Functional analysis of Rex, a sensor of the NADH/NAD^+ redox poise in *Streptomyces coelicolor*

Claire Michelle Strain-Damerell

Thesis submitted for the degree of Doctor of Philosophy

September 2010

Department of Biochemistry
School of Life Sciences
University of Sussex
Falmer
East Sussex
BN1 9QG

DECLARATION

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signed…….....

Claire M. Strain-Damerell

UNIVERSITY OF SUSSEX

Claire Michelle Strain-Damerell

Thesis submitted for the degree of Doctor of Philosophy

Functional analysis of Rex, a sensor of the NADH/NAD^+ redox poise in *Streptomyces coelicolor*

SUMMARY: Maintenance of the intracellular NADH/NAD^+ redox poise is vital for energy generation in cells. Gram-positive bacteria, including the antibiotic-producing organism, *Streptomyces coelicolor,* have evolved a regulatory protein Rex that both senses this ratio and mediates an adaptive response to changes in it. Rex is a dimeric redox-sensitive transcriptional repressor. It is capable of binding to both NAD^+ and NADH, although only NADH is an effector, causing dissociation of the protein from operator (ROP) sites. As NADH levels rise during oxygen limitation Rex dissociates from its target genes allowing expression, which helps to restore the NADH/NAD^+ ratio. Microarray-based expression studies had suggested that Rex regulated only a small number of genes. In this work, however, ChIP-on-chip analyses revealed 38 genes that are potential regulon members. Analysis of the Rex binding sites in *S. coelicolor* revealed new insights into the mode of binding and show that Rex can bind with low affinity to incomplete half sites. This work also focused on characterising two key Rex targets, *ndh* and *nuoA-N*, that encode non-proton-translocating and proton translocating NADH dehydrogenases, respectively. Whereas *nuoA-N* is not essential and was not expressed in liquid media, *ndh* was essential for growth. Depletion of NDH from growing cells led to the induction of Rex target genes confirming that *ndh* and Rex play key roles in maintaining redox homeostasis. Structure-based dissection of Rex, via a close homologue in *Thermus aquaticus*, identified a key interaction between the NADH- and DNA-binding domains of Rex. An R29-D203' salt-bridge, that traverses the NADH binding and DNA binding domains of Rex, appeared to stabilise the DNA-bound form of Rex, but is 'broken' in the presence of NADH. In the NADH-bound form of Rex, D203 alternatively interacts with Y111, which in turn interacts with the nicotinamide ring of NADH. In order to assess the importance of individual subunits in the dimeric Rex, a single-chain derivative was constructed and the NADH binding and DNA binding domains individually disrupted.

Acknowledgements

First and foremost I wish to thank my supervisor; Dr Mark Paget. He has tested my limitations and helped me overcome them, yet shown great patience with me throughout. I thank him for giving me the opportunity to work in his lab and with him. I also wish to thank my co-supervisor; Prof. Andrew Smith for his support and guidance throughout.

I thank my fellow Paget lab members past and present; Dr Phil Doughty, Dr Ben Pascoe, Dr Dimitris Kallifidas, Aline Tabib and Elodie Bentley. I have learnt from and laughed with each of them. I wish to especially thank Dr Dimitris Brekasis, although we did not always see eye-to-eye I greatly respect all the work that he has done in making the Rex project what it is today.

I would like to thank Lin Cox, our lab technician, for all those times she held the autoclave for that one last buffer of mine. She has been invaluable to our group. I must also thank Dr Giselda Bucca, Prof. Colin Smith and the rest of the team at the University of Surrey microarray facility for their assistance during this project. I also thank Dr Clara Kielkopf for her collaborations on the structural aspects of the project.

I thank the BBSRC for my funding and the University of Sussex for my DPhil position.

I wish to thank my family and friends for their support throughout, especially my dad; Andrew Strain and my mum; Angela Strain. They put me through university and encouraged me every step of the way. I also thank my in-laws; Tony and Jayne Damerell, for all that they have done for both David and I over the years.

Last but not least, I thank my considerably better half; David Damerell. He has been with me since my first day of university and is the reason I am here today. I thank him for always being there for me - no mean feat when he has also had the pressure of a DPhil thesis to deal with. I really couldn't have done this without him.

Publication

McLaughlin KJ*, Strain-Damerell CM*, Xie K, Brekasis D, Soares AS, Paget MSB and Kielkopf CL (2010) Structural Basis for NADH/NAD^+ Redox Sensing by a Rex Family Repressor. Mol Cell 38(4): 563-575.
(* These authors contributed equally to this work)

Contents

List of Figures and Tables

Abbreviations

Apr^R	Apramycin resistant
APS	Ammonium persulphate
ATP	Adenosine 5'-triphosphate
BLAST	Basic local alignment tool
bp	Base pair
B-Rex	*Bacillus subtilis* Rex
CBS	Cystathionine beta synthase
ChIP	Chromatin immunoprecipitation
Da	Daltons
(d)dH_2O	(Double) distilled water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
EMSA	Electromobility shift assay
g	Grams
GAF	cGMP-specific phosphodiesterases, adenylyl cyclases and FhlA
h	Hour(s)
IC_{50}	Inhibitory concentration that results in a 50% reduction in binding
Kan^R	Kanamycin resistant
Kan^S	Kanamycin sensitive
K_d	Dissociation constant
KEGG	Kyoto Encyclopaedia of Genes and Genomes
L	Litre
min	Minute(s)
NAD^+	Nicotinamide adenine dinucleotide, oxidised form
NADH	Nicotinamide adenine dinucleotide, reduced form
OD	Optical density (wavelength indicated in subscript)
PAS	Per, Arnt and Sim
PCR	Polymerase chain reaction
PDB	Protein data bank
PMF	Proton motive force
qPCR	Quantitative PCR
ROP	Rex operator
RT-qPCR	Reverse transcriptase qPCR
sec	Second(s)
SPR	Surface Plasmon resonance
S-Rex	*Streptomyces coelicolor* Rex
StrepDB	The Streptomyces Annotation Server
$Thio^R$	Thiostrepton resistant
T-Rex	*Thermus aquaticus* Rex
U	Unit(s)
WT	Wild type

Chapter 1

Introduction

"The greater our knowledge increases, the greater our ignorance unfolds."

John F. Kennedy (1917-1963)

Section 1.1 – Overview

Free-living bacteria are faced with a range of environmental challenges, necessitating the need for sensory regulatory pathways in order to switch on adaptive responses. Nutrient and oxygen limitation are classical problems faced by such organisms, requiring global changes in the metabolic profiles in order to thrive in these conditions. Such changes in growth conditions are detected by a wide variety of sensors that somehow transduce this signal to alter gene expression. The primary sensor and regulator might be encoded on a single polypeptide chain, or they might be encoded on separate proteins. This study focuses on the Rex regulator of *Streptomyces coelicolor* that both senses changes in the NADH/NAD^+ ratio that accompany oxygen limitation, and controls the transcription of target genes in response to these changes. This chapter discusses the biological context of this protein, along with its previous characterisation and gene target identification.

Section 1.2 – *Streptomyces coelicolor*

The actinomycete family of bacteria includes several medically and industrially important genera including the mycobacteria, corynebacteria, and the streptomycetes. The *Mycobacterium* genus is probably best known for *Mycobacterium tuberculosis*, a pathogen responsible for infecting one person per second and causing >1 million deaths per year. Corynebacteria, particularly *Corynebacterium glutamicum*, are extensively used in industry as they are fast growing and are capable of fermentation. They are most notably used for L-amino acid production. The *Streptomyces* genus is unusually morphologically and physiologically complex. Members have intricate colony structures with hyphae delving deep into the substrate and spore chains reaching into the air. Importantly, they produce an array of secondary metabolites, e.g. chloramphenicol (antibacterial), daunorubicin (antitumor) and rapamycin (immunosuppressant). *S. coelicolor* itself is so named for the blue colour of its most noted secondary metabolite, actinorhodin, with '*coelicolor*' translating as

'*sky colour*' in latin. The following section provides an overview of the genome structure, life cycle and metabolism of *S. coelicolor*.

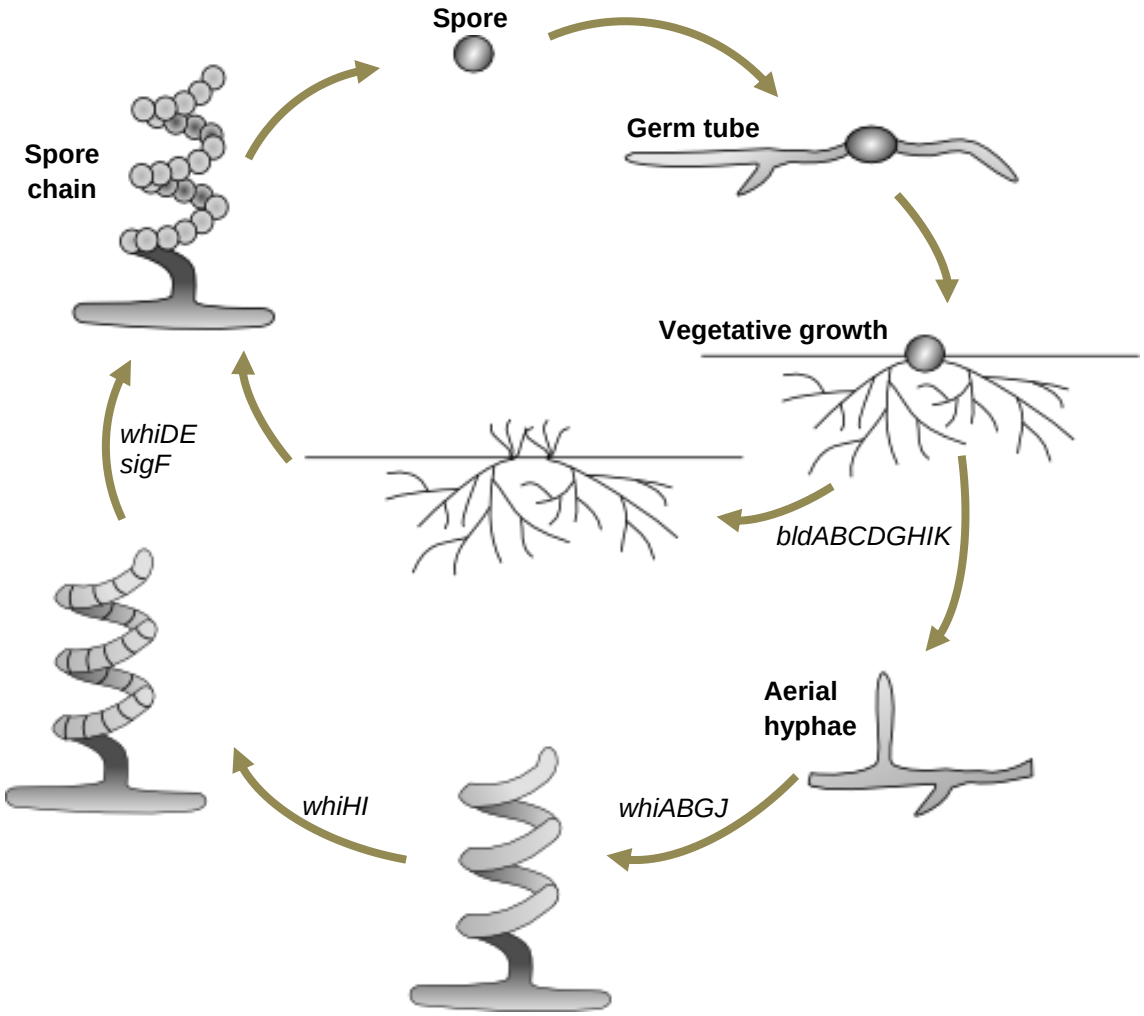


Figure 1.1: *Streptomyces coelicolor* life cycle. The involvement of different regulatory genes at each stage is also shown. Figure adapted from that of Kieser *et al.* (Kieser *et al.*, 2000).

Section 1.2.1 – Life-cycle

The life-cycle of *S. coelicolor* is complex (Figure 1.1), with aspects often more familiar to fungal or even plant development, and is heavily regulated (Flärdh and Buttner, 2009). Starting with germination, a single spore will produce

multiple germ tubes in response to certain environmental cues, e.g. nutrient supply (Flärdh and Buttner, 2009). Upon emergence, the germ tubes begin to replicate rapidly by tip extension, without forming septa (Chater and Losick, 1996). They begin to form branch points, quickly establishing vegetative growth (Chater and Losick, 1996). As they do they begin to limit the surrounding nutrient supply, which corresponds with aerial hyphae formation (Flärdh and Buttner, 2009). Branch points occur at or near the colony surface and continue to spread upwards, supplied with nutrients from its base (Kieser *et al.*, 2000). In the later stages of spore development the septa form, as does the thick cell wall and grey spore pigment that protects the mature spores from desiccation (Kieser *et al.*, 2000). Sporulation is the final stage of development and only forms a small proportion of the total colony mass (Chater, 1998). Sporulation and the mechanisms that regulate it have been studied extensively through use of a number of 'white' and 'bald' mutants, unable to form mature spores or unable to form aerial hyphae at all. Subsequent identification of the genes responsible for these phenotypes revealed a range of *whi* (white phenotype) and *bld* (bald phenotype) genes that regulate *S. coelicolor* development. The first of the *bld* genes, *bldA*, is actually a leucyl tRNA gene, which recognises the rare UUA codon, and is essential for aerial hyphae and antibiotic production (Leskiw *et al.*, 1993). Regulation by *bldA* requires the inclusion of the TTA codon within the target genes; thus their expression is dependent upon expression of *bldA*. Other *bld* genes are transcriptional regulators; for example *bldD* has been shown to recognise sites upstream of several genes, including sigma factors *bldN* and *whiG* (Elliot *et al.*, 2001). BldD is characterised as a global regulator of development, and while it is not essential for viability it is essential for sporulation and antibiotic production (Elliot *et al.*, 2001, Elliot *et al.*, 2003). It appears to act as a repressor, allowing the expression of its target genes (developmental σ factors and regulators) only at the onset of sporulation, when they are required (Elliot *et al.*, 2001). BldD is primarily dimeric, with each subunit able to coordinate one half-site of its binding site via a helix-turn-helix motif (Kim *et al.*, 2006). This system demonstrates some of the complexity commonly associated with *S. coelicolor* gene regulation, with a transcriptional regulator having targets that regulate other transcriptional regulators causing a cascade of changes throughout the colony.

Section 1.2.2 – Genes and genome

The 8,667,507 bp *Streptomyces coelicolor* genome was released in full in 2002, and contained 7,825 genes (Bentley *et al.*, 2002). Unusual for bacteria, *S. coelicolor* has a linear, not circular, chromosome. Its ends are protected by terminal proteins, attached to each 5' end, and these appear to be essential for DNA replication and propagation of the chromosome (Bao and Cohen, 2001). Just as with Eukaryotes, bacterial chromosomes also require compaction. A number of classes of nucleoid-associated proteins exist with this function in bacteria, including: Lrp, HU, Lsr2 and SMC (Luijsterburg *et al.*, 2008). Again just as with Eukaryotes these proteins are implicated in regulation of gene expression. The leucine responsive protein (Lrp) is common to both gram-positive (including *S. coelicolor*) and gram-negative bacteria, and is able to wrap the DNA around its multimeric form (Luijsterburg *et al.*, 2008). As its name would suggest Lrp is sensitive to leucine, which destabilises the DNA complex - allowing expression of targets such as *rrn* operons (Pul *et al.*, 2007). *S. coelicolor* have two copies of the HU protein encoded in their genome, one of which (*hupS*) has been shown to be developmentally regulated (Salerno *et al.*, 2009). HupS appears to localise to the nucleoid of developing spores, with mutants being pigmentation deficient and more heat labile than their wild type counterparts (Salerno *et al.*, 2009). Lsr2 of *M. tuberculosis* appears to have multiple functions, not only regulating gene expression but also protecting the cell from reactive oxygen species produced by the host (Colangeli *et al.*, 2009). Lsr2 also has two potential homologues in *S. coelicolor*. The structural maintenance of chromosomes (SMC) proteins appear to be involved in chromosome condensation during septation in sporulating aerial hyphae (Dedrick *et al.*, 2009, Kois *et al.*, 2009). Regulation of gene expression is a major theme of the *S. coelicolor* genome, with 65 sigma factors and 965 regulatory genes (Bentley *et al.*, 2004). Protein synthesis is mediated by six rRNA (16S, 23S and 5S rRNA) operons and 63 tRNAs (Bentley *et al.*, 2002). *S. coelicolor* is a high G+C gram-positive bacterium with a G+C content of ~72% (Bentley *et al.*, 2002). Interestingly its promoter regions are characterised by their AT-rich sequences aiding their identification, whereas the coding regions themselves show a clear GC-bias (Jaurin and Cohen, 1985). The *Streptomyces*

genus is also intimately associated with natural product biosynthesis; thus another notable feature of the genome is the presence of 22 gene clusters, encoding enzymes characteristic of their biosynthesis (Bentley *et al.*, 2002). As mentioned previously, *S. coelicolor* produces actinorhodin, a pH-sensitive antibiotic (Wright and Hopwood, 1976). In addition to this, they produce undecylprodigiosin and the calcium-dependent antibiotic (CDA), as well as the polyketide synthase expressed from the *whiE* cluster (Chong *et al.*, 1998, Davis and Chater, 1990, Hopwood *et al.*, 1995). The *Streptomyces* genus accounts for ~80% of all identified secondary metabolites, having various purposes and potencies (Challis and Hopwood, 2003). They are generally induced when growth slows and their functions range from antimicrobial, to limit the competition for dwindling resources, to intercellular signalling, acting as secreted hormones to influence colony development (Challis and Hopwood, 2003). Genes involved in the production of the secondary metabolites coelibactin and γ-butyrolactone were also identified in the genome of *S. coelicolor*, products whose inhibitory effect on surrounding colonies is a secondary effect (Bentley *et al.*, 2002). Coelibactin is a non-ribosomal peptide thought to function as a Zn^{2+} siderophore (Hesketh *et al.*, 2009, Kallifidas *et al.*, 2010). Expression of the cluster is regulated by both the antibiotic synthesis regulator, AbsC, and the Zn^{2+} uptake regulator, Zur, in order to keep Zn^{2+} levels at sufficient but non-toxic levels within the cells (Hesketh *et al.*, 2009, Kallifidas *et al.*, 2010). The γ-butyrolactones are signalling molecules that regulate expression of other secondary metabolites; actinorhodin, undecylprodigiosin and cryptic polyketide type I (Cpk), as well as development (Takano, 2006). One type SCB1 has been shown to inhibit the DNA-binding ability of polyketide regulator ScbR (Takano *et al.*, 2005). This is yet another example of the complexity of *Streptomyces* gene regulation.

Section 1.3 – Global control of respiration

Members of the *Streptomyces* genus are, by definition, aerobic. As soil bacteria, they have to cope with extremely variable dissolved oxygen concentrations;

fluctuating from 9.5mg/L to as low as 0.1mg/L (Alberic *et al.*, 2009). Streptomycetes have evolved to thrive in both environments, with substrate hyphae delving deep into the often anoxic conditions of the soil and aerial hyphae repelling water to reach out into the air (Claessen *et al.*, 2003, van Keulen *et al.*, 2007, van Keulen *et al.*, 2003).

