-EM structure of a full-length bEBP (Figure 9.11). Superimposition of EM maps reveals a similar architecture for the top-half of the NorRQ304E-His protein as the ADP.AlF$_x$-bound structure of NtrC (De Carlo *et al.* 2006). Importantly, the small ring-like density below the hexameric ring in the Q304E-His reconstruction corresponds to the DNA-binding domain-containing density of NtrC. Therefore, the density in the reconstruction of the Q304E variant is not likely to be due to an artefact or due to imposed symmetry constraints but rather is likely to contain the DNA-binding domains of the protein. In the NtrC structure, the density of the upper-ring correlates to the presence of the AAA+ and regulatory domains, as it does in the 3D-reconstruction of the G266D-His heptamer (Chapter 8). Since the circular-region of density below the lower-hexamer is unaccounted for in the Q304E structure, it is possible that this correlates to the bound dsDNA that wraps around the oligomer.

9.5.1 An atomic model of the DNA-bound Q304E-His hexamer

In order to confirm the predicted locations of the three NorR domains in the reconstruction of the Q304E variant, atomic models of the NorR domains were fitted into the EM-structure. For the fitting of the central AAA+ and regulatory GAF domains, the model of monomeric NorR^{R+C} (lacking the DNA-binding domain), generated in the fitting of G266D-His heptameric 3D-reconstruction (Chapter 8), was used. This model was fitted into the Q304E-His hexameric reconstruction as a rigid-body such that the AAA+ domain occupied the main EM-density of the ring. This arrangement ensured that the L1 and L2 loops were surface-exposed and therefore well-placed for σ^{54}-contact. However, the
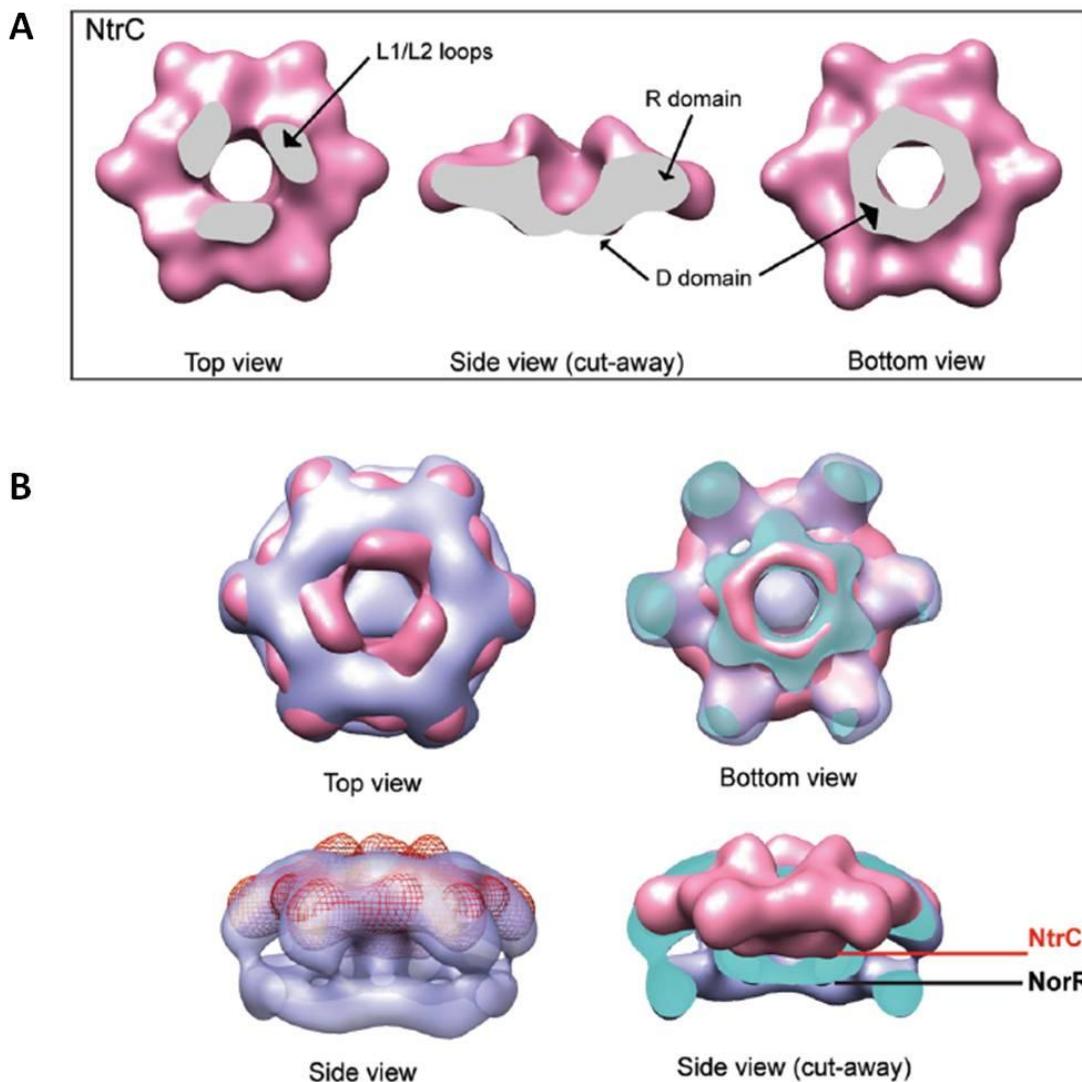
Figure 9.11 - Superimposed negative-stain EM maps of full-length, DNA-bound NorRQ304E-His and NtrC-ADP.AlF$_x$. (A) A surface representation of the 28 Å structure of NtrC (EMD 1218; (De Carlo *et al.* 2006) shown in different orientations. Positions of the N-terminal receiver (R) domain, the C-terminal 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. **Analysis was conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

EM-density at the upper-surface of the hexameric ring did not accommodate the GAFTGA-loops in an extended conformation, unlike the 3D reconstructions for NorRG266D-His (Chapter 8) and ADP.AlF$_x$-bound NtrC (De Carlo *et al.* 2006). This suggests that the L1 and L2 loops are in a collapsed state as would be expected in a structure that lacks bound nucleotide. A comparison between the two EM-maps of G266D-His and Q304E-His reveal similar diameters but significantly the G266D-His reconstruction has a much wider central channel that does not contain additional EM-density within it (Figure 9.12). Fitting of the regulatory GAF domains places them at the periphery of the AAA+ ring, as was observed for the 3D-reconstruction of G266D-His. Interestingly, the EM-density attributed to the GAF domains in the Q304E reconstruction reaches down to contact the circular structure believed to correspond to the position of the dsDNA. This would suggest that the GAF domains are well-placed to interact with the intergenic DNA upon derepression. Next, the ZraR DNA-binding domain which is predicted to have the same overall fold as the NorR DNA-binding domain was fitted manually as a rigid-body into the small ring-like density located below the hexamer. In this orientation, the predicted-recognition helix is surface-exposed and well placed to interact with enhancer DNA. After the location of the three NorR domains had been determined (Figure 9.13A, B and C), the EM-density of the lower circular-ring remained unaccounted for. This density is likely to correspond to the location of the dsDNA wrapped around the hexamer since the DNA-binding domains of the bEBP are predicted to be in close proximity. Initial fitting of a circular dsDNA model was conducted by collaborators (Figure 9.13D, E and F). The model shows the recognition helix of the NorR DNA-binding domain well-placed to bind to the major-groove of the enhancer DNA.

9.5.2 The 3D-reconstruction predicts that the Q304E-His protein is ATPase active *in vitro*

In order to assess the implications of the Q304E substitution upon the arrangement of protomers in the bEBP, inter-protomer interactions identified in the 3D-model were compared to those present in the crystal structure of the ZraR^{CD} hexamer. In particular the position of key residues located at the ATPase site, at the interface between two protomers, was examined. Since the published crystal structure of ZraR reveals bound nucleotide, an ATP molecule was positioned in the cleft between the α/β and α-helical subdomains such that the phosphate backbone wrapped around the P- loop (residues 214-222). The Walker A motif forms the P-loop and is essential for ATPase activity, contacting the phosphates of
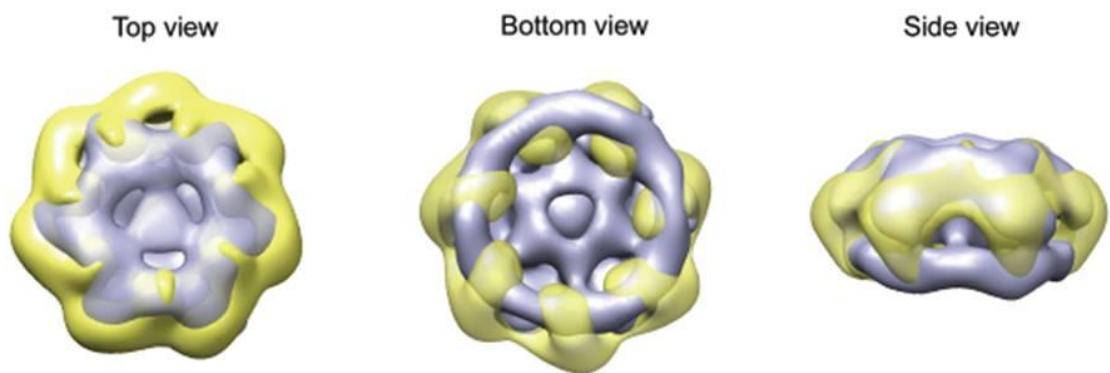
Figure 9.12 – Comparison of the EM-density maps for the Q304E-His hexamer and the G266D-His heptamer. Superimposition of the 3D maps of the full-length, unbound G266D-His variant (yellow) and the full-length DNA-bound Q304E-His variant (blue). The three-fold symmetry imposed 3D map of the Q304E nucleoprotein complex (blue) is superimposed on to the heptameric ring structure of the G266D mutant (yellow). Both structures were filtered to ~20 Å. **Analysis was conducted by Tamaswati Ghosh as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**
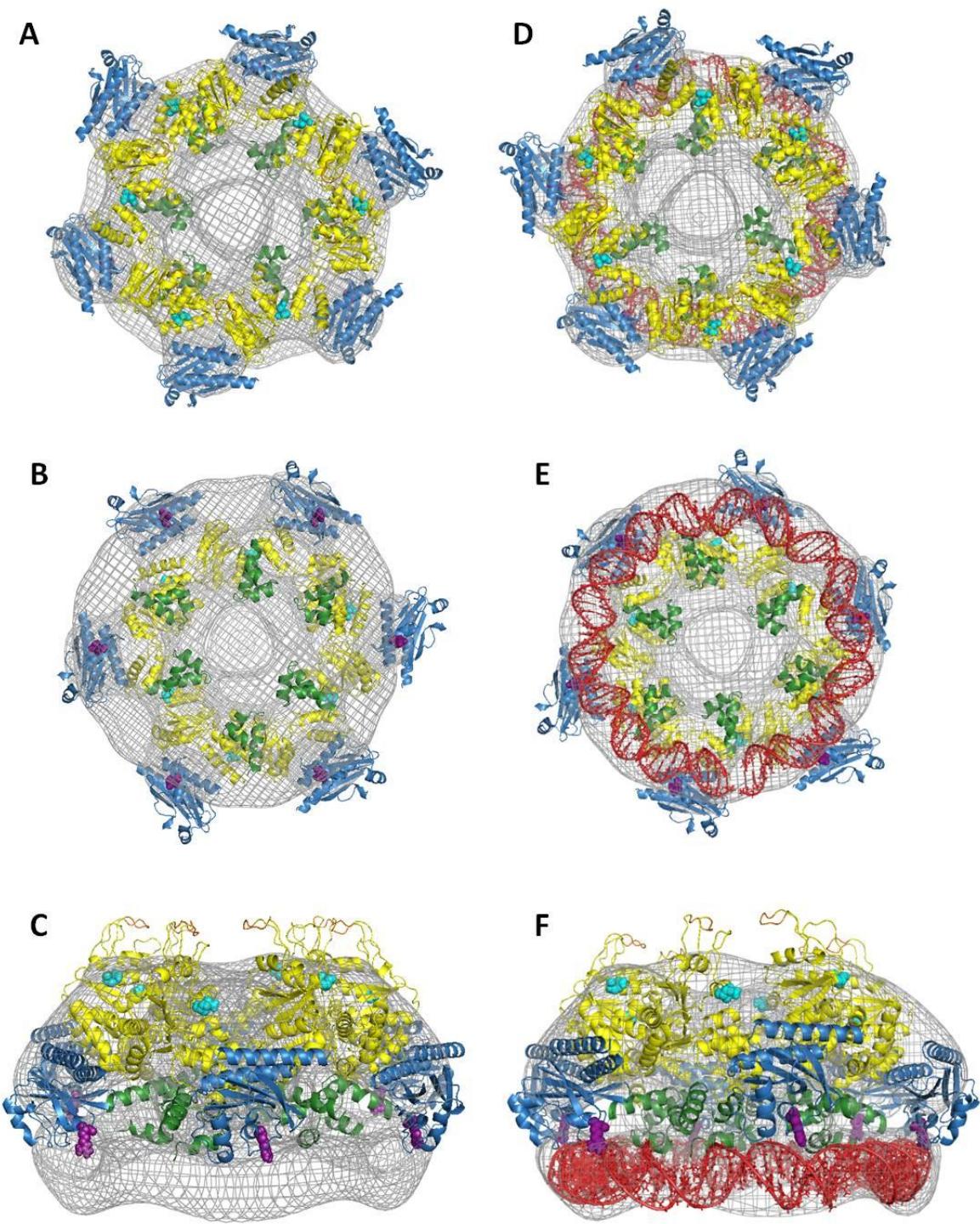
Figure 9.13 - Assignment of the NorR domains in the 3D reconstruction of NorRQ304E-His bound to enhancer DNA. Top (**A**), bottom (**B**) and side (**C**) view orientations of the model created by manual fitting of the atomic models of the three NorR domains, refined by applying three-fold symmetry restraints. The atomic models of the N-terminal GAF domain (blue), the NorR AAA+ domain (yellow), and the crystal structure of the ZraR DNA-binding domain (green; PDB code 1OJL, chain A) have been placed in the top half of the reconstruction (grey mesh). The bottom half of the map, consisting of a circular-density and displaying no symmetry, remains unoccupied. Fitting of a circular dsDNA model into the bottom half of the EM-map with top (**D**), bottom (**E**) and side (**F**) views. A representation of the DNA model is shown in red. Also shown are the positions of the GAFTGA motifs (orange) and R81 residues (purple spheres), both of which are thought to play key roles in maintaining the mechanism of repression. The Q304 residue is indicated by cyan spheres. **Analysis was conducted as part of collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

ATP (Saraste *et al.* 1990). In line with this the conserved lysine residues in both the NorRQ304E-His reconstruction (K221) and the ZraR crystal structure (K175) are suitably placed. Importantly, the conserved Walker B "DE" residues (D286, E287 in NorR; D240, E241 in ZraR) are also appropriately positioned for a role in catalysis. In addition the catalytically-active Sensor II argninine (R405 in NorR) is well placed in the 3D-model to coordinate the γ-phosphate prior to hydrolysis, in agreement with its position in the ZraR structure (R359). Finally the putative R-finger in the Q304E reconstruction (R341) has its side chain pointing towards the active site, as is the case in the ZraR structure (R295). Taken together this analysis suggests that in contrast to the 3D-reconstruction of the unbound G266D-His heptamer, the key catalytic residues are well placed in the model of the DNA-bound Q304E-His hexamer for efficient ATP hydrolysis. To test this prediction, biochemical characterisation of the Q304E-His variant was performed.

9.6   The Q304E variant shows enhancer-independent ATPase activity *in vitro*

The ability of bEBPs to hydrolyse ATP is central to their role as activators of σ^{54}-dependent transcription. Since the catalytic site is formed via the interactions of neighbouring protomers, oligomerisation of the activator is required to form the functional ATPase. This self-association is highly dependent on the individual protomers binding to the *norR-norVW* intergenic region, reflected in the enhancer-dependent ATPase activity that is observed for NorRΔGAF-His (Figure 7.4A). Analysis of the 3D-reconstruction of DNA-bound Q304E-His indicated that the protomers of the hexamer are tightly packed in the front-to-back configuration that is required for ATP-hydrolysis in related bEBPs (Figure 9.14). In order to more accurately investigate the effect of the Q304E substitution on the ability of NorR to hydrolyse ATP, ATPase assays were conducted *in vitro*. Results indicate that in the presence of DNA, the full-length Q304E-His protein was competent to hydrolyse ATP (Figure 9.15B, open squares), although compared to NorRΔGAF-His (Figure 9.15A, open squares), the curve is less cooperative in nature. Interestingly, Q304E-His also exhibited ATPase activity in the absence of DNA that contains the three NorR binding sites, indicating that, under the conditions of the assay, the protein was able to oligomerise in an enhancer-independent manner (Figure 9.15B, closed bars/closed squares). However, the addition of DNA to the reaction did cause a further increase in activity (Figure 9.15B, open bars/open squares) suggesting that the protein still needs to bind to the DNA to become fully active. In the absence of the N-terminal regulatory
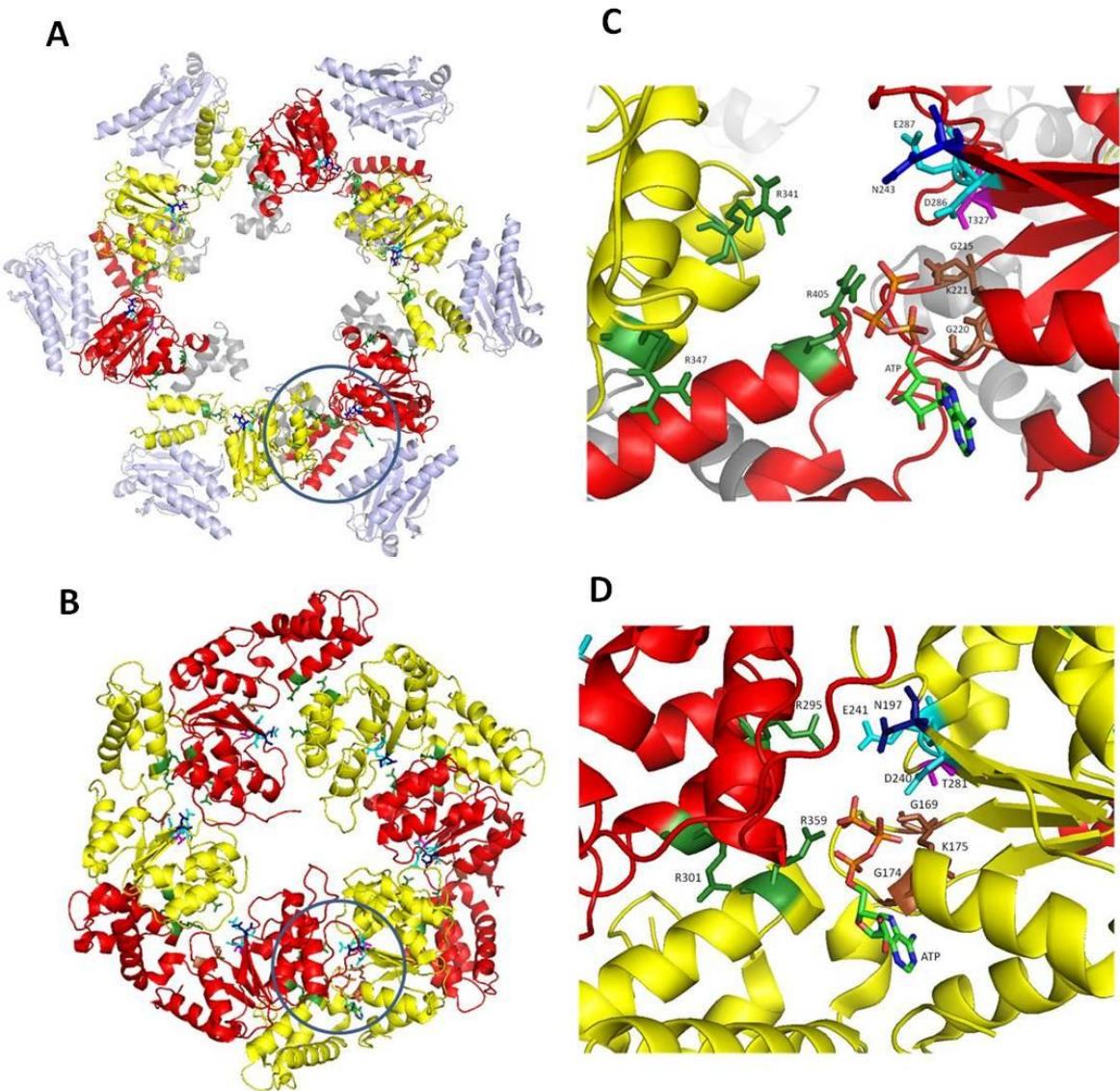
Figure 9.14 – Comparison of the ATP hydrolysis sites in the hexameric, DNA-bound NorRQ304E-His 3D reconstruction and the hexameric ZraR crystal structure. (**A**) Top view of the Q304E variant model showing the hexameric ring assembly. Consecutive AAA+ domains are shown in alternating red and yellow colours. GAF domains are in light blue and DNA binding domains in light grey. For clarity the DNA has been omitted. (**B**) The structure of the AAA+ domain of ZraR, built from PDB: 1OJL. Consecutive AAA+ domains are shown in alternating red and yellow colours. The regulatory domains are absent in the crystal structure and for clarity the DNA binding domains have been omitted. An example of the interprotomer interface at which the ATP hydrolysis site is found is shown in both models (blue circles). (**C**) Close-up of the ATPase active site located between adjacent AAA+ subunits in the atomic model of NorRQ304E-His and in the crystal structure of ZraR (**D**). Conserved residues implicated in ATP binding and hydrolysis, inter-subunit catalysis and relaying nucleotide states to the surface exposed L1/L2 loops are indicated: Walker A residues G220/G174, K221/K175, G215/G169 (brown), Walker B residues D286/D240 and E287/E241 (cyan), "switch" asparagine N243/N197 (dark blue), Sensor I T327/T281 (magenta), sensor II residue R405/R359 (green), and *trans*-acting putative R-fingers R341/R395 and R347/R301 (green). ATP molecules have been placed in the active site of both NorRQ304E-His (**C**) and ZraR (**D**). **Analysis was conducted as part of a collaboration with Prof. Xiaodong Zhang, Imperial College, London (Ghosh 2010).**

265
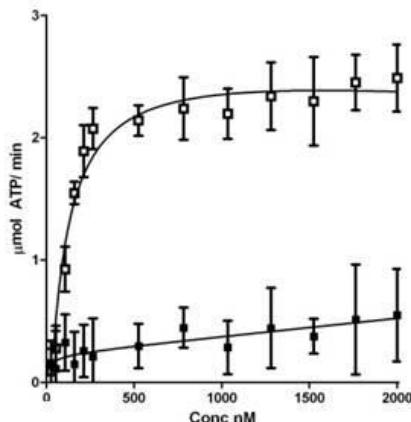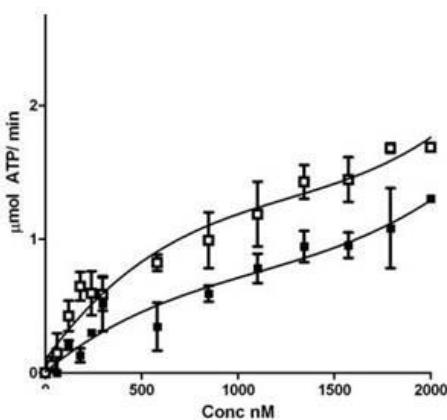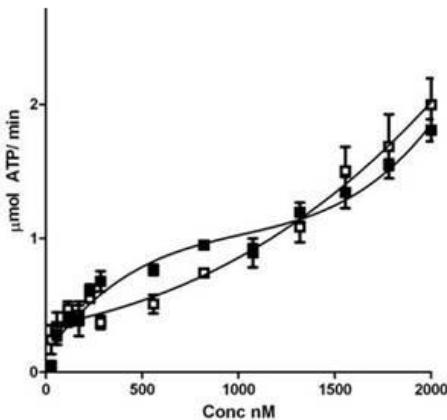
Figure 9.15 - ATPase activity of the NorRΔGAF-His (A), Q304E-His (B) and Q304EΔGAF-His (C) variants in response to protein concentration and the presence of enhancer DNA. Non-linear regression was carried out using GraphPad Prism software. Assays were conducted either in the absence (closed squares) or presence (open squares) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

domain, the Q304E variant also exhibited enhancer-independent ATPase activity, similar to that of the full-length mutant protein (Figure 9.15C), but in contrast to the truncated forms of wild-type NorR and the GAFTGA variant G266D (Figure 7.4A and B). However, compared to the NorRΔGAF-His protein, Q304EΔGAF-His was significantly less competent to hydrolyse ATP, particularly at lower concentrations. Cryo-EM indicates that like the full-length variant, Q304EΔGAF-His must bind DNA in order to oligomerise (Figure 9.8), apparently contradicting the enhancer-independent ATPase activity observed here.

In Chapter 8, it was shown that the full-length GAFTGA variant protein preparations exhibited strong DNA-independent ATPase activity that did not reflect turnover of ATP by NorR (Figure 8.12). It is likely that the presence of contaminating ATPases was responsible since the GAFTGA-variants did not show good separation from the void volume in gel filtration chromatography. In the case of Q304E-His, pure protein elutes away from the void, at a volume of around 65 ml on the Superdex 200 16/60 column (Amersham Biosciences). Furthermore, ATPase activities are in line with values that have previously been observed for bEBPs. This suggests that the ATPase activity observed for the NorR variant Q304E may represent a bona fide turnover of ATP by the NorR protein. To confirm this, an additional D286A substitution was made at the Walker B motif. This mutation renders bEBPs inactive for ATPase activity as observed for NorRΔGAF-His (Figure 7.5). Importantly, the additional substitution did not alter the gel filtration profile of the Q304E variant (data not shown); any contaminants present in the Q304E-His preparation would therefore also be expected to be present in the D286A-Q304E-His preparation. The Walker B substitution diminished the ATPase activity of the Q304E variant protein, suggesting that the ATP turnover observed for Q304E-His was not due to the presence of contaminating ATPase (Figure 9.16A). Likewise, ATPase assays confirmed that the alanine substitution in the Walker B motif rendered the Q304EΔGAF variant inactive, suggesting that the activity observed for Q304EΔGAF-His also represents a bona fide catalytic turnover of ATP (Figure 9.16B). Overall, it can be concluded that the Q304E substitution allows the NorR protein to hydrolyse ATP in an enhancer-independent manner. Whilst binding to DNA may be required for full NorR activity, it appears that functional ATPase units can be formed in the absence of the three NorR binding sites.
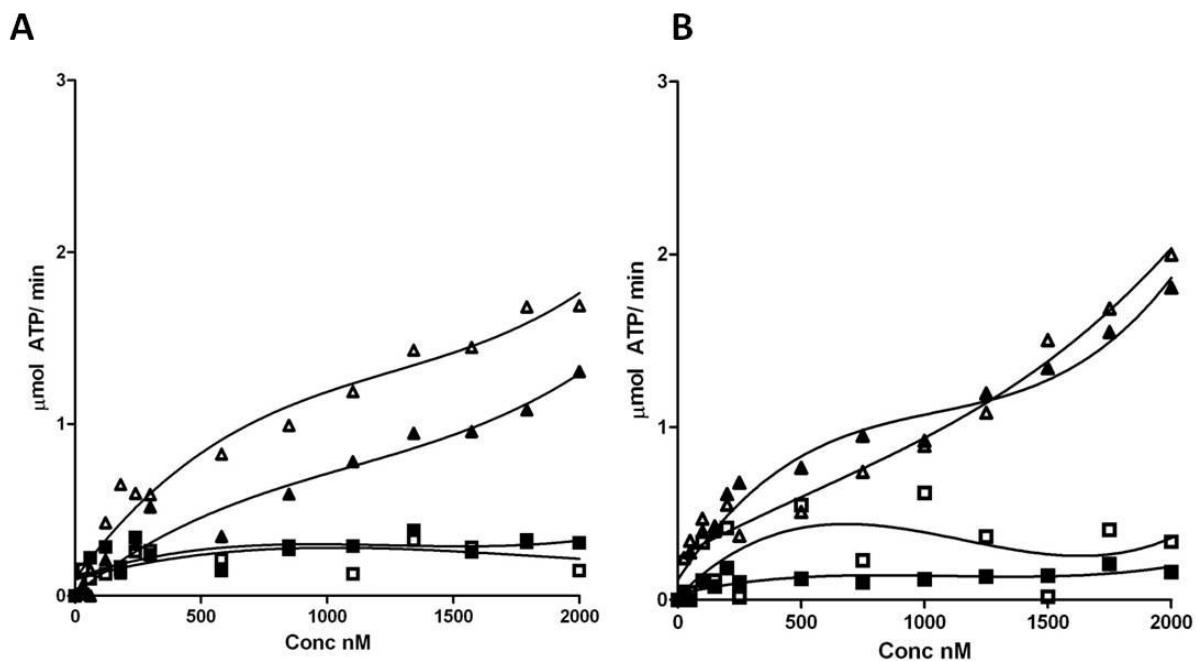
Figure 9.16 - ATPase activity of the Q304E-His (A) and Q304EΔGAF-His (B) variants when an additional D286A substitution is made in the Walker B motif, in response to protein concentration and the presence of enhancer DNA. Each data set shows the ATPase activity of Q304E variant (triangles) and the same variant with the additional D286A substitution (squares). For clarity error-bars are not included. Non-linear regression was carried out using GraphPad Prism software. All assays were conducted either in the absence (closed shapes) or presence (open shapes) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

9.7 Testing the requirement for ATPase activity in the NorR variant Q304E *in vivo*

Structural studies using the bEBP PspF bound to different nucleotide analogs has revealed the importance of the nucleotide driven-conformational change in order to expose the σ^{54} interaction surface of the AAA+ domain in the mechanism of transcriptional activation (Rappas *et al.* 2006; Schumacher *et al.* 2006). Although the Q304 residue is not expected to have a role in modulating the conformation of the σ^{54} interaction surface, the possibility that the Q304E substitution affects the position of the L1 and L2 loops, cannot be entirely ruled out. It was therefore important to confirm that the ability of the Q304E variant to activate transcription *in vivo* is dependent upon its ability to hydrolyse ATP. If the Q304E-mutant version of NorR can activate transcription without hydrolysing ATP, it would suggest that the Q304E substitution promotes a conformation of the AAA+ domain that models the transition state at the point of sigma contact. The ability of NorRQ304E to activate transcription *in vivo* was completely abolished when the additional D286A substitution was present (Figure 9.17). Similar results were observed for the both the full-length and GAF-truncated forms of wild-type NorR (Figure 9.17) as well as the GAFTGA variant G266D (Figure 7.7). Therefore like G266D, the Q304E substitution does not negate the requirement for ATPase activity and it can be concluded that the Q304E variant of NorR does not partially escape the GAF-mediated repression mechanism by modelling the post-ATP hydrolysis conformation of the AAA+ domain. This is in agreement with data which suggest that the ATP-hydrolysis machinery is not a direct target of negative control in NorR as it is in PspF (Figure 7.8).
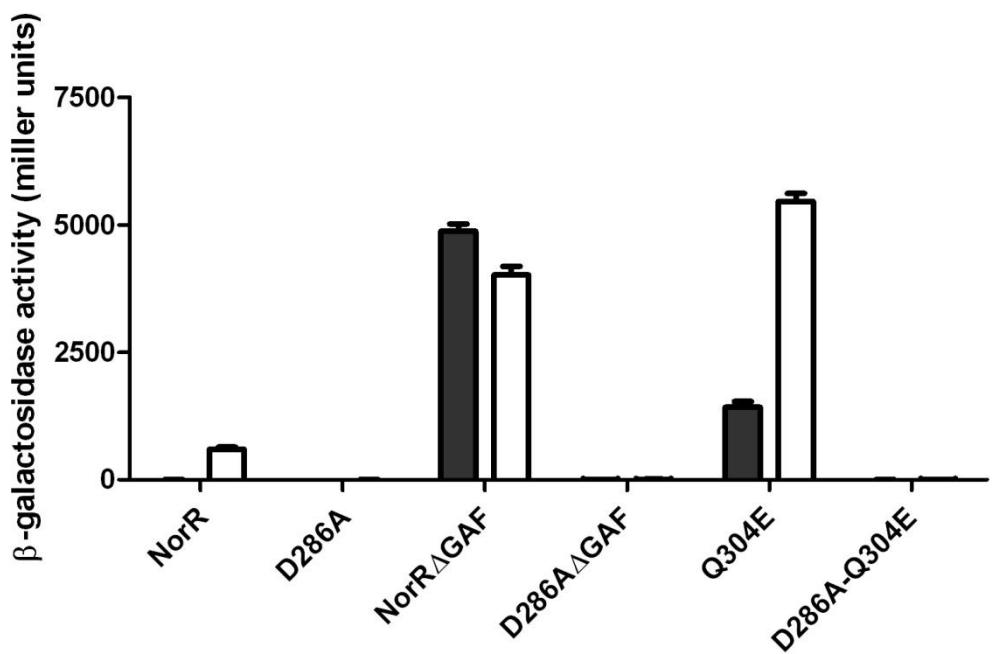
Figure 9.17 - Activities of NorR, NorRΔGAF and the NorRQ304E variant *in vivo* **when the additional D286A substitution is made at the Walker B motif in the AAA+ domain, as measured by the** *norV-lacZ* **reporter assay.** Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein 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.

9.8   Testing the requirement for enhancer binding in the NorR variant Q304E *in vivo*

The binding of NorR to the *norR-norVW* intergenic region that contains the three NorR binding sites is essential to form the ATPase-active oligomer that is competent to activate transcription. However, the Q304E-mutant version of NorR exhibited a certain level of enhancer-independent ATPase activity *in vitro* in both full-length and GAF-truncated forms (Figure 9.16), suggesting that this substitution may promote the formation of ATPase-active oligomers. Although, negative-stain EM indicated that the formation of Q304E hexamers was strictly DNA-dependent *in vitro* (Figure 9.6), it is possible that transcriptional activation by Q304E variant *in vivo* can occur independently of enhancer-binding. In order to test this, C-terminal truncations were made in the *norR* constructs to delete the helix-turn-helix (HTH)-encoding sequence of wild-type and Q304E variant forms of NorR. Activation of transcription *in vivo* by the Q304E variant was dependent on the ability to bind to the NorR enhancer sites. Each of the three C-terminal truncations made (Δ444-504, Δ442-504 and Δ436-504) rendered the NorR and Q304E proteins inactive both in the absence and presence of an NO-source (Figure 9.18A). A similar result was observed for the GAFTGA variant G266D (Figure 8.6). Western blot analysis revealed that the Δ442-504 and Δ436-504 constructs were stable in the complementation assay, although the Δ444-504 construct was shown to be unstable (Figure 9.18B). Whilst it is possible that the Q304E substitution promotes the formation of an ATPase active oligomer in the absence of DNA, the binding to the *norR-norVW* intergenic region may be required to anchor the NorR variant upstream of the promoter and orientate it relative to the holoenzyme. In order to assess the requirement of each of the three NorR binding sites but at the same time allow *in cis* activation at the *norV* promoter, the ability of the Q304E-variant to activate transcription *in vivo* was tested in strains of *E. coli* with altered NorR-binding sites. When the consensus sequence of binding site 1 (S1), site 2 (S2) or site 3 (S3) is altered from GT-(N7)-AC to GG-(N7)-CC, the NorRΔGAF protein binds to the wild-type sites but is unable to hydrolyse ATP and effectively activate open complex formation (Tucker et al., 2010). Consistent with the published result, the wild-type protein was significantly diminished in its ability to activate transcription *in vivo* in the absence of functional binding sites (Figure 9.19). Intriguingly, the Q304E variant was still active but only when endogenously-produced NO was present (Figure 9.19). This suggests that the Q304E variant does not require each of the three NorR binding sites to form the functional oligomer *in vivo* when activated by NO.

271
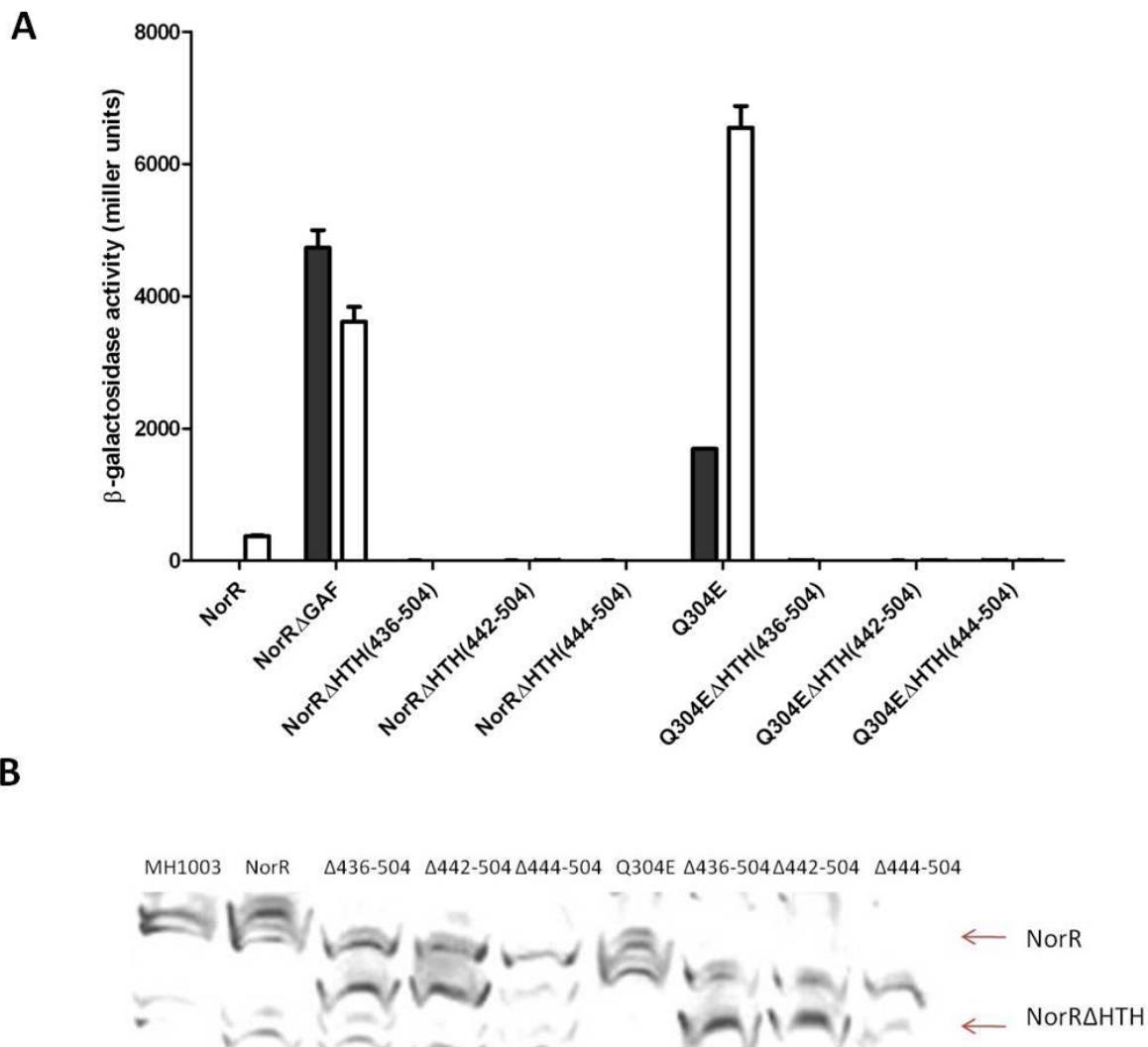
Figure 9.18 – (A) Activities of NorR and the Q304E variant when C-terminal truncations are made in the *norR* sequence. Substitutions are indicated on the x axis. "NorR" refers to the wild-type protein, "NorRΔGAF" refers to the N-truncated form lacking the GAF domain (residues 1-170) and "NorRΔHTH" refers to the C-truncated form lacking the helix-turn-helix (HTH) motif (residues 436-504, 442-504 or 444-504). 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. **(B) Western blot analysis indicating the stability of NorR variants *in vivo* when cultures are grown in the absence of potassium nitrite**. The locations of the bands corresponding to the full-length and ΔHTH constructs are indicated by red arrows. "MH1003" refers to the *E. coli* strain only. The uppermost band that is not detected in the MH1003 strain correlates to full-length NorR and its variants. Results showed that the Δ444-504 constructs were unstable *in vivo*.

Figure 9.19 – *In vivo* transcriptional activation by the NorR variant Q304E in the absence of NorR binding site 1, 2 or 3. NorR constucts 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. "NorR" refers to the wild-type protein and "NorRΔGAF" refers to the truncated form lacking the GAF domain (Δ1-170). 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.

9.9 The GAF-truncated version of the NorR variant Q304E can activate open complex formation *in vitro*

In order to further test the functionality of the Q304E variant *in vitro*, open promoter complex assays were conducted. This assay measures the conversion of the closed σ^{54}-RNA polymerase complex to an open promoter complex. Despite the fact that NorR-DNA complexes are heparin resistant, the open complex can be visualised on non-denaturing gels as it forms a distinct, supershifted species that is also heparin resistant (D'Autreaux *et al.* 2005). As has been observed previously, the formation of open complex by the NorRΔGAF-His protein was ATP-dependent (Figure 9.20A, compare lanes 2 and 3). The Q304EΔGAF-His variant was also able to form the ATP-dependent supershifted species (compare lanes 6 and 7) but the proportion of closed comples converted to open complex was significantly lower (compare lanes 3 and 7). This may reflect the lower ATPase activity of the Q304EΔGAF-His protein compared to the NorRΔGAF-His protein (Figure 9.15, compare B and C). In contrast, when the Q304E substitution was present in the full-length form of NorR, the variant protein was apparently unable to form open complexes (lanes 4 and 5). Interestingly, in the presence of ATP, the Q304E-DNA complex was less able to migrate into the non-denaturing gel, indicating the formation of a larger complex (lane 5). This has been observed previously for the GAFTGA variant G266D-His which forms a high-molecular weight complex irrespective of the presence of nucleotide (Figure 8.13C). Possibly the Q304E substitution promotes the ATP-dependent formation of a larger NorR complex but only in the presence of DNA.

In order to further probe the nature of the open complexes formed, and to confirm the results of the standard OPC assay, complexes were footprinted using potassium permanganate. This method cleaves single stranded regions of DNA and so can detect the presence of "melted" DNA at the promoter in a sequence-specific manner. For the NorRΔGAF-His and Q304EΔGAF-His proteins, enhanced cleavage was observed corresponding to T residues located between -11 to +1 at the *norV* promoter, consistent with the expected footprint (Figure 9.20B, lanes 3 and 5). The band intensity for the Q304EΔGAF-His footprint was decreased in comparison with NorRΔGAF-His, in agreement with the lower proportion of open complexes formed by the Q304EΔGAF variant in the standard open promoter complex assay (Figure 9.20A, lane 7). These results also confirmed that the full-length Q304E-His protein was unable to activate the formation
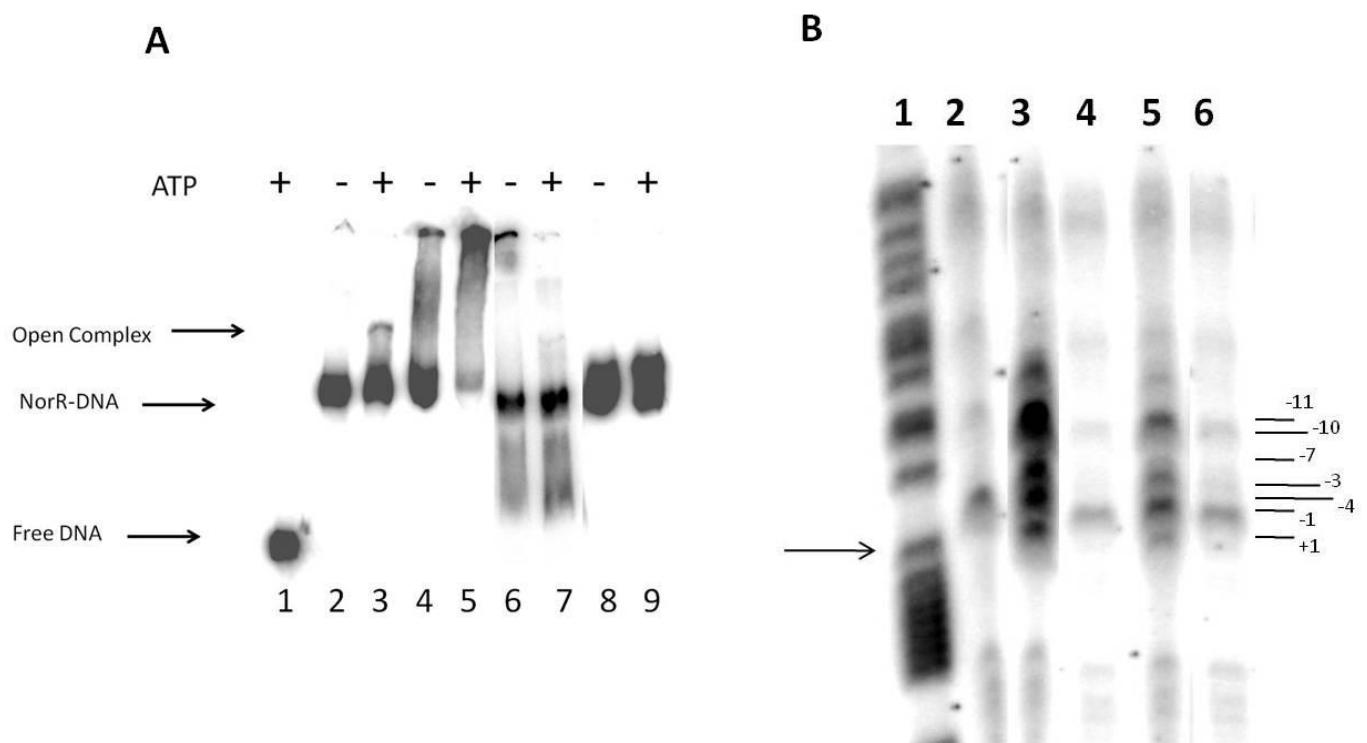
Figure 9.20 - Open promoter complex formation by AAA+ variants. (**A**) Heparin resistant complexes formed by NorRΔGAF-His, Q304E-His, Q304EΔGAF-His and C113S-Q304E-His on the 361bp DNA fragment carrying the *norR-norVW* intergenic region. In all cases the final NorR concentration was 1500 nM. Reactions contained no NorR (lane 1), NorRΔGAF-His (lanes 2 and 3), Q304E-His (lanes 4 and 5), Q304EΔGAF-His (lanes 6 and 7) and C113S-Q304E-His (lanes 8 and 9). Reactions loaded in lanes 1, 3, 5, 7 and 9 contained ATP (final concentration 5 mM), which was absent in lanes 2, 4, 6 and 8. Arrows indicate the position of free DNA, NorR bound DNA and the open promoter complexes. (**B**) Potassium permanganate footprinting of the 266bp *norR-norVW* promoter fragment after open complex formation initiated by NorR. Lane 1 is a G+A ladder. Lane 2 is a control without activator present. Lanes 3, 4, 5 and 6 show footprinting after initiation of open complexes in the presence of 1 μM (final concentration) ΔGAF-His, Q304E-His, Q304EΔGAF-His and C113S-Q304E-His respectively. The arrow marks the *norVW* transcriptional start and the positions of the enhanced cleavage at T bases are indicated.

of open complexes *in vitro*, at least under the conditions used in the assay since no enhancement of cleavage was observed compared to a reaction in which the activator was absent (Figure 9.21B compare lane 4 with lane 2).

9.10 The C113S-Q304E variant protein is ATPase active but unable to form open promoter complex *in vitro*

Open promoter complex assays and subsequent potassium permanganate footprinting have shown that the full-length Q304E-His protein is inactive *in vitro* with respect to open complex formation (Figure 9.20). *In vivo* the protein exhibits a partial-escape phenotype since the Q304E substitution does not allow NorR to fully escape GAF-mediated repression of AAA+ activity (Figure 9.1). One explanation is that the repression still exerted by the regulatory domain upon the central domain is responsible for the inability of the Q304E variant to activate the formation of open promoter complexes *in vitro*. β-galactosidase assays have shown that additional substitutions in the GAF domain increase the activity of the Q304E variant *in vivo* (Figure 9.2). Although the C113S substitution did not enable complete escape from repression, the higher activity in the non-induced state *in vivo* might give rise to an active protein *in vitro*. Therefore, the C113S-Q304E variant was overexpressed and purified by Nickel affinity chromatography, followed by gel filtration. The resulting pure-protein eluted at a similar volume to the Q304E-His protein on the Superdex 200 16/60 column (Amersham Biosciences), corresponding to a molecular weight in the dimer-trimer range (data not shown). ATPase assays revealed that the additional C113S substitution in the GAF domain of the Q304E variant did not affect the ability of the protein to hydrolyse ATP (Figure 9.21). Both the Q304E and C113S-Q304E variants exhibited ATPase activity in the absence of enhancer DNA, with maximal activity achieved upon addition of DNA containing the three NorR binding sites. Therefore, the C113S substitution was unable to stimulate ATP turnover any further, or achieve the high level of cooperativity that is observed for NorRΔGAF-His (Figure 9.15A). Furthermore, assays showed that the additional C113S substitution did not allow the Q304E-His protein to activate the formation of open complex (Figure 9.20A, lane 8 and 9). Potassium permanganate footprinting of the open complex reactions confirmed that the C113S-Q304E-His protein was inactive with respect to open promoter complex formation since the C113S-Q304E-His variant showed a similar level of enhancement of cleavage compared to a reaction that does not contain activator (Figure 9.20B, compare lanes 2 and
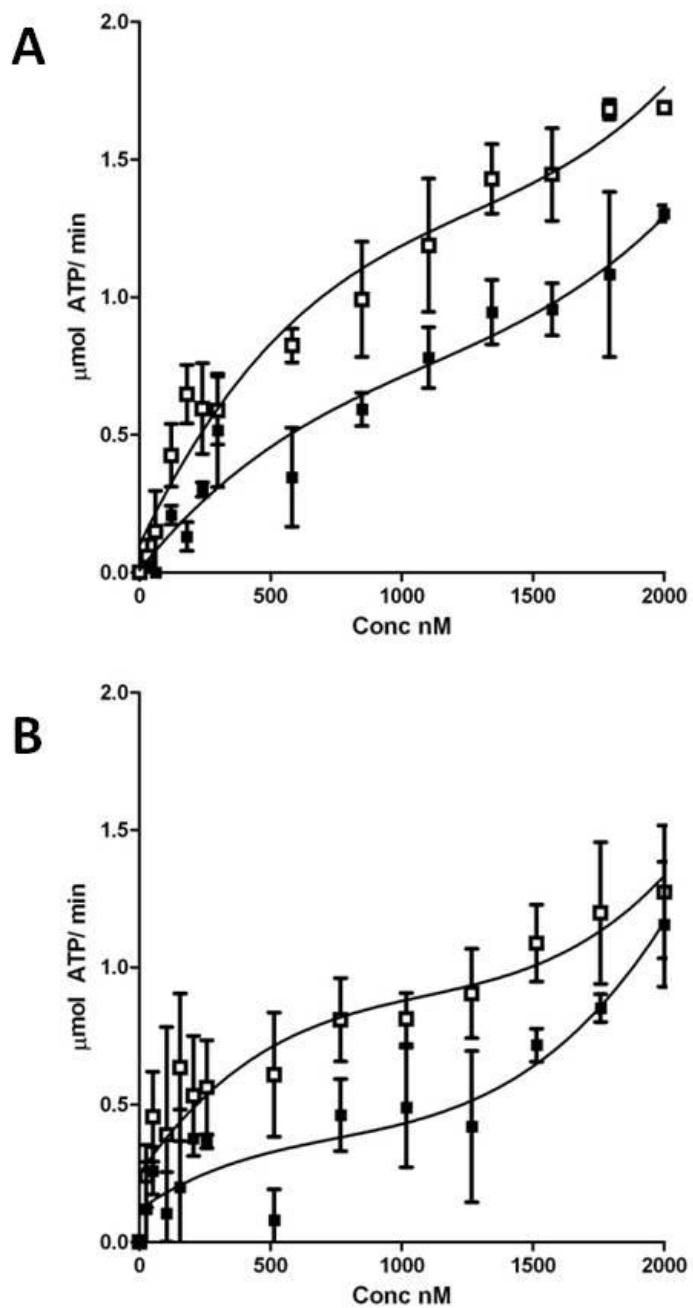
Figure 9.21 - ATPase activity of the Q304E-His (A) and C113S-Q304E-His (B) variants in response to protein concentration and the presence of enhancer DNA. Non-linear regression was carried out using GraphPad Prism software. Assays were conducted either in the absence (closed squares) or presence (open squares) of the 266bp DNA fragment (final concentration 5 nM) that includes the *norR-norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

6). It is interesting that the C113S substitution caused an increase in the level of transcriptional activation by the Q304E variant *in vivo* but that the C113S-Q304E mutant version of NorR remained inactive for open complex formation *in vitro*. Although the C113S-Q304E variant did not exhibit a complete escape phenotype, *in vivo* assays suggest that the C113S-Q304E variant is not subject to the same level of GAF-mediated repression as the Q304E protein. One explanation is that when the Q304E substitution is present, only complete escape from repression leads to open complex formation *in vitro*. In this work, this is not the first time that a NorR variant with an escape-phenotype *in vivo* has been shown to be unable to activate transcription *in vitro*. The full-length G266D-His protein was also apparently unable to activate the formation of open promoter complex (Figure 8.13 and 8.14), although this was presumably due to the formation of ATPase-inactive heptamers rather than the ATPase-active "hexamers" observed for Q304E-His.

9.11   Discussion

Random mutagenesis has identified a number of substitutions in the central domain of NorR that escape the GAF-mediated repression of AAA+ activity (Chapter 6). The majority of these residues are predicted to be located in the vicinity of the surface-exposed L1 and L2 loops or within the highly conserved GAFTGA motif that contacts σ^{54}. It is likely that such substitutions allow NorR to escape the repression mechanism by preventing the GAF domain from targeting the σ^{54}-interaction surface of the AAA+ domain (Chapter 7). Unlike the other substitutions that gave rise to strong-escape phenotypes, the Q304E substitution is at a residue not thought to have a significant role in modulating the conformation of the σ^{54}-interaction surface. In agreement with this, targeted mutagenesis at position 304 revealed that this residue is not essential in maintaining the mechanism of negative control in NorR (Figure 9.1). The requirement of a polar carboxyl group for escape is in line with the ability of glutamate and aspartate but not glutamine or asparagine changes to give rise to constitutive activity.

In the model of the NorR AAA+ fold based on the NtrC1 structure, Q304 is located in Helix 4, next to loop 2 in the AAA+ central domain (Figure 6.5B). Equivalent residues in other activators place this residue in a similar location. Interestingly, in NtrC1, DctD and PspF, the equivalent residue is glutamic acid rather than glutamine as in wild-type NorR. In the NtrC1 activated heptamer (PDB ID: 1NY6) (Lee *et al.* 2003), this glutamate is predicted to form a polar interaction with the sensor II arginine (R405 in NorR) of the adjacent protomer (Figure 9.22C). In contrast, the structure of the ZraR activated hexamer (PDB ID: 1OJL) (Sallai and Tucker 2005) reveals a glutamine residue as is present in the wild-type NorR protein. In this case, the interaction with the Sensor II arginine is not predicted to occur (Figure 9.22B). Therefore, it is possible that the Q304E mutation in NorR results in a polar contact between the 304 residue and the Sensor II arginine. In the NtrC1 inactive dimer the equivalent E256 and R357 residues are on opposite surfaces of adjacent subunits (Figure 9.22A). Upon activation and subsequent reorientation of the NtrC1 dimer from a front-to-front to a front-to-back configuration (Figure 3.11) (Doucleff *et al.* 2005a), interaction between these two residues may help drive formation of the oligomer. Therefore the Q304E substitution in NorR may facilitate the process of oligomerisation through the formation of the 304-405 polar interaction. Whilst the formation of the putative 304-405 polar contact in the Q304E variant of NorR may
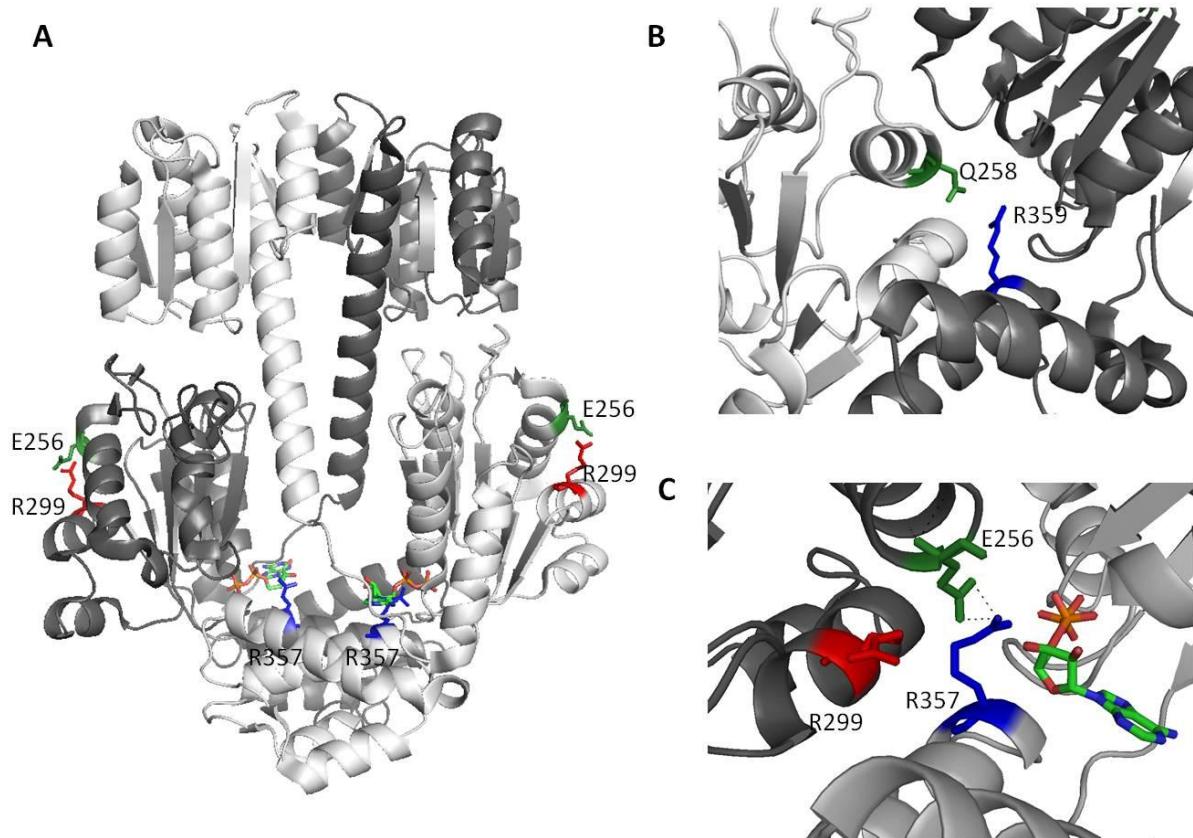
Figure 9.22 - Structures of related bEBPs suggest the Q304E substitution may influence the oligomerisation state of NorR. Adjacent protomers are indicated by light and dark shading. (**A**) Structure of NtrC1 (PDB ID: 1NY5 inactive dimer) (Lee *et al.* 2003) showing the location of the Sensor II R357 (blue), E256 (green) and proposed R-finger R299 (red) in the non-activated state. ADP is also shown bound. In the front-to-front configuration (Doucleff *et al.* 2005a), the R357-E256 polar interaction is prevented. Reorientation of the inactive dimer into a front-to-back configuration might allow formation of the polar contact that could facilitate the formation of the heptamer. (**B**) The protomer interface from the structure of the ZraR active hexamer (PDB ID: 1OJL) (Sallai and Tucker 2005) showing the Q258 residue and the R359 residue of the adjacent protomer. In ZraR, the glutamine at position 258 is not able to interact with the R359 residue. (**C**) The expected polar interaction that forms between E256 and R357 of the adjacent protomer in the ADP-bound activated heptamer of NtrC1 (PDB ID: 1NY6, B-C interface) (Lee *et al.* 2003).

stimulate oligomerisation, it is unclear how this would lead to the partial escape phenotype of the Q304E variant *in vivo.* In chapter 7, it was proposed that the regulatory domain targets the σ^{54}-interaction surface and not the oligomeric determinants in the mechanism of negative regulation. In addition, wild-type, non-activated NorR was only able to form higher oligomers in the presence of DNA (Figure 8.5), further suggesting that solely promoting constitutive hexamerisation would not lead to the escape from GAF-mediated repression of AAA+ activity. Possibly, formation of the new polar contact indirectly disrupts the interface of interdomain repression or the Q304E variant escapes negative control by some other mechanism not yet considered.

The hypothesis that the GAFTGA (G266D) and Q304E variants escape repression by different mechanisms is enhanced by the observation that substitutions in the GAF domain expected to disrupt NO-signalling at the iron-centre, led to an increased rather than a reduced escape-phenotype in NorRQ304E (Figure 9.2). In the wild-type protein, the R75K, D99A, H111Y, Y98L and C113S substitutions gave rise to a null-phenotype and the observation that these changes had no effect on the phenotype of the GAFTGA variant G266D confirmed that the NO-signal was not required for its activity (Figure 6.12). The reduction in the ability of the GAF domain to inhibit the activity of the Q304E variant when these substitutions are present suggests that the non-heme iron centre may have a role in maintaining repression of AAA+ activity. Furthermore an unexpected role for the GAF domain in the DNA-binding of NorR was revealed by EMSA assays using both NorR and the Q304E variant in full-length and ΔGAF forms. The Q304E substitution did not affect the affinity of binding to either the 266bp or 361bp fragments of the *norR-norVW* intergenic region (Figure 9.4), suggesting that the Q304E variant does not partially escape repression by altering the affinity of binding to enhancer DNA. However, there was a significant difference in the binding affinity between full-length and GAF-truncated (Δ1-170) forms of NorR. This reduction in affinity in the absence of the GAF domain was observed irrespective of the presence of the His-tag (Figure 9.5), confirming that the regulatory domain does contribute to the strength of enhancer-binding in NorR. Caution should be applied when interpreting these data since a greater reduction in the affinity of binding when the GAF domain was absent was observed for tagged compared to non-tagged proteins.

In vivo studies confirmed that the Q304E-version of NorR still required the ability to bind to the *norR-norVW* intergenic region in order to activate transcription, although this could additionally reflect the requirement of the hexamer to be brought into close proximity with the σ^{54}-RNA polymerase at the *norV* promoter (Figure 9.18). However, the Q304E protein was able to activate transcription in *E. coli* strains with altered NorR binding sites *in* vivo but only in the presence of NO (Figure 9.19). Other AAA+ variants were shown to activate transcription at the mutant promoters (data not shown) but none showed the absolute dependence of *in vivo* activity upon NO, as was observed for the Q304E variant. Possibly the Q304E substitution increases the propensity of the protein to oligomerise but only under conditions of nitrosative stress. However, to test this, the protein would need to be purified under anaerobic conditions and to-date heparin affinity chromatography has been unsuccessful in purifying non-tagged NorRQ304E (Figure 9.3A and B).

Biochemical studies of the Q304E variant in both full-length and ΔGAF forms have given conflicting data on whether the proteins display altered oligomeric properties. However, contradictions may be partly explained by the differences in the protein: DNA ratio employed in Cryo-EM, ATPase assays and open promoter complex experiments. Cryo-EM analysis showed that at a molar ratio of 12:1 monomer: DNA, the formation of hexamers was strictly dependent upon the presence of enhancer DNA, as was previously observed for NorR-His (Figure 8.5), NorRΔGAF-His (Tucker *et al.* 2010a) and G266DΔGAF-His (Figure 7.3). However, at a higher protein: DNA ratio (24:1 monomer: DNA), a small number of Q304E-His heptamers were visualised, similar in architecture to the enhancer-independent G266D-His heptamers identified in Chapter 8 (Figure 9.7). This suggests that the Q304E substitution stimulates self-association, especially at high protein concentrations, in line with structural data predicting the formation of a 304-405 polar, inter-protomer interaction. 3D-reconstruction of the G266D variant predicted that catalytic residues were not suitably positioned for ATP hydrolysis and *in vitro* ATPase assays confirmed that this was likely to be the case (Figures 8.11 and 8.12). In contrast the Q304E variant was able to form hexamers and hydrolyse ATP *in vitro*, in both full-length and GAF-truncated forms (Figure 9.15). The ability to turnover nucleotide is in line with the prediction based on the location of catalytic residues in the 3D-reconstruction of DNA-bound Q304E-His (Figure 9.14). Importantly, the Q304E substitution enabled NorR to exhibit a low level of activity in the absence of enhancer DNA (Figure 9.15). Since self-association of bEBP AAA+ domains is required for the formation of a functional oligomer,

the Q304E substitution must partially negate the requirement for DNA in oligomerisation, at least *in vitro*. In contrast, wild-type NorR lacking the N-terminal GAF domain exhibited DNA-dependent hydrolysis with activity increasing as a sigmoidal response to increasing protein concentration. A lower rate of increase was exhibited at concentrations above 250 nM (Figure 9.15A, open squares) which may reflect saturation of the enhancer sites consistent with the observed DNA binding constant (100 nM, Figure 7.2). At similar concentrations of Q304E-His, the enhancer DNA is presumably also saturated since the Q304E substitution was shown not to have a significant effect on the affinity of DNA-binding (Figure 9.4). Therefore, the increase in the ATPase activity of Q304E observed in the presence of DNA and at higher protein concentrations (Figure 9.15B, open squares) may be due to enhancer-independent self-association. In line with this, the same increase in turnover was observed when DNA was absent from the reaction (Figure 9.15B, closed squares).

The inability of the full-length Q304E variant to activate the formation of open complexes (Figure 9.20) is puzzling given the ability of the protein to turnover ATP at similarly high concentrations (Figure 9.15). In the presence of DNA and at protein concentrations closer to the kd of binding, the Q304E variant exhibits a turnover approximately four-fold less than the NorRΔGAF-His protein (compare Figures 9.15A and 9.15B, closed squares). This reduction in activity may explain why open complex formation was not observed *in vitro*. An alternative explanation is that although the Q304E substitution gives rise to a partial-bypass phenotype *in vivo*, it does not cause sufficient escape from GAF-mediated repression to enable the formation of a fully active protein *in vitro*. However, when the additional C113S substitution was introduced into the Q304E variant, the purified protein was equally unable to activate the formation of open complexes (Figure 9.20). Therefore the increased escape-phenotype of C113S-Q304E observed *in vivo* (Figure 9.2) appears insufficient to render the protein fully active *in vitro*. Another clue as to why the Q304E-His variant was inactive for open complex formation *in vitro* may be revealed in the formation of a high molecular weight species in the presence of ATP that is unable to migrate effectively into the polyacrylamide gel (Figure 9.20A, lane 5). This has been previously observed for the GAFTGA-variant G266D that migrates poorly irrespective of the presence of ATP. This large protein complex may be competent to hydrolyse ATP but not activate open complex formation. In addition, the possibility that full-length, His-

tagged proteins may not be able to catalyse open complex formation under the conditions employed in this assay cannot be eliminated.

Cryo-Electron Microscopy of Q304E has led to the first 3D-reconstruction of a bEBP bound to enhancer DNA. In this model, there are two stacked rings of EM-density with the upper ring corresponding to the AAA+ hexamer with GAF domains located at the outer-edge of the ring (Figure 9.13). In Chapter 7, biochemical analysis of the GAFTGA variant combined with genetic suppression studies suggested that the regulatory GAF-domain targets the σ^{54}-interaction surface in the mechanism of interdomain repression. Since the L1 and L2 loops are surface-exposed, the Q304E-His reconstruction likely represents an on-state of the NorR protein. As the GAF domains are located at the periphery of the AAA+ ring in both the Q304E and G266D EM-structures, this raises the question as to why the G266D variant has a full-escape phenotype *in vivo* but the Q304E substitution only enables partial escape. Significantly, the EM-density of the lower circular ring was attributed to the intergenic DNA used in the negative-stain analysis. A small ring-like density located below the hexamer and with a connection to the DNA ring was attributed to the DNA-binding domains of the protein. Therefore enhancer binding occurs on the opposite face of the bEBP compared to the σ^{54}-interaction surface. Since the EM-density attributed to the N-terminal GAF domain is connected to the EM-density attributed to the dsDNA DNA, it might be speculated that specific contacts form between the GAF domain and enhancer DNA. In particular, the positively charged R49 and R81 residues are predicted to be located close to the phosphate backbone (Ghosh 2010) and therefore may contribute to the stability of the nucleoprotein complex, at least in the on-state. The R81 residue was earlier shown to be critical for maintaining GAF-mediated repression of AAA+ activity and was suggested to be located at the interface of interdomain repression (Chapter 7). The significant movement of the GAF domain, proposed to occur upon release of repression (Figure 9.2) may allow this residue to form an interaction with DNA rather than the AAA+ domain in the on-state. Interestingly, the reconstruction predicts that ~ 130bp of the intergenic region is required to encircle the Q304E-His hexamer. This is in line with previous work that has shown the importance of DNA flanking the 3 NorR enhancer sites for NorR to activate transcription. A 66bp fragment is insufficient to stabilise the hexamerisation of NorRΔGAF reflected in the lower ATP turnover compared to a 266bp fragment (Tucker *et al.* 2010a). However, the remaining length of the 266bp DNA fragment is presumably unbound in the reconstruction and refinement of the model is

likely to have prevented its EM-density from being represented. DNA-binding studies in this chapter have shown that a 361bp fragment including the intergenic region further contributes to the strength of enhancer-binding (Figure 9.4) and it is possible that additional protein-DNA interactions are present to those identified in the reconstruction. Since three-fold symmetry was applied to generate the 3D-model, local-distortions in the enhancer DNA cannot be visualised. In this case, single-particle reconstruction, performed in the absence of symmetry constraints may help to increase the resolution of the bound DNA and the DNA-binding domains. Although the identification of the Q304E substitution has not significantly contributed to the study of interdomain repression in NorR, the subsequent structural studies have provided a greater understanding as to how NorR and other bEBPs bind to enhancer DNA to form functional oligomers capable of activating transcription.

Chapter 10 - General discussion

In contrast to σ^{70}-dependent transcription, the requirement of an activator at σ^{54}-dependent promoters imposes tight regulation of transcription which occurs primarily in response to cellular and extracellular signals that regulate the activity of the bEBP. A variety of regulatory domains exist to couple the activation of transcription to the presence of these signals, producing a highly specific response in the bacterial cell. Recent work has shed light on the mechanisms utilised by the variety of response regulator and sensory domains that regulate the activity of the enzymatic AAA+ domain responsible for the isomerisation of the closed promoter complex. Examples of both positive and negative control have been demonstrated in which the regulatory domain either stimulates or represses the activity of the central domain. Much effort has been expended to characterise the route by which these bEBPs couple signal sensing to substrate remodelling. In several bEBPs (e.g.NtrC1, DctD), the N-terminal regulatory domain regulates the activity of the AAA+ domain by controlling the oligomeric state of the activator (Figure 10.1A) (Lee *et al.* 2003; Doucleff *et al.* 2005a). The adaptation of activators to control their enzymatic activity by this mechanism is understandable, given the absolute requirement for self-association in order to form the functional ATPase (Zhang *et al.* 2002; Rappas *et al.* 2007). The well-studied PspF falls into a second class of bEBPs in which the ATP hydrolysis determinants are the target of regulation (Figure 10.1B). In the absence of an N-terminal regulatory domain, AAA+ activity is modulated *in trans* through direct interaction between the activator and the negative regulator PspA (Dworkin *et al.* 2000; Elderkin *et al.* 2002; Elderkin *et al.* 2005). The PspA-PspF regulatory complex is expected to have an altered arrangement in the key ATPase determinants that form the catalytic site at the inter-protomer interfaces of the PspF hexamer (Joly *et al.* 2009).

This work has identified for the first time, a third mechanism of regulation in bEBPs (Figure 10.1C). In NorR, a number of AAA+ variants that escape GAF-mediated repression have been identified (Chapter 6), located in a key region of the central domain that undergoes significant conformational change as ATP is hydrolysed. This invokes a model whereby the GAF domain negatively regulates the AAA+ domain by preventing access of the L1 and L2 loops to σ^{54}. Genetic and biochemical studies, using the GAFTGA-motif escape variant (G266D) suggest that the GAF domain does not regulate AAA+ activity through the control of oligomerisation or by directly targeting the ATP hydrolysis machinery (Chapter 7). Regulation at the level of σ^{54}-interaction is further

supported by suppression of the escape-phenotype via substitutions at the R81 residue. Mutagenesis has revealed that this arginine is essential in maintaining the repression mechanism and models predict it is surface-exposed and well-placed to make direct interactions with the AAA+ domain. The R81 residue is predicted to be located at the opposite end of an α-helix that also contains the R75 residue which is a proposed ligand to the hexa-coordinated non-heme iron and the most likely candidate to be displaced upon NO binding. Upon ligand displacement, significant conformational change is likely to occur in the α-helix, altering the position of R81 and potentially modifying GAF-AAA+ interactions. Whilst random screening of GAF-domain variants has been largely unsuccessful in identifying further residues at the GAF-AAA+ interface, during the completion of this work, targeted changes have identified two further amino acids, F68 and D84, with important roles in the repression mechanism (data not shown). The systematic substitution of the phenylalanine at position 68 suggests that F68, like R81 may have a direct role in maintaining GAF-AAA+ interactions in the absence of signal. Although the D84 residue does not appear essential for interdomain repression, its predicted position suggests that it may coordinate the position of other key residues involved in AAA+ contact. In the absence of high resolution structural data for NorR in the off-state, it will be difficult to further characterise the interface of repression or identify interacting partners.

Cryo-electron microscopy carried out by collaborators at Imperial College, London has led to a 3D-reconstruction of the GAFTGA variant. Studies show that the full-length form of G266D-His forms heptamers irrespective of the presence of enhancer DNA (Chapter 8). This form of the protein was shown to be inactive for ATP hydrolysis and open complex formation, in line with the hypothesis that bEBPs form hexamers in their physiologically relevant and functional form. It is likely that conditions inside the cell in which DNA acts as a scaffold for self-association of NorR, promotes the formation of a hexameric rather than a heptameric activator. Importantly, the EM model suggests that in the "on" state, the GAF domains are located at the periphery of the AAA+ ring, leaving the surface-exposed L1 and L2 loops free to facilitate the remodelling of σ^{54}. Consequently, activation of the NorR protein as a transcription factor is predicted to require a significant "swing-out" movement of the GAF domain(s) from the top to the outer edge of the oligomeric ring, facilitated by the long, flexible linker between the regulatory and central domains.

Structural studies in PspF suggest that there is significant *in cis* communication between the σ^{54} interaction determinants on the surface of the bEBP-ring and the ATP hydrolysis machinery (Rappas *et al.* 2006) as well as *in trans* interactions between subunits of the hexamer (Joly and Buck 2010). Therefore, the mechanisms of regulation that target ATP hydrolysis, oligomerisation and σ^{54}-interaction are likely to be highly interconnected. In the case of NorR, the 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. Considering the inter-relatedness of the different control mechanisms, the evolutionary advantage of each remains unclear. However, regulation at the level of ATP hydrolysis or σ^{54}-contact would potentially allow for assembly of the higher order oligomer prior to activation and this may confer a physiological advantage by enabling a rapid stress response. In PspF, the inhibition of a preassembled PspF hexamer by a PspA complex may allow the cell to rapidly respond to the dissipation of the proton motive force (PMF) at the cell membrane (Joly *et al.* 2009). In the case of NorR, the pre-assembly of a NorR hexamer, "poised" as a nucleoprotein complex at the enhancer sites, may enable the cell to rapidly respond to nitrosative stress.

Previous work has highlighted the importance of enhancer DNA for the activation of NorR as a transcription factor (Tucker *et al.* 2010a). In the absence of the regulatory domain, the requirement for the NO signal is bypassed but DNA containing the three NorR binding sites is still required for NorR to hydrolyse ATP and activate transcription. Furthermore each of the three sites has been shown to be essential for NorR activity both in *vivo* and *in vitro* (Tucker *et al.* 2010a). While multiple enhancer sites for bEBPs are not uncommon, an absolute dependency on more than one binding site is unusual. In NorR this may reflect the requirement of DNA to act as a scaffold to facilitate oligomerisation prior to receipt of the NO-signal. In contrast, NtrC dimers bind to two target sites and recruit a third dimer from solution upon activation by phosphorylation (De Carlo *et al.* 2006). Possibly, since release of the GAF-mediated repression mechanism does not stimulate self-association in NorR, DNA-binding has instead evolved to drive the process of oligomerisation. Interestingly, the number of enhancer sites is not strictly conserved between different *norR*-containing proteobacteria. For example, in the aerobic, soil-dwelling organism *Azotobacter vinelandii* only two predicted NorR binding sites exist upstream of the *hmp* gene. It would be beneficial to extend this work in such bacteria to provide insight into the conservation of NorR-mediated regulation of gene transcription.

Cryo-EM microscopy, performed in collaboration has provided further insights into the requirement of enhancer DNA for the activation of NorR (Chapter 9). In contrast to the G266D variant protein which forms heptamers in the absence of DNA, the Q304E variant forms DNA-dependent hexamers, in agreement with the expected oligomeric state of wild-type NorR. Although at a low-resolution, the resulting reconstruction represents the first structure of a bEBP bound to DNA. The 3D-model reveals an on-state protein with the regulatory domains occupying the EM-density at the periphery of the AAA+ ring, leaving the GAFTGA-motifs exposed at the top of the ring. Significantly, the structure predicts that ~130bp of the *norR-norVW* intergenic region encircles the Q304E hexamer. Since the 3 NorR binding sites cover a length of only 66bp, the structure suggests that DNA flanking the three enhancers is required to stabilise the oligomeric ring. This is in agreement with previous work that demonstrated the ability of a 266bp fragment but not the minimal 66bp fragment of the intergenic region to stimulate self-association (Tucker *et al.* 2010a).

In summary, the regulation of NorR as a transcriptional activator of σ^{54}-dependent transcription is proposed to occur by a unique mechanism (Figure 10.2). Three NorR dimers are thought to bind with high affinity to the enhancer sites of the *norR-norVW* intergenic region. The binding of DNA induces conformational changes in NorR that result in the self-association of the AAA+ domains to form a hexamer. This hexamer is stabilised by interactions with ~130bp DNA that includes the NorR binding sites. In the absence of the NO-signal, the regulatory (GAF) domains repress the activity of the central (AAA+) domains via interactions with the σ^{54}-interaction surface. In the presence of Integration Host Factor (IHF), DNA-bending brings the inactive NorR-oligomer into close proximity with σ^{54}-RNAP bound at the *norVW* promoter. Under conditions of nitrosative stress, NO binds to the non-heme iron centre, causing displacement of the R75 ligand. Conformational changes in the regulatory domain result in the disruption of GAF-AAA+ interactions, leading to a release of the GAF-mediated mechanism of interdomain repression. NorR is then free to hydrolyse ATP, leading to conformational changes in the AAA+ domain(s) that enable the hexamer to make contact with the σ^{54}-RNAP holoenzyme via the GAFTGA motifs. Remodelling of the closed complex and the subsequent expression of the *norVW* genes enables NO-detoxification by the flavorubredoxin (NorV) and its associated redox partner (NorW).
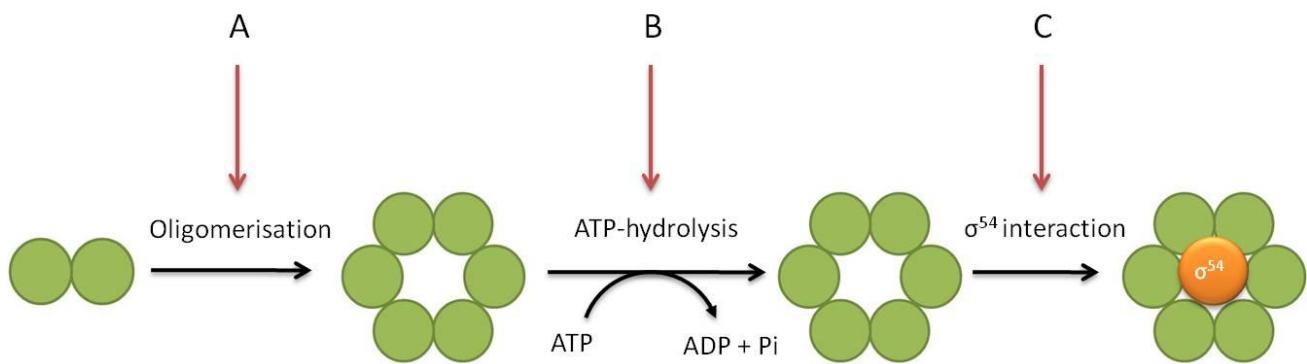
Figure 10.1 – Possible targets of regulatory domain-mediated regulation. Most commonly, the regulatory domain represses oligomerisation (**A**) e.g. NtrC1, DctD (Lee *et al.* 2003; Doucleff *et al.* 2005a) or promotes self-association in response to signal e.g. NtrC (De Carlo *et al.* 2006). (**B**) In PspF, negative regulation directly targets the nucleotide hydrolysis machinery (Joly *et al.* 2008b). Whilst the binding of ATP, releases the L1 and L2 loops to establish a weak interaction with σ^{54}, hydrolysis is required to produce a strong interaction that results in remodelling of the holoenzyme (Rappas *et al.* 2006). NorR may represent a newly identified mechanism of control in which the interaction with σ^{54} is the target of the regulatory domain (**C**). Where oligomerisation is not the target, pre-assembly of a hexamer prior to activation may have a physiological advantage e.g. rapid response to stress. The mechanisms of regulation that target ATP hydrolysis, oligomerisation and σ^{54}-interaction are likely to be highly interconnected with the enzymatic activity of the AAA+ domain the ultimate target of regulation.
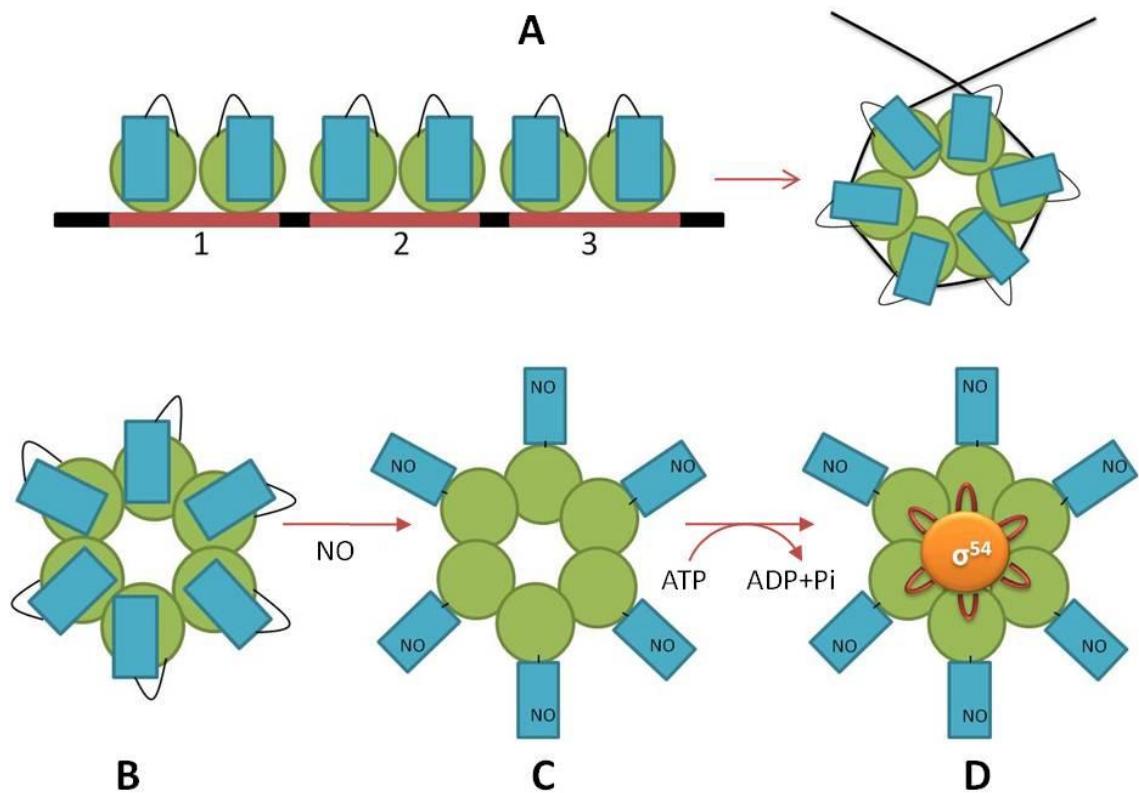
Figure 10.2 - Model of NorR-dependent activation of *norVW* (A) Binding of NorR to the *norR-norVW* intergenic region that contains the three NorR binding sites (1, 2 and 3, highlighted in red) is thought to facilitate the formation of a higher order oligomer that is most likely to be a hexamer (Tucker *et al.* 2010a). (**B**) Although bound to DNA, in the absence of NO, the N-terminal GAF domains (blue rectangles) negatively regulate the activity of the AAA+ domains (green circles) by preventing access of the surface-exposed loops to σ^{54}. (**C**) In the 'on'state, NO binds to the iron centre in the GAF domain forming a mononitrosyl iron species. The repression of the AAA+ domain is relieved, enabling ATP hydrolysis by NorR coupled to conformational changes in the AAA+ domain. (**D**) In the presence of ATP, the surface-exposed loops (red) that include the GAFTGA motifs move into an extended conformation to establish an initial interaction with σ^{54} that is strengthend upon hydrolysis, resulting in the remodelling of the closed complex. Upon phosphate release, the L1 and L2 loops compact downwards, enabling relocation of the sigma factor (Rappas *et al.* 2006). For simplicity, DNA is not illustrated in B, C or D.

11   References

Adams, H, Teertstra, W, Demmers, J, Boesten, R and Tommassen, J (2003). Interactions between phage-shock proteins in Escherichia coli. *J Bacteriol*. **185**, 1174-1180.

Ades, SE (2006). AAA+ molecular machines: firing on all cylinders. *Curr Biol*. **16**, R46-48.

Akoev, V, Gogol, EP, Barnett, ME and Zolkiewski, M (2004). Nucleotide-induced switch in oligomerization of the AAA+ ATPase ClpB. *Protein Sci*. **13**, 567-574.

Alfano, JR and Collmer, A (1997). The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *J Bacteriol*. **179**, 5655-5662.

Almeida, CC, Romao, CV, Lindley, PF, Teixeira, M and Saraiva, LM (2006). The role of the hybrid cluster protein in oxidative stress defense. *J Biol Chem*. **281**, 32445-32450.

Almiron, M, Link, AJ, Furlong, D and Kolter, R (1992). A novel DNA-binding protein with regulatory and protective roles in starved Escherichia coli. *Genes Dev*. **6**, 2646-2654.

Altuvia, S, Almiron, M, Huisman, G, Kolter, R and Storz, G (1994). The dps promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol Microbiol*. **13**, 265-272.

Andrews, SC, Robinson, AK and Rodriguez-Quinones, F (2003). Bacterial iron homeostasis. *FEMS Microbiol Rev*. **27**, 215-237.

Aravind, L and Ponting, CP (1997). The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends Biochem Sci*. **22**, 458-459.

Arciero, DM, Lipscomb, JD, Huynh, BH, Kent, TA and Munck, E (1983). EPR and Mossbauer studies of protocatechuate 4,5-dioxygenase. Characterization of a new Fe^{2+} environment. *J Biol Chem*. **258**, 14981-14991.

Arfin, SM, Long, AD, Ito, ET, Tolleri, L, Riehle, MM, Paegle, ES and Hatfield, GW (2000). Global gene expression profiling in *Escherichia coli* K12. The effects of integration host factor. *J. Biol. Chem.* **275**, 29672-29684.