Section 1.3.1 – Overview of the respiratory chain

It is a widely accepted theory that eukaryotic mitochondria have prokaryotic origins. Whilst a large amount of gene transfer and alteration has occurred during evolution the structural similarities between the respiratory chains of the distinct super-kingdoms have remained (Andersson *et al.*, 2003). Thus a number of conclusions can be made about bacterial respiration from their phylogenetically diverse cousins. The mitochondrial respiratory chain consists of 4 complexes (I, II, III and IV) which generate a proton motive force (PMF) and an ATP synthase, which uses the proton gradient to generate ATP. Complex I is a proton-translocating NADH dehydrogenase, which passes the electrons from NADH oxidation to the quinone pool. Complex II is a succinate dehydrogenase, which although not contributory to the PMF also produces a reduced quinone. Complex III is the cytochrome bc1 complex, which uses the energy from quinone oxidation to reduce the electron carrier cytochrome C, with concomitant proton translocation. Complex IV is the cytochrome C oxidase or terminal oxidase, using oxygen as the terminal electron acceptor, re-oxidising cytochrome C and further increasing the PMF. Finally Complex V or ATP synthase, reduces the electrochemical gradient across the mitochondrial membrane by allowing protons back through; it uses the energy this releases to phosphorylate ADP to ATP (Figure 1.2). Glycolysis and the TCA cycle produce reduced NADH and $FADH_2$, which act as electron carriers for oxidative phosphorylation. For every ATP generated $4H^+$ must be translocated across the membrane. Each NADH results in the translocation of $12H^+$, yielding 3 ATP, whereas each $FADH_2$ results in 2 ATP. $FADH_2$ produces a lower yield as it forms part of Complex II, which is non-proton translocating thus only contributing to the PMF indirectly by reducing quinone. The total yield possible from aerobic respiration is 38 ATP. In contrast the maximum yield stated for

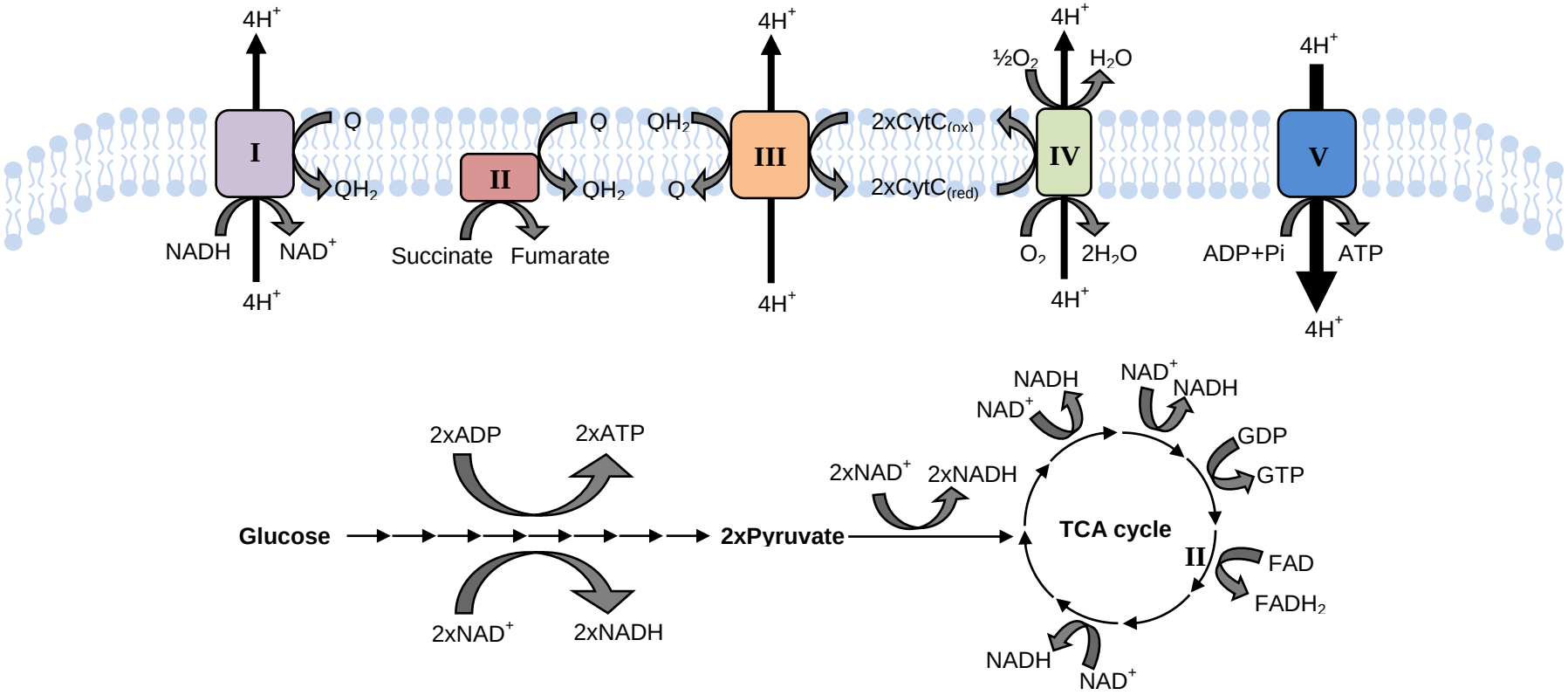


Figure 1.2: Overview of the respiratory chain. The individual reactions of Complex I (NADH dehydrogenase), II (Succinate dehydrogenase), III (cytochrome *bc* complex), IV (cytochrome *bo* terminal oxidase) and V (ATP synthase) are indicated. The glycolytic reactions are also represented in a summarised form. Note that Complex II is also shown as part of the TCA cycle in the bottom diagram. Glycolytic yields stated are correct for one molecule of glucose, the electron transport chain values are stated per NADH molecule produced by glysolysis.

fermentation is 2 molecules of ATP per molecule of glucose. This is because fermentation relies on glycolysis alone to generate ATP and subsequent reactions to re-oxidise the reduced cofactors (Figure 1.3) (Hoogerheide, 1975). Therefore there is a clear difference in the potential yield of each form of energy generation. However, in both cases the NADH, reduced during glycolysis, must be recycled to allow continued ATP synthesis. Inability to recycle the cofactor results in redox stress, i.e. a build up of reduced NADH compared to oxidised NAD^+, and is generally associated with oxygen limitation. When this occurs the cells can either adapt their aerobic respiratory chains to cope or can switch to anaerobic pathways.

Section 1.3.2 – Life with little oxygen

Bacteria have branched respiratory chains, with different components recruited for different growth conditions. Two such branches feature prominently within this study: cytochrome *bd* terminal oxidase and NADH dehydrogenase type II. The terminal oxidase has two forms *bo*- (*cyo*) and *bd*-type (*cyd*). Functionally they are very similar; both are capable of reducing oxygen to water, proton-translocation and cytochrome C oxidation, however these proteins are unrelated (Junemann, 1997). The *bo*-type have a haem-copper core, whereas the *bd*-type have multi-haem (*b*-type and *d*-type) redox centres (Junemann, 1997). The *bd*-type has a higher affinity for oxygen giving it an obvious function during oxygen limitation (Poole and Cook, 2000). It is also less sensitive to respiratory inhibitors such as cyanide, IC_{50} of 10µM (*bo*-type) and 2mM (*bd*-type) (Kita *et al.*, 1984). However, its limitation is that it translocates fewer protons than the *bo*-type ($1H^+$ instead of $2H^+$) and therefore contributes fewer H^+ to the PMF (Calhoun *et al.*, 1993). The *bo*-type is therefore not surprisingly most associated with aerated growth, whereas *cyd* tends to be induced by oxygen-limitation (Tseng *et al.*, 1996). The respiratory NADH dehydrogenases are likewise functionally similar but structurally diverse. *S. coelicolor* have at least one type I (NDH-1) and type II (NDH-2) NADH dehydrogenase. Type I is commonly referred to as Complex I of the respiratory chain and is proton-translocating. While NDH-1 is a multimeric, transmembrane complex, NDH-2 is formed of a single membrane-associated subunit. NDH-2 is incapable of protein

translocation but still able to pass electrons to the quinone pool. Thus NDH-2 can recycle reduced NADH but is less energy-efficient compared to NDH-1. However, it appears that NDH-2 may be quicker at re-oxidising the cofactor than NDH-1 (Esterhazy *et al.*, 2008, Jaworowski *et al.*, 1981). Thus NDH-2 may yield less ATP but can potentially relieve redox stress quicker than its counterpart.

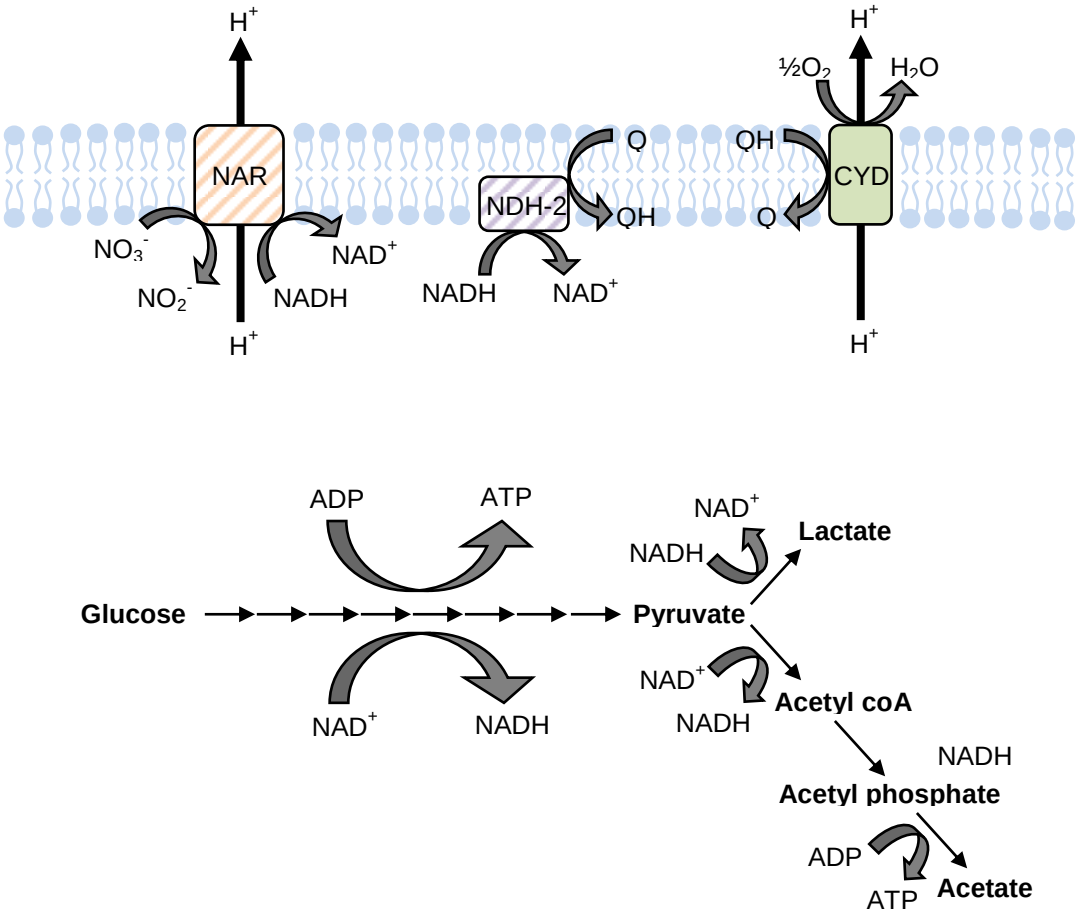


Figure 1.3: Overview of some of the different branches of the respiratory chain in prokaryotes, including the Type II NADH dehydrogenase (NDH-2), respiratory nitrate reductase (NAR) and cytochrome *bd* terminal oxidase (CYD). The glycolytic reactions are also represented in a summarised form, showing the fermentation pathways branching off from pyruvate. Note that yields are not indicated on this figure.

There are however other options when oxygen becomes limiting, for example nitrate reduction. The reduction potential ($E^{0'}$) for nitrate is lower than that of oxygen, +0.42V and +0.82V, respectively (Iuchi and Lin, 1988). However, the use of nitrate is still preferable to either fumarate or trimethylamine N-oxide (TMAO), with $E^{0'}$ values of +0.03V and +0.13V, respectively (Iuchi and Lin,

1988). As mentioned previously, glycolysis produces a net yield of 2 ATP per mole of glucose, with the limitation of producing reduced cofactors requiring re-oxidation. However, during fermentation the end product of glycolysis, pyruvate can be reduced to waste products such as lactate or acetate with both pathways re-oxidising the NADH, allowing NAD^+ to feed back into glycolysis (Figure 1.2). Although it is not as high yielding as aerobic respiration it does provide a continued source of ATP when oxygen is scarce. Conditions can vary widely and change rapidly, thus a rapid response to oxygen limitation is essential, and requires specific sensors that can either directly or indirectly sense this stress.

Section 1.3.3 – Bacterial oxygen sensors and stress regulators

Branching of the respiratory chain is regulated by a number of different sensors and regulators. The following section discusses the various regulators of aerobic and anaerobic respiration and the regulons which they control.

ArcAB: The ArcAB system of *E. coli* senses the redox poise under oxygen limitation, detecting the presence of oxidised versus reduced quinones in the cell membrane, and regulates expression of *cyd* (Green and Paget, 2004). ArcA is a response regulator, whereas ArcB is a sensor kinase. Under anaerobic conditions ArcA appears to repress expression of genes linked to aerobic respiration and metabolism (Kwon *et al.*, 2000). ArcB is a multi-domain protein (Figure 1.4), containing a membrane association domain, a PAS domain, a transmitter domain, a receiver domain and a phosphotransfer domain (Malpica *et al.*, 2004). PAS (Per, ARNT, Sim) domains are commonly associated with histidine sensor kinases and occur, without exception, on the N-terminal side of the phosphotransferase domain of the containing protein (Ponting and Aravind, 1997). ArcB is able to autophosphorylate, a common feature among proteins containing PAS domains (Georgellis *et al.*, 1997). However, the ability of ArcB to do this is determined by the presence or absence of two disulphide bonds between the PAS domains of the dimer (Malpica *et al.*, 2004). Under aerobic conditions the quinone pool will be predominantly in the oxidised form, in which state they are proposed to react with the cysteine residues in the PAS domain oxidising them to form disulphide

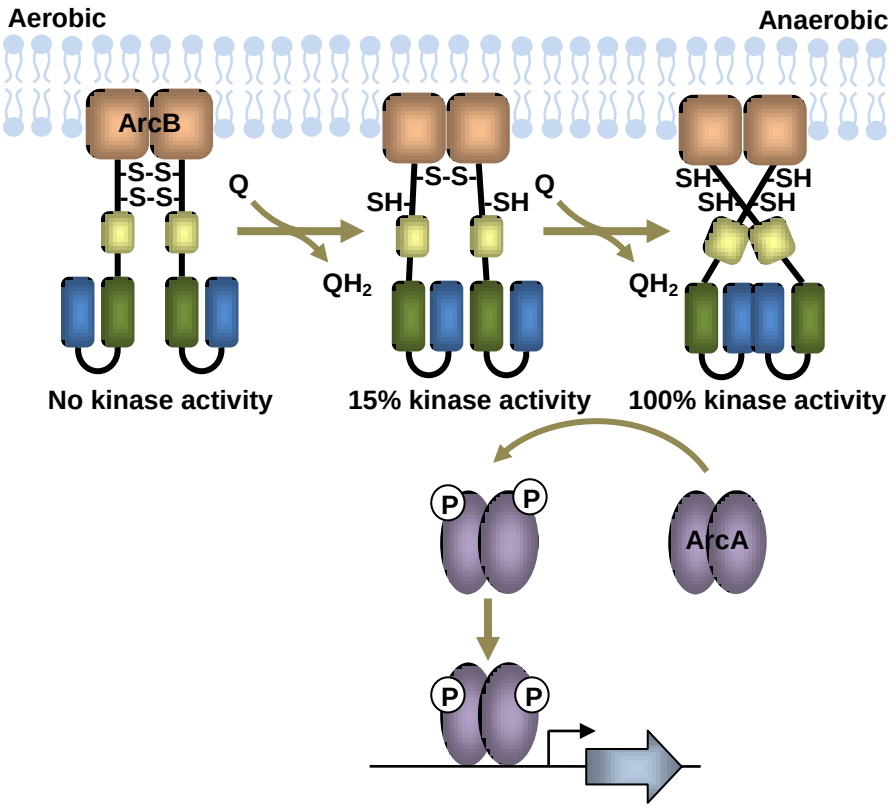


Figure 1.4 ArcB is embedded in the cell membrane, consisting of a PAS domain (indicated by S/SH), a transmitter domain (yellow), a receiver domain (green) and a phosphotransfer domain (blue). The presence of quinones determines the presence or absence of two disulphide bonds of ArcB, reduction of which results in full kinase activity. This protein is then able to phosphorylate ArcA, which recognises specific regions within the promoters that it regulates. Figure adapted from that of (Malpica *et al.*, 2004).

bridges (Malpica *et al.*, 2004). Under anaerobiosis the quinone pool changes so that the reduced form dominates, under these conditions the disulphide is reduced (Malpica *et al.*, 2004). This breakage is thought to permit a large structural change that activates the kinase activity of ArcB (Malpica *et al.*, 2004). ArcB is then able to transfer a phosphoryl group to ArcA (Iuchi *et al.*, 1990). It has recently been shown that ArcB is sensitive to both ubiquinone and menaquinone, with the former being more closely associated with regulation during high aeration (Bekker *et al.*, 2010). As the ratio of each can vary at different oxygen concentrations it gives ArcB the ability to sense changes to the quinone pool over a wide range of conditions, allowing the system to fine-tune the switch between aerobic and anaerobic growth (Bekker *et al.*, 2010). In *Salmonella*, the ArcAB system has been shown to not only be involved in regulating the switch between aerobic and anaerobic metabolism, but also appears to be important for coping with reactive oxygen species (Loui *et al.*,

2009). In a recent study it was shown that ArcA was constitutively induced if all three terminal oxidases and the quinol monooxygenase were lost from *E. coli*, which resulted in a strain unable to switch away from anaerobic metabolism regardless of oxygen abundance (Portnoy *et al.*, 2010). This is in accordance with the model for ArcB activation as under these conditions the re-oxidation of quinones would be blocked, thus the reduced form would predominate. This would keep the cysteines of ArcB reduced and thus maintain its kinase ability. ArcA itself has been shown to regulate the expression of dehydrogenases, terminal oxidases (including *cyd*) and TCA enzymes, among others (Alexeeva *et al.*, 2003). This allows the induction of branched respiratory chain components under oxygen limitation alone, as their use under aerated conditions is inefficient. It should be noted however that ArcA is not essential for anaerobic growth, its function is to modulate gene expression in the transition between aerobic and anaerobic growth (Alexeeva *et al.*, 2003).

DosR-DosS-DosT: DosR is a dormancy response regulator in *M. tuberculosis*, essential for persistence during latent infections (Leistikow *et al.*, 2010). The DosR system involves three components; DosR (response regulator), S and T (sensor kinases) (Kumar *et al.*, 2007). The dormancy regulon is composed of 48 genes (Voskuil *et al.*, 2003). Dormancy is induced by both oxygen limitation and nitric oxide (Voskuil *et al.*, 2003). Nitric oxide is a result of the host organisms' immune response, produced by the nitric oxide synthases of the macrophages (MacMicking *et al.*, 1997). At low levels NO acts as a reversible inhibitor of the cytochrome c terminal oxidase, competing with oxygen for binding to the active site (Brown, 2001). Both signals essentially indicate that the conditions have become oxygen-limited, thus a switch to alternate pathways is required. Both sensor kinases, DosS and DosT, are haem-containing proteins allowing them to sense changes in O_2, NO or CO concentration via their prosthetic groups (Figure 1.5) (Kumar *et al.*, 2007). DosS is termed the redox sensor of the Dos regulon, as its activity is modulated by a switch between Fe^{3+} and Fe^{2+} states (Kumar *et al.*, 2007). On the other hand, DosT is termed the hypoxia sensor as it is inhibited by O_2 (Kumar *et al.*, 2007). In each case the proteins exist in the Fe^{2+} state when active, being stabilised in their active forms by NO or CO ligands (Kumar *et al.*, 2007). The reason for the lack of O_2 associated with

DosS in its inactive state is thought to be that the domain organisation around the haem prevents it access (Cho *et al.*, 2009). Both DosS and DosT contain tandem GAF domains, followed by a histidine kinase domain. It is the GAF domains that co-ordinate the haem, and so confer their sensing ability – just as the PAS domain of ArcB does (Sardiwal *et al.*, 2005). Members of the dormancy regulon include genes encoding nitrate reductase (e.g. *narX*), nitrite/nitrate transport system (*narK2*), as well as an oxygen-independent form of the ribonucleotide reductase (*nrdZ*), all of which are typically associated with oxygen limitation (Voskuil *et al.*, 2003).

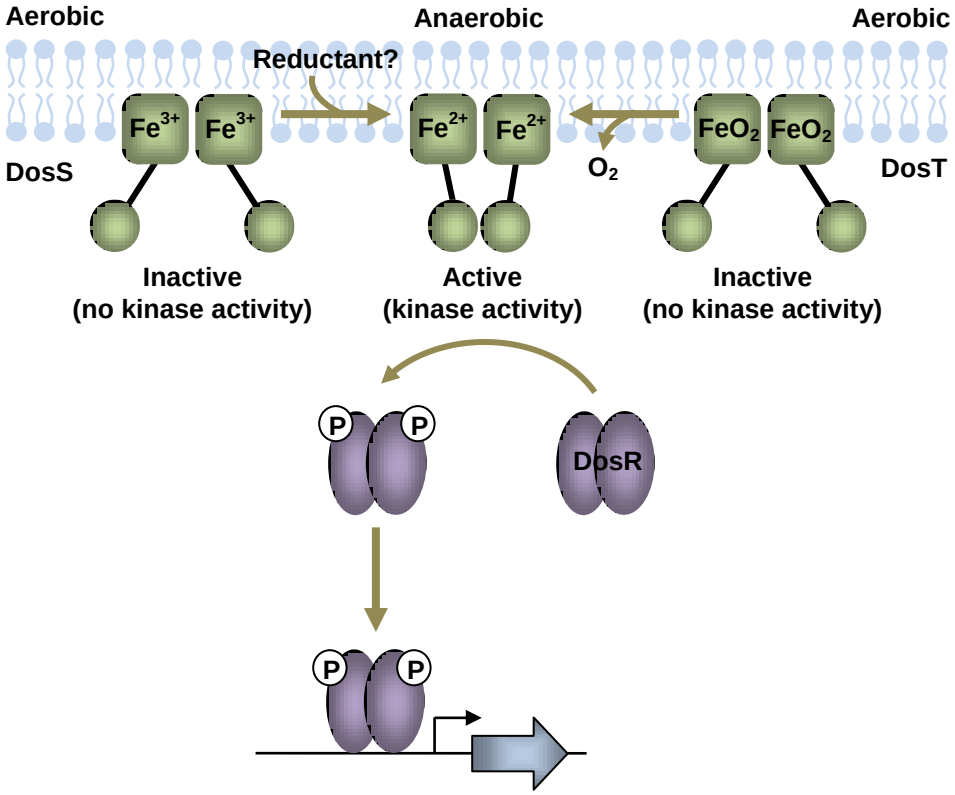


Figure 1.5: The DosRST system of *M. tuberculosis*. Sensor kinases DosS and DosT change the oxidation state of their haem centres in response to environmental cues. This then results in auto-phosphorylation and phosphotransfer to DosR. Figure adapted from that of (Kumar *et al.*, 2007).