Arnold, WP, Mittal, CK, Katsuki, S and Murad, F (1977). Nitric oxide activates guanylate cyclase and increases guanosine 3'-5'-cyclic monophosphate levels in various tissue preparations. *PNAS*. **74**, 3203-3207.

Austin, S and Dixon, R (1992). The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *EMBO J*. **11**, 2219-2228.

Bamford, VA, Angove, HC, Seward, HE, Thomson, AJ, Cole, JA, Butt, JN, Hemmings, AM and Richardson, DJ (2002). Structure and spectroscopy of the periplasmic cytochrome c nitrite reductase from *Escherichia coli. Biochemistry*. **41**, 2921-2931.

Barne, KA, Bown, JA, Busby, SJ and Minchin, SD (1997). Region 2.5 of the Escherichia coli RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J*. **16**, 4034-4040.

Barrios, H, Valderrama, B and Morett, E (1999). Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucl. Acids Res.* **27**, 4305-4313.

Batchelor, JD, Doucleff, M, Lee, CJ, Matsubara, K, De Carlo, S, Heideker, J, Lamers, MH, Pelton, JG and Wemmer, DE (2008). Structure and regulatory mechanism of Aquifex aeolicus NtrC4: variability and evolution in bacterial transcriptional regulation. *J Mol Biol*. **384**, 1058-1075.

Batchelor, JD, Sterling, HJ, Hong, E, Williams, ER and Wemmer, DE (2009). Receiver domains control the active-state stoichiometry of Aquifex aeolicus sigma54 activator NtrC4, as revealed by electrospray ionization mass spectrometry. *J Mol Biol*. **393**, 634-643.

Beaumont, HJ, Lens, SI, Reijnders, WN, Westerhoff, HV and van Spanning, RJ (2004). Expression of nitrite reductase in Nitrosomonas europaea involves NsrR, a novel nitrite-sensitive transcription repressor. *Mol Microbiol*. **54**, 148-158.

Beck, LL, Smith, TG and Hoover, TR (2007). Look, no hands! Unconventional transcriptional activators in bacteria. *Trends Microbiol*. **15**, 530-537.

Beinert, H (2000). Iron-sulfur proteins: ancient structures, still full of surprises. *J Biol Inorg Chem*. **5**, 2-15.

Berger, DK, Narberhaus, F and Kustu, S (1994). The isolated catalytic domain of NIFA, a bacterial enhancer-binding protein, activates transcription in vitro: activation is inhibited by NIFL. *Proc Natl Acad Sci U S A*. **91**, 103-107.

Berger, DK, Narberhaus, F, Lee, HS and Kustu, S (1995). In vitro studies of the domains of the nitrogen fixation regulatory protein NIFA. *J Bacteriol*. **177**, 191-199.

Berkovitch, F, Nicolet, Y, Wan, JT, Jarrett, JT and Drennan, CL (2004). Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. *Science*. **303**, 76-79.

Berks, BC, Ferguson, SJ, Moir, JW and Richardson, DJ (1995). Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions. *Biochim Biophys Acta*. **1232**, 97-173.

Bertoni, G, Perez-Martin, J and de Lorenzo, V (1997). Genetic evidence of separate repressor and activator activities of the XylR regulator of the TOL plasmid, pWW0, of *Pseudomonas putida*. *Mol Microbiol*. **23**, 1221-1227.

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

Blatter, EE, Ross, W, Tang, H, Gourse, RL and Ebright, RH (1994). Domain organization of RNA polymerase [alpha] subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell*. **78**, 889-896.

Bochman, ML and Schwacha, A (2008). The Mcm2-7 complex has in vitro helicase activity. *Mol Cell*. **31**, 287-293.

Bochtler, M, Hartmann, C, Song, HK, Bourenkov, GP, Bartunik, HD and Huber, R (2000). The structures of HsIU and the ATP-dependent protease HsIU-HsIV. *Nature*. **403**, 800-805.

Bodenmiller, DM and Spiro, S (2006). The yjeB (nsrR) gene of Escherichia coli encodes a nitric oxide-sensitive transcriptional regulator. *J Bacteriol*. **188**, 874-881.

Bordes, P, Wigneshweraraj, SR, Chaney, M, Dago, AE, Morett, E and Buck, M (2004). Communication between Esigma(54) , promoter DNA and the conserved threonine residue in the GAFTGA motif of the PspF sigma-dependent activator during transcription activation. *Mol Microbiol*. **54**, 489-506.

Bordes, P, Wigneshweraraj, SR, Schumacher, J, Zhang, X, Chaney, M and Buck, M (2003). The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: Identifying a surface that binds sigma 54. *PNAS*. **100**, 2278-2283.

Bosca, L, Zeini, M, Traves, PG and Hortelano, S (2005). Nitric oxide and cell viability in inflammatory cells: a role for NO in macrophage function and fate. *Toxicology*. **208**, 249-258.

Bose, D, Joly, N, Pape, T, Rappas, M, Schumacher, J, Buck, M and Zhang, X (2008a). Dissecting the ATP hydrolysis pathway of bacterial enhancer-binding proteins. *Biochem. Soc. Trans.* **36**, 83-88.

Bose, D, Pape, T, Burrows, PC, Rappas, M, Wigneshweraraj, SR, Buck, M and Zhang, X (2008b). Organization of an activator-bound RNA polymerase holoenzyme. *Mol Cell*. **32**, 337-346.

Bowman, WC and Kranz, RG (1998). A bacterial ATP-dependent, enhancer binding protein that activates the housekeeping RNA polymerase. *Genes Dev*. **12**, 1884-1893.

Bradford, MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. **72**, 248-254.

Brahmachary, P, Dashti, MG, Olson, JW and Hoover, TR (2004). Helicobacter pylori FlgR is an enhancer-independent activator of sigma54-RNA polymerase holoenzyme. *J Bacteriol*. **186**, 4535-4542.

Bredt, DS, Hwang, PM, Glatt, CE, Lowenstein, C, Reed, RR and Snyder, SH (1991). Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase. **351**, 714-718.

Briggs, LC, Baldwin, GS, Miyata, N, Kondo, H, Zhang, X and Freemont, PS (2008). Analysis of nucleotide binding to P97 reveals the properties of a tandem AAA hexameric ATPase. *J Biol Chem*. **283**, 13745-13752.

Brondijk, TH, Fiegen, D, Richardson, DJ and Cole, JA (2002). Roles of NapF, NapG and NapH, subunits of the Escherichia coli periplasmic nitrate reductase, in ubiquinol oxidation. *Mol Microbiol*. **44**, 245-255.

Brown, CA, Pavlosky, MA, Westre, TE, Zhang, Y, Hedman, B, Hodgson, KO and Solomon, EI (1995). Spectroscopic and Theoretical Description of the Electronic-Structure of S=3/2 Iron-Nitrosyl Complexes and Their Relation to O-2 Activation by Nonheme Tron Enzyme Active-Sites. *Journal of the American Chemical Society*. **117**, 715-732.

Browning, DF, Beatty, CM, Sanstad, EA, Gunn, KE, Busby, SJ and Wolfe, AJ (2004). Modulation of CRP-dependent transcription at the Escherichia coli acsP2 promoter by nucleoprotein complexes: anti-activation by the nucleoid proteins FIS and IHF. *Mol Microbiol*. **51**, 241-254.

Browning, DF, Beatty, CM, Wolfe, AJ, Cole, JA and Busby, SJ (2002). Independent regulation of the divergent *Escherichia coli* nrfA and acsP1 promoters by a nucleoprotein assembly at a shared regulatory region. *Mol Microbiol*. **43**, 687-701.

Browning, DF and Busby, SJ (2004). The regulation of bacterial transcription initiation. *Nat Rev Microbiol*. **2**, 57-65.

Browning, DF, Cole, JA and Busby, SJ (2000). Suppression of FNR-dependent transcription activation at the Escherichia coli nir promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol*. **37**, 1258-1269.

Browning, DF, Grainger, DC, Beatty, CM, Wolfe, AJ, Cole, JA and Busby, SJ (2005). Integration of three signals at the Escherichia coli nrf promoter: a role for Fis protein in catabolite repression. *Mol Microbiol*. **57**, 496-510.

Browning, DF, Lee, DJ, Spiro, S and Busby, SJ (2010). Down-regulation of the Escherichia coli K-12 nrf promoter by binding of the NsrR nitric oxide-sensing transcription repressor to an upstream site. *J Bacteriol*. **192**, 3824-3828.

Buck, M, Bose, D, Burrows, P, Cannon, W, Joly, N, Pape, T, Rappas, M, Schumacher, J, Wigneshweraraj, S and Zhang, X (2006). A second paradigm for gene activation in bacteria. *Biochem Soc Trans*. **34**, 1067-1071.

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

Buck, M and Hoover, TR (2010). An ATPase R-finger leaves its print on transcriptional activation. *Structure*. **18**, 1391-1392.

Buck, M, Miller, S, Drummond, M and Dixon, R (1986). Upstream activator sequences are present in the promoters of nitrogen fixation genes *Nature*. **320**, 374-378.

Burgess, RR, Travers, AA, Dunn, JJ and Bautz, EK (1969). Factor stimulating transcription by RNA polymerase. *Nature*. **221**, 43-46.

Burrows, PC, Joly, N, Nixon, BT and Buck, M (2009). Comparative analysis of activator-Esigma54 complexes formed with nucleotide-metal fluoride analogues. *Nucleic Acids Res*. **37**, 5138-5150.

Burrows, PC, Severinov, K, Buck, M and Wigneshweraraj, SR (2004). Reorganisation of an RNA polymerase-promoter DNA complex for DNA melting. *Embo J*. **23**, 4253-4263.

Burrows, PC, Severinov, K, Ishihama, A, Buck, M and Wigneshweraraj, SR (2003). Mapping sigma 54-RNA polymerase interactions at the -24 consensus promoter element. *J Biol Chem*. **278**, 29728-29743.

Busby, S and Ebright, RH (1994). Promoter structure, promoter recognition, and transcription activation in prokaryotes. *Cell*. **79**, 743-746.

Busch, A, Pohlmann, A, Friedrich, B and Cramm, R (2004). A DNA region recognized by the nitric oxide-responsive transcriptional activator NorR is conserved in {beta}- and {gamma}-proteobacteria. *J. Bacteriol*. **186**, 7980-7987.

Cabello, P, Pino, C, Olmo-Mira, MF, Castillo, F, Roldan, MD and Moreno-Vivian, C (2004). Hydroxylamine assimilation by Rhodobacter capsulatus E1F1. requirement of the hcp gene (hybrid cluster protein) located in the nitrate assimilation nas gene region for hydroxylamine reduction. *J Biol Chem*. **279**, 45485-45494.

Campbell, EA, Muzzin, O, Chlenov, M, Sun, JL, Olson, CA, Weinman, O, Trester-Zedlitz, ML and Darst, SA (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell*. **9**, 527-539.

Cannon, W, Bordes, P, Wigneshweraraj, SR and Buck, M (2003). Nucleotide-dependent triggering of RNA polymerase-DNA interactions by an AAA regulator of transcription. *J Biol Chem*. **278**, 19815-19825.

Cannon, W, Chaney, M and Buck, M (1999a). Characterisation of holoenzyme lacking sigmaN regions I and II. *Nucleic Acids Res*. **27**, 2478-2486.

Cannon, W, Claverie-Martin, F, Austin, S and Buck, M (1994). Identification of a DNA-contacting surface in the transcription factor sigma-54. *Mol Microbiol*. **11**, 227-236.

Cannon, W, Gallegos, MT and Buck, M (2001). DNA melting within a binary sigma(54)-promoter DNA complex. *J Biol Chem*. **276**, 386-394.

Cannon, W, Gallegos, MT, Casaz, P and Buck, M (1999b). Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization. *Genes Dev*. **13**, 357-370.

Cannon, WV, Chaney, MK, Wang, X and Buck, M (1997). Two domains within sigmaN (sigma54) cooperate for DNA binding. *Proc Natl Acad Sci U S A*. **94**, 5006-5011.

Cannon, WV, Gallegos, M., Buck, M. (2000). Isomerization of a binary sigma–promoter DNA complex by transcription activators. *Nature Structural Biology*. **7 pp**, 594 - 601.

Cannon, WV, Schumacher, J and Buck, M (2004). Nucleotide-dependent interactions between a fork junction-RNA polymerase complex and an AAA+ transcriptional activator protein. *Nucleic Acids Res*. **32**, 4596-4608.

Casadaban, MJ and Cohen, SN (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *Journal of Molecular Biology*. **138**, 179-207.

Chaney, M and Buck, M (1999). The sigma 54 DNA-binding domain includes a determinant of enhancer responsiveness. *Mol Microbiol*. **33**, 1200-1209.

Chaney, M, Grande, R, Wigneshweraraj, SR, Cannon, W, Casaz, P, Gallegos, MT, Schumacher, J, Jones, S, Elderkin, S, Dago, AE, Morett, E and Buck, M (2001). 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*. **15**, 2282-2294.

Chen, B, Doucleff, M, Wemmer, DE, De Carlo, S, Huang, HH, Nogales, E, Hoover, TR, Kondrashkina, E, Guo, L and Nixon, BT (2007). ATP ground- and transition states of bacterial enhancer binding AAA+ ATPases support complex formation with their target protein, sigma54. *Structure*. **15**, 429-440.

Chen, B, Sysoeva, TA, Chowdhury, S, Guo, L, De Carlo, S, Hanson, JA, Yang, H and Nixon, BT (2010). Engagement of arginine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase. *Structure*. **18**, 1420-1430.

Chen, B, Sysoeva, TA, Chowdhury, S and Nixon, BT (2008). Regulation and action of the bacterial enhancer-binding protein AAA+ domains. *Biochem Soc Trans*. **36**, 89-93.

Chen, L, Liu, MY, Legall, J, Fareleira, P, Santos, H and Xavier, AV (1993). Purification and characterization of an NADH-rubredoxin oxidoreductase involved in the utilization of oxygen by Desulfovibrio gigas. *Eur J Biochem*. **216**, 443-448.

Clarke, TA, Kemp, GL, Van Wonderen, JH, Doyle, RM, Cole, JA, Tovell, N, Cheesman, MR, Butt, JN, Richardson, DJ and Hemmings, AM (2008a). Role of a conserved glutamine residue in tuning the catalytic activity of Escherichia coli cytochrome c nitrite reductase. *Biochemistry*. **47**, 3789-3799.

Clarke, TA, Mills, PC, Poock, SR, Butt, JN, Cheesman, MR, Cole, JA, Hinton, JC, Hemmings, AM, Kemp, G, Soderberg, CA, Spiro, S, Van Wonderen, J and Richardson, DJ (2008b). Escherichia coli cytochrome c nitrite reductase NrfA. *Methods Enzymol*. **437**, 63-77.

Claverie-Martin, F and Magasanik, B (1992). Positive and negative effects of DNA bending on activation of transcription from a distant site. *J Mol Biol*. **227**, 996-1008.

Clay, MD, Cosper, CA, Jenney, FE, Jr., Adams, MW and Johnson, MK (2003). Nitric oxide binding at the mononuclear active site of reduced Pyrococcus furiosus superoxide reductase. *Proc Natl Acad Sci U S A*. **100**, 3796-3801.

Collmer, A, Badel, JL, Charkowski, AO, Deng, WL, Fouts, DE, Ramos, AR, Rehm, AH, Anderson, DM, Schneewind, O, van Dijk, K and Alfano, JR (2000). Pseudomonas syringae Hrp type III secretion system and effector proteins. *Proc Natl Acad Sci U S A*. **97**, 8770-8777.

Contreras, A and Drummond, M (1988). The effect on the function of the transcriptional activator NtrC from *Klebsiella pneumoniae* of mutations in the DNA-recognition helix. *Nucl. Acids Res.* **16**, 4025-4039.

Corbin, JD and Francis, SH (1999). Cyclic GMP phosphodiesterase-5: target of sildenafil. *J. Biol. Chem.* **274**, 13729-13732.

Corker, H and Poole, RK (2003). Nitric oxide formation by *Escherichia coli*: dependence on nitrite reductase, the NO sensing regulator Fnr, and Flavohemoglobin Hmp. *J. Biol. Chem.* **278**, 31584-31592.

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

Costa, A, Pape, T, van Heel, M, Brick, P, Patwardhan, A and Onesti, S (2006b). Structural studies of the archaeal MCM complex in different functional states. *J Struct Biol*. **156**, 210-219.

Costa, C, Macedo, A, Moura, I, Moura, JJG, Le Gall, J, Berlier, Y, Liu, M-Y and Payne, WJ (1990). Regulation of the hexaheme nitrite/nitric oxide reductase of *Desulfovibrio desulfuricans*, *Wolinella succinogenes* and *Escherichia coli* : A mass spectrometric study. *FEBS Letters*. **276**, 67-70.

Crack, JC, Le Brun, NE, Thomson, AJ, Green, J and Jervis, AJ (2008). Reactions of nitric oxide and oxygen with the regulator of fumarate and nitrate reduction, a global transcriptional regulator, during anaerobic growth of Escherichia coli. *Methods Enzymol*. **437**, 191-209.

Cramer, P, Bushnell, DA, Fu, J, Gnatt, AL, Maier-Davis, B, Thompson, NE, Burgess, RR, Edwards, AM, David, PR and Kornberg, RD (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science*. **288**, 640-649.

Cramer, P, Bushnell, DA and Kornberg, RD (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*. **292**, 1863-1876.

Cramm, R, Pohlmann, A and Friedrich, B (1999). Purification and characterization of the single-component nitric oxide reductase from *Ralstonia eutropha* H16. *FEBS Lett*. **460**, 6-10.

Cramm, R, Siddiqui, RA and Friedrich, B (1997). Two isofunctional nitric oxide reductases in Alcaligenes eutrophus H16. *J Bacteriol*. **179**, 6769-6777.

Crane, BR, Arvai, AS, Ghosh, DK, Wu, C, Getzoff, ED, Stuehr, DJ and Tainer, JA (1998). Structure of nitric oxide synthase oxygenase dimer with pterin and substrate. *Science*. **279**, 2121-2126.

Cruz-Ramos, H, Crack, J, Wu, G, Hughes, MN, Scott, C, Thomson, AJ, Green, J and Poole, RK (2002). NO sensing by FNR: regulation of the Escherichia coli NO-detoxifying flavohaemoglobin, Hmp. *Embo J*. **21**, 3235-3244.

Cunha, CA, Macieira, S, Dias, JM, Almeida, G, Goncalves, LL, Costa, C, Lampreia, J, Huber, R, Moura, JJ, Moura, I and Romao, MJ (2003). Cytochrome c nitrite reductase from Desulfovibrio desulfuricans ATCC 27774. The relevance of the two calcium sites in the structure of the catalytic subunit (NrfA). *J Biol Chem*. **278**, 17455-17465.

D'Autreaux, B, Touati, D, Bersch, B, Latour, J and Michaud-Soret, I (2002). Direct inhibition by nitric oxide of the transcriptional ferric uptake regulation protein via nitrosylation of the iron. *PNAS*. **99**, 16619-16624.

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

Dago, AE, Wigneshweraraj, SR, Buck, M and Morett, E (2007). A role for the conserved GAFTGA motif of AAA+ transcription activators in sensing promoter DNA conformation. *J Biol Chem*. **282**, 1087-1097.

Darwin, AJ (2005). The phage-shock-protein response. *Mol Microbiol*. **57**, 621-628.

Darwin, AJ and Stewart, V (1996). The Nar modulon systems: nitrate and nitrite regulation of anaerobic gene expression. *Regulation of gene expression in Escherichia coli*. Lin, ECC and Lynch, AS. Austin, TX, R.G. Landes**:** 343-360.

Davies, JM, Brunger, AT and Weis, WI (2008). Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change. *Structure*. **16**, 715-726.

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

Deckert, G, Warren, PV, Gaasterland, T, Young, WG, Lenox, AL, Graham, DE, Overbeek, R, Snead, MA, Keller, M, Aujay, M, Huber, R, Feldman, RA, Short, JM, Olsen, GJ and Swanson, RV (1998). The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. *Nature*. **392**, 353-358.

deHaseth, PL, Zupancic, ML and Record, MT, Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J Bacteriol*. **180**, 3019-3025.

Delgado, A, Salto, R, Marques, S and Ramos, JL (1995). Single amino acids changes in the signal receptor domain of XylR resulted in mutants that stimulate transcription in the absence of effectors. *J Biol Chem*. **270**, 5144-5150.

Demple, B, Ding, H and Jorgensen, M (2002). Escherichia coli SoxR protein: sensor/transducer of oxidative stress and nitric oxide. *Methods Enzymol*. **348**, 355-364.

Ding, H and Demple, B (1997). In vivo kinetics of a redox-regulated transcriptional switch. *Proc Natl Acad Sci U S A*. **94**, 8445-8449.

Ding, H and Demple, B (2000). Direct nitric oxide signal transduction via nitrosylation of iron-sulfur centers in the SoxR transcription activator. *PNAS*. **97**, 5146-5150.

Dixon, R and Kahn, D (2004). Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology*
Nat Rev Micro. **2**, 621-631.

Doucleff, M, Chen, B, Maris, AE, Wemmer, DE, Kondrashkina, E and Nixon, BT (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, LT, Pelton, JG and Wemmer, DE (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.

Doucleff, M, Pelton, JG, Lee, PS, Nixon, BT and Wemmer, DE (2007). Structural basis of DNA recognition by the alternative sigma-factor, sigma54. *J Mol Biol*. **369**, 1070-1078.

Drapier, JC (1997). Interplay between NO and [Fe-S] clusters: relevance to biological systems. *Methods*. **11**, 319-329.

Drummond, MH, Contreras, A and Mitchenall, LA (1990). The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of Klebsiella pneumoniae. *Mol Microbiol*. **4**, 29-37.

Dworkin, J, Jovanovic, G and Model, P (1997). Role of upstream activation sequences and integration host factor in transcriptional activation by the constitutively active prokaryotic enhancer-binding protein PspF. *J Mol Biol*. **273**, 377-388.

Dworkin, J, Jovanovic, G and Model, P (2000). The PspA protein of Escherichia coli is a negative regulator of sigma(54)-dependent transcription. *J Bacteriol*. **182**, 311-319.

Einsle, O, Messerschmidt, A, Stach, P, Bourenkov, GP, Bartunik, HD, Huber, R and Kroneck, PM (1999). Structure of cytochrome c nitrite reductase. *Nature*. **400**, 476-480.

Einsle, O, Stach, P, Messerschmidt, A, Simon, J, Kroger, A, Huber, R and Kroneck, PM (2000). Cytochrome c nitrite reductase from Wolinella succinogenes. Structure at 1.6 A resolution, inhibitor binding, and heme-packing motifs. *J Biol Chem*. **275**, 39608-39616.

Elderkin, S, Bordes, P, Jones, S, Rappas, M and Buck, M (2005). Molecular determinants for PspA-mediated repression of the AAA transcriptional activator PspF. *J. Bacteriol.* **187**, 3238-3248.

Elderkin, S, Jones, S, Schumacher, J, Studholme, D and Buck, M (2002). Mechanism of action of the *Escherichia coli* phage shock protein PspA in repression of the AAA family transcription factor PspF. *J Mol Biol*. **320**, 23-37.

Fang, FC (2004). Antimicrobial reactive oxygen and nitrogen species: concepts and controversies. *Nat Rev Microbiol*. **2**, 820-832.

Feng, J, Goss, TJ, Bender, RA and Ninfa, AJ (1995). Repression of the Klebsiella aerogenes nac promoter. *J Bacteriol*. **177**, 5535-5538.

Fernandez, S, de Lorenzo, V and Perez-Martin, J (1995). Activation of the transcriptional regulator XylR of Pseudomonas putida by release of repression between functional domains. *Mol Microbiol*. **16**, 205-213.

Field, SJ, Thorndycroft, FH, Matorin, AD, Richardson, DJ and Watmough, NJ (2008). The respiratory nitric oxide reductase (NorBC) from Paracoccus denitrificans. *Methods Enzymol*. **437**, 79-101.

Filenko, N, Spiro, S, Browning, DF, Squire, D, Overton, TW, Cole, J and Constantinidou, C (2007). The NsrR regulon of Escherichia coli K-12 includes genes encoding the hybrid cluster protein and the periplasmic, respiratory nitrite reductase. *J Bacteriol*. **189**, 4410-4417.

Flatley, J, Barrett, J, Pullan, ST, Hughes, MN, Green, J and Poole, RK (2005). Transcriptional responses of *Escherichia coli* to S-Nitrosoglutathione under defined chemostat conditions reveal major changes in methionine biosynthesis. *J. Biol. Chem.* **280**, 10065-10072.

Forsburg, SL (2004). Eukaryotic MCM proteins: beyond replication initiation. *Microbiol Mol Biol Rev*. **68**, 109-131.

Frazao, C, Silva, G, Gomes, CM, Matias, P, Coelho, R, Sieker, L, Macedo, S, Liu, MY, Oliveira, S, Teixeira, M, Xavier, AV, Rodrigues-Pousada, C, Carrondo, MA and Le Gall, J (2000). Structure of a dioxygen reduction enzyme from *Desulfovibrio gigas*. *Nature Structural Biology*. **7**, 1041-1045.

Fu, J, Gnatt, AL, Bushnell, DA, Jensen, GJ, Thompson, NE, Burgess, RR, David, PR and Kornberg, RD (1999). Yeast RNA polymerase II at 5 A resolution. *Cell*. **98**, 799-810.

Gai, D, Zhao, R, Li, D, Finkielstein, CV and Chen, XS (2004). Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell*. **119**, 47-60.

Gallegos, MT and Buck, M (1999). Sequences in sigmaN determining holoenzyme formation and properties. *J Mol Biol*. **288**, 539-553.

Gardner, AM and Gardner, PR (2002). Flavohemoglobin detoxifies nitric oxide in aerobic, but not anaerobic, *Escherichia coli*. Evidence for a novel inducible anaerobic nitric oxide-scavenging activity. *J Biol Chem*. **277**, 8166-8171.

Gardner, AM, Gessner, CR and Gardner, PR (2003). Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and sigma 54 In the nitric oxide stress response. *J. Biol. Chem.* **278**, 10081-10086.

Gardner, AM, Helmick, RA and Gardner, PR (2002). Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in *Escherichia coli*. *J Biol Chem*. **277**, 8172-8177.

Gardner, AM, Martin, LA, Gardner, PR, Dou, Y and Olson, JS (2000a). Steady-state and transient kinetics of *Escherichia coli* nitric-oxide dioxygenase (flavohemoglobin). The B10 tyrosine hydroxyl is essential for dioxygen binding and catalysis. *J Biol Chem*. **275**, 12581-12589.

Gardner, PR, Costantino, G and Salzman, AL (1998a). Constitutive and adaptive detoxification of nitric oxide in *Escherichia coli*. Role of nitric-oxide dioxygenase in the protection of aconitase. *J Biol Chem*. **273**, 26528-26533.

Gardner, PR, Costantino, G, Szabo, C and Salzman, AL (1997). Nitric oxide sensitivity of the aconitases. *J Biol Chem*. **272**, 25071-25076.

Gardner, PR, Gardner, AM, Martin, LA, Dou, Y, Li, T, Olson, JS, Zhu, H and Riggs, AF (2000b). Nitric-oxide dioxygenase activity and function of flavohemoglobins. sensitivity to nitric oxide and carbon monoxide inhibition. *J Biol Chem*. **275**, 31581-31587.

Gardner, PR, Gardner, AM, Martin, LA and Salzman, AL (1998b). Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proc Natl Acad Sci U S A*. **95**, 10378-10383.

Garmendia, J and de Lorenzo, V (2000). The role of the interdomain B linker in the activation of the XylR protein of Pseudomonas putida. *Mol Microbiol*. **38**, 401-410.

Gerstel, U, Park, C and Romling, U (2003). Complex regulation of csgD promoter activity by global regulatory proteins. *Mol Microbiol*. **49**, 639-654.

Ghosh, T (2010). Single Particle Cryo-Electron Microscopy of the Enhancer Binding Protein, NorR. *Thesis*, Macromolecular Structure and Function, Imperial College

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

Gilberthorpe, NJ and Poole, RK (2008). Nitric oxide homeostasis in Salmonella typhimurium: roles of respiratory nitrate reductase and flavohemoglobin. *J Biol Chem*. **283**, 11146-11154.

Gomes, CM, Giuffre, A, Forte, E, Vicente, JB, Saraiva, LM, Brunori, M and Teixeira, M (2002). A novel type of nitric-oxide reductase. *Escherichia coli* flavorubredoxin. *J Biol Chem*. **277**, 25273-25276.

Gomes, CM, Vicente, JB, Wasserfallen, A and Teixeira, M (2000). Spectroscopic studies and characterization of a novel electron-transfer chain from *Escherichia coli* involving a flavorubredoxin and its flavoprotein reductase partner. *Biochemistry*. **39**, 16230-16237.

Gonzalez, V, Olvera, L, Soberon, X and Morett, E (1998). In vivo studies on the positive control function of NifA: a conserved hydrophobic amino acid patch at the central domain involved in transcriptional activation. *Mol Microbiol*. **28**, 55-67.

Gourse, RL, Ross, W and Gaal, T (2000). UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol Microbiol*. **37**, 687-695.

Grant, RA, Filman, DJ, Finkel, SE, Kolter, R and Hogle, JM (1998). The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nat Struct Biol*. **5**, 294-303.

Green, J, Scott, C and Guest, JR (2001). Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv Microb Physiol*. **44**, 1-34.

Greenleaf, WB, Shen, J, Gai, D and Chen, XS (2008). Systematic study of the functions for the residues around the nucleotide pocket in simian virus 40 AAA+ hexameric helicase. *J Virol*. **82**, 6017-6023.

Gruber, TM and Gross, CA (2003). Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol*. **57**, 441-466.

Gu, B, Lee, JH, Hoover, TR, Scholl, D and Nixon, BT (1994). Rhizobium meliloti DctD, a sigma 54-dependent transcriptional activator, may be negatively controlled by a subdomain in the C-terminal end of its two-component receiver module. *Mol Microbiol*. **13**, 51-66.

Guest, JR, Green, J, Irvine, AS and Spiro, S (1996). The FNR modulon and FNR-regulated gene expression. *Regulation of gene expression*. Lin, ECC and Lynch, AS. New York, Chapman & Hall: 317–342.

Guo, Y and Gralla, JD (1998). Promoter opening via a DNA fork junction binding activity. *Proc Natl Acad Sci U S A*. **95**, 11655-11660.

Guo, Y, Lew, CM and Gralla, JD (2000). Promoter opening by sigma(54) and sigma(70) RNA polymerases: sigma factor-directed alterations in the mechanism and tightness of control. *Genes Dev*. **14**, 2242-2255.

Guo, Y, Wang, L and Gralla, JD (1999). A fork junction DNA-protein switch that controls promoter melting by the bacterial enhancer-dependent sigma factor. *Embo J*. **18**, 3736-3745.

Haltia, T, Brown, K, Tegoni, M, Cambillau, C, Saraste, M, Mattila, K and Djinovic-Carugo, K (2003). Crystal structure of nitrous oxide reductase from Paracoccus denitrificans at 1.6 A resolution. *Biochem J*. **369**, 77-88.

Hanahan, D (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. **166**, 557-580.

Hanson, PI and Whiteheart, SW (2005). AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol*. **6**, 519-529.

Hantke, K (2001). Iron and metal regulation in bacteria. *Curr Opin Microbiol*. **4**, 172-177.

Harley, C and Reynolds, R (1987). Analysis of *E. coli* promoter sequences. *Nucl. Acids Res*. **15**, 2343-2361.

Hastings, CA, Lee, SY, Cho, HS, Yan, D, Kustu, S and Wemmer, DE (2003). High-resolution solution structure of the beryllofluoride-activated NtrC receiver domain. *Biochemistry*. **42**, 9081-9090.

Haugen, SP, Ross, W and Gourse, RL (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol*. **6**, 507-519.

Hauser, C, Glaser, T, Bill, E, Weyhermuller, T and Wieghardt, K (2000). The electronic structures of an isostructural series of octahedral nitrosyliron complexes {Fe-NO}(6,7,8) elucidated by Mossbauer spectroscopy. *Journal of the American Chemical Society*. **122**, 4352-4365.

Hausladen, A, Gow, A and Stamler, JS (2001). Flavohemoglobin denitrosylase catalyzes the reaction of a nitroxyl equivalent with molecular oxygen. *Proc Natl Acad Sci U S A*. **98**, 10108-10112.

Helmann, JD (1999). Anti-sigma factors. *Curr Opin Microbiol*. **2**, 135-141.

Hernandez-Urzua, E, Zamorano-Sanchez, DS, Ponce-Coria, J, Morett, E, Grogan, S, Poole, RK and Membrillo-Hernandez, J (2007). Multiple regulators of the Flavohaemoglobin (hmp) gene of Salmonella enterica serovar Typhimurium include RamA, a transcriptional regulator conferring the multidrug resistance phenotype. *Arch Microbiol*. **187**, 67-77.

Hirata, A, Klein, BJ and Murakami, KS (2008). The X-ray crystal structure of RNA polymerase from Archaea. *Nature*. **451**, 851-854.

Ho, YS, Burden, LM and Hurley, JH (2000). Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *Embo J*. **19**, 5288-5299.

Hong, E, Doucleff, M and Wemmer, DE (2009). Structure of the RNA polymerase core-binding domain of sigma(54) reveals a likely conformational fracture point. *J Mol Biol*. **390**, 70-82.

Hoover, TR, Santero, E, Porter, S and Kustu, S (1990). The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell*. **63**, 11-22.

Hopper, S and Bock, A (1995). Effector-mediated stimulation of ATPase activity by the sigma 54- dependent transcriptional activator FHLA from *Escherichia coli*. *J. Bacteriol*. **177**, 2798-2803.

Hsieh, M, Hsu, HM, Hwang, SF, Wen, FC, Yu, JS, Wen, CC and Li, C (1999). The hydrophobic heptad repeat in region III of *Escherichia coli* transcription factor sigma 54 is essential for core RNA polymerase binding. *Microbiology*. **145 (Pt 11**), 3081-3088.

Huala, E and Ausubel, FM (1989). The central domain of Rhizobium meliloti NifA is sufficient to activate transcription from the R. meliloti nifH promoter. *J Bacteriol*. **171**, 3354-3365.

Huala, E, Stigter, J and Ausubel, F (1992). The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription. *J. Bacteriol.* **174**, 1428-1431.

Huo, YX, Tian, ZX, Rappas, M, Wen, J, Chen, YC, You, CH, Zhang, X, Buck, M, Wang, YP and Kolb, A (2006). Protein-induced DNA bending clarifies the architectural organization of the sigma54-dependent glnAp2 promoter. *Mol Microbiol.* **59**, 168-180.

Hutcheson, SW, Bretz, J, Sussan, T, Jin, S and Pak, K (2001). Enhancer-binding proteins HrpR and HrpS interact to regulate hrp-encoded type III protein secretion in Pseudomonas syringae strains. *J Bacteriol.* **183**, 5589-5598.

Hutchings, MI, Crack, JC, Shearer, N, Thompson, BJ, Thomson, AJ and Spiro, S (2002a). Transcription factor FnrP from Paracoccus denitrificans contains an iron-sulfur cluster and is activated by anoxia: identification of essential cysteine residues. *J Bacteriol.* **184**, 503-508.

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

Hutchings, MI, Shearer, N, Wastell, S, van Spanning, RJM and Spiro, S (2000). Heterologous NNR-mediated nitric oxide signaling in *Escherichia coli*. *J. Bacteriol.* **182**, 6434-6439.

Hutchings, MI and Spiro, S (2000). The nitric oxide regulated nor promoter of Paracoccus denitrificans. *Microbiology.* **146 (Pt 10)**, 2635-2641.

Hyduke, DR, Jarboe, LR, Tran, LM, Chou, KJ and Liao, JC (2007). Integrated network analysis identifies nitric oxide response networks and dihydroxyacid dehydratase as a crucial target in Escherichia coli. *Proc Natl Acad Sci U S A*. **104**, 8484-8489.

Ilari, A, Bonamore, A, Farina, A, Johnson, KA and Boffi, A (2002). The X-ray structure of ferric *Escherichia coli* flavohemoglobin reveals an unexpected geometry of the distal heme pocket. *J Biol Chem.* **277**, 23725-23732.

Iobbi-Nivol, C, Santini, CL, Blasco, F and Giordano, G (1990). Purification and further characterization of the second nitrate reductase of Escherichia coli K12. *Eur J Biochem*. **188**, 679-687.

Isabella, VM, Lapek, JD, Jr., Kennedy, EM and Clark, VL (2009). Functional analysis of NsrR, a nitric oxide-sensing Rrf2 repressor in Neisseria gonorrhoeae. *Mol Microbiol.* **71**, 227-239.

Ishihama, A (2000). Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol.* **54**, 499-518.

Jain, D, Nickels, BE, Sun, L, Hochschild, A and Darst, SA (2004). Structure of a ternary transcription activation complex. *Mol Cell.* **13**, 45-53.

Janssens, S, Shimouchi, A, Quertermous, T, Bloch, D and Bloch, K (1992). Cloning and expression of a cDNA encoding human endothelium-derived relaxing factor/nitric oxide synthase [published erratum appears in J Biol Chem 1992 Nov 5;267(31):22694]. *J. Biol. Chem.* **267**, 14519-14522.

Ji, X-b and Hollocher, TC (1988). Reduction of nitrite to nitric oxide by enteric bacteria. *Biochemical and Biophysical Research Communications*. **157**, 106-108.

Johnson, DC, Dean, DR, Smith, AD and Johnson, MK (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem*. **74**, 247-281.

Joly, N and Buck, M (2010). Engineered interfaces of an AAA+ ATPase reveal a new nucleotide-dependent coordination mechanism. *J Biol Chem*. **285**, 15178-15186.

Joly, N and Buck, M (2011). Single Chain Forms of the Enhancer Binding Protein PspF Provide Insights into Geometric Requirements for Gene Activation. *J Biol Chem*. **286**, 12734-12742.

Joly, N, Burrows, PC and Buck, M (2008a). An intramolecular route for coupling ATPase activity in AAA+ proteins for transcription activation. *J Biol Chem*. **283**, 13725-13735.

Joly, N, Burrows, PC, Engl, C, Jovanovic, G and Buck, M (2009). A lower-order oligomer form of phage shock protein A (PspA) stably associates with the hexameric AAA(+) transcription activator protein PspF for negative regulation. *J Mol Biol*. **394**, 764-775.

Joly, N, Rappas, M, Buck, M and Zhang, X (2008b). Trapping of a transcription complex using a new nucleotide analogue: AMP aluminium fluoride. *J Mol Biol*. **375**, 1206-1211.

Joly, N, Rappas, M, Wigneshweraraj, SR, Zhang, X and Buck, M (2007). Coupling nucleotide hydrolysis to transcription activation performance in a bacterial enhancer binding protein. *Mol Microbiol*. **66**, 583-595.

Joly, N, Schumacher, J and Buck, M (2006). Heterogeneous nucleotide occupancy stimulates functionality of phage shock protein F, an AAA+ transcriptional activator. *J Biol Chem*. **281**, 34997-35007.

Jones-Carson, J, Laughlin, J, Hamad, MA, Stewart, AL, Voskuil, MI and Vazquez-Torres, A (2008). Inactivation of [Fe-S] metalloproteins mediates nitric oxide-dependent killing of Burkholderia mallei. *PLoS One*. **3**, e1976.

Jormakka, M, Richardson, D, Byrne, B and Iwata, S (2004). Architecture of NarGH reveals a structural classification of Mo-bisMGD enzymes. *Structure*. **12**, 95-104.

Jovanovic, G and Model, P (1997). PspF and IHF bind co-operatively in the psp promoter-regulatory region of Escherichia coli. *Mol Microbiol*. **25**, 473-481.

Jovanovic, G, Rakonjac, J and Model, P (1999). In vivo and in vitro activities of the Escherichia coli sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFDeltaHTH. *J Mol Biol*. **285**, 469-483.

Jovanovic, M, James, EH, Burrows, PC, Rego, FG, Buck, M and Schumacher, J (2011). Regulation of the co-evolved HrpR and HrpS AAA+ proteins required for Pseudomonas syringae pathogenicity. *Nat Commun*. **2**, 177.

Justino, MC, Almeida, CC, Teixeira, M and Saraiva, LM (2007). Escherichia coli di-iron YtfE protein is necessary for the repair of stress-damaged iron-sulfur clusters. *J Biol Chem*. **282**, 10352-10359.

Justino, MC, Goncalves, VM and Saraiva, LM (2005a). Binding of NorR to three DNA sites is essential for promoter activation of the flavorubredoxin gene, the nitric oxide reductase of *Escherichia coli*. *Biochem Biophys Res Commun*. **328**, 540-544.

Justino, MC, Vicente, JB, Teixeira, M and Saraiva, LM (2005b). New genes implicated in the protection of anaerobically grown *Escherichia coli* against nitric oxide. *J Biol Chem*. **280**, 2636-2643.

Jyot, J, Dasgupta, N and Ramphal, R (2002). FleQ, the major flagellar gene regulator in Pseudomonas aeruginosa, binds to enhancer sites located either upstream or atypically downstream of the RpoN binding site. *J Bacteriol*. **184**, 5251-5260.

Kapanidis, AN, Margeat, E, Ho, SO, Kortkhonjia, E, Weiss, S and Ebright, RH (2006). Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science*. **314**, 1144-1147.

Kennedy, MC, Antholine, WE and Beinert, H (1997). An EPR investigation of the products of the reaction of cytosolic and mitochondrial aconitases with nitric oxide. *J Biol Chem*. **272**, 20340-20347.

Kessel, M, Maurizi, MR, Kim, B, Kocsis, E, Trus, BL, Singh, SK and Steven, AC (1995). Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. *J Mol Biol*. **250**, 587-594.

Kiley, PJ and Beinert, H (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol Rev*. **22**, 341-352.

Kiley, PJ and Beinert, H (2003). The role of Fe-S proteins in sensing and regulation in bacteria. *Curr Opin Microbiol*. **6**, 181-185.

Kim, KI, Cheong, GW, Park, SC, Ha, JS, Woo, KM, Choi, SJ and Chung, CH (2000). Heptameric ring structure of the heat-shock protein ClpB, a protein-activated ATPase in Escherichia coli. *J Mol Biol*. **303**, 655-666.

Kim, SO, Orii, Y, Lloyd, D, Hughes, MN and Poole, RK (1999). Anoxic function for the *Escherichia coli* flavohaemoglobin (Hmp): reversible binding of nitric oxide and reduction to nitrous oxide. *FEBS Lett*. **445**, 389-394.

Knowles, R (1982). Denitrification. *Microbiol Rev*. **46**, 43-70.

Koo, IC and Stephens, RS (2003). A developmentally regulated two-component signal transduction system in Chlamydia. *J Biol Chem*. **278**, 17314-17319.

Koo, MS, Lee, JH, Rah, SY, Yeo, WS, Lee, JW, Lee, KL, Koh, YS, Kang, SO and Roe, JH (2003). A reducing system of the superoxide sensor SoxR in Escherichia coli. *EMBO J*. **22**, 2614-2622.

Kwok, T, Yang, J, Pittard, AJ, Wilson, TJ and Davidson, BE (1995). Analysis of an Escherichia coli mutant TyrR protein with impaired capacity for tyrosine-mediated repression, but still able to activate at sigma 70 promoters. *Mol Microbiol*. **17**, 471-481.

Lamattina, L, García-Mata, C, Graziano, M and Pagnussat, G (2003). Nitric oxide: The versatility of an extensive signal molecule. *Annu. Rev. Plant Biol*. **54**, 109-136.

Lanzilotta, WN, Schuller, DJ, Thorsteinsson, MV, Kerby, RL, Roberts, GP and Poulos, TL (2000). Structure of the CO sensing transcription activator CooA. *Nat Struct Biol*. **7**, 876-880.

Lee, J, Lee, HJ, Shin, MK and Ryu, WS (2004). Versatile PCR-mediated insertion or deletion mutagenesis. *Biotechniques*. **36**, 398-400.

Lee, KC, Yeo, WS and Roe, JH (2008). Oxidant-responsive induction of the suf operon, encoding a Fe-S assembly system, through Fur and IscR in Escherichia coli. *J Bacteriol*. **190**, 8244-8247.

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

Lenzen, CU, Steinmann, D, Whiteheart, SW and Weis, WI (1998). Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell*. **94**, 525-536.