Fnr: The *E. coli* fumarate nitrate reduction regulator (Fnr) acts as a key sensor of oxygen limitation, regulating the switch to anaerobic pathways (Becker *et al.*, 1996). Unlike the other systems discussed in this chapter, Fnr contains both the sensory and regulatory domains within the same protein. The formation ofactive Fnr is however still reliant on other proteins, namely the *isc* operon encoded iron-sulphur cluster synthesis and incorporation proteins (Mettert *et al.*, 2008).

Fnr is an iron-sulphur cluster (Fe-S) containing protein (Lazazzera *et al.*, 1996). It is constitutively expressed but is normally kept at low levels (Tolla and Savageau, 2010). In its DNA-bound or active state the protein is dimeric and contains 2[4Fe-4S] clusters but these are subject to oxidation in the presence of oxygen, destabilising the protein (Figure 1.6) (Lazazzera *et al.*, 1996). The protein will then either be reconstituted or degraded by ClpXP, keeping the level of active Fnr low but constant in the cell (Mettert and Kiley, 2005). Under anoxic conditions the active form predominates allowing it to regulate gene expression (Tolla and Savageau, 2010). Fnr is also sensitive to nitric oxide, as well as oxygen (Crack *et al.*, 2008). In *B. subtilis* Fnr has been shown to positively regulate expression of *arfM* (anaerobic modulator), *narK* (nitrite extrusion) and *narGHJI* (nitrate reductase) (Reents *et al.*, 2006). *S. coelicolor* also has Fnr-like proteins, however these lack the ability to co-ordinate the iron-sulphur clusters that would normally permit oxygen sensing (as in their *E. coli* homologues) (van Keulen *et al.*, 2007).

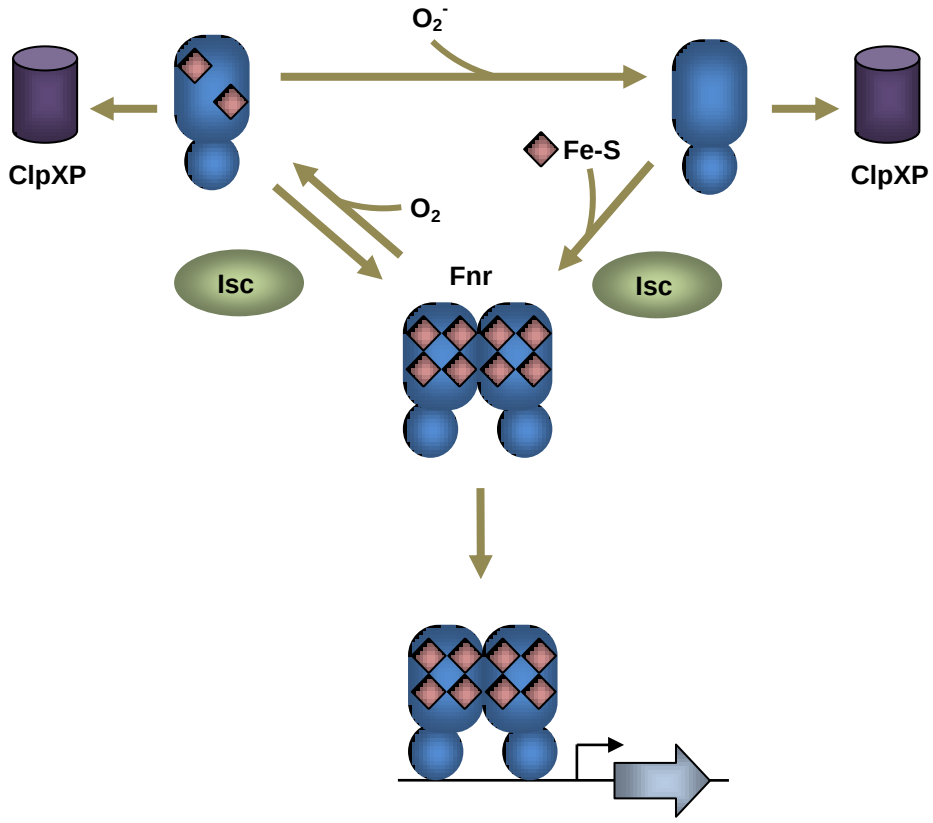


Figure 1.6: Overview of the Fnr system of *E. coli.* Active Fnr is generated when it dimerises in the presence of 2[4Fe-4S] clusters, with the aid of Isc. Inactivation is induced by oxygen, producing monomeric Fnr. The monomers are either converted back to active Fnr or are targeted for degradation via the ClpXP system. Figure derived from that of (Tolla and Savageau, 2010).

ResDE: The ResDE system of *B. subtilis* is a two component sensor kinase and response regulator of anaerobic respiration. ResD and ResE are expressed from the *resABCDE* cytochrome biogenesis operon, with the first three genes of the operon being essential (Sun *et al.*, 1996). Expression from the *resA* promoter appears to be ResD-dependent, with *resD* being expressed at low levels from its own promoter within the operon (Sun *et al.*, 1996). ResD also appears to regulate *ctaA* (haem biosynthesis) and *petCBD* (cytochrome *bf* complex) expression, as well as *fnr* (Esbelin *et al.*, 2009, Sun *et al.*, 1996). ResE senses both oxygen limitation and nitric oxide in order to modulate the activity of ResD, allowing the cells to switch from oxygen to nitrate as the terminal electron acceptor (Geng *et al.*, 2007). ResE, like ArcB, is a PAS-domain containing protein, although unlike ArcB, ResE only contains one PAS domain (Baruah *et al.*, 2004). This domain is responsible for signal sensing, however the mechanism must be different to that of ArcB as there are no cysteine residues contained in the sequence of the ResE PAS domain (Baruah *et al.*, 2004). ResE is able to both phosphorylate and dephosphorylate its target ResD in order to modulate the activity of this transcriptional regulator (Baruah *et al.*, 2004).

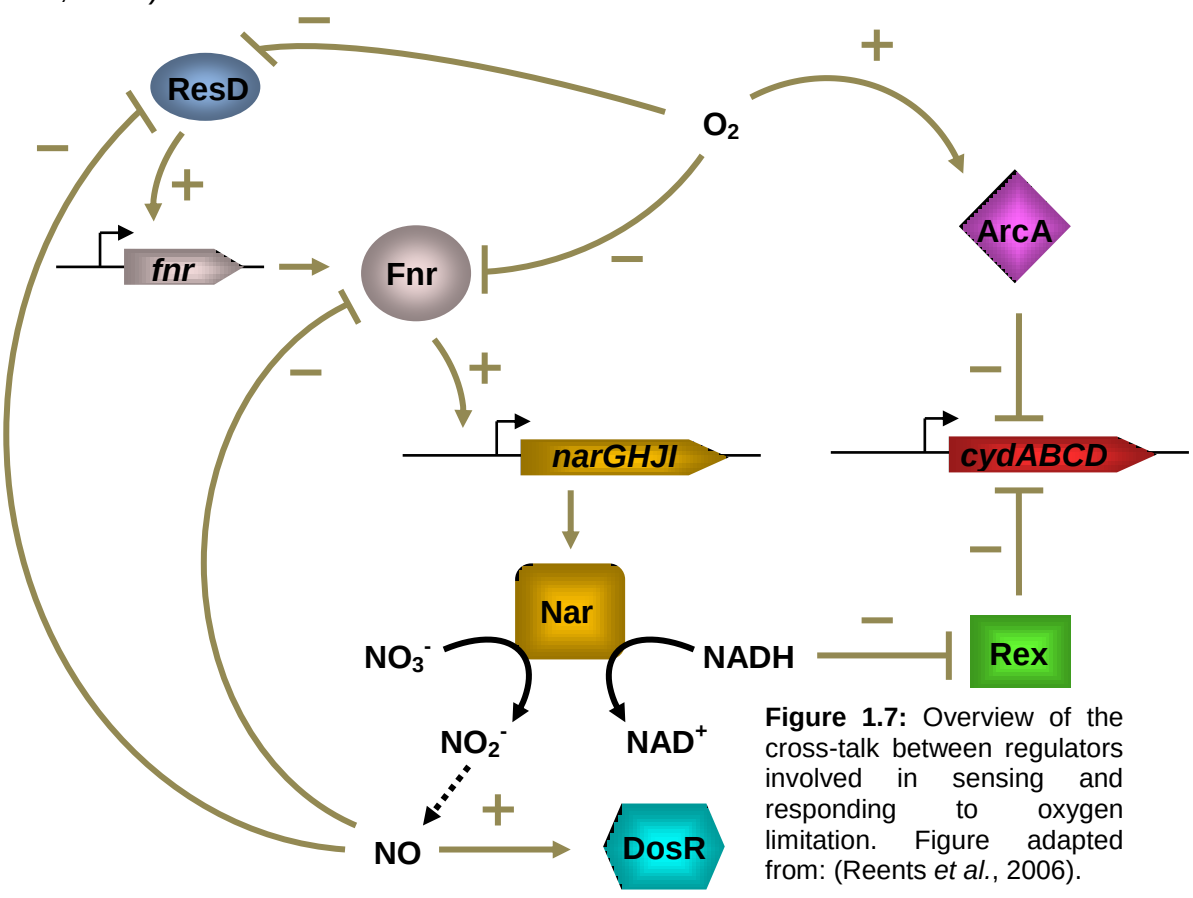


Figure 1.7: Overview of the cross-talk between regulators involved in sensing and responding to oxygen limitation. Figure adapted from: (Reents *et al.*, 2006).

Bacterial genomes contain a plethora of different sensors, with a range of sensing mechanisms. There does however appear to be a lot of cross-talk between these systems (Figure 1.7). This allows a high level of control of respiration, ensuring that the most energy-efficient option is applied to each growth condition and that branches of the respiratory chain are kept silenced when not required. The sensory proteins have evolved to incorporate natural redox sensors; such as haem (DosST), iron-sulphur clusters (Fnr) and disulphide bonds (ArcB). The Rex system of *S. coelicolor* however, appears to be unique in that it is able to detect the physical difference between the reduced and oxidised form of redox indicator NAD/H in order to modulate its own activity.

Section 1.4 – Rex – the story so far

As mentioned previously, a rapid response to oxygen limitation is vital for maintaining cell growth and a key indicator of oxygen limitation is redox stress. This section covers the identification of a novel redox sensor, Rex, in *S. coelicolor* and details current understanding of its structure, function and regulon.

Section 1.4.1 – Discovering Rex

Rex was originally identified indirectly through its ability to repress expression of the cytochrome *bd* terminal oxidase operon (Figure 1.8) (Brekasis, 2005, Brekasis and Paget, 2003). Analysis of the *cydABC* operon revealed two promoters; a constitutive and an anoxia-induced promoter (Brekasis, 2005, Brekasis and Paget, 2003). By mutating the entire promoter region of *cyd* a putative operator site was identified (TGTGAACGCGTTCACA), which caused an increase in the expression of the operon (Brekasis, 2005, Brekasis and Paget, 2003). Using this site as a template other potential operators were identified upstream of the *nuoA-N* and *hemACD* operons (Figure 1.9), with the latter also encoding a putative DNA-binding protein; SCO3320 (Brekasis, 2005, Brekasis and Paget, 2003). Overexpression of this gene produced a protein

capable of binding to the *cyd* operator site (Brekasis, 2005, Brekasis and Paget, 2003).

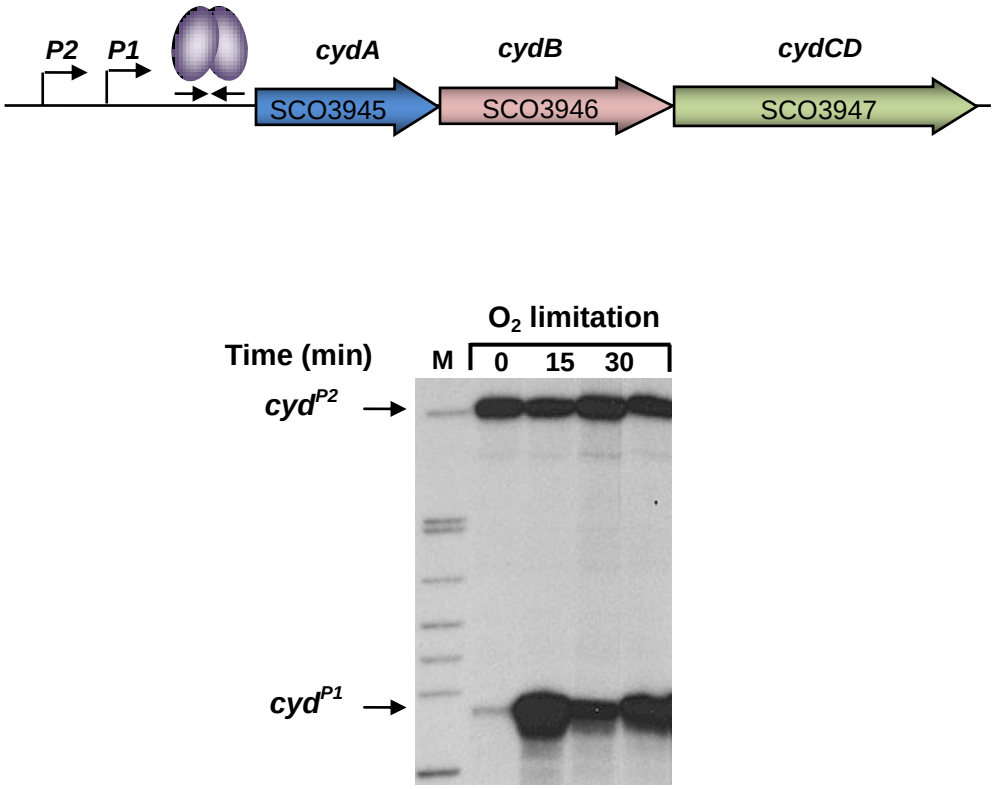


Figure 1.8: Organisation of the *cydABCD* operon of *S. coelicolor*, including the two promoter regions. The S1 nuclease protection assay for this region is also shown, taken from: (Brekasis, 2005, Brekasis and Paget, 2003).

It was further confirmed that this protein was responsible for this response as its deletion resulted in constitutive expression of the *cydA* gene (Brekasis, 2005, Brekasis and Paget, 2003). The sequence of this putative repressor contained within it a Rossmann fold, a protein fold closely associated with the binding of pyridine nucleotides. The responsiveness of the SCO3320 protein to NAD/H and NADP/H was assessed by EMSA and it was found that NADH alone could affect DNA binding, inhibiting it (Brekasis, 2005, Brekasis and Paget, 2003). As the concentration of NADH in bacterial cells has been shown to vary depending on growth conditions it seemed unlikely that SCO3320 was sensing this alone (van Keulen et al., 2007). As expected the inhibitory effect of NADH on DNA binding could be lessened by the inclusion of NAD^+ in the binding assay, allowing it to sense the redox poise of the cell via the NADH/NAD^+ ratio; the protein was therefore designated Rex (redox regulator) (Brekasis, 2005,

Brekasis and Paget, 2003). Since its discovery, Rex binding sites have been identified upstream of *ndh*, *atpI*, *wblE* and SCO4281 (Brekasis, 2005).

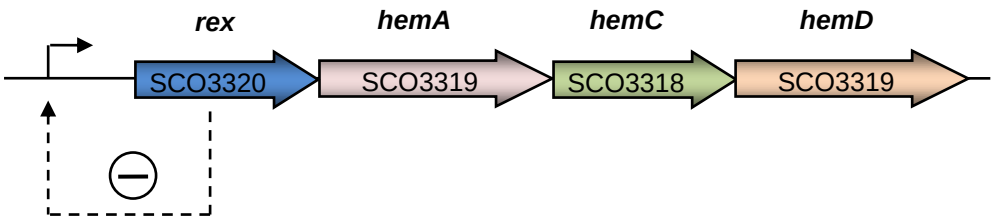


Figure 1.9: Organisation of the *rex-hemACD* operon of *S. coelicolor*. The autoregulation of Rex on its own promoter is also indicated.

Section 1.4.2 – The structure and function of Rex

There are numerous Rex homologues in other species although it is limited to gram-positive bacteria. Interestingly, Rex is not limited to aerobic bacteria but is present in facultative (e.g. staphylococci) and obligate (clostridia) anaerobes. Although the structure of the *S. coelicolor* protein (S-Rex) has not been solved, the structure of a homologue in *Thermus aquaticus* has proven extremely useful in studying S-Rex (Sickmier et al., 2005). *Thermus aquaticus* Rex (T-Rex) is a homodimer consisting of three domains; a DNA-binding domain, a NADH-binding domain and a domain-swapped helix (Figure 1.10) (Sickmier *et al.*, 2005). It has 42% identity to S-Rex and has even been shown to recognise the same S-Rex operator site (Sickmier *et al.*, 2005). The DNA binding domain is located in the N-terminal region of Rex and is characterised by a winged-helix motif, with the entire domain consisting of four α-helices and two β-strands (Sickmier *et al.*, 2005). The NADH binding domain is centrally located in the protein sequence, formed of four α-helices and seven β-strands arranged into a Rossmann fold, with characteristic GXGXXG motif (Sickmier *et al.*, 2005). The fold co-ordinates NADH and NAD^+ and therefore acts as the sensory domain for Rex (McLaughlin *et al.*, 2010, Sickmier *et al.*, 2005). It should be noted that Rex does not directly influence the redox poise, i.e. it cannot re-oxidise NADH itself, as indicated by the lack of the appropriate reactive residues and the presence of Y98 blocking the substrate access channel (Sickmier *et al.*, 2005). At the C-terminus is a final α-helix which slots in between the other two domains on the opposing subunit (Sickmier *et al.*, 2005). As previously mentioned Rex is a redox sensor, not just sensitive to the NADH concentration but to the

NADH/NAD^+ ratio (Brekasis, 2005, Brekasis and Paget, 2003, Sickmier *et al.*, 2005). How it is able to sense the charge difference between the two cofactors will be addressed during this study.

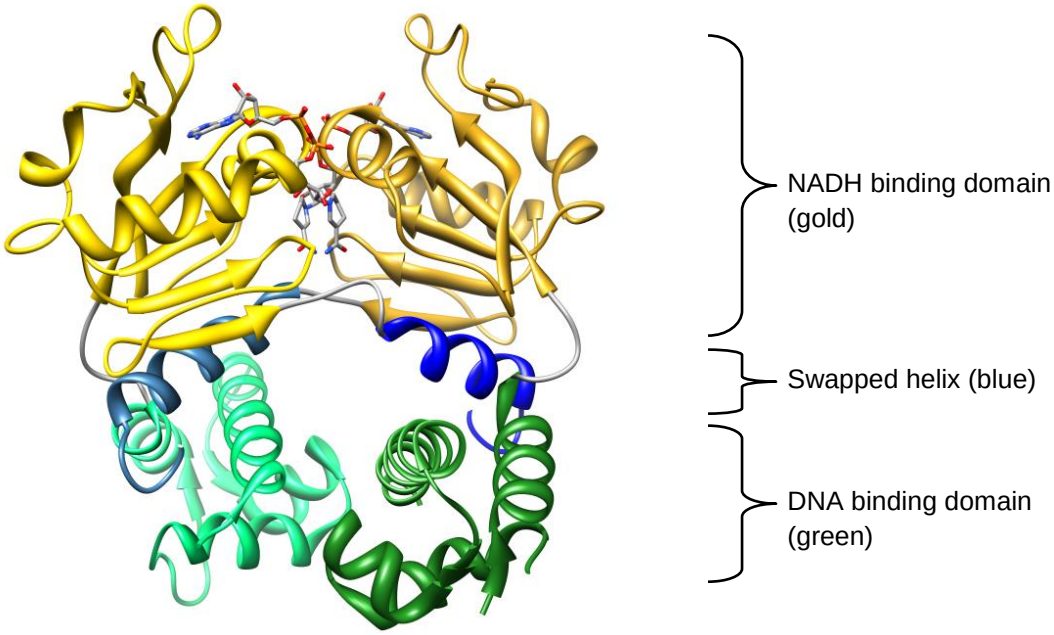


Figure 1.10: Structure of NADH-bound *Thermus aquaticus* Rex. The three functional domains are indicated above (coloured by domain), with the two NADH molecules shown in stick form at the dimer interface.