Leonhartsberger, S, Huber, A, Lottspeich, F and Bock, A (2001). The hydH/G Genes from Escherichia coli code for a zinc and lead responsive two-component regulatory system. *J Mol Biol*. **307**, 93-105.

Lepoivre, M, Fieschi, F, Coves, J, Thelander, L and Fontecave, M (1991). Inactivation of ribonucleotide reductase by nitric oxide. *Biochem Biophys Res Commun*. **179**, 442-448.

Li, J, Passaglia, L, Rombel, I, Yan, D and Kustu, S (1999). Mutations affecting motifs of unknown function in the central domain of nitrogen regulatory protein C. *J Bacteriol*. **181**, 5443-5454.

Lill, R and Muhlenhoff, U (2006). Iron-sulfur protein biogenesis in eukaryotes: components and mechanisms. *Annu Rev Cell Dev Biol*. **22**, 457-486.

Little, R and Dixon, R (2003). The amino-terminal GAF domain of Azotobacter vinelandii NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. *J Biol Chem*. **278**, 28711-28718.

Lorenz, E and Stauffer, GV (1995). Characterization of the MetR binding sites for the glyA gene of Escherichia coli. *J Bacteriol*. **177**, 4113-4120.

Lupas, AN and Martin, J (2002). AAA proteins. *Current Opinion in Structural Biology*. **12**, 746-753.

MacMicking, J, Xie, Q-w and Nathan, C (1997). Nitric oxide and macrophage function. *Annual Review of Immunology*. **15**, 323-350.

Marietou, A, Richardson, D, Cole, J and Mohan, S (2005). Nitrate reduction by Desulfovibrio desulfuricans: a periplasmic nitrate reductase system that lacks NapB, but includes a unique tetraheme c-type cytochrome, NapM. *FEMS Microbiol Lett*. **248**, 217-225.

Martinez-Argudo, I, Little, R and Dixon, R (2004a). A crucial arginine residue is required for a conformational switch in NifL to regulate nitrogen fixation in *Azotobacter vinelandii*. *PNAS*. **101**, 16316-16321.

Martinez-Argudo, I, Little, R and Dixon, R (2004b). Role of the amino-terminal GAF domain of the NifA activator in controlling the response to the antiactivator protein NifL. *Molecular Microbiology*. **52**, 1731-1744.

Martinez-Argudo, I, Little, R, Shearer, N, Johnson, P and Dixon, R (2004c). The NifL-NifA System: a Multidomain Transcriptional Regulatory Complex That Integrates Environmental Signals. *J. Bacteriol.* **186**, 601-610.

Martinez, SE, Wu, AY, Glavas, NA, Tang, X-B, Turley, S, Hol, WGJ and Beavo, JA (2002). The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding. *PNAS*. **99**, 13260-13265.

Mathew, R and Chatterji, D (2006). The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol*. **14**, 450-455.

Maxam, AM and Gilbert, W (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A*. **74**, 560-564.

Membrillo-Hernandez, J, Coopamah, MD, Channa, A, Hughes, MN and Poole, RK (1998). A novel mechanism for upregulation of the *Escherichia coli* K-12 hmp (flavohaemoglobin) gene by the 'NO releaser', S-nitrosoglutathione: nitrosation of

homocysteine and modulation of MetR binding to the glyA-hmp intergenic region. *Mol Microbiol*. **29**, 1101-1112.

Merrick, M and Chambers, S (1992). The helix-turn-helix motif of sigma 54 is involved in recognition of the -13 promoter region. *J Bacteriol*. **174**, 7221-7226.

Merrick, MJ (1993). In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). *Mol Microbiol*. **10**, 903-909.

Meyer, MG, Park, S, Zeringue, L, Staley, M, McKinstry, M, Kaufman, RI, Zhang, H, Yan, D, Yennawar, N, Yennawar, H, Farber, GK and Nixon, BT (2001). A dimeric two-component receiver domain inhibits the sigma54-dependent ATPase in DctD. *FASEB J*. **15**, 1326-1328.

Missiakas, D and Raina, S (1998). The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol*. **28**, 1059-1066.

Miyata, T, Yamada, K, Iwasaki, H, Shinagawa, H, Morikawa, K and Mayanagi, K (2000). Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J Struct Biol*. **131**, 83-89.

Model, P, Jovanovic, G and Dworkin, J (1997). The Escherichia coli phage-shock-protein (psp) operon. *Mol Microbiol*. **24**, 255-261.

Money, T, Barrett, J, Dixon, R and Austin, S (2001). Protein-protein Interactions in the complex between the enhancer binding protein NIFA and the sensor NIFL from *Azotobacter vinelandii*. *J. Bacteriol*. **183**, 1359-1368.

Mooney, RA, Darst, SA and Landick, R (2005). Sigma and RNA polymerase: an on-again, off-again relationship? *Mol Cell*. **20**, 335-345.

Moreno-Vivian, C, Cabello, P, Martinez-Luque, M, Blasco, R and Castillo, F (1999). Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. *J. Bacteriol.* **181**, 6573-6584.

Morett, E and Buck, M (1989). In vivo studies on the interaction of RNA polymerase-[sigma]54 with the *Klebsiella pneumoniae* and *Rhizobium meliloti* nifH promoters : The role of NifA in the formation of an open promoter complex. *Journal of Molecular Biology*. **210**, 65-77.

Morett, E, Cannon, W and Buck, M (1988). The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of nif promoters and can be separated from the positive control function of NifA. *Nucleic Acids Res*. **16**, 11469-11488.

Morett, E and Segovia, L (1993). The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains. *J. Bacteriol.* **175**, 6067-6074.

Mukhopadhyay, P, Zheng, M, Bedzyk, LA, LaRossa, RA and Storz, G (2004). Prominent roles of the NorR and Fur regulators in the *Escherichia coli* transcriptional response to reactive nitrogen species. *PNAS*. **101**, 745-750.

Mur, LA, Kenton, P, Lloyd, AJ, Ougham, H and Prats, E (2008). The hypersensitive response; the centenary is upon us but how much do we know? *J Exp Bot*. **59**, 501-520.

Murakami, KS and Darst, SA (2003). Bacterial RNA polymerases: the wholo story. *Curr Opin Struct Biol*. **13**, 31-39.

Murakami, KS, Masuda, S, Campbell, EA, Muzzin, O and Darst, SA (2002a). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*. **296**, 1285-1290.

Murakami, KS, Masuda, S and Darst, SA (2002b). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 A resolution. *Science*. **296**, 1280-1284.

Nathan, C (1992). Nitric oxide as a secretory product of mammalian cells. *FASEB J*. **6**, 3051-3064.

Nathan, C and Xie, QW (1994). Nitric oxide synthases: roles, tolls, and controls. *Cell*. **78**, 915-918.

Neuwald, AF, Aravind, L, Spouge, JL and Koonin, EV (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*. **9**, 27-43.

Nilavongse, A, Brondijk, TH, Overton, TW, Richardson, DJ, Leach, ER and Cole, JA (2006). The NapF protein of the Escherichia coli periplasmic nitrate reductase system: demonstration of a cytoplasmic location and interaction with the catalytic subunit, NapA. *Microbiology*. **152**, 3227-3237.

Ninfa, AJ and Magasanik, B (1986). Covalent Modification of the glnG Product, NRI, by the glnL Product, NRII, Regulates the Transcription of the glnALG Operon in *Escherichia coli*. *PNAS*. **83**, 5909-5913.

Nojiri, M, Koteishi, H, Nakagami, T, Kobayashi, K, Inoue, T, Yamaguchi, K and Suzuki, S (2009). Structural basis of inter-protein electron transfer for nitrite reduction in denitrification. *Nature*. **462**, 117-120.

North, AK and Kustu, S (1997). Mutant forms of the enhancer-binding protein NtrC can activate transcription from solution. *J Mol Biol*. **267**, 17-36.

North, AK, Weiss, DS, Suzuki, H, Flashner, Y and Kustu, S (1996). Repressor forms of the enhancer-binding protein NtrC: some fail in coupling ATP hydrolysis to open complex formation by sigma 54-holoenzyme. *J Mol Biol*. **260**, 317-331.

O'Neill, E, Wikstrom, P and Shingler, V (2001). An active role for a structured B-linker in effector control of the sigma54-dependent regulator DmpR. *Embo J*. **20**, 819-827.

Ogura, T and Wilkinson, AJ (2001). AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells*. **6**, 575-597.

Osuna, J, Soberon, X and Morett, E (1997). A proposed architecture for the central domain of the bacterial enhancer-binding proteins based on secondary structure prediction and fold recognition. *Protein Sci*. **6**, 543-555.

Outten, FW, Djaman, O and Storz, G (2004). A suf operon requirement for Fe-S cluster assembly during iron starvation in Escherichia coli. *Mol Microbiol*. **52**, 861-872.

Pacelli, R, Wink, DA, Cook, JA, Krishna, MC, DeGraff, W, Friedman, N, Tsokos, M, Samuni, A and Mitchell, JB (1995). Nitric oxide potentiates hydrogen peroxide-induced killing of *Escherichia coli*. *J Exp Med*. **182**, 1469-1479.

Paget, MS and Helmann, JD (2003). The sigma70 family of sigma factors. *Genome Biol*. **4**, 203.

Park, S, Meyer, M, Jones, AD, Yennawar, HP, Yennawar, NH and Nixon, BT (2002). Two-component signaling in the AAA + ATPase DctD: binding Mg2+ and BeF3-selects between alternate dimeric states of the receiver domain. *FASEB J*. **16**, 1964-1966.

Park, SC, Jia, B, Yang, JK, Van, DL, Shao, YG, Han, SW, Jeon, YJ, Chung, CH and Cheong, GW (2006). Oligomeric structure of the ATP-dependent protease La (Lon) of Escherichia coli. *Mol Cells*. **21**, 129-134.

Partridge, JD, Bodenmiller, DM, Humphrys, MS and Spiro, S (2009). NsrR targets in the Escherichia coli genome: new insights into DNA sequence requirements for binding and a role for NsrR in the regulation of motility. *Mol Microbiol*. **73**, 680-694.

Pelton, JG, Kustu, S and Wemmer, DE (1999). Solution structure of the DNA-binding domain of NtrC with three alanine substitutions. *J Mol Biol*. **292**, 1095-1110.

Perez-Martin, J and De Lorenzo, V (1995). Integration host factor suppresses promiscuous activation of the sigma 54-dependent promoter Pu of Pseudomonas putida. *Proc Natl Acad Sci U S A*. **92**, 7277-7281.

Perez-Martin, J and de Lorenzo, V (1996). In vitro activities of an N-terminal truncated form of XylR, a sigma 54-dependent transcriptional activator of *Pseudomonas putida*. *J Mol Biol*. **258**, 575-587.

Perez-Martin, J and Lorenzo, V (1995). The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. *PNAS*. **92**, 9392-9396.

Pittard, AJ and Davidson, BE (1991). TyrR protein of *Escherichia coli* and its role as repressor and activator. *Mol Microbiol*. **5**, 1585-1592.

Pittard, J, Camakaris, H and Yang, J (2005). The TyrR regulon. *Mol Microbiol*. **55**, 16-26.

Poggio, S, Osorio, A, Dreyfus, G and Camarena, L (2002). The four different sigma(54) factors of Rhodobacter sphaeroides are not functionally interchangeable. *Mol Microbiol*. **46**, 75-85.

Pohlmann, A, Cramm, R, Schmelz, K and Friedrich, B (2000). A novel NO-responding regulator controls the reduction of nitric oxide in *Ralstonia eutropha*. *Molecular Microbiology*. **38**, 626-638.

Ponting, CP and Aravind, L (1997). PAS: a multifunctional domain family comes to light. *Curr Biol*. **7**, R674-677.

Poock, SR, Leach, ER, Moir, JWB, Cole, JA and Richardson, DJ (2002). Respiratory detoxification of nitric oxide by the cytochrome c nitrite reductase of *Escherichia coli*. *J. Biol. Chem*. **277**, 23664-23669.

Poole, R, Anjum, M, Membrillo-Hernandez, J, Kim, S, Hughes, M and Stewart, V (1996). Nitric oxide, nitrite, and Fnr regulation of hmp (flavohemoglobin) gene expression in *Escherichia coli* K-12. *J. Bacteriol*. **178**, 5487-5492.

Poole, RK and Hughes, MN (2000). New functions for the ancient globin family: bacterial responses to nitric oxide and nitrosative stress. MicroReview. *Molecular Microbiology*. **36**, 775-783.

Porrua, O, Garcia-Gonzalez, V, Santero, E, Shingler, V and Govantes, F (2009). Activation and repression of a sigma(N)-dependent promoter naturally lacking upstream activation sequences. *Molecular Microbiology*. **73**, 419-433.

Potter, LC, Millington, P, Griffiths, L, Thomas, GH and Cole, JA (1999). Competition between Escherichia coli strains expressing either a periplasmic or a membrane-bound nitrate reductase: does Nap confer a selective advantage during nitrate-limited growth? *Biochem J*. **344 Pt 1**, 77-84.

Prentki, P and Krisch, HM (1984). In vitro insertional mutagenesis with a selectable DNA fragment. *Gene*. **29**, 303-313.

Preston, G, Deng, WL, Huang, HC and Collmer, A (1998). Negative regulation of hrp genes in Pseudomonas syringae by HrpV. *J Bacteriol*. **180**, 4532-4537.

Pullan, ST, Gidley, MD, Jones, RA, Barrett, J, Stevanin, TM, Read, RC, Green, J and Poole, RK (2007). Nitric oxide in chemostat-cultured Escherichia coli is sensed by Fnr and other global regulators: unaltered methionine biosynthesis indicates lack of S nitrosation. *J Bacteriol*. **189**, 1845-1855.

Rabin, RS and Stewart, V (1993). Dual response regulators (NarL and NarP) interact with dual sensors (NarX and NarQ) to control nitrate- and nitrite-regulated gene expression in Escherichia coli K-12. *J Bacteriol*. **175**, 3259-3268.

Raivio, TL and Silhavy, TJ (2001). Periplasmic stress and ECF sigma factors. *Annu Rev Microbiol*. **55**, 591-624.

Ramseier, TM, Figge, RM and Saier, MH, Jr. (1994). DNA sequence of a gene in *Escherichia coli* encoding a putative tripartite transcription factor with receiver, ATPase and DNA binding domains. *DNA Seq*. **5**, 17-24.

Rappas, M, Bose, D and Zhang, X (2007). Bacterial enhancer-binding proteins: unlocking sigma54-dependent gene transcription. *Curr Opin Struct Biol*. **17**, 110-116.

Rappas, M, Schumacher, J, Beuron, F, Niwa, H, Bordes, P, Wigneshweraraj, S, Keetch, CA, Robinson, CV, Buck, M and Zhang, X (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 conformational changes in the AAA+ domain of transcription activator PspF. *J Mol Biol*. **357**, 481-492.

Ray, M, Golombek, AP, Hendrich, MP, Yap, GPA, Liable-Sands, LM, Rheingold, AL and Borovik, AS (1999). Structure and magnetic properties of trigonal bipyramidal iron nitrosyl complexes. *Inorganic Chemistry*. **38**, 3110-3115.

Reid, JD, Siebert, CA, Bullough, PA and Hunter, CN (2003). The ATPase activity of the ChlI subunit of magnesium chelatase and formation of a heptameric AAA+ ring. *Biochemistry*. **42**, 6912-6920.

Reitzer, L (2003). Nitrogen assimilation and global regulation in Escherichia coli. *Annu Rev Microbiol*. **57**, 155-176.

Ren, B, Zhang, N, Yang, J and Ding, H (2008). Nitric oxide-induced bacteriostasis and modification of iron-sulphur proteins in Escherichia coli. *Mol Microbiol*. **70**, 953-964.

Revyakin, A, Liu, C, Ebright, RH and Strick, TR (2006). Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science*. **314**, 1139-1143.

Richardson, DJ, Berks, BC, Russell, DA, Spiro, S and Taylor, CJ (2001). Functional, biochemical and genetic diversity of prokaryotic nitrate reductases. *Cell Mol Life Sci*. **58**, 165-178.

Rippe, K, Guthold, M, von Hippel, PH and Bustamante, C (1997). Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of *E. coli* RNA polymerase x sigma 54 holoenzyme by scanning force microscopy. *J Mol Biol*. **270**, 125-138.

Rodionov, DA, Dubchak, IL, Arkin, AP, Alm, EJ and Gelfand, MS (2005). Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. *PLoS Comput Biol*. **1**, e55.

Rodrigues, ML, Oliveira, T, Matias, PM, Martins, IC, Valente, FM, Pereira, IA and Archer, M (2006). Crystallization and preliminary structure determination of the membrane-bound complex cytochrome c nitrite reductase from Desulfovibrio vulgaris Hildenborough. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. **62**, 565-568.

Rogers, PA and Ding, H (2001). L-cysteine-mediated destabilization of dinitrosyl iron complexes in proteins. *J Biol Chem*. **276**, 30980-30986.

Rogers, PA, Eide, L, Klungland, A and Ding, H (2003). Reversible inactivation of E. coli endonuclease III via modification of its [4Fe-4S] cluster by nitric oxide. *DNA Repair (Amst)*. **2**, 809-817.

Rohrwild, M, Pfeifer, G, Santarius, U, Muller, SA, Huang, HC, Engel, A, Baumeister, W and Goldberg, AL (1997). The ATP-dependent HslVU protease from Escherichia coli is a four-ring structure resembling the proteasome. *Nat Struct Biol*. **4**, 133-139.

Rombel, I, North, A, Hwang, I, Wyman, C and Kustu, S (1998). The bacterial enhancer-binding protein NtrC as a molecular machine. *Cold Spring Harb Symp Quant Biol*. **63**, 157-166.

Rombel, I, Peters-Wendisch, P, Mesecar, A, Thorgeirsson, T, Shin, YK and Kustu, S (1999). MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. *J Bacteriol*. **181**, 4628-4638.

Romermann, D and Friedrich, B (1985). Denitrification by Alcaligenes eutrophus is plasmid dependent. *J Bacteriol*. **162**, 852-854.

Romermann, D, Warrelmann, J, Bender, RA, Friedrich, B, Romermann, D and Friedrich, B (1989). An rpoN-like gene of Alcaligenes eutrophus and Pseudomonas facilis controls expression of diverse metabolic pathways, including hydrogen oxidation
Denitrification by Alcaligenes eutrophus is plasmid dependent. *J Bacteriol*. **171**, 1093-1099.

Rosenzweig, AC (2000). Nitrous oxide reductase from CuA to CuZ. *Nat Struct Biol*. **7**, 169-171.

Ross, W, Ernst, A and Gourse, RL (2001). Fine structure of E. coli RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev*. **15**, 491-506.

Rouiller, I, DeLaBarre, B, May, AP, Weis, WI, Brunger, AT, Milligan, RA and Wilson-Kubalek, EM (2002). Conformational changes of the multifunction p97 AAA ATPase during its ATPase cycle. *Nat Struct Biol*. **9**, 950-957.

Sallai, L and Tucker, PA (2005). Crystal structure of the central and C-terminal domain of the sigma(54)-activator ZraR. *J Struct Biol*. **151**, 160-170.

Sanderson, A, Mitchell, JE, Minchin, SD and Busby, SJ (2003). Substitutions in the Escherichia coli RNA polymerase sigma70 factor that affect recognition of extended -10 elements at promoters. *FEBS Lett*. **544**, 199-205.

Saraste, M, Sibbald, PR and Wittinghofer, A (1990). The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci*. **15**, 430-434.

Schapiro, JM, Libby, SJ and Fang, FC (2003). Inhibition of bacterial DNA replication by zinc mobilization during nitrosative stress. *PNAS*. **100**, 8496-8501.

Schlensog, V, Lutz, S and Bock, A (1994). Purification and DNA-binding properties of FHLA, the transcriptional activator of the formate hydrogenlyase system from *Escherichia coli*. *J. Biol. Chem.* **269**, 19590-19596.

Schroeder, LA, Choi, AJ and DeHaseth, PL (2007). The -11A of promoter DNA and two conserved amino acids in the melting region of sigma70 both directly affect the rate limiting step in formation of the stable RNA polymerase-promoter complex, but they do not necessarily interact. *Nucleic Acids Res*. **35**, 4141-4153.

Schroeder, LA, Gries, TJ, Saecker, RM, Record, MT, Jr., Harris, ME and DeHaseth, PL (2009). Evidence for a tyrosine-adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of Escherichia coli RNA polymerase to promoter DNA. *J Mol Biol*. **385**, 339-349.

Schroeder, LA, Karpen, ME and deHaseth, PL (2008). Threonine 429 of Escherichia coli sigma 70 is a key participant in promoter DNA melting by RNA polymerase. *J Mol Biol*. **376**, 153-165.

Schumacher, J, Joly, N, Claeys-Bouuaert, IL, Aziz, SA, Rappas, M, Zhang, X and Buck, M (2008). Mechanism of homotropic control to coordinate hydrolysis in a hexameric AAA+ ring ATPase. *J Mol Biol*. **381**, 1-12.

Schumacher, J, Joly, N, Rappas, M, Bradley, D, Wigneshweraraj, SR, Zhang, X and Buck, M (2007). Sensor I threonine of the AAA+ ATPase transcriptional activator PspF is involved in coupling nucleotide triphosphate hydrolysis to the restructuring of sigma 54-RNA polymerase. *J Biol Chem*. **282**, 9825-9833.

Schumacher, J, Joly, N, Rappas, M, Zhang, X and Buck, M (2006). Structures and organisation of AAA+ enhancer binding proteins in transcriptional activation. *J Struct Biol*. **156**, 190-199.

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.

Schuster, M and Grimm, C (2000). Domain switching between hrpR and hrpS affects the regulatory function of the hybrid genes in Pseudomonas syringae pv. phaseolicola. *Mol Plant Pathol*. **1**, 233-241.

Schwartz, CJ, Giel, JL, Patschkowski, T, Luther, C, Ruzicka, FJ, Beinert, H and Kiley, PJ (2001). IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proc Natl Acad Sci U S A*. **98**, 14895-14900.

Sears, HJ, Sawers, G, Berks, BC, Ferguson, SJ and Richardson, DJ (2000). Control of periplasmic nitrate reductase gene expression (napEDABC) from Paracoccus pantotrophus in response to oxygen and carbon substrates. *Microbiology*. **146 (Pt 11)**, 2977-2985.

Sessa, W, Harrison, J, Barber, C, Zeng, D, Durieux, M, D'Angelo, D, Lynch, K and Peach, M (1992). Molecular cloning and expression of a cDNA encoding endothelial cell nitric oxide synthase. *J. Biol. Chem.* **267**, 15274-15276.

Shingler, V (1996). Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism. *Molecular Microbiology*. **19**, 409-416.

Shingler, V (2010). Signal sensory systems that impact sigma(54) -dependent transcription. *FEMS Microbiol Rev*,

Shingler, V, Bartilson, M and Moore, T (1993). Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators. *J Bacteriol*. **175**, 1596-1604.

Shingler, V, Franklin, FC, Tsuda, M, Holroyd, D and Bagdasarian, M (1989). Molecular analysis of a plasmid-encoded phenol hydroxylase from Pseudomonas CF600. *J Gen Microbiol*. **135**, 1083-1092.

Shingler, V and Moore, T (1994). Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas sp.* strain CF600. *J Bacteriol*. **176**, 1555-1560.

Shingler, V and Pavel, H (1995). Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. *Mol Microbiol*. **17**, 505-513.

Sorenson, MK, Ray, SS and Darst, SA (2004). Crystal structure of the flagellar sigma/anti-sigma complex sigma(28)/FlgM reveals an intact sigma factor in an inactive conformation. *Mol Cell*. **14**, 127-138.

Southern, E and Merrick, M (2000). The role of region II in the RNA polymerase sigma factor sigma(N) (sigma(54)). *Nucleic Acids Res*. **28**, 2563-2570.

Spiro, S (2006). Nitric oxide-sensing mechanisms in Escherichia coli. *Biochem Soc Trans*. **34**, 200-202.

Spiro, S (2007). Regulators of bacterial responses to nitric oxide. *FEMS Microbiol Rev*. **31**, 193-211.

Spohn, G and Scarlato, V (1999). Motility of Helicobacter pylori is coordinately regulated by the transcriptional activator FlgR, an NtrC homolog. *J Bacteriol*. **181**, 593-599.

Squire, DJ, Xu, M, Cole, JA, Busby, SJ and Browning, DF (2009). Competition between NarL-dependent activation and Fis-dependent repression controls expression from the Escherichia coli yeaR and ogt promoters. *Biochem J*. **420**, 249-257.

Stahlberg, H, Kutejova, E, Suda, K, Wolpensinger, B, Lustig, A, Schatz, G, Engel, A and Suzuki, CK (1999). Mitochondrial Lon of Saccharomyces cerevisiae is a ring-shaped protease with seven flexible subunits. *Proc Natl Acad Sci U S A*. **96**, 6787-6790.

Stamler, JS, Singel, DJ and Loscalzo, J (1992). Biochemistry of nitric oxide and its redox-activated forms. *Science*. **258**, 1898-1902.

Stevanin, TM, Ioannidis, N, Mills, CE, Kim, SO, Hughes, MN and Poole, RK (2000). Flavohemoglobin Hmp affords inducible protection for *Escherichia coli* respiration, catalyzed by cytochromes bo' or bd, from nitric oxide. *J Biol Chem*. **275**, 35868-35875.

Stewart, V, Lu, Y and Darwin, AJ (2002). Periplasmic nitrate reductase (NapABC enzyme) supports anaerobic respiration by Escherichia coli K-12. *J Bacteriol*. **184**, 1314-1323.

Stock, AM, Robinson, VL and Goudreau, PN (2000). Two-component signal transduction. *Annu Rev Biochem*. **69**, 183-215.

Studholme, DJ and Dixon, R (2003). Domain architectures of sigma 54-dependent transcriptional activators. *J. Bacteriol*. **185**, 1757-1767.

Studier, WF, Rosenberg, AH, Dunn, JJ and Dubendorff, JW (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology*, 60-89.

Su, W, Porter, S, Kustu, S and Echols, H (1990). DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter. *Proc Natl Acad Sci U S A*. **87**, 5504-5508.

Svergun, DI, Malfois, M, Koch, MH, Wigneshweraraj, SR and Buck, M (2000). Low resolution structure of the sigma54 transcription factor revealed by X-ray solution scattering. *J Biol Chem*. **275**, 4210-4214.

Syed, A and Gralla, JD (1998). Identification of an N-terminal region of sigma 54 required for enhancer responsiveness. *J Bacteriol*. **180**, 5619-5625.

Takahashi, Y and Tokumoto, U (2002). A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *J Biol Chem*. **277**, 28380-28383.

Tavares, P, Pereira, AS, Moura, JJ and Moura, I (2006). Metalloenzymes of the denitrification pathway. *J Inorg Biochem*. **100**, 2087-2100.

Taylor, BL and Zhulin, IB (1999). PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev*. **63**, 479-506.

Taylor, BS and Geller, DA (2000). Molecular regulation of the human inducible nitric oxide synthase (iNOS) gene. *Shock*. **13**, 413-424.

Taylor, M, Butler, R, Chambers, S, Casimiro, M, Badii, F and Merrick, M (1996). The RpoN-box motif of the RNA polymerase sigma factor sigma N plays a role in promoter recognition. *Mol Microbiol*. **22**, 1045-1054.

Todorovic, S, Justino, MC, Wellenreuther, G, Hildebrandt, P, Murgida, DH, Meyer-Klaucke, W and Saraiva, LM (2008). Iron-sulfur repair YtfE protein from Escherichia coli: structural characterization of the di-iron center. *J Biol Inorg Chem*. **13**, 765-770.

Tucker, N, D'Autreaux, B, Spiro, S and Dixon, R (2005). DNA binding properties of the *Escherichia coli* nitric oxide sensor NorR: towards an understanding of the regulation of flavorubredoxin expression. *Biochem Soc Trans*. **33**, 181-183.

Tucker, NP (2005). Regulation of transcription by the nitric oxide sensing transcription factor NorR. *Thesis*, Department of Molecular Microbiology, John Innes Centre, University of East Anglia

Tucker, NP, D'Autreaux, B, Studholme, DJ, 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, NP, D'Autreaux, B, Yousafzai, F, Fairhurst, S, Spiro, S and Dixon, R (2007). Analysis of the nitric oxide-sensing non-heme iron center in the NorR regulatory protein. *J Biol Chem*,

Tucker, NP, Ghosh, T, Bush, M, Zhang, X and Dixon, R (2010a). Essential roles of three enhancer sites in sigma54-dependent transcription by the nitric oxide sensing regulatory protein NorR. *Nuc. Acids Res*. **38**, 1182-1194.

Tucker, NP, Hicks, MG, Clarke, TA, Crack, JC, Chandra, G, Le Brun, NE, Dixon, R and Hutchings, MI (2008). The transcriptional repressor protein NsrR senses nitric oxide directly via a [2Fe-2S] cluster. *PLoS One*. **3**, e3623.

Tucker, NP, Le Brun, NE, Dixon, R and Hutchings, MI (2010b). There's NO stopping NsrR, a global regulator of the bacterial NO stress response. *Trends Microbiol*. **18**, 149-156.

Tucker, PA and Sallai, L (2007). The AAA+ superfamily--a myriad of motions. *Curr Opin Struct Biol*. **17**, 641-652.

Tyson, K, Busby, S and Cole, J (1997). Catabolite regulation of two Escherichia coli operons encoding nitrite reductases: role of the Cra protein. *Arch Microbiol*. **168**, 240-244.

van Wonderen, JH, Burlat, B, Richardson, DJ, Cheesman, MR and Butt, JN (2008). The nitric oxide reductase activity of cytochrome c nitrite reductase from Escherichia coli. *J Biol Chem*. **283**, 9587-9594.

van Wonderen, JH, Knight, C, Oganesyan, VS, George, SJ, Zumft, WG and Cheesman, MR (2007). Activation of the cytochrome cd1 nitrite reductase from Paracoccus pantotrophus. Reaction of oxidized enzyme with substrate drives a ligand switch at heme c. *J Biol Chem*. **282**, 28207-28215.

Verger, D, Carr, PD, Kwok, T and Ollis, DL (2007). Crystal structure of the N-terminal domain of the TyrR transcription factor responsible for gene regulation of aromatic amino acid biosynthesis and transport in Escherichia coli K12. *J Mol Biol*. **367**, 102-112.

Vrentas, CE, Gaal, T, Ross, W, Ebright, RH and Gourse, RL (2005). Response of RNA polymerase to ppGpp: requirement for the omega subunit and relief of this requirement by DksA. *Genes Dev*. **19**, 2378-2387.

Walker, JE, Saraste, M, Runswick, MJ and Gay, NJ (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J*. **1**, 945-951.

Wang, H and Gunsalus, RP (2000). The nrfA and nirB nitrite reductase operons in Escherichia coli are expressed differently in response to nitrate than to nitrite. *J Bacteriol*. **182**, 5813-5822.

Wang, J, Song, JJ, Seong, IS, Franklin, MC, Kamtekar, S, Eom, SH and Chung, CH (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. *Structure*. **9**, 1107-1116.

Wang, JT, Syed, A, Hsieh, M and Gralla, JD (1995). Converting *Escherichia coli* RNA polymerase into an enhancer-responsive enzyme: role of an NH2-terminal leucine patch in sigma 54. *Science*. **270**, 992-994.

Wang, L, Guo, Y and Gralla, JD (1999). Regulation of sigma 54-dependent transcription by core promoter sequences: role of -12 region nucleotides. *J Bacteriol*. **181**, 7558-7565.

Wang, PG, Xian, M, Tang, X, Wu, X, Wen, Z, Cai, T and Janczuk, AJ (2002). Nitric oxide donors: chemical activities and biological applications. *Chem. Rev.* **102**, 1091-1134.

Wang, YK, Lee, JH, Brewer, JM and Hoover, TR (1997). A conserved region in the sigma54-dependent activator DctD is involved in both binding to RNA polymerase and coupling ATP hydrolysis to activation. *Mol Microbiol*. **26**, 373-386.

Wang, YP, Kolb, A, Buck, M, Wen, J, O'Gara, F and Buc, H (1998). CRP interacts with promoter-bound sigma54 RNA polymerase and blocks transcriptional activation of the dctA promoter. *EMBO J*. **17**, 786-796.

Watmough, NJ, Field, SJ, Hughes, RJ and Richardson, DJ (2009). The bacterial respiratory nitric oxide reductase. *Biochem Soc Trans*. **37**, 392-399.

Weiss, DS, Batut, J, Klose, KE, Keener, J and Kustu, S (1991). The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell*. **67**, 155-167.

Weiss, V and Magasanik, B (1988). Phosphorylation of nitrogen regulator I (NRI) of *Escherichia coli*. *PNAS*. **85**, 8919-8923.

Wigneshweraraj, S, Bose, D, Burrows, PC, Joly, N, Schumacher, J, Rappas, M, Pape, T, Zhang, X, Stockley, P, Severinov, K and Buck, M (2008). Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Mol Microbiol*. **68**, 538-546.

Wigneshweraraj, SR, Burrows, PC, Bordes, P, Schumacher, J, Rappas, M, Finn, RD, Cannon, WV, Zhang, X and Buck, M (2005). The second paradigm for activation of transcription. *Prog Nucleic Acid Res Mol Biol*. **79**, 339-369.

Wigneshweraraj, SR, Casaz, P and Buck, M (2002). Correlating protein footprinting with mutational analysis in the bacterial transcription factor sigma54 (sigmaN). *Nucleic Acids Res*. **30**, 1016-1028.

Wigneshweraraj, SR, Chaney, MK, Ishihama, A and Buck, M (2001). Regulatory sequences in sigma 54 localise near the start of DNA melting. *J Mol Biol*. **306**, 681-701.

Wikstrom, P, O'Neill, E, Ng, LC and Shingler, V (2001). The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation. *J Mol Biol*. **314**, 971-984.

Willows, RD, Hansson, A, Birch, D, Al-Karadaghi, S and Hansson, M (2004). EM single particle analysis of the ATP-dependent BchI complex of magnesium chelatase: an AAA+ hexamer. *J Struct Biol*. **146**, 227-233.

Wolfe, MT, Heo, J, Garavelli, JS and Ludden, PW (2002). Hydroxylamine reductase activity of the hybrid cluster protein from Escherichia coli. *J Bacteriol*. **184**, 5898-5902.

Wong, C and Gralla, JD (1992). A role for the acidic trimer repeat region of transcription factor sigma 54 in setting the rate and temperature dependence of promoter melting in vivo. *J Biol Chem*. **267**, 24762-24768.

Woodmansee, AN and Imlay, JA (2002). Reduced flavins promote oxidative DNA damage in non-respiring *Escherichia coli* by delivering electrons to intracellular free iron. *J Biol Chem*. **277**, 34055-34066.

Woodmansee, AN and Imlay, JA (2003). A mechanism by which nitric oxide accelerates the rate of oxidative DNA damage in *Escherichia coli*. *Mol Microbiol*. **49**, 11-22.

Wootton, JC and Drummond, MH (1989). The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng*. **2**, 535-543.

Wosten, MM (1998). Eubacterial sigma-factors. *FEMS Microbiol Rev*. **22**, 127-150.

Wu, G, Cruz-Ramos, H, Hill, S, Green, J, Sawers, G and Poole, RK (2000). Regulation of cytochrome bd expression in the obligate aerobe Azotobacter vinelandii by CydR (Fnr). Sensitivity to oxygen, reactive oxygen species, and nitric oxide. *J Biol Chem*. **275**, 4679-4686.

Xie, Q, Cho, H, Kashiwabara, Y, Baum, M, Weidner, J, Elliston, K, Mumford, R and Nathan, C (1994). Carboxyl terminus of inducible nitric oxide synthase. Contribution to NADPH binding and enzymatic activity. *J. Biol. Chem.* **269**, 28500-28505.

Xu, H, Gu, B, Nixon, BT and Hoover, TR (2004a). Purification and characterization of the AAA+ domain of Sinorhizobium meliloti DctD, a sigma54-dependent transcriptional activator. *J Bacteriol*. **186**, 3499-3507.

Xu, H and Hoover, TR (2001). Transcriptional regulation at a distance in bacteria. *Current Opinion in Microbiology*. **4**, 138-144.

Xu, H, Kelly, MT, Nixon, BT and Hoover, TR (2004b). Novel substitutions in the sigma54-dependent activator DctD that increase dependence on upstream activation sequences or uncouple ATP hydrolysis from transcriptional activation. *Mol Microbiol*. **54**, 32-44.

Yan, D and Kustu, S (1999). "Switch I" mutant forms of the bacterial enhancer-binding protein NtrC that perturb the response to DNA. *Proc Natl Acad Sci U S A*. **96**, 13142-13146.

Yang, CC and Nash, HA (1989). The interaction of E. coli IHF protein with its specific binding sites. *Cell*. **57**, 869-880.

Yang, W, Rogers, PA and Ding, H (2002). Repair of nitric oxide-modified ferredoxin [2Fe-2S] cluster by cysteine desulfurase (IscS). *J Biol Chem*. **277**, 12868-12873.

Yanisch-Perron, C, Vieira, J and Messing, J (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*. **33**, 103-119.

Yeo, WS, Lee, JH, Lee, KC and Roe, JH (2006). IscR acts as an activator in response to oxidative stress for the suf operon encoding Fe-S assembly proteins. *Mol Microbiol*. **61**, 206-218.

Young, BA, Anthony, LC, Gruber, TM, Arthur, TM, Heyduk, E, Lu, CZ, Sharp, MM, Heyduk, T, Burgess, RR and Gross, CA (2001). A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell*. **105**, 935-944.

Yu, A and Haggard-Ljungquist, E (1993). Characterization of the binding sites of two proteins involved in the bacteriophage P2 site-specific recombination system. *J Bacteriol*. **175**, 1239-1249.

Yu, X, VanLoock, MS, Poplawski, A, Kelman, Z, Xiang, T, Tye, BK and Egelman, EH (2002). The Methanobacterium thermoautotrophicum MCM protein can form heptameric rings. *EMBO Rep*. **3**, 792-797.

Yukl, ET, Elbaz, MA, Nakano, MM and Moenne-Loccoz, P (2008). Transcription Factor NsrR from Bacillus subtilis Senses Nitric Oxide with a 4Fe-4S Cluster (dagger). *Biochemistry*. **47**, 13084-13092.

Yura, T and Nakahigashi, K (1999). Regulation of the heat-shock response. *Curr Opin Microbiol*. **2**, 153-158.

Yurgel, S, Mortimer, MW, Rogers, KN and Kahn, ML (2000). New substrates for the dicarboxylate transport system of Sinorhizobium meliloti. *J Bacteriol*. **182**, 4216-4221.

Zhang, G, Campbell, EA, Minakhin, L, Richter, C, Severinov, K and Darst, SA (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. *Cell*. **98**, 811-824.

Zhang, N (2010). Functional Studies of the "GAFTGA" motif of *Escherichia coli* Phage Shock Protein F. *Thesis*, Division of Biology, Imperial College London

Zhang, N, Joly, N, Burrows, PC, Jovanovic, M, Wigneshweraraj, SR and Buck, M (2009). The role of the conserved phenylalanine in the sigma54-interacting GAFTGA motif of bacterial enhancer binding proteins. *Nucleic Acids Res*. **37**, 5981-5992.

Zhang, X, Chaney, M, Wigneshweraraj, SR, Schumacher, J, Bordes, P, Cannon, W and Buck, M (2002). Mechanochemical ATPases and transcriptional activation. *Molecular Microbiology*. **45**, 895-903.

Zhang, X, Shaw, A, Bates, PA, Newman, RH, Gowen, B, Orlova, E, Gorman, MA, Kondo, H, Dokurno, P, Lally, J, Leonard, G, Meyer, H, van Heel, M and Freemont, PS (2000). Structure of the AAA ATPase p97. *Mol Cell*. **6**, 1473-1484.

Zhang, X and Wigley, DB (2008). The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. *Nat Struct Mol Biol*. **15**, 1223-1227.

Zheng, L, Cash, VL, Flint, DH and Dean, DR (1998). Assembly of iron-sulfur clusters. Identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii. *J Biol Chem*. **273**, 13264-13272.

Zheng, M, Doan, B, Schneider, TD and Storz, G (1999). OxyR and SoxRS regulation of fur. *J Bacteriol*. **181**, 4639-4643.

Zhulin, IB, Taylor, BL and Dixon, R (1997). PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox. *Trends Biochem Sci*. **22**, 331-333.

Zumft, W (1997). Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev*. **61**, 533-616.