Section 1.4.3 – Project aims

The Rex regulon:

Several putative Rex binding sites had been identified by bioinformatic analysis prior to this project. However, most genes thought to be controlled by these sites were not upregulated when the transcriptome of a wild-type strain was compared to that of a *rex* mutant. It was therefore decided to take a more direct approach to define the Rex regulon. As such the aims were as follows:

- Design and characterise an epitope-tagged *rex* strain
- Optimise the ChIP-on-chip method for use with this strain
- Use ChIP-on-chip to identify ROP sites *in vivo*
- Validate binding to these target *in vitro* and, if possible, investigate their regulation by Rex

The respiratory NADH dehydrogenases:

The two types of respiratory NADH dehydrogenases (*nuoA-N*; NDH-1 and *ndh*; NDH-2) of *S. coelicolor* were present within the list of known Rex targets but only the regulation of *ndh* could be detected (Brekasis, 2005). Therefore the aims of this section of the project were as follows:

- Investigate the potential regulation of both *ndh* and *nuo* by Rex
- Generate *ndh* and *nuo* disruption strains
- Investigate the impact of these mutations on the ability of Rex to repress its targets

Rex structure/function relationship:

Previous work had characterised Rex as a redox-sensitive repressor, able to bind to DNA when the NADH/NAD^+ redox poise was low and dissociate when it was high (Brekasis, 2005, Brekasis and Paget, 2003). With the structure of a close homologue in *T. aquaticus* available it was possible to model S-Rex mutations using the T-Rex structure, but a number of questions still remained about the function of Rex:

- How does NADH binding to one domain trigger DNA-dissociation in another?
- Why doesn't NAD^+ have the same effect as NADH on DNA-binding?
- What factors are required for DNA-binding?
- Can a protein with only one functional DNA-binding domain still bind to DNA?
- Are one or both NADH molecules required to dissociate DNA-bound Rex?

The results that follow are designed to meet these objectives and hopefully further our knowledge of this unique transcriptional repressor.

Chapter 2

Materials and Methods

"Do something. If it works, do more of it. If it doesn't, do something else."

Franklin D. Roosevelt (1882-1945)

Section 2.1 – Chemicals, reagents, enzymes and strains

The suppliers for all of the chemicals, reagents and enzymes used in this study are listed in the following sections. Any specialist equipment used is also detailed. All of the primers were obtained from MWG-Biotech but the sequences and modifications, were applicable, and are listed in full within this chapter.

Section 2.1.1 – Chemicals

- Acrylamide solutions – Severn Biotech Ltd
- Ammonium persulphate – Sigma
- Ampicillin – Melford
- Apramycin – Duchefa Biochemie
- Bromophenol blue – Amersham Biosciences
- Casamino acids – Difco
- Chloramphenicol – Melford
- Chloroform – Fisher
- dNTPs – New England Biolabs
- Glycogen – Fisher
- Hepes (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) – Fisher
- IPTG (Isopropyl-β-D-thiogalactopyranoside) – Melford
- Isoamyl alcohol – Sigma
- Kanamycin – Melford
- Malt extract - Oxoid
- Nalidixic acid – Duchefa Biochemie
- Nicotinamide adenine dinucleotide (NAD^+, NADH) – Melford
- Nutrient agar - Difco
- PEG 1000 (Polyethylene glycol) – BDH
- Peptone - Difco
- Phenol – Fisher Scientific
- PMSF (Phenylmethylsufonyl fluoride) – Sigma
- Pronase – Roche
- Radionuclides ([γ-^{32}P]-ATP / [α-^{32}P]-dCTP) – Perkin Elmer
- SDS (Sodium dodecyl sulphate) – Fisher Scientific
- Spectinomycin – Duchefa Biochmie
- TEMED (Tetramethyl-ethylenediamine) – Fisher
- TES (N-[tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid) – Fisher Scientific

- Tris (2-Amino-2-hydroxymethyl-propane-1,3-diol) – Fisher Scientific
- Tryptone – Difco
- TSB (Tryptone soya broth) – Oxoid
- X-gal (5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside) – Melford Laboratories Ltd
- Yeast extract - Oxoid

Section 2.1.2 – Enzymes

DNA/RNA restriction enzymes:

Restriction endonucleases – New England Biolabs

S1 Nuclease – Invitrogen

RQ RNAse-free DNAse - Promega

Ribonuclease A – Sigma Aldrich

Polymerases:

Accuzyme – Bioline

Klenow fragment – New England Biolabs

Phusion – New England Biolabs

Reverse transcriptase (iScript) – Bio-Rad

Taq DNA polymerase – New England Biolabs

QuantiTect SYBR green PCR Kit – QIAGEN

DNA modifying enzymes:

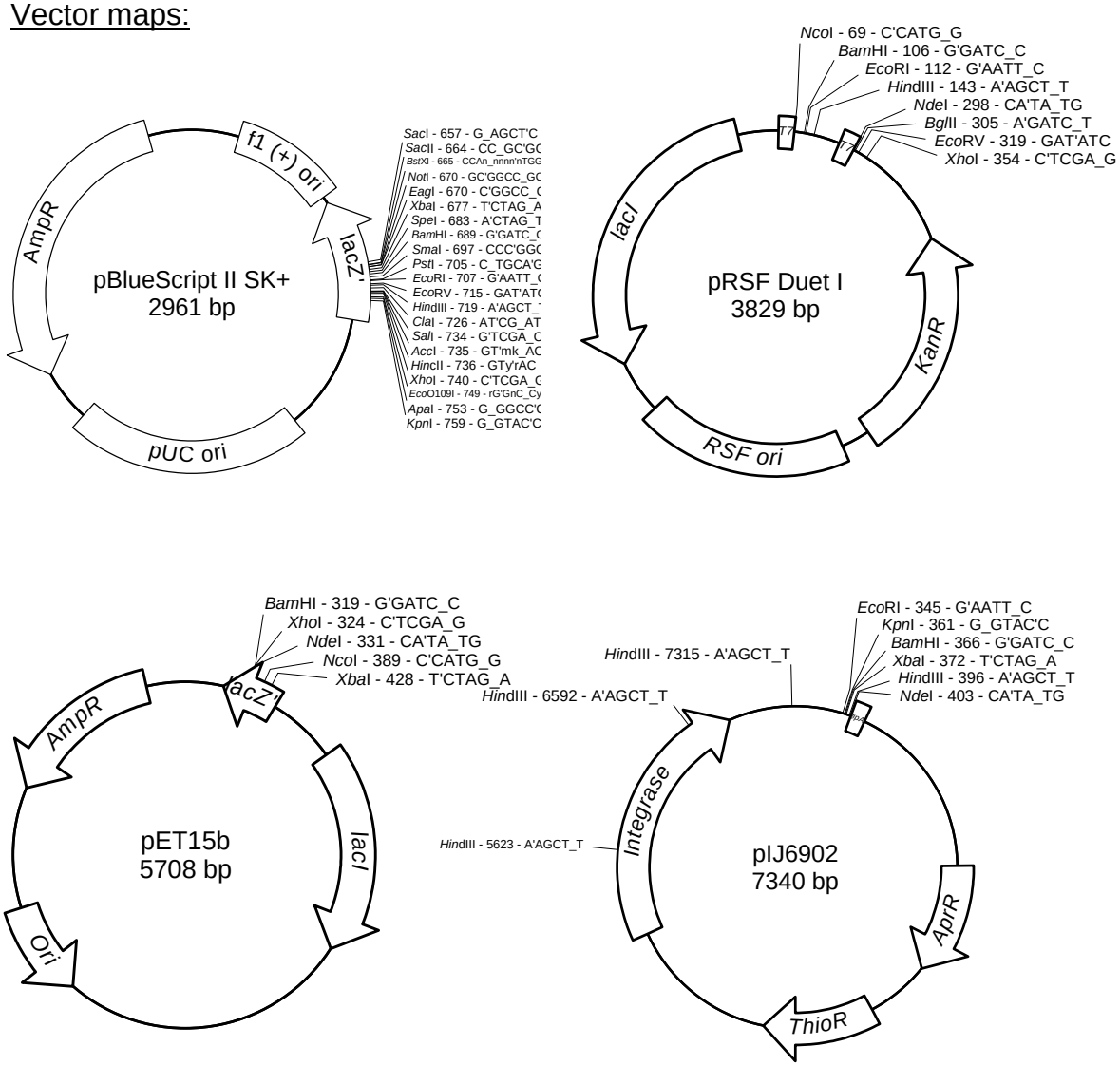
Shrimp alkaline phosphatase - Promega

T4 DNA ligase – New England Biolabs

T4 polynucleotide kinase – New England Biolabs

Section 2.1.3 – Vectors used in this study:

Vector maps:



Vectors generated in this study:

Name	Details	Section
pBS::ermE*::FLAG	pBlueScript II SK+ derivative containing the ermE* promoter and the FLAG-tag (DYKDHDGDDYKDHDI DYKDDDDK) encoding sequence, flanked by *Hin*dIII and *Xho*I sites. (M. Paget, personal communication).	
pLST920::wt P^{ndh}	pLST920 derivative containing the *ndh* promoter region, isolated from pSX414 as a *Kpn*I fragment.	**4.3.3**
pLST920::ΔROP1	pSX418 derivative containing the mutated *ndh*	**4.3.3**

P^{ndh}	promoter region from pSX415.	
pLST920::ΔROP2 P^{ndh}	pSX418 derivative containing the mutated *ndh* promoter region from pSX416.	**4.3.3**
pLST920::ΔROP1+2 P^{ndh}	pSX418 derivative containing the mutated *ndh* promoter region from pSX417.	**4.3.3**
pSX400	Suicide vector generated from pIJ6902, integrase disrupted by *Hin*dIII digest and subsequent religation. Contains part of the *nuo* operon fused to a tmRNA degradation signal, under the control of the *tipA* promoter.	**4.4.1**
pSX401	pBlueScript derivative containing the *rex* promoter and the entire coding region translationally fused to the 3XFLAG sequence (introduced into the *Eco*RV site).	**3.2.1**
pSX402	pSET152 derivative containing a *Not*I – *Xho*I (end-filled) isolate of pSX401, cloned into *Not*I – *Eco*RV cut pSET152. Conjugative vector containing the rex^{FLAG} gene.	**3.2.1**
pSX403	pBlueScript derivative containing the *ndh* gene, isolated using primers SCO3092Complete_For and _Rev, introduced as a blunt-end fragment into the vectors' *Eco*RV site.	**4.3.6**
pSX404	pSX403, containing a proteolytic cleavage tag (annealed primers Ndh_deg_a and Ndh_deg_b) ligated as a *Sph*I/*Eco*RI fragment.	**4.3.6**
pSX405	pIJ6902 derivative containing the ndh^{deg} region from pSX404, isolated as an *Nde*I/*Eco*RI fragment.	**4.3.6**
pSX406	pSX405 lacking an intact integrase gene, removed by complete HindIII digestion and subsequent self-ligation.	**4.3.6**
pSX407	pIJ6902 containing the *Nde*I-*Bam*HI *rex* fragment from pSX142.	**3.4.1**
pSX408	pRSF-Duet1 vector containing the synthetic *rex* sequence (from pUC57::rex^{synth}) introduced as a *Bgl*II/*Hin*dIII fragment into the *Bam*HI/*Hin*dIII cut vector.	**5.5.1**
pSX409	pSX408 derivative containing the *rex* sequence (from pSX142) introduced as an *Nde*I/*Bam*HI fragment.	**5.5.1**
pSX410	pSX409 derivative where the two *rex* genes had been fused together with an $(SG_4)_2$-linker sequence, introduced as a *Hin*dIII/*Nde*I fragment.	**5.5.1**
pSX411	pET15b derivative containing rex^{SC} from pSX410,	**5.5.1**

	introduced as an *Nco*I/*Hin*dIII fragment.	
pSX412	pBlueScript derivative containing the *Eco*O109I *rex*^{synth} fragment from pSX411.	**5.5.2**
pSX413	pBlueScript SK+ derivative containing PCR-amplified *nuoA*, *nuoB* and 625 bp of *nuoC,* flanked by *Nde*I and *Bam*HI/*Sph*I restriction sites.	**4.4.1**
pSX414	pBlueScript SK+ derivative containing a 414 bp segment of the *ndh* promoter introduced into the *Eco*RV site, in the reverse orientation.	**4.3.2**
pSX415	pSX414 derivative containing the *ndh* promoter region with its upstream ROP site mutated.	**4.3.2**
pSX416	pSX414 derivative containing the *ndh* promoter region with its downstream ROP site mutated.	**4.3.2**
pSX417	pSX414 derivative containing the *ndh* promoter region with both ROP sites mutated.	**4.3.2**
pSX418	The entire *ndh* open reading frame and ~105 bp of its promoter region, amplified with primers SCO3092_ROP1 and SCO3092_REV, ligated into *Eco*RV-cut pBlueScript II SK+.	**4.3.5**
pSX419	*Eco*RI fragment from pSX418 ligated into same site of pHJL401.	**4.3.5**
pSX420	pIJ6902 derivative with the apramycin resistance cassette replaced with a strep/spec cassette from pIJ778.	**4.3.6**

Table 2.1: Table of vectors generated during the course of this study and the corresponding sections in which they were used.

Section 2.1.4 – Primers used in this study:

Mutagenesis primers:

Name	Sequence (5' to 3')	Section
A56F_F2	GGTCAACTCCTTCAAGCTGCGCAAG	5.3.2
A56F_R2	CTTGCGCAGCTTGAAGGAGTTGACC	5.3.2
D203A_F	CTGCAGATCCTCGCCTTCCACGAGCAG	5.4.1
D203A_R	CTCGATGGAGAGGGCGACCTTGCGCACG	5.4.1
D203R_F	CGCAAGGTCCGCCTCTCCATCG	5.4.1
D203R_R	CGATGGAGAGGCGGACCTTGCG	5.4.1
D203S_F	GAGCTGCAGATCCTCGCCTTC	5.4.1

D203S_R	GATGGAGAGGGAGACCTTGCGC	5.4.1
FLAG_Rev	GGAAGCTTTGCCGGCATCACGGC	3.2.1
FLAG-ROP_For	GGTCTAGAGCGTGTGAACGAGGAAC	3.2.1
G102A_SC_F	GTTATTGTGGGCATTGCGAATCTGGGCG	5.5.2
G102A_SC_R	CGCCCAGATTCGCAATGCCCACAATAAC	5.5.2
K60A	GCGAAGCTGCGCGCGGACTTCTCCTAC	5.3.2
K60A_SC_F	GCCAAATTACGTGCCGATTTTAGCTATC	5.5.2
K60A_SC_R	GATAGCTAAAATCGGCACGTAATTTGGC	5.5.2
K60-R	GGAGTTGACCCCCGCGGCGGCCGCCAG	5.3.2
Ndh_deg_a	AATTCTCAGGCGGCCAGGGCGAAGGCCTGCTGGC TGCTGTCCCGCATG	4.3.6
Ndh_deg_b	CGGGACAGCAGCCAGCAGGCCTTCGCCCTGGCC GCCTGAG	4.3.6
ndh_SDM1_for	AAGTTCTTTGTAAGGAATTGGGC	4.3.2
ndh_SDM2_rev	TGTGAACTTTCCCGACGGGACGTC	4.3.2
ndh_SDM3_for	AAGGGGCGTGTGATCCACCCCCCTC	4.3.2
ndh_SDM4_rev	TGTAAGCTTCGGCGGGTGCTGTGC	4.3.2
NuoKO_F	CCCATATGAACGCGTATGCGCCCATCCTCGTACTG	4.4.1
NuoKO_R	CCGCATGCCGGATCCCAGGGGCCGGGTGATCGT CGAAGACGATG	4.4.1
R29A_F	GCTGTCCGAGCGCTCGGTGCCCACG	5.4.1
R29A_R	GCGGTCAGTGCGGCGAGGTACAGCGGAAG	5.4.1
R29D_F	GCTGTACCTCGACGCACTGACC	5.4.1
R29D_R	GGTCAGTGCGTCGAGGTACAGC	5.4.1
R59A_ii_F	CGAAGCTGGCCAAGGACTTCTC	5.3.2
R59A_ii_R	GAGAAGTCCTTGGCCAGCTTCG	5.3.2
SCO3092Complete_For	GGCATATGAGCACCACGGAGCGTCCC	4.3.6
SCO3092Complete_Rev	GGGATATCGGAGGCCTTGGCCTCGGT	4.3.6
SCO3092_KO_For	CGCATCCAGAAGAAGATGCGTTACGGCGAGGCGA CCGTCATTCCGGGGATCCGTCGACC	4.3.4
SCO3092_KO_Rev	CGGGAGTCAGGAGGCCTTGGCCTCGGTCTTCTCC TGCTTTGTAGGCTGGAGCTGCTTC	4.3.4
SCO3092_ROP1_FOR	CCGAATTCGCCATGATCTCCGTCACGTG	4.3.5
SCO3092_REV	CCGAATTCTGAGTACTGCTCAGTACTA	4.3.5
Y111F_F	GTTTCGCCTCCCGCGGGTTC	5.4.2
Y111F_R	CACCGAAGTTGGCGAGCGCG	5.4.2
Y111R_F	GCTCGCCAACCGCGGTGGTTTC	5.4.2
Y111R_R	GAAACCACCGCGGTTGGCGAGC	5.4.2

Table 2.2: The mutagenesis primers used during the course of this study. The underlined regions represent restriction sites or parts of restriction sites introduced as part of the primer design.