12   Appendix

12.1   Appendix – Materials and Methods

12.1.1 Appendix - *E. coli* strains used in this work

Strain	Description	Reference/Source
BL21(DE3)	*ompT hsdSB(rB⁻mB⁻) gal dcm* (DE3). General expression host.	(Studier *et al.* 1990)
DH10B	*mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX74 recA*1 *endA1 araD139* Δ(*ara,leu*)7697 *galU galK* λ- *rps*L *nupG*	Invitrogen
DH5α	*sipE44* Δ(*lacU169* (Φ80d*lacZ*ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	(Hanahan 1983)
XL10 Gold	Tetr Δ*(mcrA)183* Δ*(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lacIqZ*ΔM15 *Tn10* (Tetr) Camr]. Quickchange host.	Stratagene
MH1003	DH10B *norR::cat* λ*norV-lacZ*	(Hutchings *et al.* 2002b)
MC1000	*araD139* Δ (*araABC-leu*)7679 *galU galK* Δ(*lac*)X74 *rpsL*	(Casadaban and Cohen 1980)
MC100V	MC1000 λ*norV-lacZ*	(Tucker 2005)
MC101V	MC1000 λ*norV-lacZ* NorR site 1 mutant	(Tucker 2005)
MC102V	MC1000 λ*norV-lacZ* NorR site 2 mutant	(Tucker 2005)
MC103V	MC1000 λ*norV-lacZ* NorR site 3 mutant	(Tucker 2005)
MC100VΔ*norR*	MC1000 λ*norV-lacZ norR::cat*	This work
MC101VΔ*norR*	MC1000 λ*norV-lacZ norR::cat* NorR site 1 mutant	This work
MC102VΔ*norR*	MC1000 λ*norV-lacZ norR::cat* NorR site 2 mutant	This work
MC103VΔ*norR*	MC1000 λ*norV-lacZ norR::cat* NorR site 3 mutant	This work

12.1.2 Appendix - The plasmids used and engineered in this work

Construct Name	Description	Reference
pUC19	Standard cloning vector	(Yanisch-Perron *et al.* 1985)
pET21a	Expression vector. Cb resistance	Novagen
pETM11	Expression vector derived from pET24d (Novagen) by introducing a TEV protease cleavage site between the polyhistidine tag and the multiple cloning site.	EMBL
pETNdeM11	pETM11 but with NcoI site mutated to *Nde* I site to allow for easy cloning of the *norR* sequence	This Work
pHP45Ω	plasmid containing 2.0kb omega cassette between SmaI sites – confers Sm/Spc resistance	(Prentki and Krisch 1984)
pNorR	pET21a expressing *norR*	(Tucker *et al.* 2004)
pNorRΔGAF	pET21a expressing *norR*ΔGAF	(D'Autreaux *et al.* 2005)
pNH8	carries the *nifH* promoter fragment from pSMM8 in the transcription vector pTE103	(Buck *et al.* 1986)
pNPTprom	pUC19 encoding the 361bp fragment of the *norR-norVW* intergenic region between *Eco* RI and *Bam* HI sites	(D'Autreaux *et al.* 2005)
pNPTprom2	pNPTprom with the 66bp sequence that contains the 3 NorR binding sites deleted (295bp intergenic region)	This work
pMJB1	pNorR with C496T and G1341C substitutions in the *norR* sequence to create the *Mfe* I/*Mun* I and *Sst* II/*Sac* II restriction sites	This work
pMJB2	pMJB1 with Q304E substitution in the AAA+ domain	This work
pMJB3	pMJB1 with L295S substitution in the AAA+ domain	This work
pMJB4	pMJB1 with P248L substitution in the AAA+ domain	This work
pMJB5	pMJB1 with V251M substitution in the AAA+ domain	This work
pMJB6	pMJB1 with S292L substitution in the AAA+ domain	This work
pMJB7	pMJB1 with E276G substitution in the AAA+ domain	This work
pMJB8	pMJB1 with G266S substitution in the AAA+ domain	This work
pMJB9	pMJB1 with F264Y substitution in the AAA+ domain	This work
pMJB10	pMJB1 with L256F substitution in the AAA+ domain	This work
pMJB11	pMJB1 with G266D substitution in the AAA+ domain	This work
pMJB12	pMJB1 with K226E substitution in the AAA+ domain	This work
pMJB13	pMJB1 with E249K substitution in the AAA+ domain	This work
pMJB14	pMJB1 with K274R substitution in the AAA+ domain	This work
pMJB15	pMJB1 with Y305C substitution in the AAA+ domain	This work
pMJB16	pMJB1 with Y305N substitution in the AAA+ domain	This work
pMJB17	pMJB1 with V323A substitution in the AAA+ domain	This work
pMJB18	pMJB1 with H346Y substitution in the AAA+ domain	This work
pMJB19	pMJB1 with S349N substitution in the AAA+ domain	This work
pMJB20	pMJB1 with E276A substitution in the AAA+ domain	This work
pMJB21	pMJB1 with E276K substitution in the AAA+ domain	This work
pMJB22	pMJB1 with E276R substitution in the AAA+ domain	This work
pMJB23	pMJB1 with E276H substitution in the AAA+ domain	This work
pMJB24	pMJB1 with R310A substitution in the AAA+ domain	This work
pMJB25	pMJB1 with R310E substitution in the AAA+ domain	This work
pMJB26	pMJB1 with G266S substitution in the AAA+ domain	This work
pMJB27	pMJB1 with G266C substitution in the AAA+ domain	This work
pMJB28	pMJB1 with G266N substitution in the AAA+ domain	This work
pMJB29	pMJB1 with G266K substitution in the AAA+ domain	This work
pMJB30	pMJB1 with G266R substitution in the AAA+ domain	This work
pMJB31	pMJB1 with G266P substitution in the AAA+ domain	This work
pMJB32	pMJB1 with G266A substitution in the AAA+ domain	This work
pMJB33	pMJB1 with G266E substitution in the AAA+ domain	This work
pMJB34	pMJB1 with G266L substitution in the AAA+ domain	This work
pMJB35	pMJB1 with G266Q substitution in the AAA+ domain	This work
pMJB36	pMJB1 with G266V substitution in the AAA+ domain	This work
pMJB37	pMJB1 with G266T substitution in the AAA+ domain	This work
pMJB38	pMJB1 with G266H substitution in the AAA+ domain	This work
pMJB39	pMJB1 with G266I substitution in the AAA+ domain	This work
pMJB40	pMJB1 with G266M substitution in the AAA+ domain	This work
pMJB41	pMJB1 with G266W substitution in the AAA+ domain	This work
pMJB42	pMJB1 with G266Y substitution in the AAA+ domain	This work
pMJB43	pMJB1 with G266F substitution in the AAA+ domain	This work
pMJB44	pNorRΔGAF with Q304E substitution in the AAA+ domain	This work
pMJB45	pNorRΔGAF with G266D substitution in the AAA+ domain	This work
pMJB46	pNorRΔGAF with G266N substitution in the AAA+ domain	This work
pMJB47	pNorRΔGAF with F264Y substitution in the AAA+ domain	This work
pMJB48	pNorRΔGAF with E276G substitution in the AAA+ domain	This work
pMJB49	pNorRΔGAF with V251M substitution in the AAA+ domain	This work
pMJB50	pNorRΔGAF with S296L substitution in the AAA+ domain	This work

pMJB51	pNorRΔGAF with P248L substitution in the AAA+ domain	This work
pMJB52	pNorRΔGAF with L295S substitution in the AAA+ domain	This work
pMJB53	pNorRΔGAF with L256F substitution in the AAA+ domain	This work
pMJB54	pETNdeM11 expressing pMJB1-derived NorR with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB55	pETNdeM11 expressing NorRΔGAF with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB56	pETNdeM11 expressing G266D with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB57	pETNdeM11 expressing G266DΔGAF with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB58	pETNdeM11 expressing Q304E with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB59	pETNdeM11 expressing Q304EΔGAF with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB60	pETNdeM11 expressing G266N with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB61	pETNdeM11 expressing G266NΔGAF with 2.0kb omega cassette from pHP45Ω inserted at SmaI site to allow selection using Sm/Spc	This work
pMJB62	pMJB1 with D286A substitution in the AAA+ domain	This work
pMJB63	pNorR with R75K substitution	(Tucker *et al.* 2007)
pMJB64	pNorR with Y98L substitution	(Tucker *et al.* 2007)
pMJB65	pNorR with C113S substitution	(Tucker *et al.* 2007)
pMJB66	pNorR with H111Y substitution	(Tucker *et al.* 2007)
pMJB67	pNorR with D131A substitution	(Tucker *et al.* 2007)
pMJB68	pNorR with H111L substitution	(Tucker *et al.* 2007)
pMJB69	pNorR with C113G substitution	(Tucker *et al.* 2007)
pMJB70	pNorR with D96A substitution	(Tucker *et al.* 2007)
pMJB71	pNorR with D99A substitution	(Tucker *et al.* 2007)
pMJB72	pNorR with R81L substitution	(Tucker *et al.* 2007)
pMJB73	pMJB11 with additional H111L substitution in the GAF domain	This work
pMJB74	pMJB11 with additional R81L substitution in the GAF domain	This work
pMJB75	pMJB11 with additional D131A substitution in the GAF domain	This work
pMJB76	pMJB11 with additional D96A substitution in the GAF domain	This work
pMJB77	pMJB11 with additional C113G substitution in the GAF domain	This work
pMJB78	pMJB11 with additional C113S substitution in the GAF domain	This work
pMJB79	pMJB11 with additional H111Y substitution in the GAF domain	This work
pMJB80	pMJB11 with additional D99A substitution in the GAF domain	This work
pMJB81	pMJB11 with additional R75K substitution in the GAF domain	This work
pMJB82	pNorRΔGAF with D286A substitution in the AAA+ domain	This work
pMJB83	pMJB1 with G266D and D286A substitutions in the AAA+ domain	This work
pMJB84	pMJB1 with N243A substitution in the AAA+ domain	This work
pMJB85	pMJB1 with N243S substitution in the AAA+ domain	This work
pMJB86	pMJB1 with R81G substitution in the GAF domain	This work
pMJB87	pMJB1 with R81D substitution in the GAF domain	This work
pMJB88	pMJB1 with R81S substitution in the GAF domain	This work
pMJB89	pMJB1 with R81T substitution in the GAF domain	This work
pMJB90	pMJB1 with R81E substitution in the GAF domain	This work
pMJB91	pMJB1 with R81N substitution in the GAF domain	This work
pMJB92	pMJB1 with R81Q substitution in the GAF domain	This work
pMJB93	pMJB1 with R81H substitution in the GAF domain	This work
pMJB94	pMJB1 with R81F substitution in the GAF domain	This work
pMJB95	pMJB1 with R81A substitution in the GAF domain	This work
pMJB96	pMJB1 with R81Y substitution in the GAF domain	This work
pMJB97	pMJB1 with R81M substitution in the GAF domain	This work
pMJB98	pMJB1 with R81L substitution in the GAF domain	This work
pMJB99	pMJB1 with R81I substitution in the GAF domain	This work
pMJB100	pMJB1 with R81W substitution in the GAF domain	This work
pMJB101	pMJB1 with R81V substitution in the GAF domain	This work
pMJB102	pMJB1 with R81C substitution in the GAF domain	This work
pMJB103	pMJB1 with R81K substitution in the GAF domain	This work
pMJB104	pMJB1 with R81P substitution in the GAF domain	This work
pMJB105	pMJB7 with additional R81L substitution in the GAF domain	This work
pMJB106	pMJB5 with additional R81L substitution in the GAF domain	This work
pMJB107	pMJB8 with additional R81L substitution in the GAF domain	This work
pMJB108	pMJB3 with additional R81L substitution in the GAF domain	This work
pMJB109	pMJB10 with additional R81L substitution in the GAF domain	This work
pMJB110	pMJB9 with additional R81L substitution in the GAF domain	This work
pMJB111	pMJB2 with additional R81L substitution in the GAF domain	This work
pMJB112	pMJB4 with additional R81L substitution in the GAF domain	This work
pMJB113	pMJB28 with additional R81L substitution in the GAF domain	This work
pMJB114	pMJB6 with additional R81L substitution in the GAF domain	This work

pMJB115	pMJB2 with additional R81A substitution in the GAF domain	This work
pMJB116	pMJB11 with additional R81A substitution in the GAF domain	This work
pMJB117	pMJB11 with additional R81V substitution in the GAF domain	This work
pMJB118	pMJB11 with additional R81I substitution in the GAF domain	This work
pMJB119	pMJB11 with additional R81F substitution in the GAF domain	This work
pMJB120	pMJB11 with additional R81D substitution in the GAF domain	This work
pMJB121	pMJB11 with additional R81N substitution in the GAF domain	This work
pMJB122	pMJB11 with additional R81K substitution in the GAF domain	This work
pMJB123	pMJB1 encoding a truncated form of NorR with residues 444-504 deleted	This work
pMJB124	pMJB1 encoding a truncated form of NorR with residues 442-504 deleted	This work
pMJB125	pMJB1 encoding a truncated form of NorR with residues 436-504 deleted	This work
pMJB126	pMJB11 encoding a truncated form of NorR with residues 444-504 deleted	This work
pMJB127	pMJB11 encoding a truncated form of NorR with residues 442-504 deleted	This work
pMJB128	pMJB11 encoding a truncated form of NorR with residues 436-504 deleted	This work
pMJB129	pMJB1 with Q304A substitution in the AAA+ domain	This work
pMJB130	pMJB1 with Q304D substitution in the AAA+ domain	This work
pMJB131	pMJB1 with Q304N substitution in the AAA+ domain	This work
pMJB132	pMJB1 with Q304R substitution in the AAA+ domain	This work
pMJB133	pMJB1 with Q304P substitution in the AAA+ domain	This work
pMJB134	pMJB2 with additional H111Y substitution in the GAF domain	This work
pMJB135	pMJB2 with additional R75K substitution in the GAF domain	This work
pMJB136	pMJB2 with additional C113S substitution in the GAF domain	This work
pMJB137	pMJB2 with additional D99A substitution in the GAF domain	This work
pMJB138	pMJB2 with additional Y98L substitution in the GAF domain	This work
pMJB139	pMJB2 with additional D286A substitution in the AAA+ domain	This work
pMJB140	pMJB2 encoding a truncated form of NorR with residues 444-504 deleted	This work
pMJB141	pMJB2 encoding a truncated form of NorR with residues 442-504 deleted	This work
pMJB142	pMJB2 encoding a truncated form of NorR with residues 436-504 deleted	This work
pMJB143	pETNdeM11 expressing pMJB1-derived NorR	This work
pMJB144	pETNdeM11 expressing NorRΔGAF	This work
pMJB145	pETNdeM11 expressing the pMJB11-derived NorR variant G266D	This work
pMJB146	pETNdeM11 expressing the pMJB45-derived NorR variant G266DΔGAF	This work
pMJB147	pETNdeM11 expressing the pMJB28-derived NorR variant G266N	This work
pMJB148	pETNdeM11 expressing the pMJB46-derived NorR variant G266NΔGAF	This work
pMJB149	pETNdeM11 expressing the pMJB2-derived NorR variant Q304E	This work
pMJB150	pETNdeM11 expressing the pMJB44-derived NorR variant Q304EΔGAF	This work
pMJB151	pETNdeM11 expressing the pMJB62-derived NorR variant D286A	This work
pMJB152	pETNdeM11 expressing the pMJB82-derived NorR variant D286AΔGAF	This work
pMJB153	pETNdeM11 expressing the pMJB83-derived NorR variant G266D-D286A	This work
pMJB154	pETNdeM11 expressing the pMJB45-derived NorR G266DΔGAF with the additional D286A substitution	This work
pMJB155	pETNdeM11 expressing the pMJB139-derived NorR variant D286A-Q304E	This work
pMJB156	pETNdeM11 expressing the pMJB44-derived NorR variant Q304E with the additional D286A substitution	This work
pMJB157	pETNdeM11 expressing the pMJB136-derived NorR varitant C113S-Q304E	This work
pES6	Plasmid encoding wild-type *K.pneumoniae* σ^N with a C-terminal hexahistidine tag (based on pRJ21)	(Southern and Merrick 2000)
pEE2003	Plasmid encoding IHF (with His tag on the β subunit)	(Yu and Haggard-Ljungquist 1993)

12.1.3 Appendix - Mutagenic primers used in this work

Primer Name	Sequence	Description/Function	Reference
Engineering pMJB1 plasmid			
C496T Fwd	AGCAATGCGTTGCTGATTGAAC AATTGGAAAGCCAGAA	Forward mutagenic primer to engineer the *Mfe* I/*Mun* I site outside of the AAA+ domain in NorR	This work
C496T Rev	TTCTGGCTTTCCATTTGTTCAAT CAGCAACGCATTGCT	Reverse mutagenic primer to engineer the *Mfe* I/*Mun* I site outside of the AAA+ domain in NorR	This work
G1341C Fwd	CCAGAAGTGGCCGCGGTGCCC GT	Forward mutagenic primer to engineer the *Sst* II/*Sac* II site outside of the AAA+ domain in NorR	This work
G1341C Rev	ACGGGCACCGCGGCCACTTCT GG	Reverse mutagenic primer to engineer the *Sst* II/*Sac* II site outside of the AAA+ domain in NorR	This work
SDM of GAFTGA residue G266			
G266A Fwd	AGCGTTTACTGCCGCTATCAG	Forward mutagenic primer for the G266A change in the AAA+ domain of NorR	This work
G266A Rev	CTGATAGCGGCAGTAAACGCT	Reverse mutagenic primer for the G266A change in the AAA+ domain of NorR	This work
G266E Fwd	AGCGTTTACTGAGGCTATCAG	Forward mutagenic primer for the G266E change in the AAA+ domain of NorR	This work
G266E Rev	CTGATAGCCTCAGTAAACGCT	Reverse mutagenic primer for the G266E change in the AAA+ domain of NorR	This work
G266P Fwd	AGCGTTTACTCCCGCTATCAG	Forward mutagenic primer for the G266P change in the AAA+ domain of NorR	This work
G266P Rev	CTGATAGCGGGAGTAAACGCT	Reverse mutagenic primer for the G266P change in the AAA+ domain of NorR	This work
G266R Fwd	AGCGTTTACTCGCGCTATCAG	Forward mutagenic primer for the G266P change in the AAA+ domain of NorR	This work
G266R Rev	CTGATAGCGCGAGTAAACGCT	Reverse mutagenic primer for the G266R change in the AAA+ domain of NorR	This work
G266K Fwd	AGCGTTTACTAAGGCTATCAG	Forward mutagenic primer for the G266R change in the AAA+ domain of NorR	This work
G266K Rev	CTGATAGCCTTAGTAAACGCT	Reverse mutagenic primer for the G266K change in the AAA+ domain of NorR	This work
G266C Fwd	AGCGTTTACTTGCGCTATCAG	Forward mutagenic primer for the G266C change in the AAA+ domain of NorR	This work
G266C Rev	CTGATAGCGCAAGTAAACGCT	Reverse mutagenic primer for the G266C change in the AAA+ domain of NorR	This work
G266N Fwd	GAGCGTTTACTAACGCTATCA	Forward mutagenic primer for the G266N change in the AAA+ domain of NorR	This work
G266N Rev	CTGATAGCGTTAGTAAACGCT	Reverse mutagenic primer for the G266N change in the AAA+ domain of NorR	This work
G266D Fwd	AGCGTTTACTGACGCTATCAG	Forward mutagenic primer for the G266D change in the AAA+ domain of NorR	This work
G266D Rev	CTGATAGCGTCAGTAAACGCT	Reverse mutagenic primer for the G266D change in the AAA+ domain of NorR	This work
G266Q Fwd	AGCGTTTACTCAGGCTATCAG	Forward mutagenic primer for the G266Q change in the AAA+ domain of NorR	This work
G266Q Rev	CTGATAGCCTGAGTAAACGCT	Reverse mutagenic primer for the G266Q change in the AAA+ domain of NorR	This work
G266L Fwd	AGCGTTTACTCTCGCTATCAG	Forward mutagenic primer for the G266L change in the AAA+ domain of NorR	This work
G266L Rev	CTGATAGCGAGAGTAAACGCT	Reverse mutagenic primer for the G266L change in the AAA+ domain of NorR	This work
G266V Fwd	AGCGTTTACTGTCGCTATCAG	Forward mutagenic primer for the G266V change in the AAA+ domain of NorR	This work
G266V Rev	CTGATAGCGACAGTAAACGCT	Reverse mutagenic primer for the G266V change in the AAA+ domain of NorR	This work
G266S Fwd	AGCGTTTACTTCCGCTATCAG	Forward mutagenic primer for the G266S change in the AAA+ domain of NorR	This work
G266S Rev	CTGATAGCGGAAGTAAACGCT	Reverse mutagenic primer for the G266S change in the AAA+ domain of NorR	This work
G266T Fwd	AGCGTTTACTACCGCTATCAG	Forward mutagenic primer for the G266T change in the AAA+ domain of NorR	This work
G266T Rev	CTGATAGCGGTAGTAAACGCT	Reverse mutagenic primer for the G266T change in the AAA+ domain of NorR	This work
G266H Fwd	AGCGTTTACTCACGCTATCAG	Forward mutagenic primer for the G266H change in the AAA+ domain of NorR	This work
G266H Rev	CTGATAGCGTGAGTAAACGCT	Reverse mutagenic primer for the G266H change in the AAA+ domain of NorR	This work
G266I Fwd	AGCGTTTACTATCGCTATCAG	Forward mutagenic primer for the G266I change in the AAA+ domain of NorR	This work

G266I Rev	CTGATAGCGATAGTAAACGCT	Reverse mutagenic primer for the G266I change in the AAA+ domain of NorR	This work
G266M Fwd	AGCGTTTACTATGGCTATCAG	Reverse mutagenic primer for the G266M change in the AAA+ domain of NorR	This work
G266M Rev	CTGATAGCCATAGTAACGCT	Forward mutagenic primer for the G266M change in the AAA+ domain of NorR	This work
G266F Fwd	AGCGTTTACTTTCGCTATCAG	Reverse mutagenic primer for the G266F change in the AAA+ domain of NorR	This work
G266F Rev	CTGATAGCGAAAGTAAACGCT	Forward mutagenic primer for the G266F change in the AAA+ domain of NorR	This work
G266W Fwd	AGCGTTTACTTGGGCTATCAG	Reverse mutagenic primer for the G266W change in the AAA+ domain of NorR	This work
G266W Rev	CTGATAGCCCAAGTAAACGCT	Forward mutagenic primer for the G266W change in the AAA+ domain of NorR	This work
G266Y Fwd	AGCGTTTACTTACGCTATCAG	Reverse mutagenic primer for the G266Y change in the AAA+ domain of NorR	This work
G266Y Rev	CTGATAGCGTAAGTAAACGCT	Forward mutagenic primer for the G266Y change in the AAA+ domain of NorR	This work
SDM of key AAA+ residues			
P248L Fwd	CTGTGCTGCACTGCTGGAAA	Forward mutagenic primer for the P248L change in the AAA+ domain of NorR	This work
P248L Rev	TTTCCAGCAGTGCAGCACAG	Reverse mutagenic primer for the P248L change in the AAA+ domain of NorR	This work
V251M Fwd	CCGGAAAGTATGGCGGAAAG	Forward mutagenic primer for the V251M change in the AAA+ domain of NorR	This work
V251M Rev	CTTTCCGCCATACTTTCCGG	Reverse mutagenic primer for the V251M change in the AAA+ domain of NorR	This work
L256F Fwd	AGTGAGTTCTTCGGGCATGT	Forward mutagenic primer for the L256F change in the AAA+ domain of NorR	This work
L256F Rev	ACATGCCCGAAGAACTCACT	Reverse mutagenic primer for the L256F change in the AAA+ domain of NorR	This work
F264Y Fwd	AAGGAGCGTATACTGGCGCTA	Forward mutagenic primer for the F264Y change in the AAA+ domain of NorR	This work
F264Y Rev	TAGCGCCAGTATACGCTCCTT	Reverse mutagenic primer for the F264Y change in the AAA+ domain of NorR	This work
E276G Fwd	GGGAAGTTTGGAATGGCGGAT	Forward mutagenic primer for the E276G change in the AAA+ domain of NorR	This work
E276G Rev	ATCCGCCATTCCAAACTTCCC	Reverse mutagenic primer for the E276G change in the AAA+ domain of NorR	This work
S292L Fwd	GCGAGTTGTTGTTGGCATTG	Forward mutagenic primer for the S292L change in the AAA+ domain of NorR	This work
S292L Rev	CAATGCAACAACAACTCGC	Reverse mutagenic primer for the S292L change in the AAA+ domain of NorR	This work
L295S Fwd	TTGGCATCGCAGGCCAAGCT	Forward mutagenic primer for the L295S change in the AAA+ domain of NorR	This work
L295S Rev	AGCTTGGCCTGCGATGCCAA	Reverse mutagenic primer for the L295S change in the AAA+ domain of NorR	This work
Q304E Fwd	AGGGTGTTGGAGTATGGCGAT	Forward mutagenic primer for the Q304E change in the AAA+ domain of NorR	This work
Q304E Rev	ATCGCCATACTCCAACACCCT	Reverse mutagenic primer for the Q304E change in the AAA+ domain of NorR	This work
SDM of R81 residue			
R81A Fwd	TGGAAGCGATTGCCGCCGCCG	Forward mutagenic primer for the R81A change in the GAF domain of NorR	This work
R81A Rev	CGGCGGCGGCAATCGCTTCCA	Reverse mutagenic primer for the R81A change in the GAF domain of NorR	This work
R81P Fwd	TGGAGCGATTGCCCCCGCC	Forward mutagenic primer for the R81P change in the GAF domain of NorR	This work
R81P Rev	GGCGGGGGCAATCGCTTCCA	Reverse mutagenic primer for the R81P change in the GAF domain of NorR	This work
R81V Fwd	TGGAAGCGATTGCCGTCGCC	Forward mutagenic primer for the R81V change in the GAF domain of NorR	This work
R81V Rev	GGCGACGGCAATCGCTTCCA	Reverse mutagenic primer for the R81V change in the GAF domain of NorR	This work
R81I Fwd	TGGAAGCGATTGCCATCGCC	Forward mutagenic primer for the R81I change in the GAF domain of NorR	This work
R81I Rev	GGCGATGGCAATCGCTTCCA	Reverse mutagenic primer for the R81I change in the GAF domain of NorR	This work
R81F Fwd	TGGAAGCGATTGCCTTCGCC	Forward mutagenic primer for the R81F change in the GAF domain of NorR	This work
R81F Rev	GGCGAAGGCAATCGCTTCCA	Reverse mutagenic primer for the R81F change in the GAF domain of NorR	This work

R81D Fwd	TGGAAGCGATTGCCGACGCC	Forward mutagenic primer for the R81D change in the GAF domain of NorR	This work
R81D Rev	GGCGTCGGCAATCGCTTCCA	Reverse mutagenic primer for the R81D change in the GAF domain of NorR	This work
R81N Fwd	TGGAAGCGATTGCCAACGCC	Forward mutagenic primer for the R81N change in the GAF domain of NorR	This work
R81N Rev	GGCGTTGGCAATCGCTTCCA	Reverse mutagenic primer for the R81N change in the GAF domain of NorR	This work
R81S Fwd	TGGAAGCGATTGCCTCCGCC	Forward mutagenic primer for the R81S change in the GAF domain of NorR	This work
R81S Rev	GGCGGAGGCAATCGCTTCCA	Reverse mutagenic primer for the R81S change in the GAF domain of NorR	This work
R81K Fwd	TGGAAGCGATTGCCAAAGCC	Forward mutagenic primer for the R81K change in the GAF domain of NorR	This work
R81K Rev	GGCTTTGGCAATCGCTTCCA	Reverse mutagenic primer for the R81K change in the GAF domain of NorR	This work
R81E Fwd	TGGAAGCGATTGCCGAAGCC	Forward mutagenic primer for the R81E change in the GAF domain of NorR	This work
R81E Rev	GGCTTCGGCAATCGCTTCCA	Reverse mutagenic primer for the R81E change in the GAF domain of NorR	This work
R81C Fwd	TGGAAGCGATTGCCTGCGCC	Forward mutagenic primer for the R81C change in the GAF domain of NorR	This work
R81C Rev	GGCGCAGGCAATCGCTTCCA	Reverse mutagenic primer for the R81C change in the GAF domain of NorR	This work
SDM of Q304 residue			
Q304A Fwd	AGGGTGTTGGCGTATGGCGAT	Forward mutagenic primer for the Q304A change in the AAA+ domain of NorR	This work
Q304A Rev	ATCGCCATACGCCAACACCCT	Reverse mutagenic primer for the Q304A change in the AAA+ domain of NorR	This work
Q304D Fwd	AGGGTGTTGGAGTATGGCGAT	Forward mutagenic primer for the Q304D change in the AAA+ domain of NorR	This work
Q304D Rev	ATCGCCATACTCCAACACCCT	Reverse mutagenic primer for the Q304D change in the AAA+ domain of NorR	This work
Q304N Fwd	AGGGTGTTGAACTATGGCGAT	Forward mutagenic primer for the Q304N change in the AAA+ domain of NorR	This work
Q304N Rev	ATCGCCATAGTTCAACACCCT	Reverse mutagenic primer for the Q304N change in the AAA+ domain of NorR	This work
Q304R Fwd	AGGGTGTTGCGGTATGGCGAT	Forward mutagenic primer for the Q304R change in the AAA+ domain of NorR	This work
Q304R Rev	ATCGCCATACCGCAACACCCT	Reverse mutagenic primer for the Q304R change in the AAA+ domain of NorR	This work
SDM of E276G residue			
E276H Fwd	GGGAAGTTTCACATGGCGGAT	Forward mutagenic primer for the E276H change in the AAA+ domain of NorR	This work
E276H Rev	ATCCGCCATGTGAAACTTCCC	Reverse mutagenic primer for the E276H change in the AAA+ domain of NorR	This work
E276K Fwd	GCGGGAAGTTTAAAATGGCGGA	Forward mutagenic primer for the E276K change in the AAA+ domain of NorR	This work
E276K Rev	TCCGCCATTTTAAACTTCCCGC	Reverse mutagenic primer for the E276K change in the AAA+ domain of NorR	This work
E276R Fwd	GGGAAGTTTCGAATGGCGGAT	Forward mutagenic primer for the E276R change in the AAA+ domain of NorR	This work
E276R Rev	ATCCGCCATTCGAAACTTCCC	Reverse mutagenic primer for the E276R change in the AAA+ domain of NorR	This work
E276A Fwd	GGGAAGTTTGCAATGGCGGAT	Forward mutagenic primer for the E276A change in the AAA+ domain of NorR	This work
E276A Rev	GGGAAGTTTGCAATGGCGGAT	Reverse mutagenic primer for the E276A change in the AAA+ domain of NorR	This work
SDM of residues involved in nucleotide hydrolysis cycle			
R310E Fwd	ATTCAGGAGGTTGGCGATGAC	Forward mutagenic primer for the R310E change in the AAA+ domain of NorR	This work
R310E Rev	GTCATCGCCAACCTCCTGAAT	Reverse mutagenic primer for the R310E change in the AAA+ domain of NorR	This work
R310A Fwd	ATTCAGGCCGTTGGCGATGAC	Forward mutagenic primer for the R310A change in the AAA+ domain of NorR	This work
R310A Rev	GTCATCGCCAACGGCCTGAAT	Reverse mutagenic primer for the R310A change in the AAA+ domain of NorR	This work
N243A Fwd	GCTGGTCTATCTCGCCTGTG	Forward mutagenic primer for the N243A change in the AAA+ domain of NorR	This work
N243A Rev	CACAGGCGAGATAGACCAGC	Reverse mutagenic primer for the N243A change in the AAA+ domain of NorR	This work

N243S Fwd	GCTGGTCTATCTCAGCTGTG	Forward mutagenic primer for the N243S change in the AAA+ domain of NorR	This work
N243S Rev	CACAGCTGAGATAGACCAGC	Reverse mutagenic primer for the N243S change in the AAA+ domain of NorR	This work
DNA binding and Open Promoter Complex (OPC) assays			
NorRprom Fwd	GGCGATATTCGCCAGCACAT	Forward primer used to amplify a 266bp region of the *norR-norVW* intergenic region	This work
NorRprom Rev	CGTTGACCAACCCAATGAATGT	Reverse primer used to amplify a 266bp region of the *norR-norVW* intergenic region	This work
pNPTprpm 66bp_DEL Fwd	ATCTTTGCCATTAATGGGCATA ATTTT	Forward primer used in the deletion mutagenesis of the 66bp region of the *norR-norVW* intergenic region	This work
pNPTprpm 66bp_DEL Rev	CCCATTAATGGCAAAGATGAG TTTTTC	Reverse primer used in the deletion mutagenesis of the 66bp region of the *norR-norVW* intergenic region	This work
SDM of Walker B motif			
D286A Fwd	TGTTTCTGGCTGAGATCGGC	Forward mutagenic primer for the D286A change in the AAA+ domain of NorR	This work
D286A Rev	GCCGATCTCAGCCAGAAACA	Reverse mutagenic primer for the D286A change in the AAA+ domain of NorR	This work
Deletion of Helix-Turn-Helix motif			
Δ444-504 Rev	CGAAGGATCCTTACGTCGGCA ACGTCACCTCA	reverse mutagenic primer for deletion of C-T of NorR sequence that includes HTH motif	This work
Δ442-504 Rev	CGAAGGATCCTTACAACGTCA CTTCAGGAAAAGC	reverse mutagenic primer for deletion of C-T of NorR sequence that includes HTH motif	This work
Δ436-504 Rev	CGAAGGATCCTTAAGCAAAAT GTTGCGCCTCAAG	reverse mutagenic primer for deletion of C-T of NorR sequence that includes HTH motif	This work

12.1.4 Appendix - External Mutagenic and sequencing primers used in this work

Primer Name	Sequence	Description/Function	Reference
T7long	AAATTAATACGACTCACTATAGGGG	Universal primer that anneals outside the NorR sequence in pET21a	This work
T7term-long	TATGCTAGTTATTGCTCAGCGGT	Universal primer that anneals outside the NorR sequence in pET21a	This work
AAA+ Fwd	GAAGAGCTACGGCTGATTGC	Forward primer flanking the NorR AAA+ domain	This work
AAA+ Rev	GAACGCTTCTGTCGCTTCAC	Reverse primer flanking the NorR AAA+ domain	This work
GAF Rev	CCTGGCAGCATATTCTGGCT	Reverse primer flanking the NorR GAF domain	This work
Km1	TGCCTCTTCCGACCATCAAG	Forward flanking primer to the *Sma* I site in the Km resistance cassette within pET24d+ (pETM11)	This work
Km2	AACACTGCCAGCGCATCAAC	Reverse flanking primer to the *Sma* I site in the Km resistance cassette within pET24d+ (pETM11)	This work
Km3	CAGCCATTACGCTCGTCATCAAA	Alternative reverse flanking primer to the SmaI site in the Km resistance cassette within pET24d+ (pETM11)	This work
M13 Fwd	CGTTGTAAAACGACGGCCAGTG	Forward universal primer that anneals outside of the *norR-norVW* intergenic region in pNPTprom	This work
M13 Rev	GAGCGGATAACAATTTCACACAGG	Reverse universal primer that anneals outside of the *norR-norVW* intergenic region in pNPTprom	This work
GAF1	TTTATTCCGCTTGCCATCGAC	NorR GAF domain internal sequencing primer	This work
GAF2	AGGCGTGAACCTTCAGACTCT	NorR GAF domain internal sequencing primer	This work
Seq1	GCCCGATCAGTTCGAT	NorR internal sequencing primer	This work
Seq2	CCGTTGTTTGCGGGTC	NorR internal sequencing primer	This work
Seq3	TCGCCCTGCCAGCACC	NorR internal sequencing primer	This work
Seq4	ATTGCTTAACGCTCCC	NorR internal sequencing primer	This work

12.2 Appendix - Publications

Essential roles of three enhancer sites in σ⁵⁴-dependent transcription by the nitric oxide sensing regulatory protein NorR

Nicholas P. Tucker[1], Tamaswati Ghosh[2], Matthew Bush[1], Xiaodong Zhang[2],* and Ray Dixon[1],*

[1]Department of Molecular Microbiology, John Innes Center, Colney, Norwich, NR4 7UH and [2]Division of Molecular Bioscience, Imperial College London, London, SW7 2AZ UK

ABSTRACT

The bacterial activator protein NorR binds to enhancer-like elements, upstream of the promoter site, and activates σ⁵⁴-dependent transcription of genes that encode nitric oxide detoxifying enzymes (NorVW), in response to NO stress. Unique to the *norVW* promoter in *Escherichia coli* is the presence of three enhancer sites associated with a binding site for σ⁵⁴-RNA polymerase. Here we show that all three sites are required for NorR-dependent catalysis of open complex formation by σ⁵⁴-RNAP holoenzyme (Eσ⁵⁴). We demonstrate that this is essentially due to the need for all three enhancers for maximal ATPase activity of NorR, energy from which is used to remodel the closed Eσ⁵⁴ complex and allow melting of the promoter DNA. We also find that site-specific DNA binding *per se* promotes oligomerisation but the DNA flanking the three sites is needed to further stabilise the functional higher order oligomer of NorR at the enhancers.

INTRODUCTION

Nitric oxide (NO) is a highly reactive radical species that is toxic to micro-organisms. It can be encountered exogenously, as a consequence of pathogen invasion, or endogenously through the process of respiratory denitrification (1,2). *Escherichia coli* is known to possess at least three enzymes capable of directly detoxifying NO, by utilising either NO reductase or NO dioxygenase activities (3–5). One of these systems comprises the enzyme flavorubredoxin (encoded by *norV*) and its associated NADH oxidoreductase (encoded by *norW*), which reduces the NO radical to nitrous oxide under anaerobic conditions (4,6).

Immediately upstream of the *norVW* genes is the divergently transcribed *norR* gene, which encodes an NO sensing σ⁵⁴-dependent transcriptional regulator (7,8). This protein is essential for activating transcription of the *norVW* operon and shares ~40% sequence homology with *Ralstonia eutropha* NorR, which is responsible for the transcriptional regulation of heme b_3-iron NO reductase in response to NO (9). NorR has a modular domain architecture typical of bacterial σ⁵⁴-dependent enhancer binding proteins (bEBPs) (10) and consists of three key domains: an N-terminal regulatory GAF (for cGMP-specific and cGMP-regulated cyclic nucleotide phosphodiesterase, *Anabaena* Adenylyl cyclase and *E. coli* transcription factor FhlA) domain containing a mononuclear ferrous iron centre that detects NO (11,12), a central AAA+ (for ATPase associated with various cellular activities) domain that interacts with σ⁵⁴ and couples ATP hydrolysis to promoter DNA melting by RNA polymerase, and a C-terminal helix-turn-helix DNA-binding domain.

Previously we reported that *E. coli* NorR binds to three sites upstream of the *norV* promoter that contain inverted repeats with core consensus GT-(N7)-AC (13). These NorR-binding sites are conserved amongst the proteobacteria and are found upstream of genes encoding NorV (e.g. *Salmonella typhimurium*), Hmp (e.g. *Pseudomonas aeruginosa*) and NorA (*Ralstonia eutropha*). In *E. coli*, integration host factor (IHF) binds to the region upstream of *norVW* between the *norV* transcription start

---

site and the NorR enhancer sites (13). In common with other σ^{54}-dependent systems, the bending of DNA induced by IHF at this location may encourage interactions between NorR and σ^{54}-RNA polymerase by DNA looping (14,15).

bEBPs function by coupling the energy yielded from ATP hydrolysis to the isomerisation of σ^{54}-RNA polymerase from the closed promoter complex to the open promoter complex that is competent for transcription initiation (16–18). Oligomerisation of the AAA+ domain of bEBPs to form hexameric rings is required for the formation of a functional ATPase (19–21) but not for ATP binding (18,22). In some cases the ATPase activity of bEBPs is not only regulated in response to environmental signals but is also responsive to interaction with specific enhancers (23–25).

We previously demonstrated that the ATPase activity of NorR is enhancer DNA-dependent *in vitro* (11). Here we report that each of the three NorR enhancer sites upstream of the *norV* promoter is essential for transcriptional activation. We demonstrate that this is a consequence of the stringent requirement for the three enhancers for maximal ATPase activity of NorR through promoting and stabilising a functional higher order oligomer of the activator. We use a GAF domain deleted form of NorR (NorRΔGAF), which can activate transcription in the absence of NO, to correlate ATP hydrolysis with higher order oligomer formation. We also suggest that the NorR oligomer formed at the three enhancers is stabilised by the wrapping of flanking DNA around the higher order oligomer identified in negative-stain electron microscopy images.

MATERIALS AND METHODS

Protein purification

E. coli NorRΔGAF was over-expressed and purified as described previously (11). $NorR_{178–452}$ AAA+ domain was generated from plasmid pNorR (13) that expresses NorR residues from 178 to 452 with a N-terminal 6-histidine tag in pET28b. The protein was purified by nickel affinity chromatography and gel filtration. The His-tag was removed by thrombin cleavage for 3 h at $23°C$ (38). Purified proteins were stored in buffer containing 100 mM Tris–HCl pH 8, 150 mM NaCl, 4 mM DTT and 5% glycerol at $-80°C$. Protein concentrations were determined by the Folin-Lowry method (26).

Construction of plasmids and site-directed mutagenesis

The pNPTfus series plasmids designed for use with *lacZ* fusion experiments contain either a wild type or a mutagenised *norR*/*norV* intergenic region cloned into the *Sma*I site of pUC19 (Table 1). Mutations were generated using a PCR based mutagenesis method (27). Wild type constructs were PCR amplified using the external primers only:

norRfusion (external) 5′-GGCGCTGAAAACGATC CTGG-3′,

norVfusion (external) 5′-TCACGCACTTCCCAGTC ACG-3′.

Mutant constructs were amplified using combinations of the external primers and the following mutagenic internal primers:

Site1− 5′-TAATGAGTAGGCAAAATGCCTATCAATC-3′,
Site1+ 5′-GATTGATAGGCATTTTGCCTACTCATTA-3′,
Site2− 5′-ATCAAATGGGCGATATGCCAATATCT-3′,
Site2+ 5′-AGATATTGGCATATCGCCCATTTGAT-3′,
Site3− 5′-ATCTATAGGCAAATTGCCAGTGAGGCAAAG-3′,
Site3+ 5′-CTTTGCCTCACTGGCAATTTGCCTATAGAT-3′.

To construct *lacZ* fusions in the *E. coli* chromosome, wild type and mutant *norR-norV* intergenic region constructs were cloned from pNPTfus series plasmids into pRS551 (28) using EcoRI and BamHI to generate *lacZ* fusions. Derivatives of pRS551 were transformed into *E. coli* strain MC1000. The *lacZ* fusion constructs were then crossed into phage λRS45 by homologous recombination and transferred into the MC1000 chromosome at the phage λ attachment site as described previously (28).

β-galactosidase assays

Derivatives of *E. coli* strain MC1000 containing either the wild type promoter or mutant promoters upstream of the *norV-lacZ* reporter (Table 1) were grown either in LB medium aerobically or in LB medium supplemented with 1% glucose when grown anaerobically. The cultures were grown to an OD_{600} of ~0.6 nm before being induced with potassium nitrite to a final concentration of 4 mM. Cultures were then grown for a further 2 h to allow for expression of the *norV-lacZ* promoter fusion construct.

Methylation protection footprinting

DNA fragments for footprinting reactions were prepared as described previously (13) but using plasmids pNPTfusV, pNPTfus1V, pNPTfus2V or pNPTfus3V. Binding reactions were carried out in DMS buffer (50 mM sodium cacodylate, 1 mM EDTA) with 0.5 μg of 5′ end-labelled EcoRI–BamHI restriction fragments and the indicated concentration of NorRΔGAF in a final volume of 200 μl. Salmon sperm DNA 2 μg was also present in reaction. Reactions were incubated for 10 min at room temperature, then 5 μl of 10% dimethyl sulphate (Sigma, in ethanol) was added to each binding reaction and incubation continued for a further 5 min. Reactions were stopped with 50 μl of ice cold DMS stop buffer (1 M β-mercaptoethanol, 1.5 M sodium acetate, 1 mg/ml glycogen) and the DNA was then precipitated by the addition of 750 μl of ice cold ethanol, followed by centrifugation at 13 000 rpm in bench-top centrifuge. After a wash with 0.3 M sodium acetate, the DNA was subjected to a second ethanol precipitation and then treated with 1 M piperidine (Sigma) for 30 min at 100°C. The samples were then lyophilised and subjected to two cycles of resuspension in 20 μl of sterile water followed by further lyophilisation. After the final lyophilisation, 10 μl of formamide loading dye was added to each

Table 1. Strains and plasmids

	Description	Reference/source
Plasmid		
pRS551	Vector for construction of *lacZ* promoter fusions.	(28)
pNorR	Derivative of pET21a expressing *norR*	(13)
pNPTfusV	pUC19 carrying the *norRV* intergenic region for *norV* promoter fusions	This work
pNPTfus1V	As pNPTfusV but NorR site 1 mutated to GG(N7)CC	This work
pNPTfus2V	As pNPTfusV but NorR site 2 mutated to GG(N7)CC	This work
pNPTfus3V	As pNPTfusV but NorR site 3 mutated to GG(N7)CC	This work
pRS551-wtV	Fusion vector constructed from pNPTfusV	This work
pRS551-1V	Fusion vector constructed from pNPTfus1V	This work
pRS551-2V	Fusion vector constructed from pNPTfus2V	This work
pRS551-3V	Fusion vector constructed from pNPTfus3V	This work
Strain		
DH10B	*mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara, leu*)7697 *galU galK* λ- *rps*L *nupG*	Invitrogen
MC1000	*araD139* Δ (*araABC-leu*)7679 *galU galK* Δ(*lac*)X74 *rpsL*	(44)
MH1003	DH10B *norR::cat* λ*norV-lacZ*	(7)
MC100V	MC1000 λ*norV-lacZ*	This work
MC101V	MC1000 λ*norV-lacZ* NorR site 1 mutant	This work
MC102V	MC1000 λ*norV-lacZ* NorR site 2 mutant	This work
MC103V	MC1000 λ*norV-lacZ* NorR site 3 mutant	This work

sample followed by resolution on a 6% polyacrylamide sequencing gel.

Coupled ATPase activity assay

For experiments with low protein concentrations, ATPase activity was measured using an assay in which production of ADP is coupled to the oxidation of NADH by lactate dehydrogenase and pyruvate kinase (29). The oxidation of NADH was monitored at 340 nm at 37°C. All reaction mixtures contained ATP (30 mM), phosphoenolpyruvate (1 mM), NADH (0.3 mM), pyruvate kinase (7 U, Roche), lactate dehydrogenase (23 U, Roche) in 50 mM Tris–HCl (pH 8.0), 100 mM KCl, 2 mM $MgCl_2$ and 300 nM of NorRΔGAF. Either wild type pNPTfusV or a plasmid carrying a mutation in one of the NorR-binding sites were added to the reaction mixtures, since NorRΔGAF ATPase activity is enhancer-DNA dependent (30). ATPase activity was measured by observing the change in absorbance at 340 nm.

Open promoter complex and band-shift assays

Template DNA for open complex assays was obtained by digesting the plasmid pNPTfusV, pNPTfus1V, pNPTfus2V or pNPTfus3V with EcoRI and BamHI to yield a DNA fragment including the *norVW* promoter and upstream activator sequences. The DNA fragments were 5′ end-labelled with ^{32}P as described above for use in gel retardation and DNA footprinting assays. Open complex formation was assayed in TAP buffer (50 mM Tris–acetate, 100 mM potassium acetate, 8 mM magnesium acetate, 3.5% polyethylene glycol 8000, 1 mM DTT, pH 7.9) and contained 1 nM template DNA, 200 nM core RNA polymerase (Epicentre Biotechnologies), 200 nM σ^{54}, 130 nM IHF, 5 mM ATP and 0.5 mM CTP. The reaction components were pre-incubated for 10 min at 30°C, and reactions were initiated by adding NorRΔGAF to a final concentration

of between 115 and 460 nM. After a further 20 min incubation at 30°C, samples were mixed with 3 μl of dye mixture containing 50% glycerol, 0.05% bromophenol blue, 0.1% xylene cyanol and 2 μg of heparin and immediately loaded onto a 4% (wt/vol) polyacrylamide gel (acrylamide/bisacrylamide ratio, 80 : 1) in 25 mM Tris–400 mM glycine, pH 8.6, which had been pre-run at 180 V at room temperature down to a constant power of 2 W. Gels were run at 150 V and were dried and exposed to autoradiograph film or a phosphorimager screen. NorRΔGAF-DNA band-shift experiments were carried out using either a ^{32}P labelled 266 bp PCR product (primer norvfus 5′-GGCGCTGAAAACGATCCTGG-3′ and primer norRpromR 5′-GGTTGACCAACCCAATG AATG-3′) or a 66 bp fragment generated by hybridisation of oligonucleotide primers spanning the NorR-binding region (5′-TCACTGTCAATTTGACTATAGATATTGT CATATCGACCATTTGATTGATAGTCATTTTGACT ACTC-3′ and its reverse complemented partner). Reactions were prepared in the same way as for the open complex assay but with NorRΔGAF and DNA only in TAP buffer. The same dilution series of a fresh preparation of NorRΔGAF was used for both experiments, which were run at the same time. Gels were quantified with a Fujix BAS1000 phosphorimager and the data was plotted using Graphpad Prism as described previously (13).

Radioactive ATPase activity assay

Reactions were performed in a 10 μl final volume in ATPase buffer (50 mM Tris–HCl pH 8.0, 50 mM NaCl, 15 mM $MgCl_2$, 0.01 mM DTT) and different concentrations of NorRΔGAF in complex with DNA, where indicated. The mix was preincubated at 23°C for 10 min and the reaction was started by adding 3 μl of an ATP mixture [1 mM ATP and 0.6 μCi/μl of [α-^{32}P]ATP (3000 Ci/mmol)] and incubated for different times at

37°C. We examined the enhancer DNA-dependency of NorRΔGAF ATPase using double-stranded (ds) DNA of three different lengths (21, 66 and 266 bp) containing different regions of the *norR-norV* intergenic region. The 21 bp fragment was generated by hybridisation of oligonucleotide primers spanning the NorR-binding site 1 (5′-GATAGTCATTTTGACTACTCA-3′ and its reverse complemented partner). Complexes tested were composed of different protein to DNA molar ratios (as indicated in figure legends). Reactions were stopped by adding five volumes of 2 M formic acid. [α-^{32}P]ADP was separated from [α-^{32}P]ATP using thin-layer chromatography (Polygram Cel 300 PEI). Radioactivity was detected by PhosphorImager (Fuji Bas-5000) and quantified using the AIDA image analyzer software, version 3.52 (Raytest, Straubenhardt, Germany). Reactions were stopped when ~20% of total ATP was hydrolysed to ensure similar proportion of ADP in all reactions. The ATPase activity is expressed in turnover per minute for all experiments. The curves were fitted using the Origin 7.0 software (OriginLab Corp.). All experiments were performed at least in triplicate. Moreover, we established (data not shown) that the rate of ATP hydrolysis was linear under assay conditions.

Analytical gel filtration

NorRΔGAF (at different concentrations) was incubated with different DNA fragments where specified (at concentrations specific to the desired molar stoichiometry of the complex) for 10 min at 23°C in buffer containing 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 15 mM MgCl$_2$ and 1 mM ATP where indicated. In the presence of ATP, samples were incubated at 4°C. Samples (100 µl) were then injected onto a Superose 6 column (10 × 300 mm, 24 ml) (GE Healthcare) installed on an AKTA system (GE Healthcare), which was pre-equilibrated with the sample buffer. Chromatography was performed at 4°C at a flow rate of 0.5 ml min^{-1}, and the column was calibrated with globular proteins: apoferritin (443 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa) and carbonic anhydrase (29 kDa). NorR$_{178-452}$ (100 µM) was filtered through a Superdex 200 column (10 × 300 mm, 24 ml; GE Healthcare) in the presence and absence of 1 mM ATP at 4°C. Chromatography conditions were similar to that of NorRΔGAF.

Negative-stain electron microscopy and image processing

Two microlitre of fractions containing the NorRΔGAF-266 bp DNA complex eluted from the gel filtration column (elution peak at 9.3 ml) was adsorbed onto glow-discharged continuous carbon grids (TAAB) and stained with 2% uranyl acetate. Data were collected at 50 000× magnification using a Phillips CM200 FEG electron microscope operating at 200 kV. Micrographs were recorded directly on a 4k × 4k CCD camera (F415 from Tietz Video and Imaging Processing GmbH), giving a pixel size of 1.76 Å. Digitised images were then coarsened by a factor of two giving a pixel size of 3.52 Å per pixel. Ten-thousand particles were picked automatically using the IMAGIC-5 software (31). Particles were

windowed into 128 by 128 pixel boxes, extracted and band-pass filtered between 170 and 20 Å. Poor-quality particles were removed before reference free alignment to a total sum of the dataset was carried out. Initial class averages were generated by classification based on multi-variate statistical analysis (MSA). Strong class-averages were then used as references for multi-reference alignment (MRA) (32,33) using selected class averages as new references. The quality of the alignment was assessed by the class averages produced and the individual aligned images in each class. Multiple iterations of MRA, MSA and classification were performed with the selected new class-averages used as references for subsequent rounds of MRA. Poor-quality particles were removed throughout the alignment procedure based on alignment shifts and visual inspection of the particles within each class. The final class averages were generated from 5000 particles which were classified into 500 classes.

RESULTS

All three NorR enhancer sites are required for activation of *norV* expression *in vivo*

In order to assess the importance of each NorR-binding site, the enhancers were individually altered from the consensus sequence GT-(N7)-AC to GG-(N7)-CC and then introduced as *norV-lacZ* promoter fusions into the *E. coli* chromosome at the phage lambda insertion site. To activate NorR, cultures grown either under aerobic or anaerobic conditions were treated with potassium nitrite for 2 h to induce endogenous NO production. β-Galactosidase assays were then performed to determine the level of *norV-lacZ* expression. In agreement with previous microarray data (34) and the observation that NorR is competent to activate transcription in the presence of oxygen, we observed almost identical levels of *norV* expression in cultures grown either in aerobic or
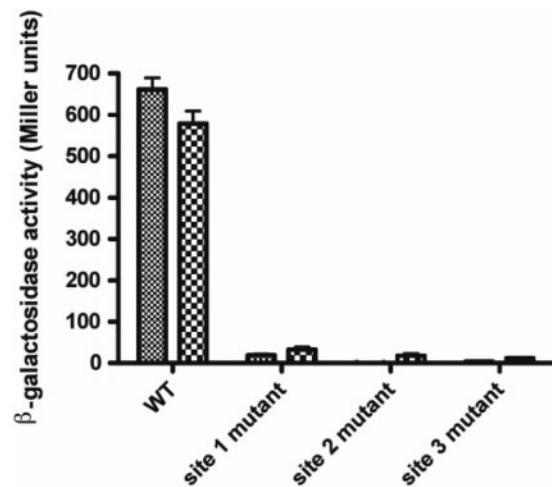
**Figure 1.** Effect of NorR-binding sites on activation of a *norV-lacZ* reporter. Strains MC100V, MC101V, MC102V and MC103V (Table 1) were grown under either aerobic or anaerobic conditions and induced with potassium nitrite (4 mM). *norV-lacZ* expression was then determined by measuring β-galactosidase activity. β-galactosidase activity was minimal in MC1000 lacking a promoter fusion construct.
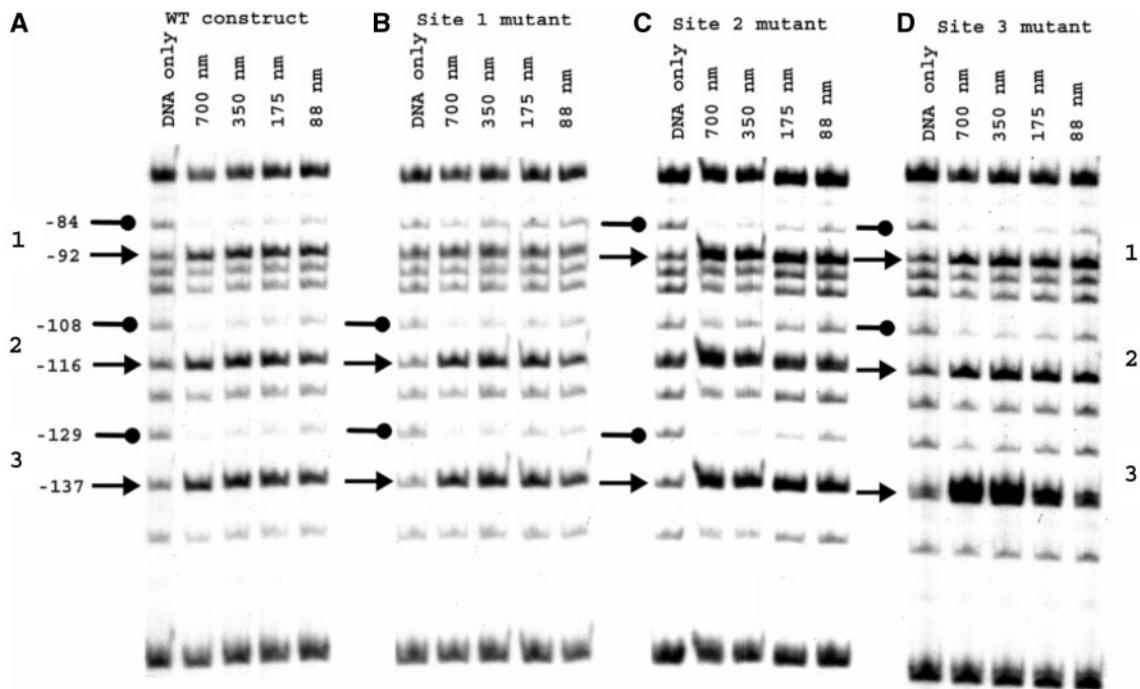
Figure 2. Methylation protection footprinting of wild type and mutant *norR-norV* intergenic region constructs with NorRΔGAF. Footprinting reactions contained ^{32}P-labelled 362-bp DNA fragments from pNPTfusV series plasmids (Table 1) spanning the *norR-norV* intergenic region 5′ end-labelled at the *Eco*RI end. Fragments encoded either the wild type promoter (**A**) or promoter constructs with NorR-binding site mutations at site 1 (**B**), site 2 (**C**) or site 3 (**D**). NorRΔGAF-DNA-binding reactions were treated with dimethyl sulphate. In each case, reactions carried out in the absence of protein are labelled 'DNA only'. Other reactions contained between 88 and 700 nM NorRΔGAF as indicated. Residues are numbered relative to the *norV* transcript start site. Protected G residues are marked with lollipops on the left-hand side of each series. Arrows denote enhanced methylation at G residues.

anaerobic conditions (Figure 1). Disruption of any one of the three NorR-binding sites was sufficient to completely prevent *norVW* expression. Thus, all three enhancer sites must be intact to facilitate NorR-dependent activation of *norVW* transcription *in vivo*.

NorR binding is disrupted at mutant, but not wild type enhancer sites *in vitro*

To determine the influence of the enhancer mutations on the interaction of NorR with the promoter, methylation protection footprinting was carried out with a 266 bp DNA fragment that contains all three NorR-binding sites. For these experiments and all subsequent biochemical assays described here, we used a truncated derivative of NorR (NorRΔGAF) that retains the AAA+ and the DNA-binding domains, but lacks the NO responsive GAF domain. We have demonstrated previously that NorRΔGAF activates *norVW* transcription in the absence of NO both *in vitro* and *in vivo* (11). The DNA-binding characteristics of NorRΔGAF are similar to that of wild-type NorR (data not shown). As observed previously, the interaction of NorR with the three enhancer sites in the wild-type promoter was manifested by protection and enhancement of G residues between position −137 and −84 relative to the *norV* transcript initiation site (13). In the case of the site 1 mutant, protection was observed at sites 2 and 3 but was absent at site 1 as anticipated (Figure 2B). Similarly, with either the site

2 or 3 mutant, protection by NorRΔGAF was lost at the mutant site but was maintained at the two remaining wild type enhancer sites (Figure 2C and D). However, some methylation enhancement was detectable in most cases at the mutant sites. If NorR binding to the three sites is strongly co-operative it would be expected that disruption of a single site would influence binding to the wild-type sites. As this was not evident from the methylation protection experiments, the mutations apparently prevent binding of NorR to each of the mutant sites, without discernable loss of binding to the remaining wild-type sites.

All three enhancer sites are required for stimulating the ATPase activity of NorRΔGAF

Our previous investigations have demonstrated that activation of the ATPase activity of NorR not only requires the binding of NO to the Fe(II) centre in the GAF domain but also requires specific DNA containing the three enhancer sites in the *norR-norV* intergenic region [(11) and data unpublished]. This suggests that in the absence of NO, the GAF domain represses the ATPase activity of the AAA+ domain and that binding to specific DNA targets is also required for ATPase activation. In the case of the truncated NorRΔGAF protein, the requirement for NO-dependent signal activation is relieved, but the presence of DNA containing the *norR-norV* intergenic region is still required for significant ATPase activity (11).

This further suggests that binding to the enhancer sites is necessary to activate the ATPase activity of NorR.

In order to investigate the role of individual enhancer sites in activation of the ATPase activity, we assayed ADP release in the presence and absence of DNA fragments carrying mutations in each of the NorR-binding sites. In the absence of promoter DNA, NorRΔGAF exhibited a low level of ATP hydrolysis (<50 mM ATP min^{-1} mol $^{-1}$ NorR) (Figure 3A, solid bars). This activity was stimulated ~40-fold in the presence of a DNA fragment containing all three enhancer sites as observed previously (11). However, the stimulation of the ATPase activity of NorRΔGAF was significantly reduced in the presence of DNA fragments carrying mutations in each of the NorR-binding sites (Figure 3A, open bars), and in each case the DNA stimulation was <2-fold. These results imply that under these conditions, at relatively low protein concentrations (300 nM), NorRΔGAF must be bound to all three enhancer binding sites to catalyse ATP hydrolysis.

Influence of NorR-binding sites on the formation of open promoter complexes by σ^{54}-RNA polymerase

Conformational changes in bEBPs derived from ATP hydrolysis are coupled to the restructuring of the σ^{54}-RNA polymerase–promoter complex to drive the transition from closed to open complexes. In order to correlate the effects of enhancer-binding on ATP hydrolysis with the ability of NorRΔGAF to remodel the σ^{54}-RNA polymerase, we performed open complex assays with wild-type and mutant *norV* promoter DNA templates. We assayed open promoter complex formation by NorRΔGAF on wild type and mutant promoter templates in the presence of σ^{54}-RNA polymerase, IHF, CTP and ATP (11) and quantified the heparin resistant species resolved on non-denaturing polyacrylamide gels (35). Although the binding of NorR to enhancer binding sites is itself heparin resistant, the mobility of nucleoprotein complexes formed upon open complex formation is considerably slower and such complexes can be detected as a super-shifted species (11). At low concentrations of NorRΔGAF (115 and 230 nM), open complex formation was only observed on the DNA fragment with the three wild type NorR enhancer sites (Figure 3B, lanes 3 and 4). In contrast, reactions containing DNA fragments with mutant NorR sites only exhibited open complex formation at concentrations of 460 nM NorRΔGAF or above (Figure 3B, lanes 10, 15 and 20). In this case, activation may occur in *trans* since open complex formation is possible in the absence of IHF at relatively high NorRΔGAF concentrations (data not shown). In control reactions carried out with 460 nM NorRΔGAF, but lacking the σ^{54} subunit, no open complex formation was observed, as expected (Figure 3B, lanes 2, 7, 12 and 17).

Super-shifted species were quantified as a percentage of the open complexes formed at the wild type *norV* promoter at 460 nM NorRΔGAF (Figure 3C). This analysis confirmed that the level of open complexes formed at the mutant promoter constructs was minimal
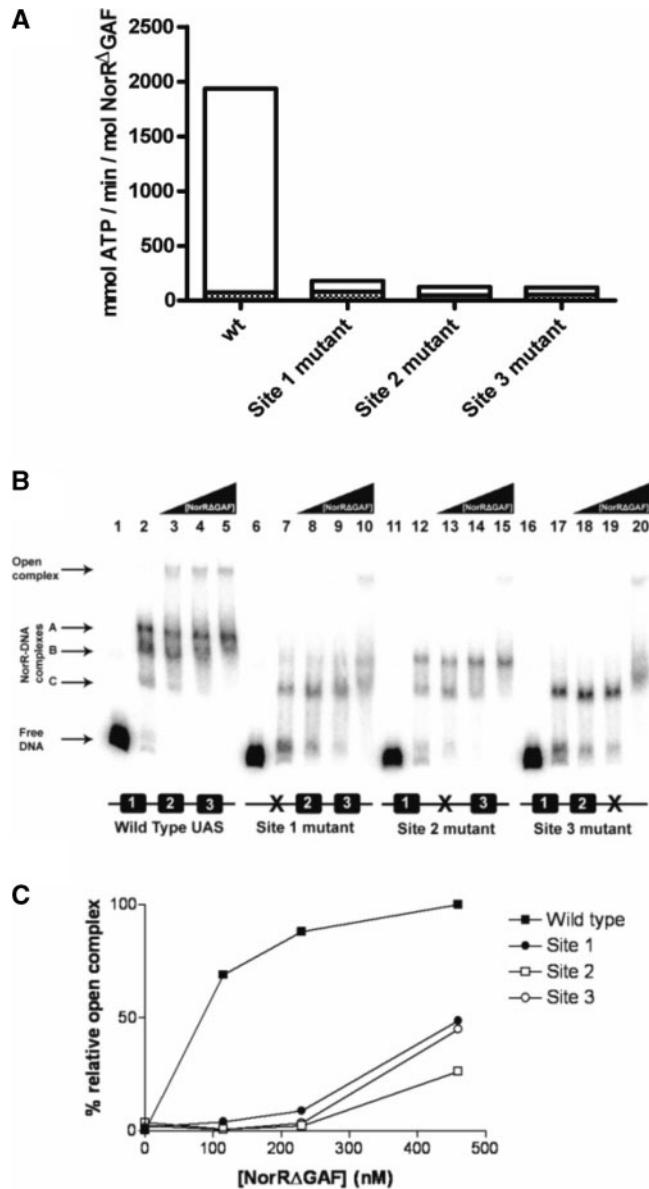

Figure 3. ATPase and open complex stimulating activities of NorRΔGAF associated with wild type and mutant promoter DNA constructs. (**A**) Rates of ATP hydrolysis by NorRΔGAF were monitored at 340 nm for 20 min at 37°C (filled bars) after which 5 nM of the wild type *norR-norVW* fragment was added, and the rates were monitored for a further 20 min at 37°C (empty bars). ATPase activities of NorRΔGAF in the presence of DNA constructs carrying a mutant NorR-binding site were compared with the wild type construct. ATPase activities are expressed as specific activity relative to protein concentration in mmol ATP min^{-1} mol NorR^{-1}. The bar representation is a sum of the values obtained without and with DNA. (**B**) Each open complex assay reaction contained 1 nM of a ^{32}P-labelled DNA fragment encoding either the wild type or mutant *norR-norVW* intergenic region. All lanes contained the components required for open complex formation except lanes 1, 6, 11 and 16, which contained DNA only and lanes 2, 7, 12 and 17, which contained 115 nM NorRΔGAF but lacked σ^{54}. NorRΔGAF concentrations were 115 nM (lanes 3, 8 13, 18), 230 nM (lanes 4, 9, 14, 19) and 460 nM (lanes 5, 10, 15, 20). Free DNA, NorRΔGAF/DNA and open complexes are indicated by the arrows to the left of the figure. (**C**) Heparin resistant open complex species were quantified using a Fujix BAS 1000 phosphorimager. Bands were quantified by their intensity relative to the open complex band formed with the wild type DNA construct at 460 nM NorRΔGAF, which was assumed to be 100%.

compared to the wild type at relatively low protein concentrations (<200 nM) but increased to 50% or less of the wild type at 460 nM NorRΔGAF. At this relatively high concentration it is possible that activation of promoter bound σ^{54}-RNAP by NorRΔGAF occurs from solution as observed with other σ^{54}-dependent activators (17,36).

Three shifted bands (other than the super-shifted species) were visible in reactions containing the wild type DNA fragment (Figure 3B, lanes 2–5). These bands were observed in the absence of IHF and σ^{54}-RNA polymerase (data not shown) and are presumably heparin resistant NorR–DNA complexes. The presence of more than one band is indicative of partial occupancy of the enhancer sites at relatively low protein concentrations under these buffer conditions. We assume that the three bands result from the binding of NorR to one, two or all three sites (labelled as NorR–DNA complexes A, B or C in Figure 3B). It is notable that bands B and C were evident in reactions carried out with the site 2 mutant DNA fragment whereas band C was the main species visible with fragments containing the site 1 or site 3 mutations (Figure 3B, lanes 7–10, 12–15 and 17–20, respectively). Therefore the latter mutations apparently influence the binding of NorR to adjacent wild-type enhancer sites in these heparin challenge experiments, thus providing some evidence for cooperative binding.

NorR functions as an oligomer *in vitro*

The prerequisite of all three enhancer sites for maximal ATP hydrolysis by NorR can be interpreted as the requirement of three functional dimers to form an active hexamer, given the dyad axes of symmetry of the individual sites. A number of well characterised bEBPs form hexameric ring assemblies in their active state, and this oligomerisation itself can depend upon nucleotide (ATP or ADP) binding or self-association at high protein concentrations (37). To determine if higher order oligomer formation is necessary for NorR ATPase activity, we assayed radioactively labelled ADP release from [α-^{32}P]ATP at different concentrations of NorRΔGAF. The relationship between ATP turnover and the NorRΔGAF concentration is shown in Figure 4A. We obtained a concentration dependent sigmoidal activity curve (Figure 4A), with a Hill coefficient of 2 (Figure 4B), implying that cooperativity between monomers is required for maximum ATPase activity of NorRΔGAF. The maximum turnover (k_{cat}), expressed in terms of NorRΔGAF monomer, was $3.8\,min^{-1}$, and the amount of NorRΔGAF required to achieve half-maximal ATPase activity, K_{eff}, was 12.5 μM (Figure 4A).

To examine the effect of physiological concentrations of ATP on oligomer formation, we performed analytical gel filtration experiments with various concentrations of NorRΔGAF in the presence and absence of 1 mM ATP at 4°C. Based on reference elution volumes obtained with different protein standards NorRΔGAF exhibited a concentration-dependent elution profile, with the dimer form predominating at high protein concentrations (Figure 4C, upper panel). However, in the presence of ATP, the dimer peak broadened towards higher
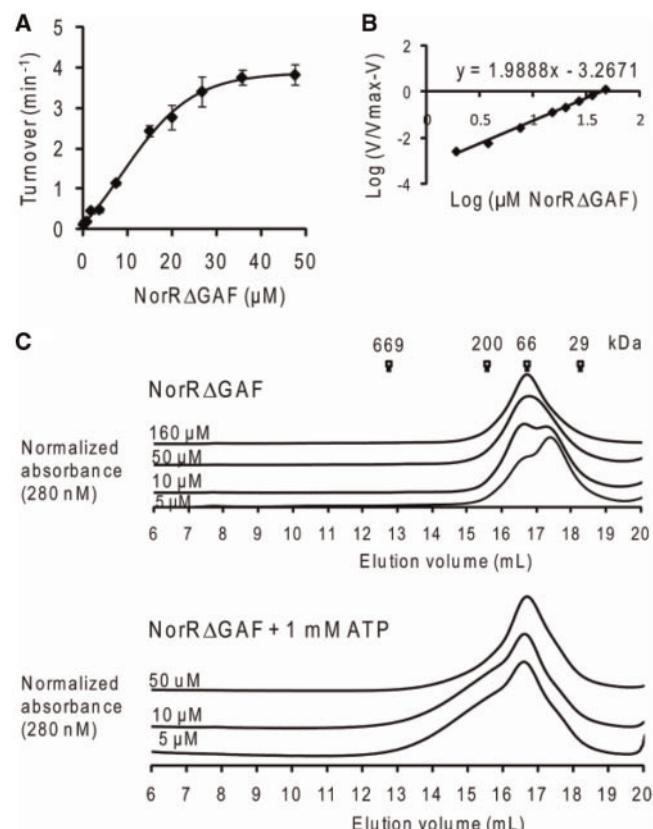

Figure 4. ATPase activity and gel filtration profile of NorRΔGAF. (**A**) Plot of ATP turnover versus protein concentration, measured at a fixed concentration of ATP as substrate (1 mM). (**B**) Log of relative ATP hydrolysis rate V (compared to maximum rate Vmax at 47 μM) was plotted against the log of NorRΔGAF concentration. The slope of the linear regression slope was used to determine the Hill coefficient (1.98). (**C**) Different concentrations of NorRΔGAF were chromatographed at 4°C (upper) or preincubated with 1 mM ATP and chromatographed in the presence of 1 mM ATP at 4°C (lower). Corresponding molecular weight of standard globular proteins were indicated at their elution volume.

oligomeric forms, independent of the protein concentration (Figure 4C, lower panel). These observations suggest that under these conditions, ATP binding promotes self association of NorRΔGAF as inferred from gel filtration profile and concentration-dependent ATPase activity (Figure 4).

Binding to a single enhancer site can stimulate NorRΔGAF ATPase activity *in vitro*

Our data so far demonstrate that the ATPase activity of NorRΔGAF is strongly stimulated by binding to DNA containing all three enhancer sites. This influence of DNA on ATP hydrolysis can be either due to DNA binding *per se* and/or due to increased protein concentration and propensity for stable oligomer formation at the enhancer sites.

To determine the effects of DNA binding *per se* on the state of association and ATPase activity of NorRΔGAF, we used a 21 bp oligonucleotide containing site 1 (NorR1), which apparently has a stronger binding affinity for

NorR than either site 2 or site 3 (13). To determine the optimal molar ratio between NorRΔGAF and DNA, ATP turnover was measured at different protein:DNA molar ratios while maintaining NorRΔGAF at 35.6 μM and ATP at 1 mM (Supplementary Figure S1A, left panel). A 1:1 molar ratio gave maximal activity and was chosen for subsequent experiments in which we measured ATPase activity at various NorRΔGAF concentrations. In the presence of the NorR1 oligonucleotide, the ATPase activity displayed the same sigmoidal kinetics with respect to NorRΔGAF concentration (Figure 5A) while the K_{eff} reduced from 12.5 μM with NorRΔGAF alone to 8.4 μM in the presence of the 21 bp DNA fragment (Figure 5A), suggesting that DNA binding increases the tendency of NorRΔGAF to self-associate. Interestingly, binding of NorRΔGAF to the 21 bp DNA fragment also increased the k_{cat} from 3.8 to 5.2 min^{-1}, implying allosteric stimulation of the intrinsic ATPase activity.

To confirm whether binding to the 21 bp oligonucleotide indeed promotes self-association, we analysed the gel filtration profiles of NorRΔGAF:NorR1 complexes in the presence and absence of nucleotide. The presence of the 21 bp oligonucleotide shifted the NorRΔGAF associated peak towards higher molecular mass species, independent of the presence of ATP (Figure 5B), suggesting that binding to the single enhancer site promotes oligomerisation, in agreement with the increased ATPase activity observed in Figure 5A.

We next investigated if the DNA-binding domain of NorR controls the ATPase activity of the AAA+ domain in a manner analogous to the intramolecular repression exerted by the GAF domain. In this case the stimulatory effect on the ATPase upon DNA binding might simply be due to the removal of this inhibition, similar to the activation of the GAF domain by NO binding. We compared the ATPase activity of AAA+ domain alone with that of NorRΔGAF under the same protein concentrations. Interestingly, the isolated AAA+ domain of NorR had negligible ATPase activity compared to that of NorRΔGAF (Figure 5C, left panel). This result may suggest that DNA binding actively promotes ATPase activity (probably through promoting oligomerisation of AAA+ domain) rather than simply relieving an inhibition imposed by the DNA-binding domain. Indeed the AAA+ domain fails to form higher order oligomers at high concentration and in the presence of ATP (Figure 5C, right panel).

Binding to a 66 bp fragment containing three enhancer sites (NorR123) marginally increases the ATPase activity relative to binding to a single NorR1 site

Our data show that DNA binding *per se* promotes self-association and ATPase activity. We next tested whether availability of the three consecutive sites further enhances the ATPase activity of NorRΔGAF. We used a 66 bp DNA fragment (NorR123), which is the minimum DNA length containing the three NorR-binding sites based on previous DNA footprinting results. ATPase activity assays were performed with increasing NorRΔGAF concentrations while maintaining a 3:1
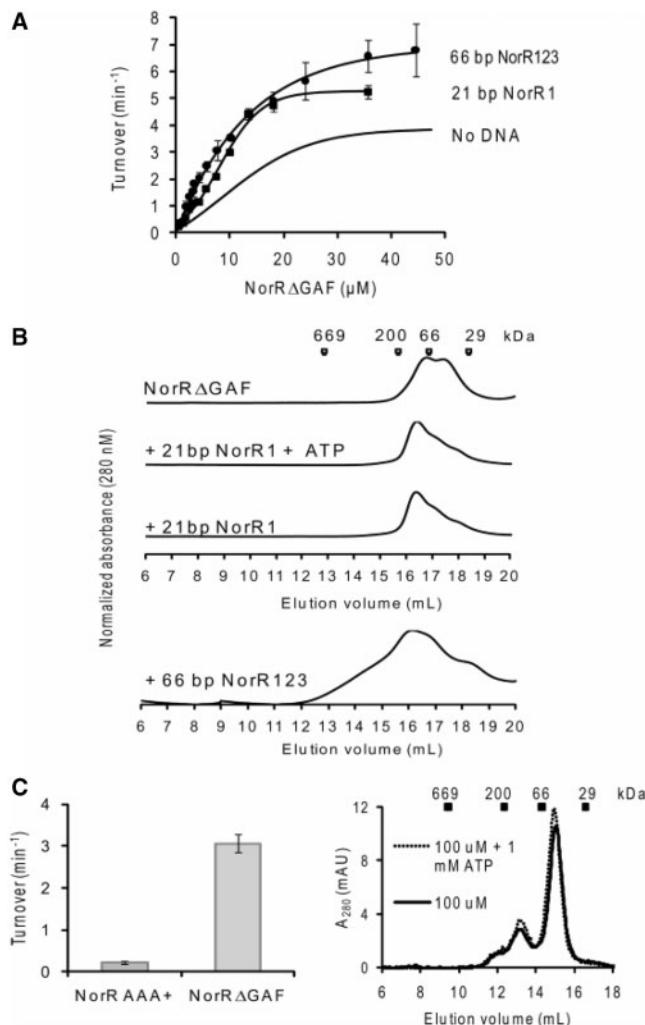

Figure 5. Effect of individual enhancer sites on ATPase activity and oligomerisation states of NorRΔGAF. (**A**) Comparison of the ATPase activity of NorRΔGAF in the absence (*cf*. Fig 4A) and presence of NorR1 or NorR123 dsDNA. A molar ratio of 1:1 activator monomer: DNA was maintained for complexes formed with NorR1 and NorR123 fragments, respectively. (**B**) Gel filtration studies of NorRΔGAF (20 μM) in the absence or presence of dsDNA as in A. Complexes were also chromatographed in the presence of 1 mM ATP, with no significant effect observed on the elution profile. (**C**) left: ATPase activities of 26.5 μM NorRΔGAF and NorR$_{178-452}$ AAA+ domain compared. The assays were performed at 37°C, and turnover in min^{-1} was calculated when [α-^{32}P]ADP formed was ~20 % of total radiolabeled nucleotide. Right: the elution profile of NorR AAA+ domain (100 μM) gel filtered through a Superdex 200 column in the presence and absence of 1 mM ATP at 4°C. Standard globular proteins were used for calibration: thyroglobulin (669 kDa), β-amylase (200 kDa), bovine serum albumin (66 kDa) and carbonic anhydrase (29 kDa).

activator monomer: DNA molar stoichiometry in the complex. This optimal DNA to protein molar stoichiometry was determined as described above by fixing protein and ATP concentrations and varying DNA: protein ratios for optimal ATP turnover (Supplementary Figure S1A, right panel). In the presence of this 66 bp DNA fragment, ATPase activity increased marginally to a k_{cat} of 6.8 min^{-1} when compared to that with the 21 bp NorR1 oligonucleotide (Figure 5A), although the K_{eff}

remained unchanged. Size exclusion chromatography indicated that the complexes elute as multiple peaks (Figure 5B), implying a highly heterogeneous population of oligomers. Our results so far indicate that the increased ATPase activity of NorRΔGAF upon binding to enhancer DNA is largely due to DNA binding *per se* rather than increased local protein concentration at the three enhancer sites.

Binding to a 266 bp DNA fragment, containing the three enhancer sites, strongly stimulates the ATPase activity of NorR

As stimulation of ATP hydrolysis by the 66 bp oligonucleotide was less than anticipated for a DNA fragment containing all three NorR enhancer sites (Figure 3A), we measured the ATPase activity of NorRΔGAF in complex with a longer (266 bp) DNA fragment. We performed ATPase assays with this fragment using a protein:DNA stoichiometry of 6 : 1, which was again chosen based on the optimal ATPase activity at fixed protein and ATP concentrations (Supplementary Figure S1B). ATP turnover increased significantly at low protein concentrations and the K_{eff} fell significantly to 1.4 μM when compared to that of protein alone or in complex with the 21 and 66 bp dsDNA (Figure 6A). Furthermore, the sigmoidal nature of the curve indicates clear evidence for cooperativity. Size exclusion chromatography of NorRΔGAF in complex with the 266 bp DNA fragment revealed a significant shift towards a higher molecular mass species (elution peak at 9.3 ml; Figure 6B), suggesting that this DNA fragment stabilises a higher order oligomeric form of NorRΔGAF (possibly a hexamer), in contrast to the complexes observed on the 21 and 66 bp oligonucleotides (compare Figures 5B with 6B). The presence of DNA in the peak fraction was confirmed by measuring the absorption at 280 and 260 nm wavelengths. We infer from these observations that stabilisation of a higher order oligomer on the 266 bp fragment may be responsible for the increased stimulation of the ATPase activity of NorRΔGAF.

In order to further assess the differences between the 66 and 266 bp DNA, we used a gel retardation assay to compare the affinity of NorRΔGAF for these DNA fragments. Whereas only a single shifted species was observed on the 266 bp DNA (Figure 7A), partial occupancy of the enhancer sites was evident at low protein concentrations with the 66 bp DNA (Figure 7B). Quantitation of the fully shifted species showed that NorRΔGAF has a 2-fold higher affinity for the 266 bp DNA fragment (K_D, 81 nM) compared to the 66 bp DNA (K_D, 174 nM). Moreover, the NorRΔGAF-DNA-binding curve for the 266 bp DNA fragment exhibited increased positive cooperativity with a Hill coefficient of 3.4 compared with 1.3 for the 66 bp oligonucleotide (Figure 7C). This increase in affinity and cooperativity is consistent with a more stable hexameric ring formation, in agreement with the increased ATP turnover observed with the longer DNA fragment (Figure 6A). Although we cannot rule out the possibility that the increase in affinity is due to thermodynamic
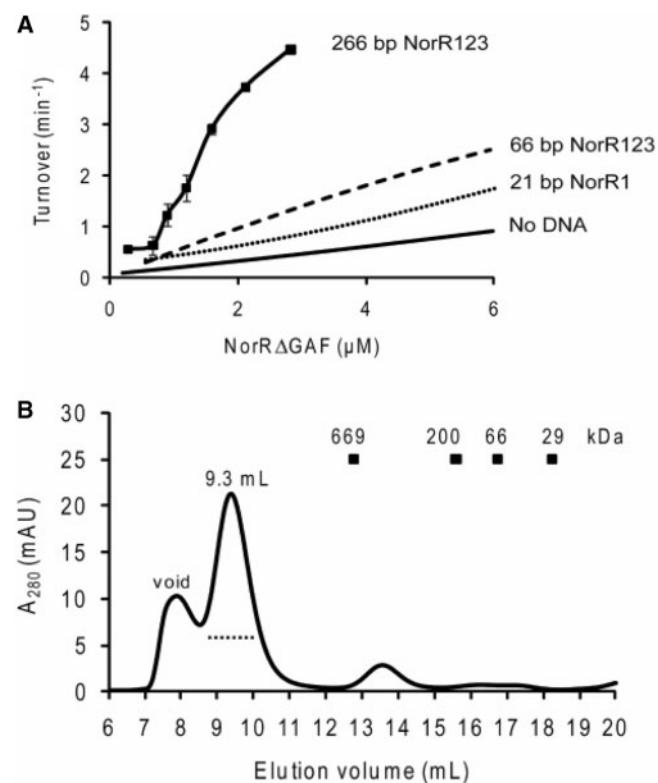

Figure 6. The ATPase activity and oligomerisation state of NorRΔGAF in the presence of the 266 bp dsDNA that contains all three enhancer sites. (**A**) Plot of ATP turnover versus NorRΔGAF concentration in the presence of 266 bp dsDNA, containing all three enhancer sites. Activity curves are also included for comparison of ATPase turnover at low concentrations in the presence and absence of shorter DNA fragments. (**B**) Gel filtration chromatography of 9 μM NorRΔGAF in complex with 0.75 μM 266 bp dsDNA (a molar ratio of 12 : 1 monomer: DNA) performed at 4°C using a Superose 6 column. The dotted line below the 9.3 ml elution peak represents the fractions analysed by negative-stain electron microscopy.

requirements for DNA recognition, it is highly likely that the affinity increase is due to additional interactions between protein and DNA, such as those encountered through the DNA wrapping around NorR. Nevertheless, under the conditions at which the ATPase activity and gel filtration were measured (micromolar concentrations), the DNA should be fully saturated with protein, thus the differences in gel filtration profile and ATPase activity (Figures 5 and 6) are likely due to the different nature of the nucleoprotein complexes formed rather than the different affinities for the protein–DNA interaction.

EM studies of NorRΔGAF bound to 266 bp NorR123 DNA

To investigate the mechanism behind the increased ATPase activity stimulated by the longer 266 bp DNA fragment, we analysed these protein–DNA complexes, prepared either *in situ* or after purification by gel filtration chromatography, using negatively-stained electron microscopy. As shown in Figure 8B, NorRΔGAF does in fact form high order oligomers in the presence of 266 bp DNA. Image analysis and classification of 5000 particles into class averages of 7–10 particles per class (Figure 8C)
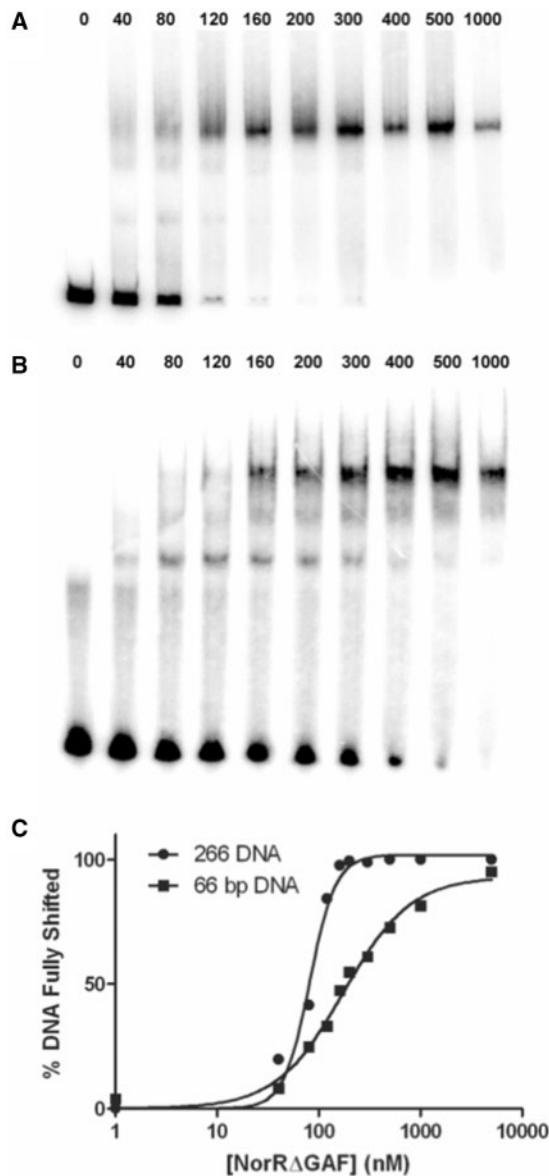
Figure 7. NorRΔGAF has a higher affinity for the 266 bp DNA fragment than for the 66 bp DNA fragment NorR123. NorRΔGAF concentrations between 0 and 1000 nM (as indicated above the gels) were incubated with either the 266 bp DNA fragment (**A**) or the 66 bp DNA fragment (**B**). The percentage of fully shifted DNA was quantified using a Fujix BAS 1000 phosphorimager (**C**).

allowed visualisation of ring-shaped particles with dimensions of 126 Å in diameter, consistent with a hexameric ring observed in cryo-EM studies of other bEBPs such as PspF$_{1-275}$ and NtrC (20,38). These ring shaped particles were not observed in the presence of the 21 or 66 bp oligonucleotides (Figure 8A), in agreement with the observations from size exclusion chromatography that the stable higher order oligomer is only observed upon binding of NorRΔGAF to the longer 266 bp DNA fragment (Figure 5B and 6B).

DISCUSSION

It has now been experimentally confirmed that all three NorR-binding sites are required for transcriptional activation of the *E. coli norV* promoter. Similar conclusions have been derived from deletion analysis of this promoter (39) and also the *Ralstonia eutropha norB* promoter (40). Our data extend these observations and characterise the requirement for the three enhancers *in vitro*. The presumed physiological role of bacterial enhancers is to tether activators at high local concentration close to the promoter and to facilitate the formation of higher oligomeric forms that are active for transcriptional activation. While multiple enhancers are common in σ^{54}-dependent promoters, an absolute dependency on more than one target site is unusual. For example, two enhancers are sufficient to assemble higher order oligomers of NtrC, in which some protomers are bound by protein–protein interactions and apparently do not contact DNA (41). In contrast, our data for NorR indicate that all three enhancer sites are necessary for the formation of an active oligomeric species. When present at high concentration, bEBPs can activate transcription from solution in the absence of enhancer DNA *in vitro* and some σ^{54}-dependent activators naturally lack a DNA-binding domain (42). Moreover, in many cases the DNA-binding domain does not appear to be essential for transcriptional activation. For example, PspF lacking the enhancer-binding domain (PspF$_{1-275}$) can form high-order oligomers and activate transcription *in vitro*.