SPR and EMSA primers:

Name	Sequence (5' to 3')	Section
1930_EMSA_F	GACGTCCTCCAGCACGGCGC	3.3.2
1930_EMSA_R	CAGCGCGCGGCGATGAAG	3.3.2
1a	*TCCGCTGCGTCCTGTGACCTGCTTCACAGGGCGCCTT	3.3.5
1b	AAGGCGCCCTGTGAAGCAGGTCACAGGACGCAGCGGA	3.3.5
2370_GS_F	CCACGCGCGTGCGCGAGG	3.3.2
2370_GS_R	CCACGCTGTTCCTTCGCTGTCAG	3.3.2
3101_EMSA_F	GTGATCCGCTGGTCGCCGAC	3.3.2
3101_EMSA_R	GACATCCGGCGCACGCTCCAG	3.3.2
3137_EMSA_F	GGATGTCGCCCTCGACGAAC	3.3.2
3137_EMSA_R	GCAAGCTGAAGTTCCTCGCG	3.3.2
3547_GS_F	GCATTCACCCATGTGTCACCCGG	3.3.2
3547_GS_R	CCATTCGTCCTCCTTGACGCTTGG	3.3.2
3615_GS_F	GCACCGCACGCCCCGGG	3.3.2
3615_GS_R	CCACGTGCGCTCCTCGCTC	3.3.2
3790_GS_F	GGAGAGTGAGAATGCCCATG	3.3.2
3790_GS_R	CCAGGATGCGAACCGGACG	3.3.2
5032_GS_F	GCACGTGCTCTCCAAAAACGCAGC	3.3.2
5032_GS_R	CTGGACACTAATAGCTACCTCCGAT	3.3.2
5207_GS2_F	CGACTCGCCCTCCGCGCCCCCTTGTTG	3.3.2
5207_GS2_R	CCTGGCGGAGTGTGTGGGCGGGACCGATG	3.3.2
5408/9_EMSA_F	CTTGAAGTCGGAACATCGCCCAC	3.3.2
5408/9_EMSA_R	GAGCAGACGAGCAGGAGGAG	3.3.2
5435/6_EMSA_F	CTTGGCTTTGGGTGCGGCAG	3.3.2
5435/6_EMSA_R	CTGAACAGCGCGTACCCGAC	3.3.2
5797_EMSA_F	CGTGTTCGACGCGGTGAGCG	3.3.2
5797_EMSA_R	GCTTGGCGTCGGCCTCGTCC	3.3.2
6168_GS2_F	GGAGGGCGCCGGAGAGCCCGGCGCGTT	3.3.2
6168_GS2_R	CACCTTCGGTCGTTCCCCAGCCAGAACCAG	3.3.2
6218_half_For	GACATTGTGAAGATTGCATGAGAAAT	3.3.5
6218_half_Rev	ATTTCTCATGCAATCTTCACAATGTC	3.3.5
6239_half_For	GCGATAGTGAATGGAGGAGGAACGCC	3.3.5
6239_half_Rev	GGCGTTCCTCCTCCATTCACTATCGC	3.3.5
6280_GS_F	CCTCAGGGTCACCGACGCTC	3.3.2
6280_GS_R	CGAAAGCGCATAACTCCCCCAG	3.3.2

6917_EMSA_F	CACCGGTTGGCACACAGGCT	3.3.2
6917_EMSA_R	GTTTCTGAGCTTCGCTCGGCCC	3.3.2
6917_half_For	GACTTTGTGAATTGAAACCGCCGAGG	3.3.5
6917_half_Rev	CCTCGGCGGTTTCAATTCACAAAGTC	3.3.5
7697_GS_F	CGTCAACTCCCCGAGGGGCAGG	3.3.2
7697_GS_R	GGTGAACACCAGGGTCGCCG	3.3.2
E68.18ci	GGCGACGGTGGCCTCGGGAATC	3.3.2
Full_26_For	TCCATTGTGAACTGCTGCACATGGTT	3.3.5
Full_26_Rev	AACCATGTGCAGCAGTTCACAATGGA	3.3.5
Half_26_For	TCCATTGTGAACTATCATGTGCGGTT	3.3.5
Half_26_Rev	AACCGCACATGATAGTTCACAATGGA	3.3.5
Nca	CTGCGTCCCTAGGCACCCGGCGGTGGGC	3.3.4
Ncb	GCCCACCGCCGGGTGCCTAGGGACGCAG	3.3.4
ndh_414_for	CGGTACCGGTAGCGGTCTGAGCAGGAC	4.3.2
ndh_414_rev	CATCGAGTATCCACCCGGTTGAGG	4.3.2
Nuo-4F	CTGCGTCCTGTGACCTGCTTCACAGGGC	3.3.4
Nuo-4R	GCCCTGTGAAGCAGGTCACAGGACGCAG	3.3.4
NuoF	CTGCGTCTTGTGACCTGCTTCACATGGC	3.3.4
NuoR	GCCATGTGAAGCAGGTCACAAGACGCAG	3.3.4
Ran_Bio_F	*GGAGCGCCGCCTTCGGCCCCCCTGCCGCCGGCGACGT	3.3.5
NUOROP1	AGTTGGGCTTGTGACCTGCTTCACATGTTCGCGATCT	3.3.4
NUOROP2	*AGATCGCGAACATGTGAAGCAGGTCACAAGCCCAACT	3.3.4
Random_R	ACGTCGCCGGCGGCAGGGGGGCCGAAGGCGGCGCTCC	3.3.5
Random26_F	TCCAATGGGGCTGGCCGACCTCGGTT	3.3.5
Random26_R	AACCGAGGTCGGCCAGCCCCATTGGA	3.3.5
rexGSrev2	CGATGAGTGAGGAACGAGAGTACG	3.3.2
SCO4461_For	CTTCGGCGATCCTGCGGAAG	3.3.2
SCO4461_Rev	GTCACGGATGTCCGGAGTGC	3.3.2
SCO5013_For	CTCACCACCCGGCACGACAA	3.3.2
SCO5013_Rev	CTGCGCTGCGGCGAGGGC	3.3.2
SCO5810_For	GAGTACGACCATGAGTTGGCAG	3.3.2
SCO5810_Rev	GTTCCAGCCGACGGTCCCG	3.3.2
SCO6218_For	CTTCTCGGTAGCGGGACGT	3.3.2
SCO6218_Rev	GAGGTGGATTCCGGCCGTCAG	3.3.2
SCO6239_For	GATGTCGCTCTGCCATGACTG	3.3.2
SCO6239_Rev	CACTCTGGCTGACCGCGGAG	3.3.2
SCO6383_For	CGAGTTCAGCGCGCACAACC	3.3.2
SCO6383_Rev	CAACTGTCGAGGTGTCTAC	3.3.2

Table 2.3: Primers used to amplify EMSA probes and primers annealed to generate SPR test fragments. The *indicates the presence of a biotin addition at the 5' end of a primer.

RT/qPCR primers:

Name	Sequence (5' to 3')	Section
16S_QF	CACTAGGTGTGGGCAACATTC	3.4.1
16S_QR	GTCGAATTAAGCCACATGCTC	3.4.1
AHPC_RT_F	TCCAGCAGATCAACCACAAG	3.4.1
AHPC_RT_R	GCACACGAAGGTGAAGTCCT	3.4.1
Cyd_qPCR_F	CTGTTGAAGTTCCGGCAGAG	3.2.3
Cyd_qPCR_R	GGTTTGAGCGTGTTTGTCAC	3.2.3
Cyd_RT_F	CTGTCGGCGTACTTCATCCT	3.4.1
Cyd_RT_R	CTGGTTGAGGGTGGTGTTCT	3.4.1
HrdB1a	CCGGTCAAGGACTACCTCAA	3.2.3
HrdB1b	TGGATGAGGTCCAGGAAGAG	3.2.3
Ndh_qPCR_F	GGTACCGGTAGCGGTCTGAG	3.2.3
Ndh_qPCR_R	GAAACCCCAAAAGGGTCAAC	3.2.3
Ndh_RT_F	GGTCTACCTGTCCACCTCCA	3.4.1
Ndh_RT_R	CACACGATGGTGTTGGAGTC	3.4.1
ResA2_RT_F	GATTACAAGGGCAAGGTCGT	3.4.1
ResA2_RT_R	GGTCCTTGACGTCCTGGTAG	3.4.1

Table 2.4: The primers used for qPCR and RT-qPCR during this study.

Section 2.1.5 – Cell-lines and strains:

Genus/species	Strain	Genotype	Reference
Escherichia coli	DH5α	*F' Φ80 dlacZ Δ*M15 *Δ*(*lacZYA-argF*) *U169 recA1 endA1 hsdR17 (r^- m^-) supE44 medλ^- thi-1 gyrA relA1*	Invitrogen
	BL21(DE3)pLysS	*F' ompT hsdS_B (r_B^- m_B^-) dcm gal λ(DE3) pLysS Cm^r*	Novagen
	BW25113	*Δ*(*araD-araB*)*567 ΔlacZ4787*(::*rrnB-4*) *lacIp-4000*(lacIQ) *λ^- rpoS369*(Am) *rph-1 Δ*(*rhaD-rhaB*)568 *hsdR514*	(Datsenko and Wanner, 2000)
	ET12567	*dam-13:: Tn9 dcm-6 hsdM Cm^r*	(MacNeil *et al.*, 1992)
Streptomyces coelicolor A3(2)	M145	$SCP1^-$ $SCP2^-$	(Bentley *et al.*, 2002)
	S106	*Δrex* M145	(Brekasis and Paget, 2003)

Table 2.5: List of *E. coli* and *S. coelicolor* strains used in this study.

Section 2.2 – Common media, buffers and solutions

Within the following section is all the recipes of growth media, buffers and solutions used in this study. Note that where the buffer used was provided by a manufacturer it is not listed below.

Section 2.2.1 – Growth media:

Lennox broth (LB):

The following were dissolved in 1L of dH_20:

10g Difco bacto tryptone

5g Oxoid yeast extract

5g NaCl

1g Glucose

Difco nutrient agar (DNA):

Each 250ml Erlenmeyer flask contained 2.3g Difco Nutrient agar in 100ml distilled water.

NMMP:

The following were made up in 800ml distilled water:

2g $(NH_4)_2SO_4$

5g Difco Casamino acids

0.6g $MgSO_4.7_2O$

50g PEG 6000

1ml NMMP Minor Elements

NMMP Minor Elements:

1g/L $ZnSO_4.7H_2O$

1g/L $FeSO_4.7H_2O$

1g/L $MnCl_2.4H_2O$

1g/L anhydrous $CaCl_2$

Prior to use this media was supplemented with 15mM phosphate buffer (from 100mM NaH_2PO_4/K_2HPO_4, pH6.8) and 0.5% glucose (from 20% glucose).

SOB:

20g tryptone

5g yeast extract

2.5ml of 1M KCl

10ml of 1M $MgCl_2$

10ml of 1M $MgSO_4$

Made up to 1L with distilled water

Tryptone soya broth (TSB):

30g tryptone soya broth powder

Made up to 1L in distilled water

YEME liquid medium:

3g yeast extract

5g Difco bacto-peptone

3g malt extract

10g glucose

340g sucrose

Made up to 1L with distilled water

2xYT:

16g Difco bacto tryptone

10g yeast extract

5g NaCl

Made up to 1L in distilled water

Mannitol Soya Flour (MS) agar:

To each 250ml Erlenmeyer flask the following were added and twice autoclaved:

1g agar

1g soya flour

100ml 2% mannitol (dissolved in tap water)

Minimal Media (MM) Agar:

The following were combined and pH adjusted to 0.7-0.72:

0.5g L-asparagine

0.5g K_2HPO_4

0.2g $MgSO_4.7H_2O$

0.01g $FeSO_4.7H_2O$

200ml of this solution were then added to 250ml Erlenmeyer flask, containing 2g agar. After autoclaving each flask was supplemented with 1% glucose and 200µl NMMP minor elements.

MYMTE media:

4g Maltose

4g Yeast extract

10g Malt extract

20g Difco bacto agar

2ml R2YE trace elements

Made up to 1 litre with RO water

R2YE trace elements:

40mg $ZnCl_2$

200mg $FeCl_3.6H_2O$

10mg $CuCl_2.2H_2O$

10mg $MnCl_4.4H_2O$

10mg $Na_2B_4O_7.10H_2O$

10mg $(NH_4)_6Mo_7O_{24}.4H_2O$

Made up to 1 litre with RO water

R5 agar:

The following were made up 1L in distilled water, and 100ml aliquoted into each 250ml Erlenmeyer flasks containing 2.2g Difco bacto agar:

103g sucrose

0.25g K_2SO_4

10.12g $MgCl_2.6H_2O$

10g glucose

0.1g difco casamino acids

2ml R2YE trace elements

5g yeast extract

5.73g TES buffer

At time of use the following were added to each flask:

1ml KH_2PO_4 (0.5%)

0.4ml $CaCl_2$.2H2O (5M)

1.5ml L-proline (20%)

0.7ml NaOH (1M)

Supplemented Minimal Medium Solid (SMMS):

The following solution was made-up and 200ml were added to 250ml Erlenmeyer flasks, containing 3g agar:

2g Difco casamino acids

5.73g TES buffer

Made up in 1L of water and pH adjusted to 7.2

At time of use the following were added to each flask:

2ml 50mM NaH_2PO_4/K_2HPO_4*

1ml 1M $MgSO_4$

3.6ml 50% glucose

200µl trace elements

*The volume of this was altered for phosphate limitation studies.

Trace elements (per litre):

0.1g $ZnSO_4.7H_2O$

0.1g $FeSO_4.7H_2O$

0.1g $MnCl_2.4H_2O$

0.1g $CaCl_2.6H_2O$

0.1g NaCl

Section 2.2.2 – Antibiotics and growth media additives

Name	Stock concentration	Stock solution made up in:	Working concentration (liquid)	Working concentration (solid)
Ampicillin	100mg/ml	50% ethanol	100µg/ml	100µg/ml
Apramycin	50mg/ml	dH_2O (filter sterilised)	20-50µg/ml	20-50µg/ml
Chloramphenicol	25mg/ml	80% ethanol	25µg/ml	25µg/ml
IPTG	1M	dH2O (filter sterilised)	1mM	1mM
Kanamycin	50mg/ml	dH_2O (filter sterilised)	25-50µg/ml	25-50µg/ml
Nalidixic acid	25mg/ml	150mM NaOH	N/A	25g/ml
Spectinomycin	25mg/ml	dH_20 (filter sterilised)	50µg/ml	50µg/ml
Thiostrepton	50mg/ml	DMSO	1-12.5µg/ml	0.5-12.5µg/ml
X-gal	40mg/ml	N, N-dimethyl-formamide	N/A	40µg/ml

Table 2.6: List of additives for growth media, including their stock and usage concentrations.

Section 2.2.3 – Buffers and solutions:

TE Buffer:

10mM Tris/HCl, pH 8

1mM EDTA

10 x Primer annealing buffer:

100mM Tris-HCl, pH8.0

150mM $NaCl_2$

10mM EDTA

STE buffer:

75mM NaCl

25mM EDTA, pH8.0

20mM Tris-HCl, pH7.5

P-buffer:

103g sucrose

0.25g K_2SO_4

2.02g $MgCl_2.6H_2O$

2ml R2YE trace elements

Made up to 800ml in distilled water and split into 80ml aliquots for autoclaving.

At time of use the following were added:

1ml KH_2PO_4 (0.5%)

10ml $CaCl_2.2H_2O$ (3.68%)

10ml TES buffer (5.73%, pH7.2)

Phenol/chloroform:

25ml phenol

24ml chloroform

1ml isoamyl alcohol

Kirby buffer:

1% (w/v) sodium-triisopropylnapthalene sulphonate (TPNS)

6% (w/v) sodium 4-amino salycilate

6% (v/v) phenol

Made up in 50mM Tris-HCl, pH8.3

Section 2.3 – Methods

All of the methods used in this study are listed in the following section. Any method with an unedited protocol is listed in full, whereas methods with changeable parameters are mentioned briefly here and detailed in their appropriate sections.

Section 2.3.1 – General DNA manipulation methods

Annealing primers destined for vectors:

Equimolar amounts of each primer mixed with 10 x annealing primer buffer (diluted to 1 x in reaction). Made up to volume with ddH20, heated to >95ºC for 2 min then allowed to cool to 50ºC slowly. Placed on ice until use or stored at -20ºC.

Polymerase chain reaction:

Each reaction consisted of 1µl 50ng/µl template DNA, 5µl 10x polymerase buffer, 1.5µl 10mM dNTPs, 5µl 10pmol/µl of each primer, 1µl polymerase made up to 50µl with ddH_20. Reactions were also supplemented with either 2.5µl

100% DMSO or 5µl 50% glycerol when appropriate. The polymerase used was determined by the intended purpose of the product.

Thermocycler conditions:

Initial denaturation:	2 min at 95°C	
Denaturation:	30 sec at 95°C	30 cycles
Annealing:	30 sec at 50-68°C	
Extension:	~30 sec/kb of product at 72°C	
Final extension:	5 min at 72°C	
	Hold at 4°C	

The conditions of the annealing and extension phases were varied to optimise the reaction for each primer and template combination.

Restriction digest:

All restriction digests in this study contained approximately 1µg DNA, 2µl 10X restriction buffer, ~10U of each enzyme, in a total volume of 20µl. Digestion was allowed to progress for 2h at the temperature recommended by the manufacturer of the enzyme.

Partial digests:

Where a partial digest was necessary to isolate an intact fragment the reaction was carried out as follows: 5µg vector DNA, ~10U restriction enzyme, 4µl 10X restriction buffer, made up to 40µl with ddH_2O. The reaction was started by the addition of the restriction enzyme, taking 4.5µl samples into 1µl of 0.5M EDTA (on ice) across a 60 min time-course. All samples were analysed via agarose gel electrophoresis, the samples producing the desired fragment sizes were combined and used to isolate the segment via gel extraction.

Klenow 3' end filling:

To the cut vector the following were added: 10µl 5X Klenow buffer, 1µl Klenow fragment, 4µl 1mM dNTPs, to 50µl with ddH_2O. The reaction mixture was incubated at room temperature for 15 min, 50µl ddH2O were added and immediately stopped with 50µl phenol:chloroform. The samples were then extracted and purified by isopropanol precipitation.

Phosphorylation of DNA primers:

Each primer was phosphorylated as follows: 4µl 200pmol/µl primer, 2µl 10X kinase buffer, 2µl 10mM ATP, 1µl T4 polynucleotide kinase, made up to 20µl with ddH_2O. Incubated at 37°C for 20 min and then used as normal for PCR.

Dephosphorylation of DNA:

To inhibit self-ligation of digested vectors the 5' ends were dephosphorylated as follows: 1µl shrimp alkaline phosphatase, cut vector, 3.5µl phosphatase buffer, made up to 35µl with ddH_2O. Incubated at 37°C for 30 min, then heat inactivated at 65°C for 20 min.

Ligation:

Each ligation reaction consisted of 100ng vector DNA, the appropriate quantity of insert DNA (see equation below), 1µl 10X T4 DNA ligase buffer, 1µl T4 DNA ligase, in a total volume of 10µl. The reaction was then allowed to proceed at either 4°C for 16h or 16°C for 4h.

Equation 1: Calculating the amount of insert to use per ligation reaction, with a 3:1 insert:vector ratio.

$$Insert(ng) = \frac{100ng \times size\ of\ insert(kb)}{size\ of\ vector(kb)} \times \frac{3}{1}$$

Section 2.3.2 – DNA extraction methods

Small scale plasmid DNA extraction:

For each purification a 3ml cell-pellet was resuspended in 200µl 50mM Tris/HCl, pH 8, 10mM EDTA. Immediately 400µl of 200mM NaOH, 1% SDS were added, mixed by inverting, followed by 300µl of 3M potassium acetate, pH 5.5. The contaminating RNA was degraded by the addition of 1µl 10mg/ml RNAse, followed by a 10 min incubation at room temperature. The samples were centrifuged at 16,100 x *g* for 5 min and the supernatant extracted with 150µl phenol/chloroform. The mixtures were vortexed for 2 min, centrifuged as before and precipitated with 600µl isopropanol. The samples were incubated on ice for 10 min and centrifuged as before. The pellet was washed with 70% ethanol and finally resuspended in 30µl TE buffer.

Large scale plasmid DNA extraction:

For large-scale DNA purifications the QIAgen Plasmid Midiprep kit was utilised, as detailed by the manufacturer. Each preparation required a cell pellet from 50ml of culture, producing ~100µg of plasmid.

Chromosomal DNA extraction from *S. coelicolor*:

For each strain 1ml of culture was harvested by centrifugation at 16,100 x *g* for 3 min. The pellets were washed with 1ml 10.3% sucrose and centrifuged as before. The final pellet was resuspended in 250µl STE buffer, containing 2mg/ml lysozyme, and incubated at 37°C for 30 min. To this mixture 330µl kirby buffer were added, vortexed and centrifuged for 5 min at 16,100 x *g*. The upper phase was extracted with 250µl phenol/chloroform, vortexed and spun as before. The DNA was precipitated with 400µl isopropanol, 40µl 3M sodium

acetate and incubated at -20°C for 2h. The DNA was pelleted by centrifugation, washed with 70% ethanol and resuspended in TE buffer. The DNA was treated with 10µg/ml RNAse at 37°C for 30 min, extracted with 100µl phenol/chloroform, then precipitated with an equal volume of isopropanol and 1/10 volume of 3M sodium acetate. Following centrifugation, as before, the pellet was washed and finally resuspended in TE buffer.

Section 2.3.3 – RNA extraction methods

RNA extraction from liquid cultures:

For each sample 15ml of culture were centrifuged at 3,824 x *g* for 1 min and immediately resuspended in 800µl Kirby buffer. The samples were sonicated for 2 x 3 sec at 30% (Vibracell, Sonics & Materials Inc.) and 600µl phenol/chloroform added. The samples were vortexed and centrifuged for 5 min at 16,100 x *g*. The upper phase was extracted with 800µl phenol/chloroform, vortexed for 2 min and centrifuged for 5 min at 16,100 x *g*. The upper phase (900 µl) was mixed with 90µl 3M sodium acetate (pH5.2) and 900µl isopropanol, and placed at -20˚C for 1h. The samples were centrifuged for 10 min at 16,100 x *g*, the supernatant discarded and pellets washed with 100µl 70% ethanol. The pellets were resuspended in 200µl 1x DNAse buffer, 0.5µl DNAse was added and samples were incubated at 37˚C for 30 min. To each sample 200µl RNAse-free water were added, along with 200µl phenol/chloroform. This was vortexed for 2 min and centrifuged for 5 minutes at 16,100 x *g*. The upper phase was mixed with 40µl sodium acetate and 400µl isopropanol. The samples were placed at -20˚C for 1h and centrifuged for 10 min at 16,100 x *g*. The supernatant was removed and each pellet washed with 100µl 70% ethanol. The pellets were air dried and resuspended in 50µl RNAse-free water. Each sample was quantified on a Nanodrop™ and the 260/280 ratio recorded. The samples were stored at -80˚C prior to use.