The increased ATPase activity of NorRΔGAF observed upon binding to the three enhancer sites could be due to one or all of the following three reasons: (i) DNA binding induces conformational changes that promote self-association and therefore increases ATPase activity, (ii) DNA binding stimulates ATPase activity *per se*, (iii) since there are three NorR-binding sites, binding to DNA increases the local protein concentration and thus promotes self-association, which increases the ATPase activity. Our data suggest that the significant increase in ATPase activity upon binding to enhancer sites is not likely to be a consequence of self-association resulting from an increase in local protein concentration. It is more likely due to the conformational changes induced by DNA binding that promote hexameric ring formation and ATP hydrolysis *per se*, therefore increasing the ATPase activity. Hence in the case of NorR, in addition to GAF domain activation, the enhancer DNA sites provide an important ligand to promote assembly of the active oligomeric form of the activator. This is in stark contrast to the activation mechanism of other bEBPs such as NtrC, where activation of the receiver domain is key to hexamer formation and ATPase activity (20). Our results are consistent with the observation that all three NorR-binding sites are required for catalytic activity. Mutating any one of the sites abolishes binding of one NorR dimer to the DNA, therefore preventing proper hexameric ring formation. However, in the absence of DNA, ATPase activity increases with protein concentration, in agreement with the observation that open complex formation can be achieved in the absence of enhancer DNA at relatively high concentrations of NorR. This demonstrates that although the native conformation in the absence of DNA is not optimal for formation of higher order oligomers (as shown in gel filtration
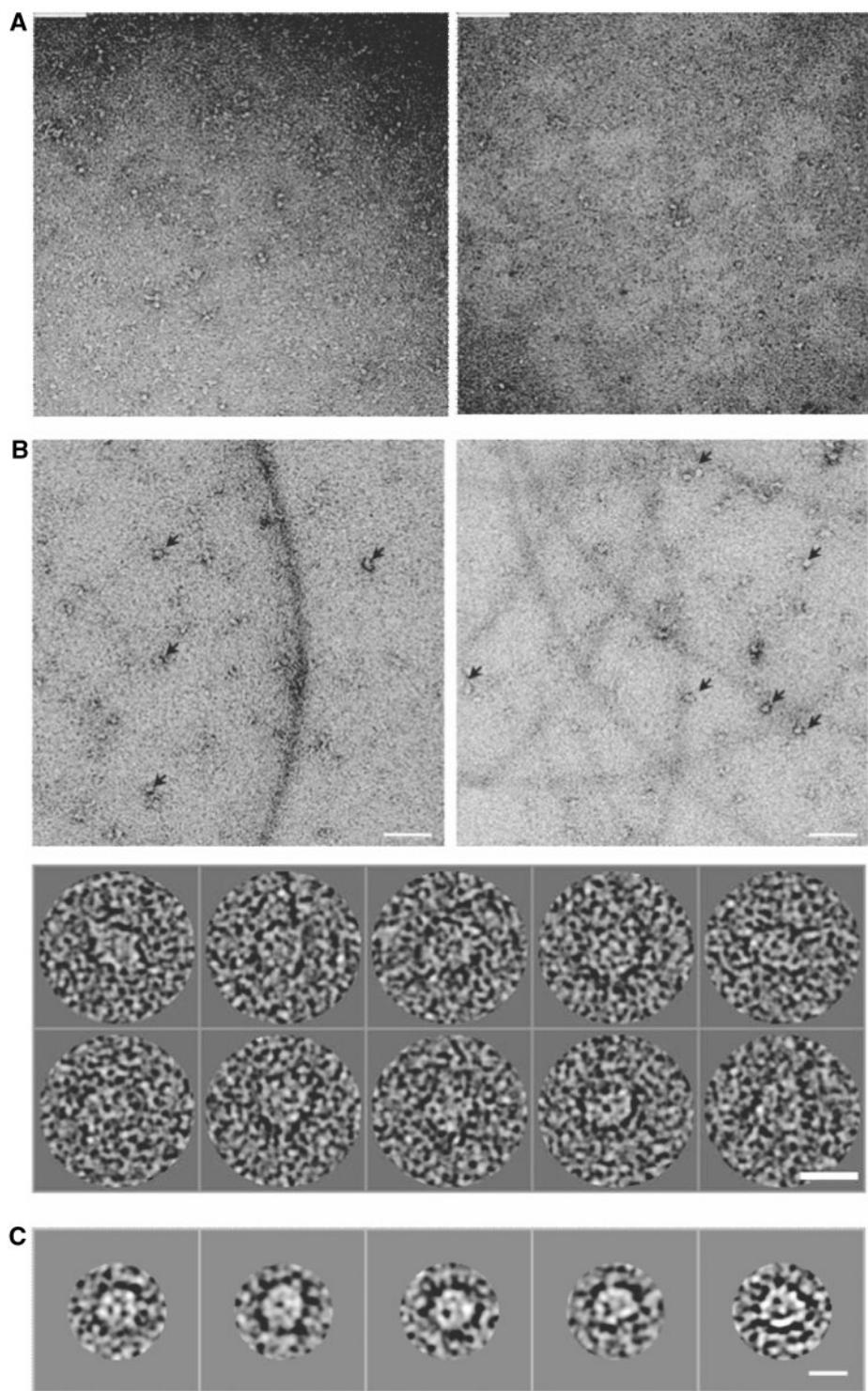
Figure 8. Negative-stain electron microscopy of NorRΔGAF in complex with 266 bp dsDNA, containing all three enhancer sites. (**A**) Raw micrographs of NorRΔGAF in complex with 21 bp NorR1 (right panel) or 66 bp NorR123 (left panel) DNA fragments. Scale bar 90 nm. (**B**) Top, raw micrograph of NorRΔGAF in complex with 266 bp dsDNA, with arrows showing some of the higher order oligomers. Scale bar 90 nm. Bottom, gallery of picked particles. Scale bar 15 nm. (**C**) Selected class averages of 7–10 particles per class generated from 5000 particles show ring-shaped particles with a diameter of 126 Å. Scale bar 11 nm.

experiments), at higher protein concentrations and in the presence of ATP, oligomerisation can occur to enable catalytic activity. However, the higher order complexes formed under these conditions are likely to be far less stable than the nucleoprotein complexes formed in the presence of the three enhancer sites. In the absence of the DNA-binding domain, the isolated AAA+ domain of NorR fails to form higher order oligomers and the

ATPase activity is further reduced compared to NorRΔGAF, consistent with the hypothesis that the DNA-binding domain can promote oligomerisation, and hence ATPase activity. DNA binding appears to shift the conformation of NorRΔGAF to a form that favours hexameric ring formation at physiological protein concentrations, thus providing the necessary ATPase activity required for open complex formation.

The requirement for three consecutive enhancer sites is reminiscent of the EBP-related protein TyrR, which is proposed to bind as a dimer to three TyrR boxes, forming a hexameric species that is active in transcriptional repression (43). Another unusual feature of the NorR–DNA interaction is the heparin resistance of the protein–DNA complexes, which implies that NorR makes extensive DNA contacts, possibly forming a topologically distinct nucleoprotein complex. Our EM data show that in the presence of the 266 bp DNA fragment carrying the three enhancer sites, NorRΔGAF forms oligomeric rings, similar to other bEBPs in their active functional states. The precise mechanism whereby the enhancer DNA stabilises ring formation is currently unclear. One possible model is that the DNA wraps around the hexameric ring, making extensive contacts that help to stabilise it. Assuming that the diameter of the hexameric ring assembly of the AAA+ domain is ~120 Å (38), a minimum of 450 Å (the total diameter to the centre of duplex DNA is $120 + 25$ Å) or ~130 bp DNA is required to wrap around the ring. The 66 bp NorR123 oligonucleotide would therefore be insufficient, consistent with our finding that the stimulatory effect of this DNA fragment, which contains all three UAS sites, is similar that of the 21 bp oligonucleotide, which carries only a single UAS site.

The assembly of bEBPs into at least a hexamer is necessary for activation of the ATPase activity required to drive the transition of the σ^{54}-RNA polymerase promoter complex from the closed to the open DNA-melted state. Although the experiments described here have been performed with a constitutive form of NorR lacking the regulatory GAF domain, the ATPase activity of wild-type NorR is also enhancer dependent, and in addition requires the binding of NO to the ferrous iron centre to activate the catalytic activity of the AAA+ domain. Given that the NorR apoprotein is fully competent for DNA binding, the three enhancers clearly provide a scaffold for the assembly of a stable heparin-resistant NorR nucleoprotein complex that is poised at the promoter, ready to perceive the NO signal. In contrast to other EBPs such as NtrC and DctD which are dimeric in their inactive forms and are regulated through control of the oligomerisation state, the activity of wild-type NorR is apparently regulated when bound to DNA as a higher order oligomer.

In summary, our data support a unique activation mechanism for NorR. Three NorR dimers readily bind to the three consecutive UAS sites. DNA binding by NorR induces conformational changes that stimulate hexameric ring formation. NorR then forms a hexameric ring with extensive DNA interactions (possibly through DNA wrapping) for increased stability. In the presence of the NO signal, intramolecular repression of the AAA+ domain by the GAF domain is released, activating ATPase activity and allowing the NorR hexamer to interact with RNAP-σ^{54} and activate transcription.

REFERENCES

1. Zumft,W. (1997) Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.*, **61**, 533–616.
2. MacMicking,J., Xie,Q.-w. and Nathan,C. (1997) Nitric oxide and macrophage function. *Annu. Rev. Immunol.*, **15**, 323–350.
3. Poole,R.K. and Hughes,M.N. (2000) New functions for the ancient globin family: bacterial responses to nitric oxide and nitrosative stress. MicroReview. *Mol. Microbiol.*, **36**, 775–783.
4. Gardner,A.M., Helmick,R.A. and Gardner,P.R. (2002) Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in *Escherichia coli. J. Biol. Chem.*, **277**, 8172–8177.
5. Poock,S.R., Leach,E.R., Moir,J.W.B., Cole,J.A. and Richardson,D.J. (2002) Respiratory detoxification of nitric oxide by the cytochrome c nitrite reductase of *Escherichia coli. J. Biol. Chem.*, **277**, 23664–23669.
6. Gomes,C.M., Giuffre,A., Forte,E., Vicente,J.B., Saraiva,L.M., Brunori,M. and Teixeira,M. (2002) A novel type of nitric-oxide reductase. *Escherichia coli* flavorubredoxin. *J. Biol. Chem.*, **277**, 25273–25276.
7. 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.
8. Gardner,A.M., Gessner,C.R. and Gardner,P.R. (2003) Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and sigma54 in the nitric oxide stress response. *J. Biol. Chem.*, **278**, 10081–10086.
9. Pohlmann,A., Cramm,R., Schmelz,K. and Friedrich,B. (2000) A novel NO-responding regulator controls the reduction of nitric oxide in Ralstonia eutropha. *Mol. Microbiol.*, **38**, 626–638.
10. Studholme,D.J. and Dixon,R. (2003) Domain architectures of sigma54-dependent transcriptional activators. *J. Bacteriol.*, **185**, 1757–1767.
11. 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.
12. 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.

13. 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.

14. Hoover,T.R., Santero,E., Porter,S. and Kustu,S. (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell*, **63**, 11–22.

15. Perez-Martin,J. and De Lorenzo,V. (1995) Integration host factor suppresses promiscuous activation of the sigma 54-dependent promoter Pu of *Pseudomonas putida*. *Proc. Natl Acad. Sci. USA*, **92**, 7277–7281.

16. Weiss,D.S., Batut,J., Klose,K.E., Keener,J. and Kustu,S. (1991) The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. *Cell*, **67**, 155–167.

17. Cannon,W.V., Gallegos,M.T. and Buck,M. (2000) Isomerization of a binary sigma-promoter DNA complex by transcription activators. *Nat. Struct. Biol.*, **7**, 594–601.

18. 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.

19. Sallai,L. and Tucker,P.A. (2005) Crystal structure of the central and C-terminal domain of the sigma54-activator ZraR. *J. Struct. Biol.*, **151**, 160–170.

20. 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.

21. Rappas,M., Schumacher.Jr, Niwa,H., Buck,M. and Zhang,X. (2006) Structural basis of the nucleotide driven conformational changes in the AAA+ domain of transcription activator PspF. *J. Mol. Biol.*, **357**, 481–492.

22. Rombel,I., Peters-Wendisch,P., Mesecar,A., Thorgeirsson,T., Shin,Y.K. and Kustu,S. (1999) MgATP binding and hydrolysis determinants of NtrC, a bacterial enhancer-binding protein. *J. Bacteriol.*, **181**, 4628–4638.

23. Austin,S. and Dixon,R. (1992) The prokaryotic enhancer binding protein NTRC has an ATPase activity which is phosphorylation and DNA dependent. *EMBO J.*, **11**, 2219–2228.

24. Porter,S., North,A., Wedel,A. and Kustu,S. (1993) Oligomerization of NTRC at the *glnA* enhancer is required for transcriptional activation. *Gen. Dev.*, **7**, 2258–2273.

25. Perez-Martin,J. and de Lorenzo,V. (1996) *In vitro* activities of an N-terminal truncated form of XylR, a sigma 54-dependent transcriptional activator of *Pseudomonas putida*. *J. Mol. Biol.*, **258**, 575–587.

26. Lowry,O.H., Rosenberg,N.J., Farr,A.L. and Randall,R.J. (1951) Protein measurement with the folin phenol reagent. *J.Biol.Chem.*, **193**, 265–275.

27. Ito,W., Ishiguro,H. and Kurosawa,Y. (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene*, **102**, 67–70.

28. Simons,R.W., Houman,F. and Kleckner,N. (1987) Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85–96.

29. Norby,J.G. (1988) Coupled assay of Na+,K+-ATPase activity. *Methods Enzymol.*, **156**, 116–119.

30. D'Autreaux,B., Tucker,N., Spiro,S. and Dixon,R. (2008) Characterization of the nitric oxide-reactive transcriptional activator NorR. *Methods Enzymol.*, **437**, 235–251.

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

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

33. Grant,T. (2007) *Advances in Single Particle Electron Microscopy*. Imperial College London, London.

34. Mukhopadhyay,P., Zheng,M., Bedzyk,L.A., LaRossa,R.A. and Storz,G. (2004) Prominent roles of the NorR and Fur regulators in the *Escherichia coli* transcriptional response to reactive nitrogen species. *Proc. Natl Acad. Sci. USA*, **101**, 745–750.

35. Eydmann,T., Söderbäck,E., Jones,T., Hill,S., Austin,S. and Dixon,R. (1995) Transcriptional activation of the nitrogenase promoter in vitro: adenosine nucleosides are required for inhibition of NIFA activity by NIFL. *J. Bacteriol.*, **177**, 1186–1195.

36. North,A.K. and Kustu,S. (1997) Mutant forms of the enhancer-binding protein NtrC can activate transcription from solution. *J. Mol. Biol.*, **267**, 17–36.

37. Joly,N., Schumacher,J. and Buck,M. (2006) Heterogeneous nucleotide occupancy stimulates functionality of phage shock protein F, an AAA+ transcriptional activator. *J. Biol. Chem.*, **281**, 34997–35007.

38. Rappas,M., Schumacher,J., Beuron,F., Niwa,H., Bordes,P., Wigneshweraraj,S., Keetch,C.A., Robinson,C.V., Buck,M. and Zhang,X. (2005) Structural insights into the activity of enhancer-binding proteins. *Science*, **307**, 1972–1975.

39. Justino,M.C., Gonçalves,V.M.M. and Saraiva,L.M. (2005) Binding of NorR to three DNA sites is essential for promoter activation of the flavorubredoxin gene, the nitric oxide reductase of *Escherichia coli*. *Biochem. Biophys. Res. Commun.*, **328**, 540–544.

40. Busch,A., Pohlmann,A., Friedrich,B. and Cramm,R. (2004) A DNA region recognized by the nitric oxide-responsive transcriptional activator NorR is conserved in beta- and gamma-proteobacteria. *J. Bacteriol.*, **186**, 7980–7987.

41. Wyman,C., Rombel,I., North,A.K., Bustamante,C. and Kustu,S. (1997) Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science*, **275**, 1658–1661.

42. Beck,L.L., Smith,T.G. and Hoover,T.R. (2007) Look, no hands! Unconventional transcriptional activators in bacteria. *Trends Microbiol.*, **15**, 530–537.

43. Wilson,T.J., Maroudas,P., Howlett,G.J. and Davidson,B.E. (1994) Ligand-induced self-association of the *Escherichia coli* regulatory protein TyrR. *J. Mol. Biol.*, **238**, 309–318.

44. Casadaban,M.J., Chou,J. and Cohen,S.N. (1980) In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translation initiation signals. *J. Bacteriol.*, **143**, 971–980.

Nitric oxide-responsive interdomain regulation targets the σ^{54}-interaction surface in the enhancer binding protein NorR

Matthew Bush,[1] Tamaswati Ghosh,[2] Nicholas Tucker,[1,3] Xiaodong Zhang[2] and Ray Dixon[1]*

[1]*Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Colney NR4 7UH, UK.*
[2]*Division of Molecular Bioscience, Imperial College London, London SW7 2AZ, UK.*
[3]*Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow G4 0RE, UK.*

Summary

Bacterial enhancer binding proteins (bEBPs) are specialized transcriptional activators that assemble as hexameric rings in their active forms and utilize ATP hydrolysis to remodel the conformation of RNA polymerase containing the alternative sigma factor σ^{54}. Transcriptional activation by the NorR bEBP is controlled by a regulatory GAF domain that represses the ATPase activity of the central AAA+ domain in the absence of nitric oxide. Here, we investigate the mechanism of interdomain repression in NorR by characterizing substitutions in the AAA+ domain that bypass repression by the regulatory domain. Most of these substitutions are located in the vicinity of the surface-exposed loops that engage σ^{54} during the ATP hydrolysis cycle or in the highly conserved GAFTGA motif that directly contacts σ^{54}. Biochemical studies suggest that the bypass mutations in the GAFTGA loop do not influence the DNA binding properties of NorR or the assembly of higher order oligomers in the presence of enhancer DNA, and as expected these variants retain the ability to activate open complex formation *in vitro*. We identify a crucial arginine residue in the GAF domain that is essential for interdomain repression and demonstrate that hydrophobic substitutions at this position suppress the bypass phenotype of the GAFTGA substitutions. These observations suggest a novel mechanism for negative regulation in bEBPs in which the GAF domain targets the σ^{54}-interaction surface to prevent access of the AAA+ domain to the sigma factor.

Introduction

The promoter specificity of bacterial RNA polymerase is determined by the binding of an additional subunit, the sigma factor (σ). In contrast to the prototypical σ^{70} class of bacterial sigma factors, transcription by the σ^{54} class requires activation by bacterial enhancer binding proteins (bEBPs) that utilize nucleotide triphosphate hydrolysis to drive conformational rearrangements in the σ^{54}-RNA polymerase holoenzyme. The central AAA+ domain of bEBPs is responsible for ATP hydrolysis and the consequent remodelling of the σ^{54}-RNA polymerase that enables isomerization of the promoter DNA complexes from the closed to the open form (Wedel and Kustu, 1995, Cannon *et al.*, 2000; Schumacher *et al.*, 2004). As in the case of other AAA+ proteins, the σ^{54}-interaction domain of bEBPs is competent for ATP hydrolysis when assembled as a hexameric ring (Rappas *et al.*, 2007 and references therein). The bEBP subfamily of AAA+ domains contain specific structural features that enable nucleotide-dependent interactions with σ^{54}. Most conserved amongst these is the GAFTGA motif, which forms a loop on the surface of the AAA+ domain that contacts σ^{54} during the ATP hydrolysis cycle (Bordes *et al.*, 2003). Structural studies demonstrate that the GAFTGA loop (also known as the L1 loop), assisted by a second surface-exposed loop, L2, is in an extended conformation in the ATP bound transition state and is thus competent to engage with σ^{54}. However, in the ADP bound state both loops are compacted towards the surface of the AAA+ domain enabling σ^{54} relocation, crucial to the conversion from the closed to the open complex (Rappas *et al.*, 2006; Chen *et al.*, 2007; Bose *et al.*, 2008). The GAFTGA loop thus performs a crucial role in the 'power stroke' of bEBPs in coupling ATP hydrolysis to conformational rearrangements of the σ^{54}-RNA polymerase.

Many bEBPs contain an amino-terminal regulatory domain that stringently controls the activity of the central AAA+ domain either negatively or positively in response to environmental cues (Studholme and Dixon, 2003). Most

bEBPs also contain a helix–turn–helix DNA binding domain that binds to enhancer-like sequences upstream of promoters. In several well-characterized examples, allosteric control by the regulatory domain is exerted by controlling the oligomeric state of the AAA+ domain. In the 'off' state, the regulatory domain holds the AAA+ domain in an inactive dimeric form (Lee *et al.*, 2003). Conformational changes in the regulatory domain induced by the signal enable transition to the 'on' state in which the AAA+ domain is released to form an active hexameric ring that is competent to activate transcription (Doucleff *et al.*, 2005; De Carlo *et al.*, 2006).

The nitric oxide (NO)-responsive bEBP, NorR, is required for transcriptional activation of the *norVW* genes in *Escherichia coli* that encode a flavorubredoxin and its associated NADH-dependent oxidoreductase respectively (Hutchings *et al.*, 2002). These enzymes provide a detoxification system that reduces the NO radical to nitrous oxide under anaerobic conditions (Gardner *et al.*, 2002; Gomes *et al.*, 2002). Transcriptional activation by NorR is controlled by intramolecular interactions between an N-terminal regulatory GAF domain and the central AAA+ domain. The GAF domain contains a mononuclear non-haem iron centre that responds to NO through the formation of a mononitrosyl iron complex (D'Autreaux *et al.*, 2005). In the absence of the NO signal, the GAF domain inhibits the activity of the AAA+ domain via interdomain repression (Gardner *et al.*, 2003). Upon receipt of the signal and formation of the mononitrosyl iron species, repression of the AAA+ domain is relieved, activating ATP hydrolysis by NorR coupled to conformational remodelling of the σ^{54}-RNA polymerase (D'Autreaux *et al.*, 2005). In addition to allosteric control exerted by the GAF domain, our studies indicate that the C-terminal DNA binding domain of NorR plays a major role in the assembly of the functional AAA+ oligomer. Three enhancer sites located upstream of the *norVW* promoter are essential for transcriptional activation by NorR and provide a scaffold for the assembly of higher order oligomers (Tucker *et al.*, 2010).

To investigate mechanisms of interdomain regulation in NorR, we have used a random mutagenesis approach to screen for mutations in the AAA+ domain that enable escape from GAF domain-mediated repression. Surprisingly, we find that substitutions within the highly conserved GAFTGA motif and in residues predicted to influence nucleotide-dependent conformational changes in this loop prevent intramolecular repression by the GAF domain in the absence of the NO signal. We demonstrate that the GAFTGA substitutions neither influence the DNA binding function of NorR nor the enhancer DNA-dependent oligomerization of the AAA+ domain and that variant proteins remain competent to catalyse open complex formation by σ^{54}-RNA polymerase. Our results suggest that the σ^{54}-

interaction surface in the AAA+ domain is a target for intramolecular repression by the GAF domain.

Results

Mutations in the GAFTGA motif of NorR give rise to constitutive activity

To explore the mechanism of interdomain repression in NorR, error-prone polymerase chain reaction (PCR) mutagenesis was employed to create mutations that potentially disrupt repression of the AAA+ domain by the N-terminal (NO-sensing) GAF domain. This strategy produced mutant versions of NorR that had significant activity in cultures grown in the absence of an NO source, in contrast to wild-type NorR, which is activated by endogenous NO generated in the presence of potassium nitrate (Fig. 1A). In some cases (e.g. G266D, S292L) activity in the absence of the signal was similar to that exhibited by a truncated version of NorR lacking the GAF domain (NorR∆GAF). This phenotype suggests that repression by the GAF domain has been bypassed, resulting in loss of regulation upon the AAA+ domain. In other cases (e.g. F264Y, Q304E) some repression in the absence of NO was evident, indicative of a partial bypass phenotype. In structural models of the AAA+ domain of NorR based on the structure of NtrC1 (Lee *et al.*, 2003), the majority of the substitutions are located in either helix 3 (H3), helix 4 (H4) or loop 1 (L1) (Fig. 1B). These are the structural features in the AAA+ domain that undergo nucleotide-dependent conformational changes during the ATPase cycle to promote engagement with σ^{54}. For example, the equivalent of E276 in PspF (E97) is located close to the base of the L1 loop and forms nucleotide-dependent interactions with R131 (equivalent to NorR R310) in the L2 loop that co-ordinate loop movements during ATP hydrolysis (Rappas *et al.*, 2006). The only substitution predicted to be located outside this region of nucleotide-induced conformational change is Q304; where the equivalent residue in NtrC1 is most probably involved in inter AAA+ domain subunit interactions. Significantly, three substitutions were identified within the highly conserved GAFTGA motif itself. The most notable of these was the G266D mutation, located in the second glycine of the motif, which allowed full escape from the GAF-mediated repression of NorR activity (Fig. 1A). This is surprising given that this loop is required to contact σ^{54} to drive open complex formation (Buck *et al.*, 2006) and that substitutions at G266 are likely to influence the conformational flexibility of this loop. In order to examine which amino acids at the G266 position give rise to constitutive activity, we substituted this residue for each of the other 19 natural amino acids (Fig. S1). In addition to the aspartate substitution that gives rise to constitutive activity; asparagine,
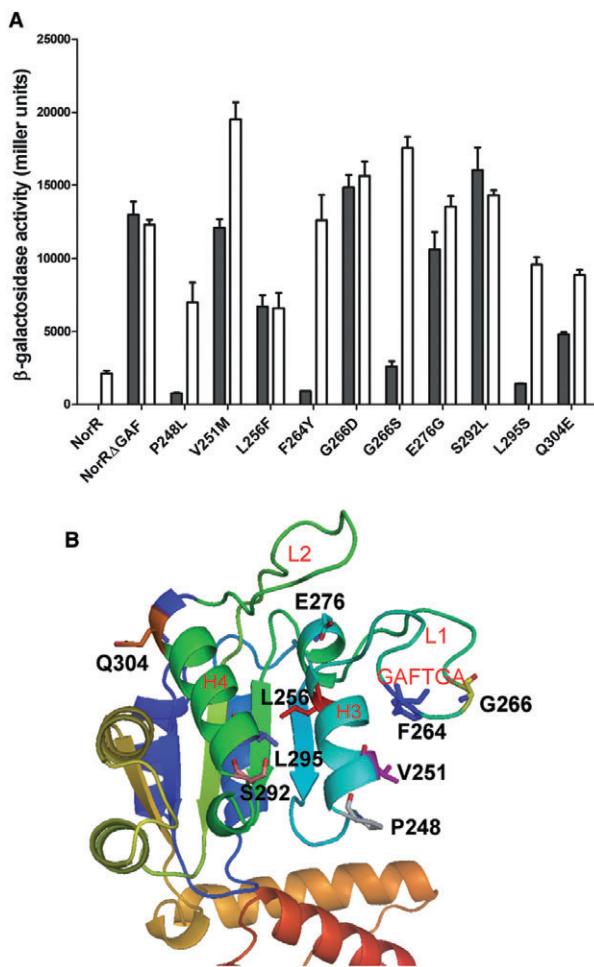
Fig. 1. A. Transcriptional activation by NorR AAA+ domain variants *in vivo* as measured by the *norV–lacZ* reporter assay. Substitutions are indicated on the *x*-axis. 'NorR' refers to the wild-type protein and 'NorRΔGAF' refers to the truncated form lacking the GAF domain (residues 1–170). Cultures 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.
B. Structural model of the AAA+ domain of NorR based on the NtrC1 structure (Lee *et al.*, 2003) (1NY5 chain A). The helices and loops (H3 and H4, L1 and L2) involved in nucleotide-dependent conformational changes in bEBPs are labelled in red. Residues that were substituted as a consequence of the PCR mutagenesis of the AAA+ domain are indicated. The F264 and G266 residues form part of the GAFTGA motif that contacts σ[54].

glutamine, serine, cysteine and methionine all gave activity in the absence of an NO source. Asparagine and aspartate changes gave fully constitutive phenotypes whereas the other changes were still partially subject to regulation by the GAF domain. Surprisingly, the glutamate substitution did not produce a functional NorR protein. The remaining amino acid changes all resulted in non-functional proteins and Western blotting confirmed that this is not due to instability (data not shown). The non-

functional nature of most substitutions at this position is not unexpected, given the importance of the GAFTGA motif and its high conservation in bEBPs.

The apparent loss of regulation in the G266D and G266N variants suggest that these mutations completely bypass the repressive function of the N-terminal GAF domain. To confirm that the NO-sensing function of the GAF domain no longer contributes to the phenotype of the G266D variant, targeted substitutions were made at residues known to disrupt the non-haem iron centre in the GAF domain (Tucker *et al.*, 2008). When the R75K, Y98L, C113S, H111Y or D99A substitutions were combined with G266D, no reduction in the ability to activate transcription by NorR was observed (Fig. S2). To further test the influence of the GAF domain in this variant, the sequence encoding the first 170 residues of NorR was deleted in constructs containing an additional N-terminal, hexa-histidine tag. The resulting G266DΔGAF–His protein was comparable with the G266D–His variant in its ability to activate transcription *in vivo* (Fig. S3). This was also true for the G266N–His protein. The Q304E variant in contrast showed a partial bypass phenotype (Fig. 1A) and removal of the GAF domain led to constitutive activity as anticipated (Fig. S3).

The G266D mutation does not affect enhancer binding or oligomerization of NorR in vitro

Since the oligomerization state and hence the activity of the AAA+ domain of bEBPs is often controlled by regulatory domains, we questioned whether the NorR GAFTGA substitutions 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 GAFTGA mutations influence DNA binding. For this and subsequent biochemical experiments we used GAF domain deleted forms of NorR (NorRΔGAF) and utilized N-terminal histidine tags as an aid to protein purification. The presence of this tag does not significantly affect the activity of wild-type NorRΔGAF or its variants *in vivo* (data not shown). We observed that the affinity of NorRΔGAF for a 361 bp DNA fragment containing the three enhancer sites upstream of the *norV* promoter was not significantly influenced by the presence of the G266D and G266N substitutions (Fig. S4). Dissociation constants (Kd) were calculated as 100 nM in each case. To determine the effect of the G266D substitution on enhancer-dependent NorR oligomer formation (Tucker *et al.*, 2010), we performed analytical gel filtration experiments in the absence and presence of a 266 bp DNA fragment containing the three enhancer sites. Based on reference elution volumes obtained with different protein

standards, unbound G266D∆GAF–His eluted as an apparent monomer/dimer species (Fig. 2A). The presence of the 266 bp DNA fragment shifted the protein peak towards a higher molecular mass species (Fig. 2A) indicating formation and stabilization of a higher order nucleoprotein complex. These elution profiles are similar to that reported recently for wild-type NorR∆GAF (Tucker *et al.*, 2010). Analysis of the purified protein–DNA complex using negatively stained electron microscopy, allowed visualization of higher order ring-shaped particles with dimensions of 125 Å in diameter (Fig. 2B) consistent with a hexameric ring observed for NorR∆GAF in complex with the 266 bp DNA fragment (Tucker *et al.*, 2010). No oligomeric particles were seen in the electron micrographs for protein alone (Fig. 2C). We conclude from these studies that the G266D mutation does not apparently influence the oligomeric assembly of the AAA+ domain or the requirement for enhancer sites to stabilize the formation of a higher order oligomer.

G266 bypass variants show enhancer-dependent ATPase activity *in vitro*

In bEBPs the ATP hydrolysis site is configured through interactions between adjacent AAA+ protomers in the hexameric ring (Schumacher *et al.*, 2008). Since the GAFTGA motif relays nucleotide-dependent interactions at this site to enable contact with σ^{54}, we were interested to examine if the G266 substitutions influence ATPase activity. We have already established that enhancer DNA is required for ATP hydrolysis by NorR and that the three binding sites upstream of the *norV* promoter are necessary for activation of ATPase activity, consistent with the requirement for DNA for formation of a functional higher order oligomer (Tucker *et al.*, 2010). Using concentrations of NorR∆GAF–His within the anticipated physiological range, we observed low levels of ATP hydrolysis in the absence of enhancer DNA. This was also a property of the G266D and G266N variants (Fig. 3B and C, black bars). Consistent with our previous studies with a non-histagged form of NorR∆GAF, ATPase activity was strongly stimulated by the presence of promoter DNA. Under these conditions ATP hydrolysis by NorR∆GAF–His increased as a sigmoidal reponse to increasing protein concentration indicative of positive cooperativity, with a lower rate of increase exhibited at concentrations above 250 nM (Fig. 3A, white bars). The absence of increased activity at higher protein concentrations may reflect saturation of the enhancer sites consistent with the observed DNA binding constant (100 nM as reported above, Fig. S4). Although
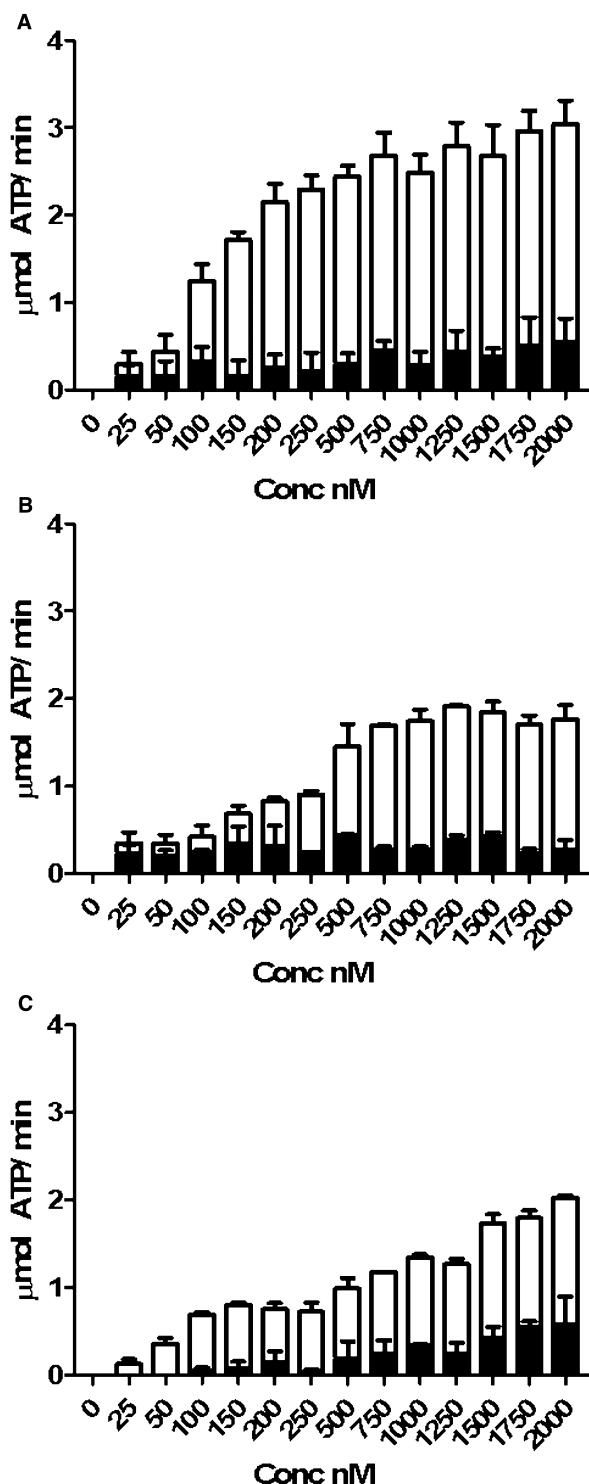
Fig. 3. ATPase activity of the NorRΔGAF–His (A), G266DΔGAF–His (B) and G266NΔGAF–His (C) variants in response to protein concentration and the presence of enhancer DNA. Assays were conducted either in the absence (closed bars) or presence (open bars) of the 266 bp DNA fragment (final concentration 5 nM) that includes the *norR–norVW* intergenic region and each of the three NorR binding sites. Data are shown as the mean from at least two experiments.

ATP hydrolysis by the G266D and G266N variants was also stimulated by the enhancer sites, the response to protein concentration was less cooperative than observed with NorRΔGAF–His and activities were lower than those of the wild-type protein even at a relatively high protein concentration (2 uM). Since the enhancer DNA is likely to be fully saturated with protein at concentrations above 300 nM, the G266 substitutions may alter the stability of the nucleoprotein complexes, perhaps by influencing protomer interactions that impact upon the ATP hydrolysis site.

The GAFTGA variants can activate open complex formation *in vitro*

To further test the functionality of the G266 variants *in vitro*, we conducted assays to measure their ability to catalyse the conversion of the σ54-RNA polymerase closed complexes to open promoter complexes. Although NorR–DNA complexes exhibit heparin resistance, open promoter complexes can be visualized as heparin-resistant super-shifted species on non-denaturing gels (D'Autreaux *et al.*, 2005). In the presence of all the components required for open complex formation, the G266D and G266N variants were competent to form the super-shifted species, as in the case of NorRΔGAF (Fig. 4A compare lanes 3, 5, 7 and 9). Open complex formation was ATP-dependent as expected (Fig. 4A lanes 2, 4, 6 and 8). In order to probe the nature of the open complexes formed, we footprinted complexes with potassium permanganate, which targets cleavage to single stranded DNA regions, hence providing sequence-specific information. In all cases, we observed enhanced cleavage corresponding to T residues located between −11 and +1 in the *norV* promoter, consistent with the expected footprint. Notably, the band intensity observed with the G266 variants was decreased in comparison with NorRΔ-GAF or NorRΔGAF–His, perhaps reflecting the lower ATPase activities exhibited by the GAFTGA variants when compared with the wild type. These results confirm that the G266 variants are competent to interact with σ54 and can activate transcription *in vitro*, even though they exhibit altered ATPase activities.

Evidence for direct intramolecular interaction between the GAF domain and the σ54-interaction surface

From the biochemical results presented thus far, it seems likely that the GAFTGA mutations do not bypass intramolecular repression solely on the basis of changes in oligomerization state. To gain more insight into the nature of the interactions between the GAF and AAA+ domains, we followed a genetic suppression strategy. In previous work,

Fig. 4. Open promoter complex formation by AAA+ variants.
A. Heparin-resistant complexes formed by NorRΔGAF, NorRΔGAF–His, G266DΔGAF–His and G266NΔGAF–His, on the 361 bp DNA fragment carrying the *norR–norVW* intergenic region. In all cases, the final NorR concentration was 1500 nM. Reactions contained no NorR (lane 1), NorRΔGAF (lanes 2 and 3), NorRΔGAF–His (lanes 4 and 5), G266DΔGAF–His (lanes 6 and 7) and G266NΔGAF–His (lanes 8 and 9). Reactions loaded in lanes 1, 3, 5, 7 and 9 contained ATP (final concentration 5 mM), which was absent in lanes 2, 4, 6 and 8. Arrows indicate the position of free DNA, NorR bound DNA and the open promoter complexes.
B. Potassium permanganate footprinting of the 266 bp *norR–norVW* promoter fragment after open complex formation initiated by NorR. Lane 1 is a G+A ladder. Lane 2 is a control without activator present. Lanes 3, 4, 5 and 6 show footprinting after initiation of open complexes in the presence of 1 μM (final concentration) ΔGAF, ΔGAF–His, G266DΔGAF–His and G266NΔGAF–His respectively. The arrow marks the *norVW* transcriptional start and the positions of the enhanced cleavage at T bases are indicated.

mutagenesis of conserved residues in the GAF domain identified the R81L change that allows partial escape from interdomain repression in NorR (Tucker *et al.*, 2008). To further investigate the role of this residue in the regulation of AAA+ activity, a number of other changes were made at this position (Fig. S5). *In vivo* assays for transcriptional activation by NorR showed that the R81 residue is critical for the negative regulation of the AAA+ domain by the GAF domain. Hydrophobic changes (including R81L) result in significant constitutive activity. Negatively charged residues and serine substitutions not only prevent negative control but also stimulate NorR activity beyond wild-type levels. R81D, R81N and R81E give rise to twofold to threefold more activity than NorRΔGAF.

Since the R81 residue appears to be critical for interdomain repression, we decided to investigate whether R81 is required for positioning the GAF domain in the vicinity of the GAFTGA motif. We observed that the R81L substitution suppresses the constitutive activity of the G266D mutant so that repression of the AAA+ domain is almost completely restored (Fig. 5). Interestingly, the R81L mutation has a similar effect on other constitutively active variants located in the key region of the AAA+ domain that is predicted to undergo conformational changes upon ATP hydrolysis (Fig. S6A). As mentioned above, the Q304 residue is predicted to be at the base of helix 4 in the AAA+ domain of NorR and is not expected to
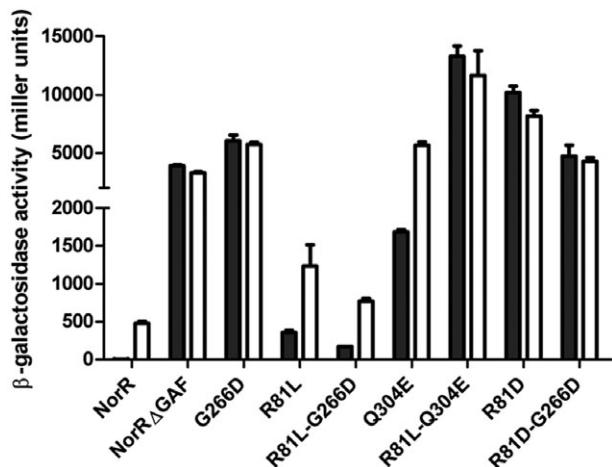
Fig. 5. Suppression of the G266D variant phenotype by the R81L mutation as measured by the *norV–lacZ* reporter assay *in vivo*. Substitutions are indicated on the *x*-axis. 'NorR' refers to the wild-type protein and 'NorRΔGAF' refers to the truncated form lacking the GAF domain (Δ1–170). Cultures were 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.
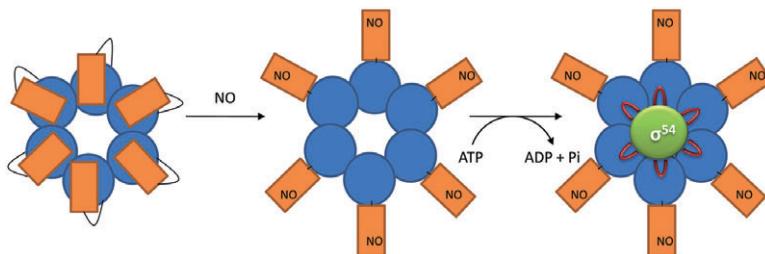
Fig. 6. Model for regulation of σ^{54}-dependent transcription by the EBP NorR. Binding of NorR to the *norR–norVW* intergenic region that contains the three NorR binding sites (not shown) is thought to facilitate the formation of a higher order oligomer that is most likely to be a hexamer (Tucker *et al.*, 2010). In the 'off' state, the N-terminal GAF domains (orange rectangles) negatively regulate the activity of the AAA+ domains (blue circles) by preventing access of the L1 and L2 loops to σ^{54} (left). In the 'on' state, NO binds to the iron centre in the GAF domain forming a mononitrosyl iron species. The repression of the AAA+ domain is relieved (centre), enabling ATP hydrolysis by NorR coupled to conformational changes in the AAA+ domain. During the nucleotide hydrolysis cycle, the surface-exposed loop that includes the GAFTGA motif moves into an extended conformation to allow σ^{54}-interaction and remodelling (right).

have a role in co-ordinating movements in the GAFTGA loop upon transition to the 'on' state. In accordance with this, the Q304E mutation was not suppressed by the R81L substitution. Instead, when combined with Q304E, the R81L substitution enabled complete escape from interdomain repression (Fig. 5).

Next, we wanted to determine whether the suppression of the G266D phenotype was dependent on the substitution made at the R81 position. Results show that only hydrophobic changes including R81L, V, I and F enable suppression (Fig. S7). It is possible that such changes introduce a new hydrophobic contact that helps restore interactions between the GAF and AAA+ domains. Western blotting analysis shows that reduction in activity is not due to a decrease in the stability of these double mutants (data not shown). Moreover other substitutions such as R81D have no effect on the constitutive activity of the G266D NorR variant (Fig. 5). Overall, the constitutive activity of the G266D variant and the specific suppression of this phenotype by hydrophobic changes at the R81 position suggest that the GAF domain may target the GAFTGA motif to prevent σ^{54} contact in the absence of the NO signal. Furthermore, the R81 residue is critical in maintaining repression and may be a key residue in mediating the transition from the 'off' to the 'on' state.

Discussion

The lack of NO-responsive regulation in truncated forms of NorR that lack the GAF domain (D'Autreaux *et al.*, 2005), clearly places NorR in the class of bEBPs that are negatively regulated. The substitutions we have identified in the AAA+ domain that bypass negative control by the GAF domain, cluster in regions that modulate the conformation of the σ^{54}-interaction surface or in the conserved GAFTGA motif itself. This invokes a model whereby the GAF domain negatively regulates

the AAA+ domain by preventing access of the L1 and L2 loops to σ^{54} (Fig. 6). This mode of repression might also serve to lock the loops in a restrained conformation that feeds back to the nucleotide binding site to prevent ATP hydrolysis. According to this model, substitutions in the σ^{54}-interaction surface bypass negative regulation either by altering the conformation of this surface to restrict access by the GAF domain or by directly disrupting GAF–AAA+ domain interactions. The alternative explanation that these substitutions bypass negative control by locking the AAA+ domain in a constitutive hexameric oligomerization state seems unlikely given that the GAFTGA substitutions exhibited no major changes in oligomerization properties when examined in the context of the NorR∆GAF protein. Although the full-length NorR apoprotein is competent to bind enhancer DNA this nucleoprotein complex is inactive with respect to ATP hydrolysis and transcriptional activation (D'Autreaux *et al.*, 2005). This suggests that in the absence of the NO signal, the GAF domain maintains the nucleoprotein complex in an inactive state by preventing access to σ^{54}-RNA polymerase.

It is remarkable that substitutions in the surface-exposed GAFTGA loop are able to prevent negative regulation by the GAF domain but still retain the ability to interact with σ^{54} and activate open complex formation. In the majority of bEBPs, substitutions in the GAFTGA motif cause a severe defect on the ability of the protein to activate transcription (Zhang *et al.*, 2002 and references cited therein). In PspF, the conserved threonine in this loop plays a critical role in contacting σ^{54} (Bordes *et al.*, 2003; Dago *et al.*, 2007) and only substitution of tyrosine for the highly conserved phenylalanine permits transcriptional activation (Zhang *et al.*, 2009). We observe that the equivalent aromatic substitution in NorR, F264Y, allows partial escape from repression by the GAF domain. Few studies have been carried out to explore the role of the second glycine in the GAFTGA

motif. In NtrC, the G219K variant has improved DNA binding activity and ATPase activity is 50% of the activated wild type (North *et al.*, 1996). However, it fails to activate transcription (a property exhibited by the equivalent mutation in NorR, G266K), suggesting that this mutation may prevent the interaction with σ^{54}. In contrast, the G219C variant of NtrC is competent to form open complexes but intriguingly can only do so in the absence of enhancer DNA (Yan and Kustu, 1999). This defect may be explained by changes in the relative juxtaposition of the DNA binding and ATPase domains observed during the ATPase cycle (De Carlo *et al.*, 2006). Overall, positively charged or aromatic residues are apparently not tolerated at this position in NorR, which may reflect a requirement for the σ^{54}-interaction.

The role of the σ^{54}-interaction surface in negative regulation by the GAF domain is further supported by our suppression data. The R81 residue in the GAF domain apparently plays a critical role in the mechanism of interdomain repression since an alanine substitution at this position leads to constitutive activation, whereas hydrophobic substitutions, particularly leucine, restore repression only when combined with specific bypass mutations in the AAA+ domain, including those in the GAFTGA loop. Structural modelling of the GAF domain, suggests that the R81 residue is surface-exposed (Tucker *et al.*, 2008). It is located at the opposite end of an α-helix to the R75 residue (Fig. S8), which is postulated to be a ligand to the hexa-coordinated iron and is the most suitable candidate to be displaced upon NO binding (Tucker *et al.*, 2008). Therefore, it is possible that formation of the mononitrosyl iron complex would displace the R75 ligand causing a conformational change in the helix that repositions R81. Interactions between the R81 residue and residue(s) in the AAA+ domain may thus facilitate the switch from the 'off' to the 'on' state.

The results presented here suggest a novel mechanism for negatively regulating bEBPs in which the σ^{54}-interaction surface is a target for repression, rather than the assembly of the active higher order oligomer. In the response regulator bEBPs NtrC1 and DctD, interactions between the receiver domain and the AAA+ domain maintain these proteins as inactive dimers in the absence of a regulatory signal. Extensive interdomain contacts, established via a coiled-coil linker, hold the ATPase domains in a dimeric front-to-front configuration to prevent oligomerization. Upon phosphorylation the dimerization interface in the receiver domain, which includes the coiled-coil linker is disrupted, allowing the AAA+ domain to reorient into the front-to-back configuration required for assembly into the active oligomeric ring (Lee *et al.*, 2003; Doucleff *et al.*, 2005). The linker region between the GAF and AAA+ domains of NorR is not predicted to form a coiled-coil helix, a structural feature that is also absent in negatively regulated NtrC4 and positively regulated NtrC (Batchelor *et al.*, 2008). NtrC4 has a partially disrupted receiver–AAA+ domain interface and can assemble into active oligomers at high protein concentrations independent of phosphorylation, a process that does not occur with NtrC1 (Batchelor *et al.*, 2008). The activated receiver domain has been shown to stabilize the hexameric form of NtrC4, thus functioning as an intermediate between the negative mechanism of NtrC1/DctD and positive mechanism of NtrC (Batchelor *et al.*, 2008; Batchelor *et al.*, 2009). In some bEBPs, the activity of the AAA+ domain is controlled by another regulatory protein, rather than by intramolecular repression (e.g. NifA, PspF, HrpR/S). In the case of PspF, which does not contain an amino-terminal regulatory domain, the activity of the AAA+ domain is negatively controlled by the PspA protein. In this case, repression is neither achieved by controlling the assembly of the ATPase subunits nor by preventing access of PspF to σ^{54}, but rather by inhibition of ATP hydrolysis (Joly *et al.*, 2009). Inhibition is mediated by the interaction of PspA with a surface-exposed tryptophan residue (W56) on PspF, which is likely to communicate with the ATP hydrolysis site. Structural studies have identified N64 in the AAA+ domain of PspF as being the key residue that translates nucleotide hydrolysis to conformational changes (Rappas *et al.*, 2006) and links ligand binding to ATPase activity (Zhang and Wigley, 2008). Although N64 variants are still able to bind PspA, their ATPase activity is no longer inhibited (Joly *et al.*, 2008) suggesting that negative regulation by PspA at the W56 residue is directly signalled to the nucleotide machinery via N64 to prevent ATPase hydrolysis by PspF. NorR represents another mechanism of negative regulation in which the N-terminal regulatory domain targets the σ^{54}-interacting region of the AAA+ domain that includes the GAFTGA motif. The evolutionary and physiological advantages of these different modes of regulation in bEBPs remain to be elucidated. In the case of NorR, we speculate that pre-assembly of an inactive oligomeric NorR species, poised as a nucleoprotein complex at the enhancer sites, enables the cell to rapidly respond to NO stress.

Experimental procedures

Plasmids and site directed mutagenesis

The pMJB1 plasmid was constructed from the pNorR plasmid (Tucker *et al.*, 2004) by making two silent mutations within the *norR* sequence. The C496T mutation produced the *Mfe*I/*Mun*I restriction site (CAATTG) upstream of the AAA+ domain and the G1341C mutation produced the *Sst*II restriction site (CCGCGG) downstream. In all other cases, targeted mutagenesis of the *norR* sequence was carried out using a PCR method (Ito *et al.*, 1991) with pMJB1 as a template.

Random mutagenesis

Random PCR mutagenesis was carried out with *Taq* DNA polymerase under standard reaction conditions. Reaction mixtures contained 75 ng of template pMJB1, 100 ng of each primer (AAA+Fwd 5′-GAAGAGCTACGGCTGATTGC-3′ and AAA+Rev 5′-GAACGCTTCTGTCGCTTCAC-3′), 0.2 mM dNTPs, 1.5 mM MgCl$_2$ and 5 units of enzyme in a final volume of 50 ml. The PCR products were purified, digested with *Mfe*I and *Sst*II and subsequently recloned into pMJB1 digested with the same enzymes. Ethanol precipitation followed by electroporation of the ligated mutant plasmid sample into DH5α was conducted and plasmid purification carried out before transformation of the sample into MH1003 (Hutchings *et al.*, 2002). Transformants were screened on Luria–Bertani (LB) supplemented with Xgal (40 μg ml^{-1}), chloramphenicol (30 μg ml^{-1}), carbenicillin (100 μg ml^{-1}) and kanamycin (50 μg ml^{-1}). Constitutive mutants were identified based on the ability of the *norR* gene to produce a protein that can activate expression of a *norV–lacZ* fusion. In the absence of inducer, constitutive mutants activate expression of β-galactosidase that cleaves the Xgal substrate to produce a blue product.

Assaying NorR activity *in vivo*

Transcriptional activation by NorR *in vivo* was measured by introducing wild-type and mutant plasmids into MH1003 a *nor*. : *cat* derivative of *E. coli* strain MC1000 with a *lacZ* reporter fusion to the *norVW* promoter inserted at the phage λ attachment site (Hutchings *et al.*, 2002). Cultures were grown with shaking in 50 ml of LB medium at 37°C until the OD650 reached 0.3, at which point glucose was added to the culture to a final concentration of 1%. Cultures were then split into 8 ml Bijou bottles and were grown anaerobically overnight at 37°C with or without potassium nitrite (4 mM). Under the latter conditions, NorR is activated by the NO that is generated endogenously by nitrite reduction in *E. coli* (Hutchings *et al.*, 2002). 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

Escherichia coli K12 NorRΔGAF was overexpressed and purified as described previously (D'Autreaux *et al.*, 2005). NorRΔGAF and mutant derivatives were additionally purified via an N-terminal TEV cleavable His-tag. Proteins were overexpressed from the pET-M11 construct but with the NcoI site altered to an NdeI site, to allow easy cloning of the *norR* sequence. BL21(DE3) transformed with the relevant construct was grown shaking at 250 r.p.m. at 30°C to an OD600 of 0.6. IPTG was then added to a final concentration of 1 mM and the cells left for 2–3 h before harvesting at 5000 r.p.m. Pellets were resuspended in buffer A (100 mM Tris-Cl, 50 mM NaCl, 50 mM imidazole, 5% glycerol, pH 8.5) containing EDTA-free protease inhibitors (Roche) and the cells were broken by French pressure disruption (1000 psi) in two passes. The insoluble material was then removed by centrifugation at 15 000 r.p.m. for 30 min. The clarified supernatant was loaded onto two 1 ml HiTrap chelating HP columns, connected in series and charged with 100 mM nickel chloride. The columns were equilibrated with NorR buffer A. Protein was then eluted using NorR buffer B (100 mM Tris-Cl, 50 mM NaCl, 500 mM imidazole, 5% glycerol, pH 8.5). To remove imidazole and to prevent precipitation, NorR containing fractions were loaded as quickly as possible onto a Superdex 200 16/60 column (Amersham Biosciences), pre-equilibrated in buffer C (100 mM Tris-Cl, 200 mM Nacl, 8 mM DTT, 5% glycerol). NorR containing fractions were concentrated using Amicon Ultra (Millipore) centrifugal devices with a 30 kDa MWC, aliquoted and stored in buffer containing 100 mM Tris-Cl, 100 mM NaCl, 4 mM DTT and 40% glycerol at −80°C, until required.

Open promoter complex and gel retardation assays

Open complex and gel retardation assays were carried out as described previously (Tucker *et al.*, 2010) using fragments derived from the pNPTprom plasmid that contains the *norR–norVW* region blunt-end cloned into the SmaI site of pUC19 (Tucker *et al.*, 2004).

Potassium permanganate footprinting of open complexes

Open complexes were probed using potassium permanganate as described previously (Whitehall *et al.*, 1992). Following potassium permanganate treatment, samples were resuspended in sodium acetate before ethanol precipitation. Samples were then subjected to chemical cleavage using the Maxam and Gilbert method. A G+A sequencing ladder was prepared by treatment with formic acid prior to the same cleavage treatment. The footprinting fragments were dried and dissolved in sequencing dye before being loaded on a sequencing gel.

ATPase assays

ATPase activities were measured using an assay in which production of ADP is coupled to the oxidation of NADH by lactate dehydrogenase and pyruvate kinase (Norby, 1988). The oxidation of NADH was monitored at 340 nm at 37°C. All reaction mixtures contained ATP (30 mM), phosphoenolpyruvate (1 mM), NADH (0.3 mM), pyruvate kinase (7 U, Roche), lactate dehydrogenase (23 U, Roche) in 50 mM Tris-HCl (pH 8.0), 100 mM KCl, 2 mM MgCl. Increasing volumes of NorR–His and its variants were added and the ATPase activity was measured by observing the change in absorbance at 340 nm. Total activity (μmol ATP min^{-1}) at each concentration was calculated using the equation: $[(\Delta OD340/\Delta t)/6220]*1 \times 10^6$ where t is the time-course of the experiment in minutes. Reactions were carried out both in the absence and presence of 5 nM of a 266 bp fragment of the *norR–norVW* intergenic region generated from the pNPTprom plasmid (Tucker *et al.*, 2004) using the norRpromF (5′-GGCGATATTCGCCAGCAC AT-3′) and norRpromR (5′-CGTTGACCAACCCAATGA ATGT-3′) primers.

Analytical gel filtration

Gel filtration chromatography of G266DΔGAF–His protein alone and in complex with a 266 bp DNA fragment, contain-

ing all three enhancer sites, was performed using a Superose 6 column (10 × 300 mm, 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

Samples (2 μl) from fractions eluted from gel filtration columns containing either G266DΔGAF–His alone or in complex with 266 bp dsDNA were absorbed onto glow-discharged continuous carbon grids (TAAB) and stained with 2% uranyl acetate. Data were collected at 35 000× magnification using a FEI Tecnai 12 electron microscope operating at 120 kV. Micrographs were recorded directly on a 1 k × 1 k CCD camera (TVIPS, Germany).

Acknowledgements

References

Batchelor, J.D., Doucleff, M., Lee, C.-J., Matsubara, K., De Carlo, S., Heideker, J., *et al.* (2008) Structure and regulatory mechanism of aquifex aeolicus NtrC4: variability and evolution in bacterial transcriptional regulation. *J Mol Biol* **384:** 1058–1075.

Batchelor, J.D., Sterling, H.J., Hong, E., Williams, E.R., and Wemmer, D.E. (2009) Receiver domains control the active-state stoichiometry of aquifex aeolicus [sigma]54 activator NtrC4, as revealed by electrospray ionization mass spectrometry. *J Mol Biol* **393:** 634–643.

Bordes, P., Wigneshweraraj, S.R., Schumacher, J., Zhang, X., Chaney, M., and Buck, M. (2003) The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: identifying a surface that binds sigma 54. *Proc Natl Acad Sci USA* **100:** 2278–2283.

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

Buck, M., Bose, D., Burrows, P., Cannon, W., Joly, N., Pape, T., *et al.* (2006) A second paradigm for gene activation in bacteria. *Biochem Soc Trans* **34:** 1067–1071.

Cannon, W.V., Gallegos, M.T., and Buck, M. (2000) Isomerization of a binary sigma-promoter DNA complex by transcription activators. *Nat Struct Biol* **7:** 594–601.

Chen, B., Doucleff, M., Wemmer, D.E., De Carlo, S., Huang, H.H., Nogales, E., *et al.* (2007) ATP ground- and transition states of bacterial enhancer binding AAA+ ATPases support complex formation with their target protein, σ^{54}. *Structure* **15:** 429–440.

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.

Dago, A.E., Wigneshweraraj, S.R., Buck, M., and Morett, E. (2007) A role for the conserved GAFTGA motif of AAA+

transcription activators in sensing promoter DNA conformation. *J Biol Chem* **282:** 1087–1097.

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. (2005) 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.

Gardner, A.M., Helmick, R.A., and Gardner, P.R. (2002) Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in *Escherichia coli*. *J Biol Chem* **277:** 8172–8177.

Gardner, A.M., Gessner, C.R., and Gardner, P.R. (2003) Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and sigma54 in the nitric oxide stress response. *J Biol Chem* **278:** 10081–10086.

Gomes, C.M., Giuffre, A., Forte, E., Vicente, J.B., Saraiva, L.M., Brunori, M., and Teixeira, M. (2002) A novel type of nitric-oxide reductase. *Escherichia coli* flavorubredoxin. *J Biol Chem* **277:** 25273–25276.

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.

Ito, W., Ishiguro, H., and Kurosawa, Y. (1991) A general method for introducing a series of mutations into cloned DNA using the polymerase chain reaction. *Gene* **102:** 67–70.

Joly, N., Burrows, P.C., and Buck, M. (2008) An intramolecular route for coupling ATPase Activity in AAA+ proteins for transcription activation. *J Biol Chem* **283:** 13725–13735.

Joly, N., Burrows, P.C., Engl, C., Jovanovic, G., and Buck, M. (2009) A lower-order oligomer form of phage shock protein A (PspA) stably associates with the Hexameric AAA(+) transcription activator protein pspF for negative regulation. *J Mol Biol* **394:** 764–775.

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.

Norby, J.G. (1988) Coupled assay of Na+,K+-ATPase activity. *Methods Enzymol* **156:** 116–119.

North, A.K., Weiss, D.S., Suzuki, H., Flashner, Y., and Kustu, S. (1996) Repressor forms of the enhancer-binding protein NtrC – some fail in coupling ATP hydrolysis to open complex-formation by sigma 54- holoenzyme. *J Mol Biol* **260:** 317–331.

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

Schumacher, J., Joly, N., Claeys-Bouuaert, I.L., Aziz, S.A., Rappas, M., Zhang, X., and Buck, M. (2008) Mechanism of homotropic control to co-ordinate hydrolysis in a hexameric AAA+ ring ATPase. *J Mol Biol* **381:** 1–12.

Studholme, D.J., and Dixon, R. (2003) Domain architectures of sigma54-dependent transcriptional activators. *J Bacteriol* **185:** 1757–1767.

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 σ54-dependent transcription by the nitric oxide sensing regulatory protein NorR. *Nucleic Acids Res* **38:** 1273–1283.

Wedel, A., and Kustu, S. (1995) The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes Dev* **9:** 2042–2052.

Whitehall, S., Austin, S., and Dixon, R. (1992) DNA supercoiling response of the sigma 54-dependent *Klebsiella pneumoniae nifL* promoter *in vitro. J Mol Biol* **225:** 591–607.

Yan, D., and Kustu, S. (1999) 'Switch I' mutant forms of the bacterial enhancer-binding protein NtrC that perturb the response to DNA. *Proc Natl Acad Sci USA* **96:** 13142–13146.

Zhang, X., and Wigley, D.B. (2008) The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. *Nat Struct Mol Biol* **15:** 1223–1227.

Zhang, X., Chaney, M., Wigneshweraraj, S.R., Schumacher, J., Bordes, P., Cannon, W., and Buck, M. (2002) Mechanochemical ATPases and transcriptional activation. *Mol Microbiol* **45:** 895–903.

Zhang, N., Joly, N., Burrows, P.C., Jovanovic, M., Wigneshweraraj, S.R., and Buck, M. (2009) The role of the conserved phenylalanine in the σ54-interacting GAFTGA motif of bacterial enhancer binding proteins. *Nucleic Acids Res* **37:** 5981–5992.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Transcriptional regulation by the dedicated nitric oxide sensor, NorR: a route towards NO detoxification

Matthew Bush*[1], Tamaswati Ghosh†, Nicholas Tucker*‡, Xiaodong Zhang† and Ray Dixon*

*Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Colney NR4 7UH, U.K., †Division of Molecular Bioscience, Imperial College London, London SW7 2AZ, U.K. and ‡Strathclyde Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral Street, Glasgow G4 0RE, U.K.

Abstract

A flavorubredoxin and its associated oxidoreductase (encoded by *norV* and *norW* respectively) detoxify NO (nitric oxide) to form N_2O (nitrous oxide) under anaerobic conditions in *Escherichia coli*. Transcription of the *norVW* genes is activated in response to NO by the σ^{54}-dependent regulator and dedicated NO sensor, NorR, a member of the bacterial enhancer-binding protein family. In the absence of NO, the catalytic activity of the central ATPase domain of NorR is repressed by the N-terminal regulatory domain that contains a non-haem iron centre. Binding of NO to this centre results in the formation of a mononitrosyl iron species, enabling the activation of ATPase activity. Our studies suggest that the highly conserved GAFTGA loop in the ATPase domain, which engages with the alternative σ factor σ^{54} to activate transcription, is a target for intramolecular repression by the regulatory domain. Binding of NorR to three conserved enhancer sites upstream of the *norVW* promoter is essential for transcriptional activation and promotes the formation of a stable higher-order NorR nucleoprotein complex. We propose that enhancer-driven assembly of this oligomeric complex, in which NorR apparently forms a DNA-bound hexamer in the absence of NO, provides a 'poised' system for transcriptional activation that can respond rapidly to nitrosative stress.

Introduction

NO (nitric oxide) is an intermediate of respiratory denitrification [1], and is one of the toxic species released by macrophages of the immune system in the defence against invading pathogenic bacteria [2]. As a consequence, bacteria have evolved mechanisms in order to survive nitrosative-induced stress. The *Escherichia coli* flavorubredoxin and its associated oxidoreductase function under anaerobic conditions to convert NO into N_2O (nitrous oxide) and are encoded by the *norV* and *norW* genes respectively [3]. Expression of these genes is controlled by *norR*, which is divergently transcribed upstream of the *norVW* transcriptional unit. NorR is a bEBP (bacterial enhancer-binding protein) of the AAA$^+$ (ATPase associated with various cellular activities) class of proteins and activates the σ^{54}-dependent transcription of *norVW* in response to NO [4,5]. NorR and the global regulator, NsrR, are the only known dedicated NO sensors [6].

In contrast to the housekeeping σ^{70}-class of bacterial σ factors, transcription by the σ^{54}-class requires the bEBP to use the energy generated from ATP hydrolysis to drive

conformational rearrangements in the σ^{54}-RNA polymerase holoenzyme. As is the case for other AAA$^+$ proteins, bEBPs are competent to hydrolyse ATP when assembled as a hexamer, since the catalytic site is formed by residues from adjacent protomers ([7] and references therein). The central (AAA$^+$) domain (Figure 1A) contains specific structural features that enable nucleotide-dependent interactions with σ^{54}. The highly conserved GAFTGA motif forms part of a surface-exposed loop that contacts σ^{54} during the ATP hydrolysis cycle [8–11]. In the ATP-bound transition state, the GAFTGA-containing loop 1, assisted by a second loop (loop 2) is in an extended conformation and is therefore able to establish contact with region I of the alternative σ factor, σ^{54}. After phosphate release, the surface-exposed loops compact down towards the surface of the AAA$^+$ domain, enabling relocation of the σ factor and formation of the open promoter complex [12–14].

Many bEBPs have an additional N-terminal regulatory domain (Figure 1A) that stringently controls the activity of the AAA$^+$ domain either positively or negatively in response to various environmental cues [15]. NorR contains an N-terminal regulatory GAF (cGMP-specific and -stimulated phosphodiesterases, *Anabaena* adenylate cylases and *Escherichia coli* FhlA) domain that has been predicted to bind NO. It has previously been shown that an N-terminally truncated form of NorR, lacking the regulatory GAF domain (NorRΔGAF), is competent to activate transcription in the absence of NO [4,16]. Therefore NorR falls into a

Figure 1 | Structural features of NorR

(**A**) Domain architecture of the bEBP NorR showing the N-terminal regulatory GAF domain (purple) containing a non-haem iron centre, the central ATPase-active domain (red) and the C-terminal DNA-binding domain (green) that contains an HTH motif. (**B**) Proposed model of the NO-sensing non-haem iron centre in the NorR regulatory domain. Structural model of the GAF domain of NorR based on the GAF-B domain of 3′,5′-cyclic nucleotide phosphodiesterase [26] showing the iron centre (magenta) and proposed ligands Cys[113], Asp[96], Asp[99], Arg[75] and Asp[131] (labelled). The Arg[75] residue is the most likely to undergo ligand displacement upon NO binding. Also shown is the Arg[81] residue, at the opposite end of an α-helix that also contains the Arg[75] ligand. The model predicts that Arg[81] is surface-exposed and well placed to make contact with the NorR AAA$^+$ domain. (**C**) The σ^{54}-interaction surface of the AAA$^+$ domain of NorR is the target of GAF-mediated repression. Structural model of the AAA$^+$ domain of NorR based on the NtrC1
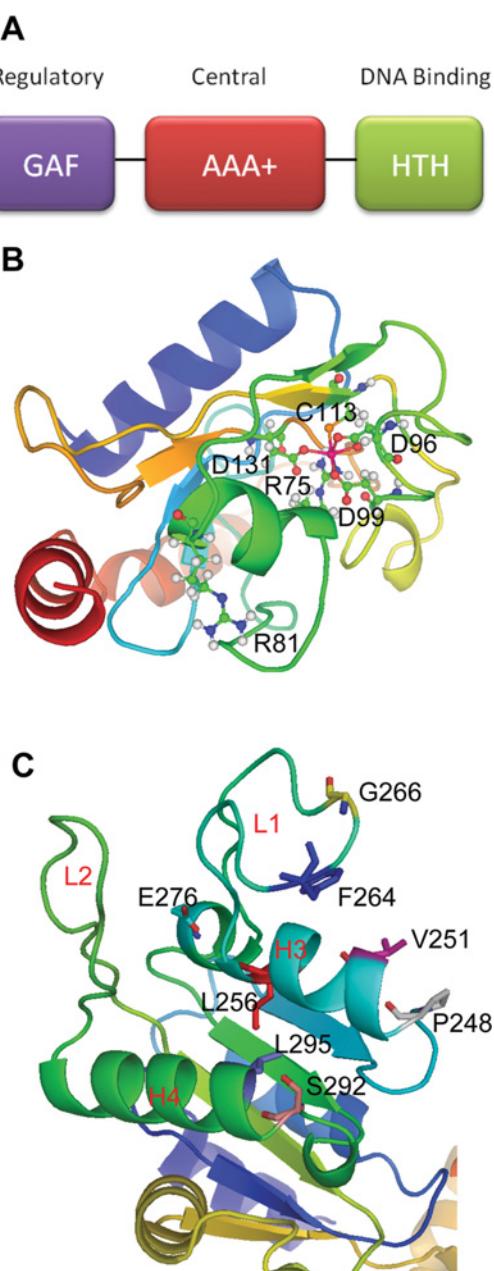
structure [23] (PDB code 1NY5, chain A). The helices and loops (H3 and H4, L1 and L2) involved in nucleotide-dependent conformational changes in bEBPs are labelled in red. Residues that were substituted as a consequence of the PCR mutagenesis of the AAA$^+$ domain are indicated. The Phe[264] and Gly[266] residues form part of the GAFTGA motif that contacts σ^{54}.

category of bEBPs in which the activity of the central AAA$^+$ domain is negatively regulated by the N-terminal domain [17]. In the absence of NO, the GAF domain represses the activity of the AAA$^+$ domain via interdomain repression [4]. The detection of the NO signal causes this repression to be relieved, allowing ATP hydrolysis by NorR and thus driving transcriptional activation. Other activators of σ^{54}-dependent transcription that are controlled by similar mechanisms of interdomain repression include XylR [18,19], DmpR [20,21], DctD [22] and NtrC1 [23].

In addition to the central (AAA$^+$) and regulatory domains, bEBPs also have a C-terminal HTH (helix–turn–helix) domain (Figure 1A) that binds to conserved enhancer-like sequences 80–150 bp upstream of the bacterial promoter. This ensures that activation of transcription occurs only at the specific promoter(s) with which the bEBP can associate. Interactions between the bEBP upstream of the promoter and the holoenzyme at the transcriptional start site are facilitated by DNA bending, assisted by IHF (integration host factor) [18,24].

In the present paper, we review recent developments in our understanding of the mechanism of NO-dependent activation of transcription by NorR. We address various aspects of this mechanism including: (i) a novel mechanism of NO sensing by the GAF domain, (ii) how the DNA-binding properties of the activator control NorR activity, and (iii) how the GAF domain negatively regulates the AAA$^+$ domain.

Mechanism of NO sensing by NorR

In order to investigate the mechanism of NO sensing in NorR, EPR spectroscopy was carried out on whole cells of *E. coli* exposed to NO [25]. A new EPR signal was observed in the $g = 4$ region only when the cells expressed NorR and were exposed to NO. This indicates that NorR contains a non-haem iron centre, since similar spectra have been observed for several non-haem iron enzymes when exposed to NO ([25] and references [15–19] therein). This characteristic EPR signal was observed in cells expressing the isolated GAF$_{NorR}$ (GAF domain of NorR) but not in cells that expressed a form of the protein lacking the regulatory domain (NorRΔGAF), indicating that the non-haem iron centre is present within the N-terminal GAF domain. Purification and reconstitution of NorR and GAF$_{NorR}$ with ferrous iron gave identical NO spectra to those observed with whole-cell EPR, confirming that the NorR GAF domain contains the non-haem iron centre. This is the first example of a GAF domain using a transition metal as a mechanism of sensing and, to our

knowledge, reveals a novel biological role for the activation of a non-haem iron centre to form a high-spin $\{Fe(NO)\}^7$ ($S = 3/2$) complex.

The spectroscopic features of this paramagnetic mononitrosyl iron complex suggest that the iron centre has distorted octahedral symmetry and is co-ordinated by five or six ligands within the GAF domain. In order to study the co-ordination of the iron centre in NorR, targeted mutagenesis was carried out at conserved residues within the regulatory domain [26]. As a result, five candidate ligands were proposed: Asp^{99}, Asp^{131}, Cys^{113}, Arg^{75} and Asp^{96} (Figure 1B). Variant forms of NorR containing substitutions at these positions gave proteins that were unable to bind iron or did not exhibit the characteristic $g=4$ EPR signal after reconstitution *in vitro*. Therefore these residues are likely to have a role in iron co-ordination. These EPR data are in agreement with our hexa-co-ordinated structural model (Figure 1B). In the model, Asp^{96} is proposed to be a bidentate ligand. Although it is possible that a water molecule instead provides a sixth ligand, we believe this is unlikely since the iron-binding site appears to be solvent inaccessible. Although, arginine is not an ideal ligand for transition metals, examples have been reported elsewhere such as for biotin synthase [27]. The predicted hexa-co-ordination of the iron centre suggests that one of the five predicted ligands would need to be displaced in order to form the mononitrosyl iron complex. Arg^{75} is the most likely candidate to relinquish a binding site for NO and may also stabilize the NO-bound form of the iron through hydrogen bonding [26].

Role of enhancer-DNA in NorR-dependent activation of transcription

In previous work, purified NorR has been shown to bind to three sites in the *norVW* promoter region [28]. To assess the importance of each of these sites, the enhancers were individually altered from the consensus GT-(N$_7$)-AC to GG-(N$_7$)-CC and introduced upstream of *norV–lacZ* promoter fusions on the *E. coli* chromosome. Disruption of any one of the three sites completely abolished the ability of NorR to activate transcription of *norVW in vivo* [29]. Biochemical experiments have demonstrated that the ATPase activity of NorR is dependent not only on the presence of NO but also on the enhancer DNA that contains the three NorR-binding sites [25]. In the absence of the regulatory GAF domain (NorRΔGAF), the requirement for the NO signal is relieved, but enhancer DNA is still required to stimulate activity. When any of the three binding sites was individually altered from the consensus, the enhancer-dependent ATPase activity of NorRΔGAF was significantly diminished [29]. Efficient open complex formation by NorR *in vitro* also required the three enhancer sites, consistent with the requirement for bEBPs to use the energy from ATP-hydrolysis to remodel the σ^{54}-RNA polymerase holoenzyme [29]. The prerequisite for the three enhancer sites for transcriptional activation by NorR both *in vivo* and *in vitro* might reflect a requirement for three NorR dimers to assemble to form an inactive hexamer

on the promoter DNA, given the dyad symmetry present at each of the sites. However, binding of NorR to a 21 bp sequence encoding one of the enhancer sites (NorR site 1) stimulated both the ATPase activity and oligomerization state to a certain extent, indicating that DNA binding *per se* promotes self-association and ATPase activity. Binding to a 66 bp DNA fragment that contained all three enhancer sites stimulated ATPase activity and oligomerization further, although not to the levels observed when a longer 266 bp fragment containing the intergenic region was used. This implies that the DNA flanking the enhancer sites has an important role in stabilizing the NorR oligomer, possibly by wrapping around the hexamer. In agreement with this, EMSAs (electrophoretic mobility-shift assays) revealed a significant increase in the affinity and co-operativity of binding when the longer DNA fragment was present. Furthermore, negative-stain electron microscopy revealed the formation of protein–DNA complexes with the expected diameter of a NorR hexamer in the presence of the 266 bp DNA, but not with the 66 bp or 21 bp fragments. Overall, these data support a model in which three NorR dimers bind to the enhancer sites, inducing conformational changes that stimulate the formation of a higher-order oligomer, most probably a hexamer. The results suggest that this higher-order NorR species is stabilized by extensive DNA interactions, possibly by wrapping around the hexamer to form a stable nucleoprotein complex.

Negative regulation of NorR activity

To investigate the mechanism of interdomain repression in NorR, we used a random mutagenic approach to search for mutations in the AAA$^+$ domain that enable escape from GAF domain-mediated repression [30]. This approach generated mutant versions of NorR that were able to activate the transcription of a *norV–lacZ* fusion *in vivo* in the absence of NO, generated endogenously by the addition of potassium nitrite. In some cases, the NorR variants exhibited activity in the absence of NO signal that was similar to a truncated version of NorR lacking the GAF domain (NorRΔGAF). This phenotype indicates a complete bypass of the repression of AAA$^+$ activity by the GAF domain. In structural models of the NorR AAA$^+$ domain based on the crystal structure of the NtrC1 AAA$^+$ domain (Figure 1C), the substitutions identified are located in H3 (helix 3), H4 (helix 4) or L1 (loop 1). This is the region of the AAA$^+$ domain that undergoes nucleotide-dependent conformational changes before engagement with σ^{54} [9,12]. Significantly, two bypass mutations were identified in the GAFTGA motif, within surface-exposed loop 1. The G266D and G266N mutations within the second glycine of this motif enabled complete escape from the negative control exerted by the GAF domain. Furthermore, when additional substitutions were made to disrupt NO signalling at the iron centre, both GAFTGA variants retained the ability to activate transcription *in vivo*, suggesting that NO sensing is not required for their activity. The ability of these variants to

Figure 2 | Model of NorR-dependent activation of *norVW*

(**A**) Binding of NorR to the *norR–norVW* intergenic region that contains the three NorR binding sites (1, 2 and 3, highlighted in red) is thought to facilitate the formation of a higher-order oligomer that is most likely to be a hexamer [29]. (**B**) Although the hexamer is bound to DNA (not shown), in the absence of NO, the N-terminal GAF domains (blue rectangles) negatively regulate the activity of the AAA⁺ domains (green circles) by preventing access of the surface-exposed loops to σ^{54}. (**C**) In the 'on' state, NO binds to the iron centre in the GAF domain forming a mononitrosyl iron species. The repression of the AAA⁺ domain is relieved, enabling ATP hydrolysis by NorR coupled to conformational changes in the AAA⁺ domain. (**D**) During the nucleotide hydrolysis cycle, the surface-exposed loops (red) that include the GAFTGA motifs move into an extended conformation to allow σ^{54}-interaction and remodelling.
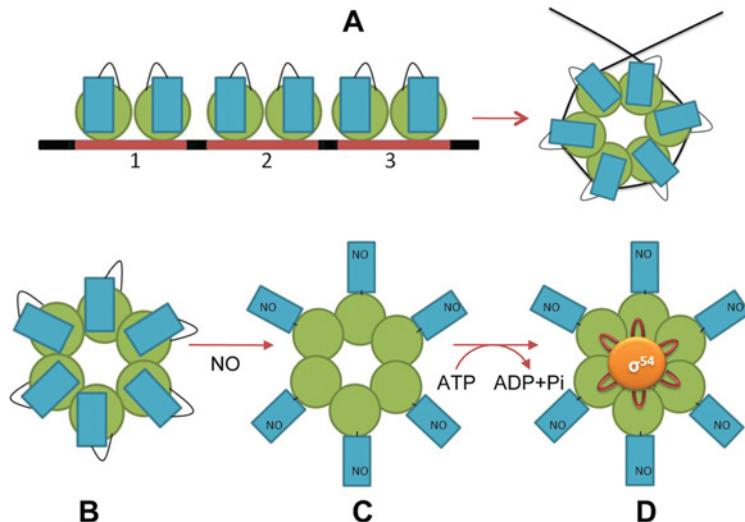
activate transcription is surprising given the high conservation of the GAFTGA motif and that substitutions at this position are likely to affect the conformational flexibility of the loop.

In several bEBPs (e.g. NtrC1 and DctD), the N-terminal regulatory domain regulates the activity of the AAA⁺ domain by controlling the oligomerization state of the activator [22,23]. In order to further study how the Gly²⁶⁶ variants escape repression by the GAF domain, the variant proteins were purified using an N-terminal His-tag. DNA-binding studies demonstrated that both wild-type NorR and the G266D and G266N substitutions had similar affinities for the *norVW* promoter DNA (K_d ∼100 nM) indicating that these Gly²⁶⁶ substitutions do not influence binding of NorR to the three enhancer sites. Furthermore, gel filtration studies indicated that, like NorR, the Gly²⁶⁶ variants eluted at a volume corresponding to a monomer or dimer in the absence of DNA, but formed stable higher-order oligomers in the presence of the promoter DNA. Analysis of eluted fractions by negative-stain electron microscopy showed the formation of particles of the size expected for a NorR hexamer. Furthermore, the ATPase activity of the substitutions was enhancer DNA dependent as is the case with wild-type NorR. Since the GAFTGA substitutions do not exhibit any changes in enhancer-dependent oligomerization, it seems unlikely that interdomain repression by the regulatory GAF domain is exerted by changing the oligomerization state of the AAA⁺ domain. Rather, the properties of the bypass mutations in the GAFTGA loop suggest that the regulatory

domain negatively regulates the activity of the AAA⁺ domain by preventing access of the surface-exposed L1 and L2 loops to σ^{54} (Figure 2). This model is further supported by genetic suppression studies. Previous work identified the R81L substitution that results in partial escape from GAF-mediated repression [26]. Targeted mutagenesis at the Arg⁸¹ position confirms that this residue is critical in maintaining the mechanism of interdomain repression. Our structural model (Figure 1B) places it on the surface of the GAF domain, at the opposite end of an α-helix that also contains the Arg⁷⁵ residue. Displacement of the Arg⁷⁵ ligand upon NO binding would lead to significant conformational changes along this helix and so the Arg⁸¹ residue may have a key role in the transmission of the NO signal to the AAA⁺ domain. Significantly, hydrophobic changes at Arg⁸¹ are able to specifically suppress the escape phenotype of AAA⁺ bypass variants including the GAFTGA mutant G266D. This suggests that the GAF domain targets the GAFTGA motif to prevent interaction with σ^{54} in the absence of NO signal. Binding of NO to form the mononitrosyl iron complex is expected to lead to conformational changes in the GAF domain which release repression, allowing the GAFTGA motif to contact the σ factor.

Conclusions

Our current model for the NO-dependent activation of *norVW* transcription by NorR is shown in Figure 2. The

work reviewed in the present paper suggests that enhancer DNA induces conformational changes in NorR that allow the formation of an inactive hexamer. This is in stark contrast to bEBPs such as NtrC1 where the signal drives the process of oligomerization [23]. We suggest that in the case of NorR, the pre-formation of a NorR hexamer on enhancer DNA, 'poised' to respond to NO, allows the cell to rapidly respond to NO-induced stress. We have shown that this dedicated NO sensor detects the signal through the formation in a novel mononitrosyl complex at the non-haem iron centre of the regulatory domain. Furthermore, we have identified a novel mechanism of interdomain repression in which the regulatory domain targets the σ^{54}-interaction surface in the absence of NO to prevent ATP hydrolysis.

Funding

References

1 Zumft, W. (1997) Cell biology and molecular basis of denitrification. Microbiol. Mol. Biol. Rev. **61**, 533–616

2 MacMicking, J., Xie, Q.-W. and Nathan, C. (1997) Nitric oxide and macrophage function. Annu. Rev. Immunol. **15**, 323–350

3 Gardner, A.M., Helmick, R.A. and Gardner, P.R. (2002) Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in *Escherichia coli*. J. Biol. Chem. **277**, 8172–8177

4 Gardner, A.M., Gessner, C.R. and Gardner, P.R. (2003) Regulation of the nitric oxide reduction operon (norRVW) in *Escherichia coli*: role of NorR and σ^{54} in the nitric oxide stress response. J. Biol. Chem. **278**, 10081–10086

5 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

6 Bodenmiller, D.M. and Spiro, S. (2006) The yjeB (nsrR) gene of *Escherichia coli* encodes a nitric oxide-sensitive transcriptional regulator. J. Bacteriol. **188**, 874–881

7 Rappas, M., Bose, D. and Zhang, X. (2007) Bacterial enhancer-binding proteins: unlocking σ^{54}-dependent gene transcription. Curr. Opin. Struct. Biol. **17**, 110–116

8 Zhang, X., Chaney, M., Wigneshweraraj, S.R., Schumacher, J., Bordes, P., Cannon, W. and Buck, M. (2002) Mechanochemical ATPases and transcriptional activation. Mol. Microbiol. **45**, 895–903

9 Bordes, P., Wigneshweraraj, S.R., Schumacher, J., Zhang, X., Chaney, M. and Buck, M. (2003) The ATP hydrolyzing transcription activator phage shock protein F of *Escherichia coli*: identifying a surface that binds σ^{54}. Proc. Natl. Acad. Sci. U.S.A. **100**, 2278–2283

10 Rappas, M., Schumacher, J., Beuron, F., Niwa, H., Bordes, P., Wigneshweraraj, S., Keetch, C.A., Robinson, C.V., Buck, M. and Zhang, X. (2005) Structural insights into the activity of enhancer-binding proteins. Science **307**, 1972–1975

11 Schumacher, J., Joly, N., Rappas, M., Zhang, X. and Buck, M. (2006) Structures and organisation of AAA$^+$ enhancer binding proteins in transcriptional activation. J. Struct. Biol. **156**, 190–199

12 Rappas, M., Schumacher, J., Niwa, H., Buck, M. and Zhang, X. (2006) Structural basis of the nucleotide driven conformational changes in the AAA$^+$ domain of transcription activator PspF. J. Mol. Biol. **357**, 481–492

13 Chen, B., Doucleff, M., Wemmer, D.E., De Carlo, S., Huang, H.H., Nogales, E., Hoover, T.R., Kondrashkina, E., Guo, L. and Nixon, B.T. (2007) ATP ground- and transition states of bacterial enhancer binding AAA$^+$ ATPases support complex formation with their target protein, σ^{54}. Structure **15**, 429–440

14 Bose, D., Joly, N., Pape, T., Rappas, M., Schumacher, J., Buck, M. and Zhang, X. (2008) Dissecting the ATP hydrolysis pathway of bacterial enhancer-binding proteins. Biochem. Soc. Trans. **36**, 83–88

15 Studholme, D.J. and Dixon, R. (2003) Domain architectures of σ^{54}-dependent transcriptional activators. J. Bacteriol. **185**, 1757–1767

16 Pohlmann, A., Cramm, R., Schmelz, K. and Friedrich, B. (2000) A novel NO-responding regulator controls the reduction of nitric oxide in *Ralstonia eutropha*. Mol. Microbiol. **38**, 626–638

17 Shingler, V. (1996) Signal sensing by σ^{54}-dependent regulators: derepression as a control mechanism. Mol. Microbiol. **19**, 409–416

18 Perez-Martin, J. and Lorenzo, V. (1995) The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. Proc. Natl. Acad. Sci. U.S.A. **92**, 9392–9396

19 Fernandez, S., de Lorenzo, V. and Perez-Martin, J. (1995) Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. Mol. Microbiol. **16**, 205–213

20 Shingler, V. and Pavel, H. (1995) Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. Mol. Microbiol. **17**, 505–513

21 Wikstrom, P., O'Neill, E., Ng, L.C. and Shingler, V. (2001) The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation. J. Mol. Biol. **314**, 971–984

22 Doucleff, M., Chen, B., Maris, A.E., Wemmer, D.E., Kondrashkina, E. and Nixon, B.T. (2005) 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

23 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

24 Hoover, T.R., Santero, E., Porter, S. and Kustu, S. (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. Cell **63**, 11–22

25 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

26 Tucker, N.P., D'Autreaux, B., Yousafzai, F., Fairhurst, S., Spiro, S. and Dixon, R. (2007) Analysis of the nitric oxide-sensing non-heme iron center in the NorR regulatory protein. J. Biol. Chem. **283**, 908–918

27 Berkovitch, F., Nicolet, Y., Wan, J.T., Jarrett, J.T. and Drennan, C.L. (2004) Crystal structure of biotin synthase, an *S*-adenosylmethionine-dependent radical enzyme. Science **303**, 76–79

28 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

29 Tucker, N.P., Ghosh, T., Bush, M., Zhang, X. and Dixon, R. (2010) Essential roles of three enhancer sites in σ^{54}-dependent transcription by the nitric oxide sensing regulatory protein NorR. Nucleic Acids Res. **38**, 1182–1194

30 Bush, M., Ghosh, T., Tucker, N., Zhang, X. and Dixon, R. (2010) Nitric oxide-responsive interdomain regulation targets the σ^{54}-interaction surface in the enhancer binding protein NorR. Mol. Microbiol. **73**, 1278–1288