RNA extraction from solid-media culture:

In all cases MYMTE agar was overlaid with twice boiled cellophane discs and inoculated with 50µl water containing $5x10^7$ spores. The cells were harvested in liquid nitrogen and were cryogenically ground for 1.5 min at full power (Mixer Mill MM301, Reitsch). This stage was repeated for at least 5 cycles, with 2 min in liquid nitrogen in between cycles. The material was removed with 5ml modified Kirby buffer, and a further 1ml Kirby was used to rinse the remaining matter from the cylinder. This suspension was then sonicated for 30 sec at 35% (Vibracell, Sonics & Materials Inc.) and 5ml phenol/chloroform were added. This was then vortexed for 2 min and centrifuged at 3,824 x *g* for 10 min. The upper phase was extracted with 5ml phenol/chloroform, vortexed and centrifuged as before. The upper phase was then precipitated with 7ml isopropanol and 700µl 3M sodium acetate. This was then placed at -20°C for at least one hour prior to centrifugation at 3,824 x *g* for 10 min. The supernatant was discarded and the pellet washed with 1ml 70% ethanol. The pellet was resuspended in 875µl ddH_2O, 100µl 10x DNAse buffer and 25U DNAse were added. The samples were then incubated at 37°C for 1h. 400µl of phenol/chloroform were added, vortexed for 2 min and centrifuged for 10 min at 3,824 x *g*. The upper phase (900µl) was taken into 900µl isopropanol and 90µl 3M sodium acetate, this was kept at -20°C for at least 1h. The RNA was pelleted at 16,100 x *g* for 10 min. The pellet was washed with 200µl 70% ethanol, air dried and resuspended in 300µl ddH_2O.

Section 2.3.4 – Southern blot

The southern blot method is a means to detect the presence of a specific DNA fragment, by means of a sequence-specific radiolabelled probe, in a sample of restriction endonuclease treated chromosomal DNA. This method can be used to confirm an insertion or deletion event in the chromosome of an organism targeted for mutagenesis.

20x SSC:

175.32g NaCl

88.23g trisodium citrate

Made up to 1L with distilled water

Denaturing solution:

10g NaOH

43.8g NaCl

Made up to 500ml in distilled water

Neutralising solution:

38.5g ammonium acetate

0.4g NaOH

Made up to 500ml in distilled water

Pre-hybridisation buffer:

12.5ml 20x SSC

5ml 10% blocking agent

250µl 20% SDS

Made up to 50ml with distilled water

Southern digests:

Each digest contained the following:

~2.5µg chromosomal DNA

3µl 10x NEB restriction buffer

>2.5U of restriction enzyme

Made up to 30µl with ddH_2O and incubated at 37°C for ~16h.

Southern gel:

All samples were run on a large 0.8% agarose/TBE gel for 10 min at 150V, followed by ~16h at 20V. The gel also included 5 or 10µl hyperladder. The gel was photographed with ruler adjacent as a marker of ladder band positions.

Labelling the probe:

The probe was generated using 1µl Klenow polymerase, 120ng of denatured PCR product (or λ-DNA for labelling the ladder), 1µl dATP/dTTP/dGTP mixture, 2µl BSA, 10µl Klenow buffer and 4µl [α-P^{32}] dCTP. This reaction was incubated at room temperature for 1h, heated at 95°C for 2 min before placing on ice. The labelled probe was then purified using a 450 sephadex column and used immediately.

Southern blot:

The blot was assembled as follows: glass plate, gel, Hybond N membrane, two wet Whatman papers, two dry Whatman papers, 8-10 cm worth of paper hand towels, glass plate and 250g weight. This was left for 4h at room temperature before UV cross-linking (120 Joules) the DNA to the nitrocellulose membrane. The membrane was placed at 65°C in 25ml prehybridisation buffer for 1h before the labelled probe was added (along with 5µl of the λ-DNA) and incubated for a further 16h. The membrane was washed once with 2xSSC, 0.1% SDS and three times with 0.1xSSC, 0.1% SDS (30 min per wash). The membrane was then exposed to x-ray film for 16h and the results analysed.

Section 2.3.5 – Protein purification

Binding buffer:

20mM Tris-HCl, pH7.9

500mM NaCl

0.1mM PMSF

Charge buffer:

50mM $NiCl_2$

Wash buffer:

20mM Tris-HCl, pH 7.9

0.5M NaCl

60mM imidazole

Elution buffer:

20mM Tris-HCl, pH 7.9

0.5M NaCl

0.5M or 1M imidazole

Strip buffer:

20mM Tris-HCl, pH7.9

0.5M NaCl

100mM EDTA

Gel filtration dialysis buffer:

20mM Tris, pH7.9

100mM NaCl

10mM EDTA

5% glycerol

Thrombin cleavage buffer:

200mM Tris-HCl, pH8.0

1.5M NaCl

25mM CaCl2

Gel filtration running buffer:

20mM Tris, pH7.9

50mM NaCl

10mM EDTA

5% glycerol

Overexpression:

For large-scale protein purifications 500ml of LB were inoculated with the pellet of a 5ml overnight culture. The culture was grown at 37°C 250rpm to anOD_{600nm} of 0.5-0.7. The cultures were then either immediately induced with 1mM IPTG (final concentration) and placed at 37°C, or were placed in an ice-water bath for 10 min prior to induction and then grown at 30°C (cold-shock). After 3h the induced cells were harvested by centrifugation at 3,824 x *g* for 10 min.

Ni^{2+}-affinity chromatography:

The cell pellets were resuspended in 20ml binding buffer and disrupted by sonication at 35% for 6 x 15 sec (Vibracell, Sonics & Materials Inc.). The cell debris was pelleted by centrifugation at 23,426 x *g* for 20 min (4°C), and the clear cell lysate decanted immediately. A Ni^{2+} column was generated by applying ~1ml IDA-sepharose to a 5ml syringe, containing a glass-wool bung. Note that all buffers used on the column were made up in ddH_2O and filter sterilised. Using a peristaltic pump, driven at a flow rate of ~2ml/min, the column was washed with 3 column volumes of ddH_2O and charged with 5 column volumes of charge buffer. The column was then equilibrated with 5 column volumes of binding buffer prior to the addition of the clear cell lysate. The lysate was chased with 10 column volumes of binding buffer and 5 column volumes of wash buffer. The protein was finally eluted in two steps with ~3ml of 0.5M and 1M imidazole elution buffers (~6ml total eluate). The column was then cleared with 6 column volumes of strip buffer, rinsed with ddH_20 and stored in 20% ethanol. All fractions from the column were then analysed by SDS-polyacrylamide gel electrophoresis. Fractions containing the highest protein concentration were combined and dialysed, in twice boiled dialysis tubing, into 3L of gel-filtration dialysis buffer at 4°C for ~16h.

Thrombin cleavage:

Where the 6xHis tag was to be removed from the purified protein the following protocol was used: The concentration of the crude protein was estimated using an A_{280nm} reading (E_{280nm} for Rex of 14650 M^{-1} cm^{-1}). This value was then converted to mg/ml (MW of Rex 26.7kDa) and finally to total mg of protein. For each mg of protein 1U of thrombin was added. The appropriate volume of thrombin cleavage buffer was also added and the reaction incubated at room temperature for 6h. The reaction was terminated by the addition of 100µM PMSF and cleavage confirmed by SDS-polyacrylamide gel electrophoresis.

Gel filtration:

The samples were concentrated down to ~1ml and injected onto a Superdex 200 HiLoad 16/60 gel filtration column. Prior to loading samples the column was rinsed with 2 column volumes of ddH_2O, then pre-equilibrated with 2 column volumes gel filtration running buffer, run at 0.5-1ml/min. After elution the column was washed with 2 column volumes of each both running buffer and ddH_2O. The column was stored in 20% ethanol. All solutions were filter sterilised prior to use and the column was kept at 4°C during use. The peaks were detected at 280nm, collected and analysed by SDS-polyacrylamide gel electrophoresis. Fractions containing the desired protein were pooled and concentrated once more. The concentration of the protein was assessed at 280nm and 340nm on the Nanodrop™ prior to snap-freezing in liquid nitrogen and storing at -80°C.

Section 2.3.6 – Electromobility shift assay (EMSA)

EMSA analysis is a method used to detect the formation of a complex between multiple components, in this case between DNA and protein. One or more of the components are radiolabelled, allowed to equilibrate with the other components and then run through a gel. The individual radiolabelled components are also run individually. Complex formation is then detected as a difference in the migration through the gel when compared to the migration of the individual components.

5 x Binding buffer for EMSAs:

100mM Tris-HCl, pH8.0

25% glycerol

5mM $MgCl_2$

200mM KCl

6x Loading dye:

1% bromophenol blue

50% glycerol

6% polyacrylamide gel:

7.925ml 1x TBE

2ml 30% acrylamide

50µl 10% APS

25µl TEMED

Generating the radiolabelled probe:

PCR was used to generate a DNA probe, which was then purified using the QIAgen Gel Extraction kit (as per the manufacturers' instructions). The labelling reaction consisted of 100ng probe, 1.11MBq of $\gamma[^{32}P]$-ATP, 1µl T4 polynucleotide kinase, 2µl 10 x kinase buffer, made up to 20µl with dH_2O. This reaction was allowed to proceed at 37˚C for 30 min and the labelled probe purified using the QIAgen PCR purification kit. Each probe was quantified on an agarose gel and 1ng used in the binding reaction.

EMSA reaction:

Each probe-only sample consisted of 1ng of $\gamma[^{32}P]$-labelled DNA, 2µl of 5x binding buffer and 2µl of 6x loading dye. Each binding reaction additionally contained a defined amount of Rex, 1µg herring-sperm DNA and NAD/H as indicated. The reaction mixtures were incubated at room temperature for 20 min prior to running on a 6% TBE-polyacrylamide gel at 120V for 1h 20 min. After vacuum drying the gels were analysed either by X-ray film or using a storage-

phosphor screen. The length of the exposure was dependent on the strength of the radio-labelled probe.

Section 2.3.7 – Surface plasmon resonance (SPR):

SPR analysis is similar to EMSA analysis in that it too is able to detect protein-protein or protein-DNA interaction, however SPR is more sensitive and allows real-time detection of the interaction. In this case the DNA was attached to the sensor surface and the protein injected over the sensor chip. The interaction was then indicated by an increase in the response units from the sensor surface.

5 x HBS:

50mM Hepes, pH 7.4

750mM NaCl

17mM EDTA

1 x HBS (SPR running buffer):

10mM Hepes, pH 7.4

150mM NaCl

3.4mM EDTA

0.005% TWEEN 20

Annealing primers for SPR:

Biotinylated primers:

In each case only one primer was biotinylated, annealing the primers with a 4-fold excess of the non-biotinylated primer. This was made up in

1 x HBS buffer (lacking the TWEEN), heated to >95ºC for 2 min, allowed to cool slowly to 50ºC and was finally stored at -20ºC.

Non-biotinylated primers:

The non-biotinylated (for competing fragments) were annealed in equimolar quantities with the same buffer and reaction conditions as for the biotinylated fragments.

SPR running conditions:

SPR detects changes in the refracted light caused by alterations to the surface Plasmon resonance. SPR is a natural phenomenon that results from fluctuations in the localisations of the electron clouds of atoms. This resonance can be altered by changing the interactions that occur on the sensor surface. This method allows the interactions between two molecules to be studied and quantified. For all SPR assays a BIAcore 2000 system was used in conjunction with streptavidin sensor chips. Each new sensor chip was first washed three times with 30µl 1M NaCl, 50mM NaOH at a flow rate of 20µl/min. Each chip contained 4 flow cells, the first was always left blank for background subtraction, the second was used for the non-specific DNA control and the test fragment was attached to lane 3. Each 1ng/µl biotinylated fragment was injected onto the appropriate lane until the response units had increased by ~250. The chip was then twice washed with 30µl of 1M NaCl. The SPR assays were run at a flow rate of 30µl/min and the injection constituents varied. The sensor surface was then regenerated with 30µl 2M $MgCl_2$. For competitive SPR an additional DNA fragment (non-biotinylated) was included in the protein injection and the reduction in the response units, compared to protein alone, recorded.

Section 2.3.8 – Transcriptome analysis methods

RT-qPCR:

RT-qPCR is a PCR-based method to detect the expression level of a gene relative to the expression of a reference gene. RNA samples were converted to

cDNA and used as the template for PCR. Each reaction was monitored in real-time on an Applied Biosystem 7500 instrument using SYBR-green as the fluorescent dye. The number of transcripts was then determined through use of a gDNA standard curve.

Additional DNAse step:

The following were added to an RNAse-free PCR tube:

1µg RNA

1µl DNAse buffer

1µl RNAse-free DNAse

Made up to 10µl with RNAse-free ddH_2O.

The reaction was incubated at 37˚C for 30 min and stopped by the addition of 1µl DNAse stop solution, followed by a 10 min 65˚C heat-inactivation.

Reverse transcriptase step:

To the above solution the following were added:

4µl 5x iScript buffer

1µl iScript reverse transcriptase

4µl ddH_2O / 9µl ddH_2O (no RT controls)

This was then incubated as follows:

25˚C 5 min

42˚C 30 min

85˚C 5 min

Hold at 4˚C

qPCR:

To each optical PCR tube the following were added:

12.5µl 2x QuantiTect SYBR Green PCR Master Mix

1µl 10pmol/µl forward primer

1µl 10pmol/µl reverse primer

5µl template (gDNA or 2µl cDNA+3µl ddH_2O)

5.5µl ddH_2O

Cycling conditions (on Applied Biosystems 7500):

95˚C 15 min

94˚C 15 sec
55˚C 30 sec
72˚C 33 sec* } Cycled 45 times

95˚C 15 sec
55˚C 1 min
95˚C 15 sec } melting point analysis

*Data acquisition stage. Note that the detection threshold was lowered to 0.02, as per the manufacturer's instructions (QIAGEN QuantiTect SYBR Green PCR Kit in combination with the Applied Biosystems 7500 thermal cycler).

Data analysis:

All qPCR assays included a M145 gDNA standard curve for each set of primers, which was used to convert Ct values into copy numbers. When the starting material was DNA this was converted to copy number per pg. However, when the amount of starting material was unknown, as was the case for cDNA due to the conversion from RNA not being 100% efficient, the values were kept as copy numbers alone but were normalised to a control gene (16s rRNA). All samples were background corrected using a "no template" control for each primer set. The samples that had undergone a reverse transcriptase step also

had an additional data validation step, using the no RT controls to subtract the signal from any gDNA carry-over that had occurred.

S1 nuclease protection assay:

The S1 nuclease protection assay is a method used to determine the expression level of a gene by means of a sequence-specific radiolabelled probe. The probe anneals to the transcript of the target gene protecting it from S1 nuclease digestion and indicating its position on a polyacrylamide gel. The expression level is then indicated by the strength of the signal from that fragment on the gel, as this is determined by the number of transcripts that were originally present for that gene.

2 x S1 hybridisation buffer:

2.63g PIPES (made up in ~90ml dH_2O)

1.67ml 0.5M EDTA

Adjusted to pH7.0 with 5M NaOH

93.1g NaTCA dissolved in above solution

Made up to 167.4ml with RNase-free H_2O

5 x S1 digestion buffer:

1.4M NaCl

150mM Sodium Acetate, pH4.4

22.5mM Zinc Acetate

100µg/ml RNase-free herring sperm DNA

For use: 150U of S1 Nuclease added per 300µl of 1 x digestion buffer

S1 stop Mixture:

2.5M Ammonium Acetate

0.05M EDTA

Formamide loading buffer:

80% formamide

10mM NaOH

1mM EDTA

0.1% Xylene Cyanol

0.1% Bromophenol Blue

6% denturing polyacrylamide gel:

30ml 6% bis-acrylamide (urea/TBE)

175µl 10% APS

27.5µl TEMED

Labelling the ΦX174 DNA/*Hin*fI ladder:

The labelling reaction consisted of 1µl of ΦX174 DNA/*Hin*fI ladder, 1µl of 10x T4 polynucleotide kinase buffer, 1µl of γ[^{32}P]-ATP, 1µl of T4 polynucleotide kinase and 6µl of ddH_2O. This was incubated at 37˚C for 30 min, 200µl of loading dye were added and the ladder was stored at -20˚C prior to use.

Generating the radiolabelled probe:

30pmol of the reverse primer (located within the gene) was labelled with 1.85MBq of γ[^{32}P]-ATP, using ~10U of T4 polynucleotide kinase, in a final reaction volume of 40µl. The labelling reaction was allowed to proceed for 30 min at 37˚C, after which time 4µl 3M sodium acetate, pH6.0, and 80µl 100% ethanol were added. The labelled primer was left to precipitate at –80˚C for ~16h. The primer was pelletted by centrifugation, washed with 100µl 75%

ethanol and air-dried. The pellet was then resuspended in PCR reaction mixture, containing 20pmol of the forward primer (located within the promoter region), and amplified using appropriate cycling conditions for each probe. The PCR product was subsequently purified using the QIAgen PCR purification kit, eluting in 20µl of RNase-free water, 2µl of which was run on an agarose gel in order to quantify the probe.

Hybridisation:

The amount of RNA used for the hybridisation was varied between 30-40µg, depending on the amount available, but was kept constant within each S1 nuclease experiment. The RNA was mixed with ~10ng of the purified probe, and 10µl of 1 x hybridisation mixture, in a flip-top eppendorf tube. The probe was denatured at 65˚C for 20 min and allowed to cool slowly to 45˚C, during which time the probe would anneal to the RNA.

S1 nuclease digestion:

To each hybridisation mixture 300µl of 1 x S1 Nuclease digestion buffer was added and incubated at 37˚C for 45 min. The reaction was stopped by the addition of 75µl S1 stop solution. The RNA was pelleted, with 1µl 20mg/ml glycogen and 400µl isopropanol, at -20˚C for 1h. The pellet was washed with 150µl 70% ethanol, air-dried and finally resuspended in 6µl formamide loading dye. The samples were denatured at 95˚C for 2 min prior to running on a 6% sequencing gel. Gels were run at 600V and 52.5˚C for 1h 20 min. After vacuum drying the gels were either analysed on x-ray film or by phosphor-imager.

Section 2.3.9 – ChIP-on-chip

ChIP-on-chip is a method of detecting the *in vivo* binding sites of a DNA-binding protein using an antibody specific to the protein of interest. In brief the protein is

formaldehyde cross-linked to the DNA, samples sonicated to fragment the chromosomal DNA, and the protein-DNA complexes purified using a specific antibody. The cross-linking was reversed by heat-treatment, the protein protease treated and the purified DNA labelled with Cy3 or Cy5 for use on a DNA microarray. The same method was also performed on a control sample, lacking the antigen for the antibody, and labelled with the opposite Cy-dye to the test sample. Any spots on the array that corresponded to a binding site for the protein would then be enriched compared to the control sample.

(method adapted from (Efromovich *et al.*, 2008):

Tris buffered saline (TBS):
20mM Tris-HCl, pH7.5
150mM NaCl

IP lysis buffer:
10mM Tris-HCl, pH8
20% sucrose
50mM NaCl
10mM EDTA
20mg/ml lysozyme

IP buffer:
50mM Hepes-KOH, pH7.5
150mM NaCl
1mM EDTA
1% Triton X-100
0.1% Na-deoxycholate
0.1% SDS

IP salt buffer:

50mM Hepes-KOH, pH7.5
500mM NaCl
1mM EDTA
1% Triton X-100
0.1% Na-deoxycholate
0.1% SDS

IP wash buffer:
10mM Tris-HCl, pH 8.0
250mM LiCl
1mM EDTA
0.5% Nonidet P-40 alternative
0.5 % Na deoxycholate

IP elution buffer:
50mM Tris, pH 7.5
10mM EDTA
1% SDS

Cryogenic grinding method:
An alternate means of cell disruption is cryogenic grinding, which was tested during the optimisation of the ChIP-on-chip protocol. For this the cell pellet was resuspended in 1ml of TBS, containing 1mM PMSF. The solution was then dripped into liquid nitrogen and the pellets placed into a grinding cylinder. This was then cycled 3 times (Mixer Mill MM301, Reitsch), grinding for 90 sec at full power then placing in liquid nitrogen for 2 min. Samples were then removed from the cylinders in 4ml IP buffer and sonicated for 11 x 15 sec at 35% (Vibracell, Sonics & Materials Inc.). The samples were phenol/chloroform extracted, isopropanol precipitated and resuspended in ddH_2O before analysing on an agarose gel.

Formaldehyde cross-linking:

The spores were germinated for 2h before being used to inoculate 50ml NMMP, which was then grown at 30°C 300rpm to an OD_{450nm} of 0.8-1. When the cultures were ready to harvest 1.35ml of 37% formaldehyde were added and the flasks were incubated at 30°C 300rpm for a further 20 min. The formaldehyde was quenched by the addition of 8ml 2.5M glycine, followed by a further 5 min incubation. The cells were finally harvested by centrifugation at 2,245 x *g* for 5 min at 4°C. The cell pellet was washed twice with TBS, 10ml then 5ml, and the final pellet was resuspended in 1ml IP lysis buffer. The lysis reaction was left to progress for 30 min at 37°C, vortexing every 10 min to ensure complete digestion. Once completed 4ml of IP buffer were added, along with 50µl 100mM PMSF (final 1mM). The samples were sonicated at 35% for 13 x 15 sec with 1 min on ice in between cycles. The cell debris was pelleted by centrifugation at 16,100 x *g* for 30 min at 4°C and the supernatant transferred to a fresh tube for the immunoprecipitation step.

Immunoprecipitation:

25µl of protein A/G resin (Ultralink) were washed 3 times in 125µl of TBS, pelleting the beads at 830 x *g* after each wash. The supernatant was completely removed and the resin resuspended in 800µl of the formaldehyde cross-linked sheared chromatin. Finally 5µl of anti-FLAG (Sigma) were added and the immunprecipitation left to proceed on a rotating wheel at 4°C for ~16h. The beads were harvested at 830 x *g* for 1 min, resuspended in 750µl IP buffer and returned to the wheel for 3 min. The solution was transferred to a Spin-X column (Corning Life Sciences) and centrifuged as before. The resin was resuspended in 500µl IP buffer, returned to wheel for 3 min and centrifuged as before. The column was washed with 1ml IP salt buffer, 1ml wash buffer and finally 1ml TE buffer, each time mixing on the rotating wheel for 3 min and centrifuging for 1 min. The column was finally transferred to a fresh tube and the resin resuspended in 100µl elution buffer. This was incubated at 65°C for 20 min, to release the chromatin from the protein A/G beads. The chromatin was then eluted by centrifugation at 830 x *g* for 1 min. The sample was transferred

to a PCR tube, with 10µl of 40mg/ml Pronase, and heated to 42°C for 2h, then 65°C for 6h. A total chromatin control was generated by mixing 70µl of unused chromatin, 20µl of 5 x elution buffer, and 10µl of 40mg/ml Pronase made up in TBS (Roche). The control was heated using the same cycle conditions as the immunoprecipitated samples. All samples were purified using the QIAgen PCR purification kit (as per the manufacturers' instructions) eluting in 30µl buffer EB.

Labelling the chromatin for microarray studies:

Approximately 150ng of chromatin was used for each labelling reaction. The reaction mixtures were made up with 20µl 2.5 x random primer (BioPrime kit), 20µl of sample and 0.25µl dH_2O. These were then mixed well and denatured at 94°C for 3 min. Once this step was completed 5µl dNTP mix (2mM dATP, 2mM dGTP, 2mM dTTP and 0.5mM dCTP) were added, followed by 3.75µl of 1mM Cy3-dCTP or Cy5-dCTP and 1.5µl of Klenow (BioPrime kit). The labelling reaction was left for ~16h at 37°C in the dark. The labelled samples were purified using the MinElute PCR purification kit (QIAgen) with a slightly modified protocol. The 50µl labelling mixture was mixed with 250µl buffer PB and applied to the column. This was then centrifuged at 16,100 x *g* for 1 min and the flow-through discarded. The column was washed twice with 500µl, then 250µl of buffer PE, followed by a third centrifugation step to remove the residual ethanol. The columns were transferred to fresh tubes and the samples eluted in two 15µl elution steps. Light exposure was limited throughout the purification to limit dye bleaching.

Hybridisation:
(This stage was performed at the *Streptomyces* microarray facility at the University of Surrey, Guildford)

The immunoprecipitated samples were hybridised to an OGT 4 x 44 K 60mer slides, with 4 separate hybridisation chambers (Bucca *et al.*, 2009). For each hybridisation chamber 150ng of both Cy3 and Cy5 labelled samples (denatured at 94°C for 3 min) were mixed in a total volume of 120µl hybridisation buffer (50mM MES, pH7, 1M NaCl, 20% formamide, 1% Triton X-100) (Bucca *et al.*,

2009). Each solution was then pipetted onto one chamber of an Agilent gasket slide and the OGT array placed face down onto these solutions. The gasket was sealed and placed in an Agilent hybridisation chamber for 60h at 55°C (Bucca *et al.*, 2009). The arrays were washed and analysed as detailed by Bucca *et al.* (Bucca et al., 2009). The final data was represented as fold-enrichment, based on comparison to the no antigen control signal at each position on the array.

Section 2.3.10 – Culturing methods

Mycelial preps:

5ml of YEME:TSB (50:50), containing the appropriate antibiotic selection, were inoculated with 5µl of a fresh high density spore stock and grown for 48h at 30ºC (300rpm). The 5ml culture was combined with 5ml sterile water and centrifuged for 10 min at 1,698 x *g*. The supernatant was removed and the mycelia pellet resuspended in 15ml 10.3% sucrose, which was centrifuged as above. The supernatant was again discarded and the pellet washed with 800µl 20% glycerol. The solution was transferred to a 2ml eppendorf and centrifuged at 16,100 x *g* for 3 min. The final pellet was resuspended in 500µl 20% glycerol and stored at -80ºC prior to use.

S. coelicolor protoplast generation:

25ml of YEME, containing 2.5ml of 20% glycine and 200µl 2.5M $MgCl_2$, were inoculated with 50µl spores and grown at 30ºC for 48h. 15ml of the culture were mixed with 30ml of H_2O and centrifuged at 2,653 x *g* for 10 min. The supernatant was discarded and the pellet was washed twice with 20ml of 10.3% sucrose. The pellet was finally resuspended in 4ml P-buffer, containing 1mg/ml lysozyme, and incubated at 30ºC for 1h. A further 5ml of p-buffer were added and the solution filtered through cotton-wool. The filtrate was centrifuged at 955 x *g* for 7 min and the supernatant discarded. The pellet was resuspended in the residual volume and an additional 1ml P-buffer were added.

Section 2.3.11 – *S. coelicolor* genetic manipulation methods

Transformation of *S. coelicolor* protoplasts:

For each transformation 50µl of protoplast were added to 1ml P-buffer, this was centrifuged at 955 x *g* for 7 min. The supernatant was discarded and the pellet resuspended in the residual volume. 5µl of unmethylated DNA were added to the protoplast immediately followed by 500µl 25% PEG (made up in P-buffer). 100µl of this were plated onto R5 agar and incubated at 30 ºC for ~16h. The plates were then overlaid with 1ml of the appropriate antibiotic and incubated at 30ºC for a further 3 days.

Conjugation of *S. coelicolor*:

For conjugation into *S. coelicolor* the E. coli strain ET12567/pUZ8002 was used. 100µl of overnight culture were used to inoculate 10ml LB, containing 25µg/ml chloramphenicol, 25µg/ml kanamycin and selection for the conjugative vector. This was grown to an OD_{600nm} of 0.4 and cells pelleted by centrifugation at 3,824 x *g* for 5 min. The pellet was washed twice with fresh LB to remove any residual antibiotics and centrifuged as before. The final pellet was resuspended in 1ml LB and 500µl used for each conjugation. 4µl of spores were added to 500µl 2xYT media and heat shocked at 50ºC for 10 min. The spores were allowed to cool and were then mixed with the *E. coli* cells. The samples were centrifuged briefly at 16,100 x *g* and the supernatant discarded. The pellet was resuspended in the residual volume and a dilution series was generated in sterile water. 100µl of each dilution was plated onto MS agar, containing 10mM $MgCl_2$, and incubated at 30ºC for ~16h. The plates were then overlaid with 1ml of water containing the appropriate antibiotic selection as well as 0.5mg nalidixic acid to select against the *E. coli* cells.

Section 2.3.12 – Generating an in-frame disruption strain in *S. coelicolor*

This method followed the protocol described by Gust *et al.* (Gust *et al.*, 2004, Gust, 2002). The first stage was the isolation of the cosmid, containing the gene

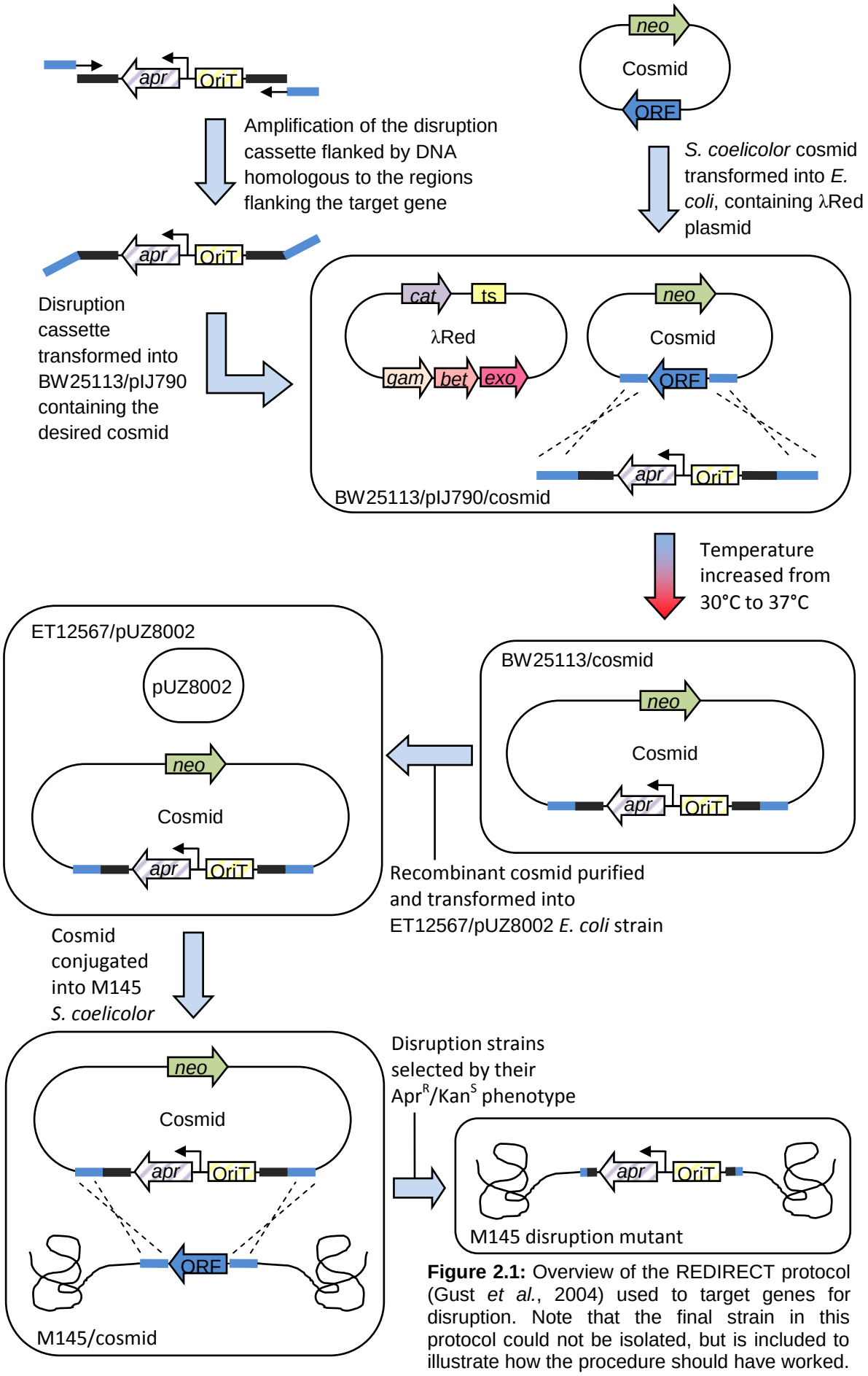


Figure 2.1: Overview of the REDIRECT protocol (Gust *et al.*, 2004) used to target genes for disruption. Note that the final strain in this protocol could not be isolated, but is included to illustrate how the procedure should have worked.

targeted for disruption, from the *S. coelicolor* cosmid library. This was transformed into electrocompetent *E. coli* cells, cultured and subsequently purified. The second stage was generation of a disruption cassette. In all instances during this study the apramycin resistance cassette was chosen, contained on pIJ773. Primers were designed to amplify the apramycin resistance cassette of pIJ773, also incorporating 39 bp of the DNA surround the open reading frame of the targeted gene. The cosmid was transformed into the BW25113 *E.coli* strain containing the λ-Red recombination plasmid (pIJ790). The transformant was then grown overnight and used to inoculate 10ml SOB media 100µg/ml ampicillin, 50µg/ml kanamycin, 25µg/ml chloramphenicol and 10mM arabinose to induce λ-Red genes. The λ-Red plasmid contained genes to allow recombination of the disruption cassette with the target gene on the cosmid. The cultures were grown to an OD_{600nm} ~0.4 at 30ºC. The cells were harvested by centrifugation at 1,698 x *g* for 5 min and the pellets washed twice with 10ml then 5ml of 10% glycerol, finally resuspended in the residual volume. 100ng of the disruption cassette was mixed with 50µl of cells and electroporated at 2.5kV. The cells were plated onto L-agar, containing 50µg/ml apramycin, 50µg/ml kanamycin and 100µg/ml ampicillin, and grown at 37ºC. The temperature at this stage was crucial for inducing the loss of the λ-Red plasmid, which is temperature sensitive, in order to prevent the disruption cassette being recombined back out of the cosmid. The recombinant cosmid was purified from the strain and confirmed by restriction analysis. This was then used to transform the ET12567/pUZ8002 *E. coli* strain, which was then conjugated with *S. coelicolor*. The conjugation plates were overlaid with 1ml of water, containing 0.5mg nalidixic acid and 1.25mg apramycin. The ex-conjugants were then screened for Apr^R/Kan^S colonies by replica plating onto DNA media.

Chapter 3

Results I: The Rex Regulon

“Look deep into nature, and then you will understand everything better.”

Albert Einstein (1879-1955)

Section 3.1 – Overview

As a transcriptional repressor Rex is able to sense changes in the NADH/NAD^+ redox poise of the cell and alter gene expression accordingly, which ensures the maintenance of redox homeostasis. During aerobic growth this ratio should remain low, as NADH is continually recycled by the respiratory chain components. However, under oxygen limitation this pathway slows in the absence of the terminal electron acceptor; oxygen. This results in an increase in the NADH/NAD^+ ratio resulting in redox stress as the NAD^+ becomes limiting for other cellular processes. Redox stress responses can vary from species to species, especially so considering that the respiratory pathways also vary between organisms. Pathogenic bacteria are constantly exposed to oxygen limitation as a consequence of the hosts' defences. Under these conditions bacteria have been shown to switch to alternate terminal electron acceptors; such as nitrate, and have even been shown to enter a dormant stage to lessen the energy demands on the cells (Rustad *et al.*, 2009, Unden and Bongaerts, 1997). As a soil-dwelling bacterium *S. coelicolor* is frequently oxygen limited but this organism cannot grow anaerobically (van Keulen *et al.*, 2007, van Keulen *et al.*, 2003). One of the coping mechanisms of *S. coelicolor* has already been identified; the Rex-regulated induction of the cytochrome *bd* terminal oxidase (Brekasis, 2005, Brekasis and Paget, 2003). In other bacteria this enzyme has been shown to have an increased affinity for oxygen and is therefore able to ensure continued electron flow at low oxygen concentrations (Poole and Cook, 2000). Rex is not however limited to obligate aerobes, it is also present in facultative and obligate anaerobes – presenting alternative options for coping with oxygen limitation. For example *Staphylococcus aureus* also contains a Rex regulator but in this species the Rex regulon appears to include genes involved in fermentative pathways; such as alcohol dehydrogenase and lactate dehydrogenase (Pagels *et al.*, 2010). These enzymes also exist in *S. coelicolor* but have not been shown to be Rex-regulated. Other *S. coelicolor* Rex targets have been identified, including *nuoA-N* and *ndh* (Brekasis, 2005). However, it was thought that the regulon was still not fully characterised in *S. coelicolor*, potentially missing such genes capable

of utilising other energy sources, thus a transcriptomics approach was taken. Whilst the method itself proved successful it did not reveal new Rex targets. Bioinformatics approaches had however revealed several potential Rex binding sites in the *S. coelicolor* genome, which although absent using the transcriptomic approach, were supported by EMSA (D. Brekasis and M. Paget, personal communication). Thus in order to fully understand the biological role of Rex, this chapter focuses on defining the Rex regulon using a ChIP-chip approach to directly identify binding sites. Several new Rex binding sites were revealed, which provide new insights into the biological role of Rex. Furthermore, the results provide important information on the structure of Rex binding sites and reveal that Rex can bind to half-sites, albeit with weaker affinity. The potential problems of studying changes in gene expression resulting from de-repression alone are discussed.

Section 3.2 – Genome-wide identification of ROP sites

In recent years, ChIP-on-chip (chromatin immunoprecipitation-on-chip) has emerged as a powerful technique to globally identify targets of DNA binding proteins (Negre *et al.*, 2006, Pillai and Chellappan, 2009, Sala *et al.*, 2009). The premise of this technique is that DNA binding proteins are chemically cross-linked to the chromosomal DNA *in vivo* and the protein of interest is selectively immunoprecipitated using a specific antibody. The co-immunoprecipitated DNA can then be labelled and used to probe a genome-scale microarray. In the absence of poly- or mono-clonal antibodies against the protein of interest, the gene that encodes the protein can be modified by the addition of an epitope tag. Thus, Rex was engineered with a 3xFLAG tag, allowing immunoprecipitations to be performed using an anti-FLAG antibody.

Section 3.2.1 – Rex^{FLAG} construction

Generation of the Rex^{FLAG} construct was done in several stages (Figure 3.1).

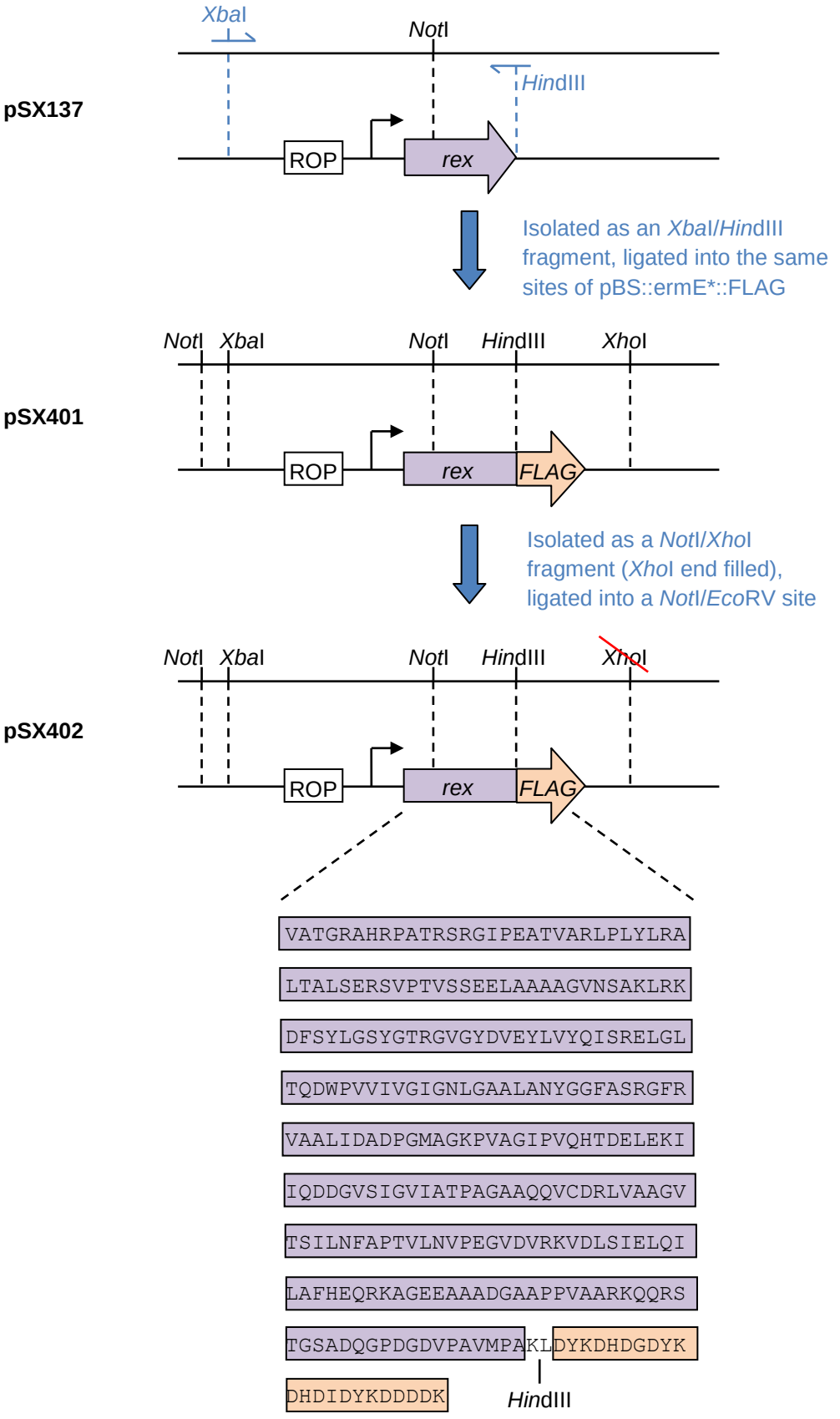


Figure 3.1: Diagrammatic view of the Rex^{FLAG} construction for use in ChIP-on-chip experiments. The Rex sequence is shown here in purple, while the in-frame FLAG tag is depicted in orange.

The first stage was to amplify the *rex* gene and its corresponding promoter from the pBlueScript-derived vector pSX137. The forward (FLAG-ROP_For) and reverse primers (FLAG_Rev) included restrictions sites, *Xba*I and *Hin*dIII, respectively; the reverse primer also had the function of removing the stop codon to allow fusion to the FLAG-tag. The resulting product contained the entire coding region of *rex*, with the exception of the stop codon only, and also included 326bp of the promoter region, which contained the ROP site (located at -88bp). This ensured that all of the upstream regulatory elements were included in the final construct. The PCR product was blunt-end ligated into *Eco*RV-cut pBlueScript II SK+ and the entire region re-isolated as an *Xba*I-*Hin*dIII fragment. This fragment was introduced into a pBlueScript derivative that contained a 3xFLAG-tag sequence (pBS::ermE*::FLAG). Translational fusion was mediated by an in-frame *Hin*dIII site preceding the tag. This construct was then digested with *Xho*I, end-filled with Klenow, and partially digested with *Not*I to release the entire rex^{FLAG} fragment. This fragment was cloned into *Not*I-*Eco*RV-digested pSET152 and the resultant construct (pSX402) introduced into *S. coelicolor* S106 (Δ*rex*) strain. The pSET152 vector is an integrative plasmid that recombines into the ϕC31 chromosomal attachment site, usually at single copy (Bierman *et al.*, 1992). Initially, experiments were performed to ensure that the Rex^{FLAG} protein was functional *in vivo.* RNA was harvested over an oxygen limitation time-course, including an aerated sample (time zero), and was used for an S1 nuclease mapping study (Figure 3.2). M145 (pSET152) and S106 (pSET152) were used as positive and negative controls, respectively. The *cydA* promoter region was used as the S1 probe as this region contains both a Rex-regulated promoter (cyd^{P1}) and a constitutive promoter (cyd^{P2}), providing an internal reference control. In the case of S106 (pSET152) the cyd^{P1} promoter remained highly active throughout the time-course, whereas for the M145 (pSET152) and S106 (pSX402), expression from cyd^{P1} was induced by oxygen-limitation. This confirmed that the Rex-FLAG is functional *in vivo* and could therefore be used for subsequent ChIP-chip experiments.

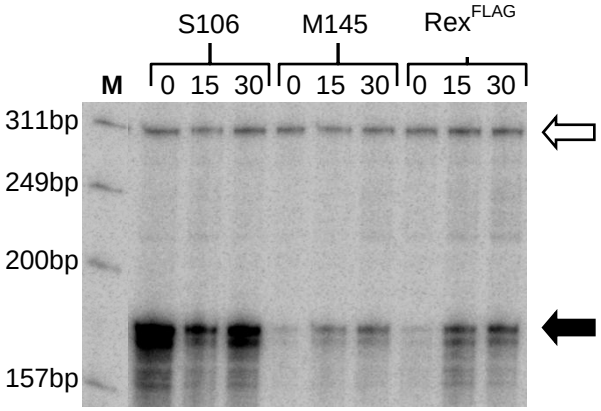


Figure 3.2: S1 nuclease protection assay on the *cydA* promoter region. RNA was harvested from S106+pSET152, M145+pSET152 and S106+pSX402 (Rex^{FLAG}) *S. coelicolor* strains. The times indicated above each well represent the period of oxygen-limitation prior to harvesting, with 0 minutes representing an aerated sampling. The closed arrow shows the position of the cyd^{P1} transcript, which is Rex-dependent, whereas the open arrow indicates that of the constitutive cyd^{P2} promoter.

Section 3.2.2 – Optimising the ChIP method for *S. coelicolor*

The chromatin immunoprecipitation method was performed as described by Grainger *et al.* (Efromovich et al., 2008), however the cell disruption and sonication steps required optimisation for *S. coelicolor*. Two methods of cell disruption were attempted; cryogenic grinding and lysozyme treatment (see Section 2.3.9 for details). The purified fragments from each method were analysed by gel electrophoresis (data not shown), looking for clean fragment ranges and a strong signal. The amounts of nucleic acid attained from the grinding method, as assessed on a Nanodrop™, appeared to be significantly higher than those of the lysis method (data not shown), however visual inspection of the samples via agarose gel electrophoresis revealed that this was not the case. The lysozyme-treated samples were much cleaner and stronger than the ground samples. This would suggest that the concentration readings taken on the ground samples were enhanced by contamination – most likely due to more RNA surviving this method. Given that the main difference between the two methods was the presence or absence of an incubation step (lysozyme treatment) one explanation is that lysozyme treatment allows time for other enzymes to degrade contaminating RNA. The lysis method was used from this point forward.

The next step for optimisation was sonication – too few cycles and the resolution on the array is too low, too many and there is a risk of degrading the samples. Sonication was therefore performed on the lysozyme-treated sample, taking aliquots after each round and analysing the fragment size ranges on an agarose gel. After a single cycle at 35% for 15 seconds (Sonicator model Vibracell, Sonics & Materials Inc.) the fragments ranged from >10kb (maximum marker size) to 1kb. With successive cycles this gradually dropped to a range of 1.5-0.5kb but did not appear to drop any further after this point. The sonication used for the arrays was therefore set at 10 cycles, the point at which the smallest fragment ranges were obtained. Nevertheless, experimental samples were checked after sonication to confirm that it had been effective.

Section 3.2.3 – Chromatin immunoprecipitation

When designing the ChIP-chip experiment a number of control strains were considered, including “no antibody” controls and “no antigen” controls. “No antibody” controls act as indicators of chromatin contamination, i.e. how much DNA is non-specifically purified along with actual IP samples. “No antigen” controls also provide this indicator but in addition reveal possible cross-reactivity of the antibody. For this reason the “no antigen” control method was chosen and, as our antigen was the 3XFLAG tag, the control would be a strain lacking Rex^{FLAG}. The control chosen was S106 (pSET152). *S. coelicolor* S106 (pSX402) and S106 (pSET152) were grown to late exponential phase (OD_{450nm} ~0.7-0.8), formaldehyde cross-linked and sonicated as described in Section 2.3.9. The samples were then split into either immunoprecipitated (IP) or total DNA control, with the former being purified via the anti-FLAG antibody. After all the IP washes were performed, both samples were de-crosslinked by heat treatment and the DNA purified.

Before launching into a full scale microarray experiment, the efficiency of the cross-linking and specificity of immunoprecipitation was tested using quantitative real-time PCR (qPCR). Two known binding regions (*cydA* and *ndh)* and a negative control (*hrdB)* were tested for enrichment in the S106 (pSX402) test samples. The total DNA control samples were used to generate standard

curves from which to determine copy numbers for the test samples. The copy numbers for the no-antigen controls were then subtracted from the IP samples, and that value divided by the total amount of DNA to give copy numbers per pg of DNA (Figure 3.3). The results showed that whilst the *hrdB* region was not significantly enriched in the IP samples (~8 copies per pg), the *ndh* region was present at ~1500 copies and the *cyd* region was present at ~9000 copies per pg of sample. This confirmed that the cross-linking had worked and that the regions with ROP sites were enriched compared to non-target regions. The average value obtained for *hrdB* was 8 copies per pg of DNA, this value fell outside the detection limits of the experiment, which only went as low as 30-copies of *hrdB* per pg. The signal for *hrdB* is therefore not considered significant.

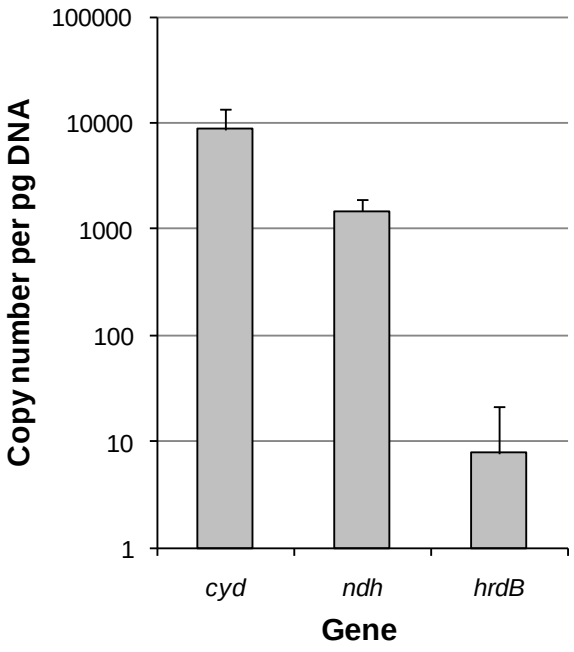


Figure 3.3: Enrichment values determined from qPCR experiment on the immunoprecipitated DNA using primers for the *cydA* (Cyd_qPCR_F/R), *ndh* (Ndh_qPCR_F/R) and *hrdB* (hrdB1a/b (Pascoe, 2009).

Section 3.2.4 – ChIP-on-chip

The OGT microarray slides used for ChIP-on-chip consisted of 4 arrays per slide, hybridising each within a separate chamber. The density of each array was 44K with 60mer probes covering the complete *S. coelicolor* genome (Bucca *et al.*, 2009). Biological replicates were performed for each strain and each resulting DNA sample was labelled with Cy3-dCTP or Cy5-dCTP according to the scheme in Table S1. This arrangement of labelling constituted a dye-match, and allowed normalisation for any differences observed due to the differing labelling efficiencies and intensities of the two Cy-dyes. For each oligonucleotide probe on the array the signal from the test strain, S106 (pSX402), was divided by that of the control, S106 (pSET152), with the ratios then indicating the enrichment values at each position. The values across all four chambers were then compared to generate average enrichment values and to determine the p-values for each hit (Section 2.3.9). A two-way cut-off was applied on the data requiring all hits to have a >3-fold enrichment compared to the control and to have a p-value <0.05, hopefully eliminating false-positives from the dataset. The targets were also manually inspected to remove those

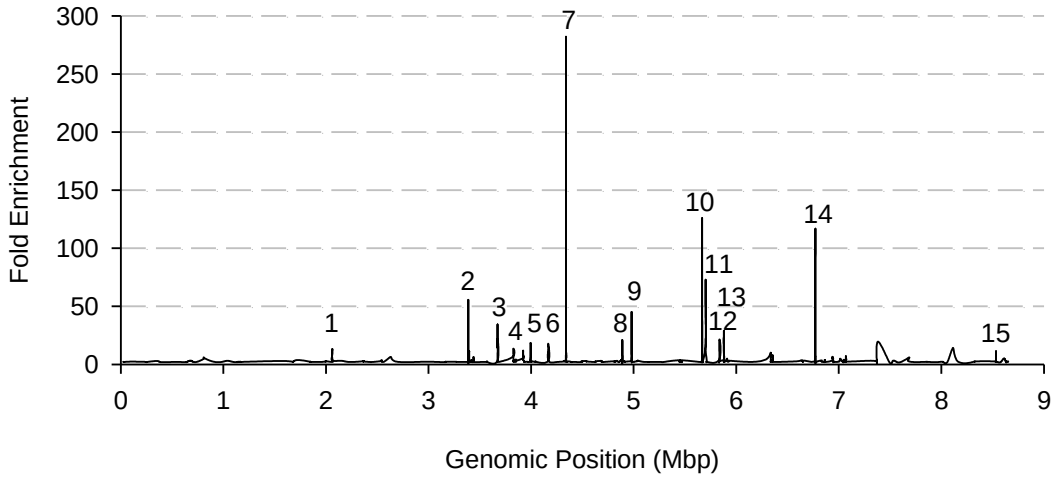


Figure 3.4: Genomic localisation of Rex binding sites. The peaks indicated are as follows; 1, SCO1930; 2, SCO3092 (*ndh*); 3, SCO3320 (*rex*); 4, SCO3547; 5, SCO3615/6; 6, SCO3790/1; 7, SCO3945 (*cydA*); 8, SCO4472 (*resA*); 9, SCO4562 (*nuoA*); 10, SCO5207; 11, SCO5240 (*wblE*); 12, SCO5366 (*atpI*); 13, SCO5408/9; 14, SCO6168; and 15, SCO7697 (*phyC*).

SCO number	Gene/comments	Fold-enrichment	Gel shift?
SCO3945	*cydA*; cytochrome bd terminal oxidase	281.9	Yes
SCO5207	Conserved hypothetical protein	125.9	Yes
SCO6168	Hypothetical protein	113.9	Yes
SCO5240	*wblE*; WhiB related protein	72.35	Yes
SCO3092	*ndh*; NADH dehydrogenase type II	55.21	Yes
SCO4562	*nuo*; NADH ubiquinone oxidoreductase type I	45.27	Yes
SCO3320	*rex*; redox sensing transcriptional repressor	34.2	Yes
SCO5366	*atpI*; ATP synthase operon	21.55	Yes
SCO4472	*resA*; cytochrome biogenesis operon	20.84	Yes
SCO3615#	*ask*; aspartokinase	18.5	No
SCO3616#	Hypothetical protein	18.5	No
SCO3790#	Conserved hypothetical protein	16.6	No
SCO3791#	Conserved hypothetical protein	16.6	No
SCO1930	ABC transport protein	13.1	Yes
SCO3547	*hppA*; H^+-translocating pyrophosphatase	11.9	Yes
SCO3101	Lipoprotein	11.3	Yes
SCO7697	*phyC*; phytase	11.3	No
SCO5408#	Conserved hypothetical protein	11.1	Yes
SCO5409#	*"ydzA"*; conserved membrane protein	11.1	Yes
SCO5797	Serine protease lipoprotein	10.0	No
SCO5810#	Transmembrane efflux protein	7.7	No
SCO5811#	Transcriptional regulator	7.7	No
SCO6280	*cpkO*; SARP (cryptic type I polyketide)	6.7	No
SCO3137	*galE1*; UDP-glucose epimerase	6.4	No
SCO6917	Hypothetical protein	5.0	No
SCO5435#	*dcuS*; sensor kinase	4.9	No
SCO5436#	*dctA*; sodium:dicarboxylate symporter	4.9	No
SCO6383	Membrane protein	4.1	No
SCO4461#	Transcriptional regulator (TetR-family)	3.9	No
SCO4462#	Membrane protein	3.9	No
SCO6239	Sigma factor	3.8	No
SCO5032#	*ahpC*; alkyl hydroperoxidase	3.6	No
SCO5033#	*oxyR*; peroxide stress regulator	3.6	No
SCO5013	Secreted protein	3.5	No
SCO2370#	Hypothetical protein	3.4	No
SCO2371#	*aceE2*; pyruvate dehydrogenase	3.4	No
SCO6218#	Phosphatase	3.4	No
SCO6219#	Serine threonine protein kinase	3.4	No

Table 3.1: List of potential target genes identified by ChIP-on-chip, including their annotated functions and fold enrichment compared to the no-antigen control. Positive results for gel retardation assays are also indicated. The # indicates divergent genes, which prevented the determination of which gene the peak was for.

where the enrichment appeared to come from one probe alone and to remove any where the peak did not correspond with the start of a gene, including the removal of genes only represented because they neighboured genes with strong peaks. The final data set consisted of 29 peaks, encompassing 38 genes, and is shown as a genomic region view in Figure 3.4 and with corresponding enrichment values in Table 3.1. Note that due to the vast differences in enrichment values, not all peaks are clearly visible in Figure 3.4, therefore all of the sites that met the selection criteria described above are also shown as region views in Figure S1 (appendix).

Section 3.3 – Binding to target sites

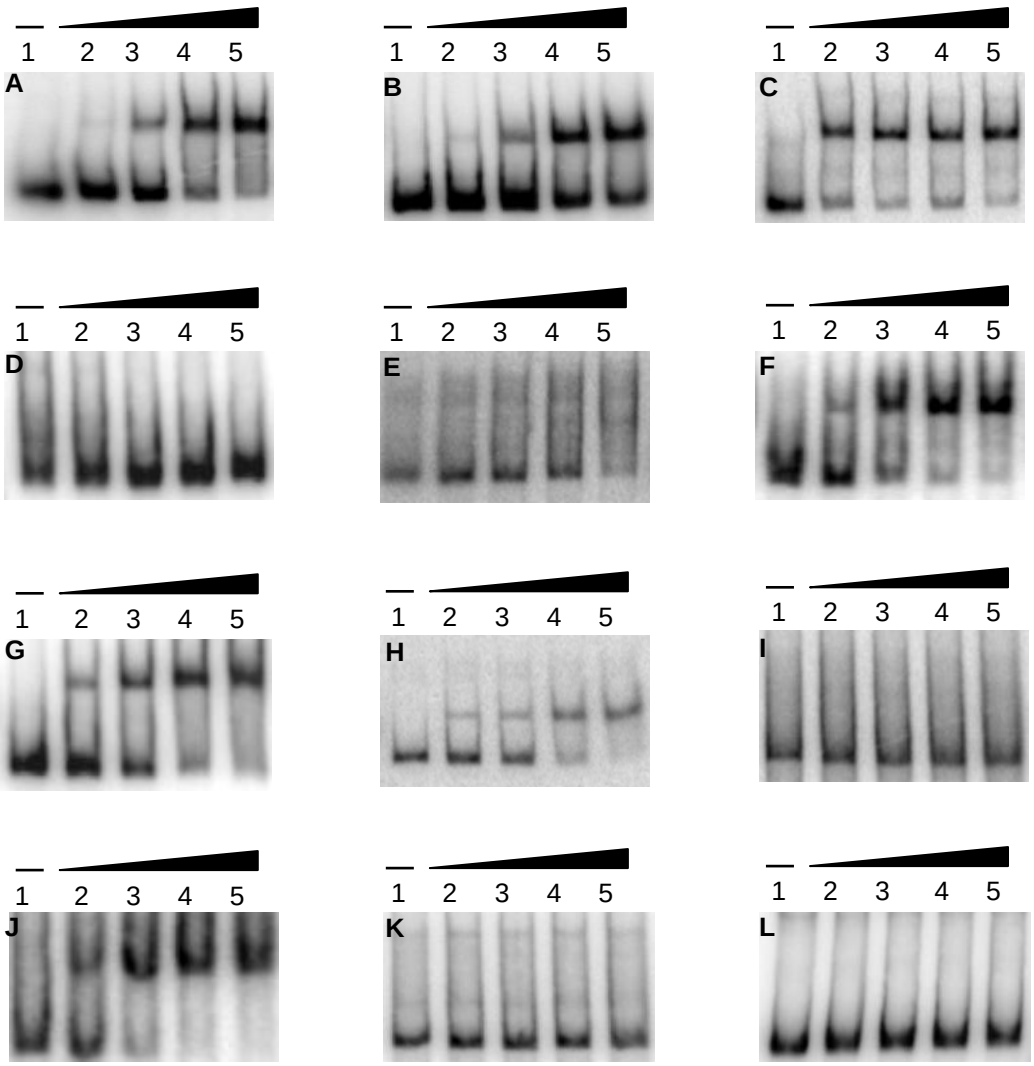
The ChIP-on-chip experiment identified several potential *in vivo* Rex binding regions. Despite, the use of controls, the identification of false positives remained a possibility and so it was necessary to confirm Rex binding sites by *in vitro* DNA binding assays. In addition, bioinformatic analyses of sequences around the peaks should narrow down the sequences responsible for binding and provide insights into the nature of ROP sites. For the purpose of this study any genes that had already been identified as Rex targets (Brekasis, 2005) have been ignored but are mentioned briefly in the following subsection.

Section 3.3.1 – Previously identified targets

ChIP-on-chip signals were detected in the promoter regions of several previously characterised members of the Rex regulon: *cydA*; *wblE*; *ndh*; *nuoA*; *rex*; *atpI* and *resA* (Table S2) (Brekasis, 2005, Brekasis and Paget, 2003). All of these genes featured prominently within the list, especially *cydA* which headed the list with a fold-enrichment value of 281.9. Interestingly one gene previously shown to have a ROP site by EMSA, SCO4281, was absent in the ChIP list. This gene failed to meet the >3-fold cut-off, only having an enrichment value of 2.3 shared across the intergenic region between SCO4281 and SCO4280.

Section 3.3.2 – EMSAs on new targets

After removing the known sites from the ChIP-on-chip list there were still 22 promoter regions that had not previously been shown to have ROP sites (Table S2). These regions were therefore subjected to an RSAT (Regulatory sequence analysis tool; (Thomas-Chollier *et al.*, 2008) search and visually inspected to identify potential ROP sites. This method identified a number of potential sites; however the majority of the genes appeared to lack full ROP sites. To ensure that no sites had been overlooked during the search, the whole intergenic region of each gene, encompassing the enriched DNA, was used for EMSA analysis. In each case the region was amplified by PCR, end-labelled with γ^{32}P-ATP, mixed with Rex and run on a 6% polyacrylamide gel (Figure 3.5).



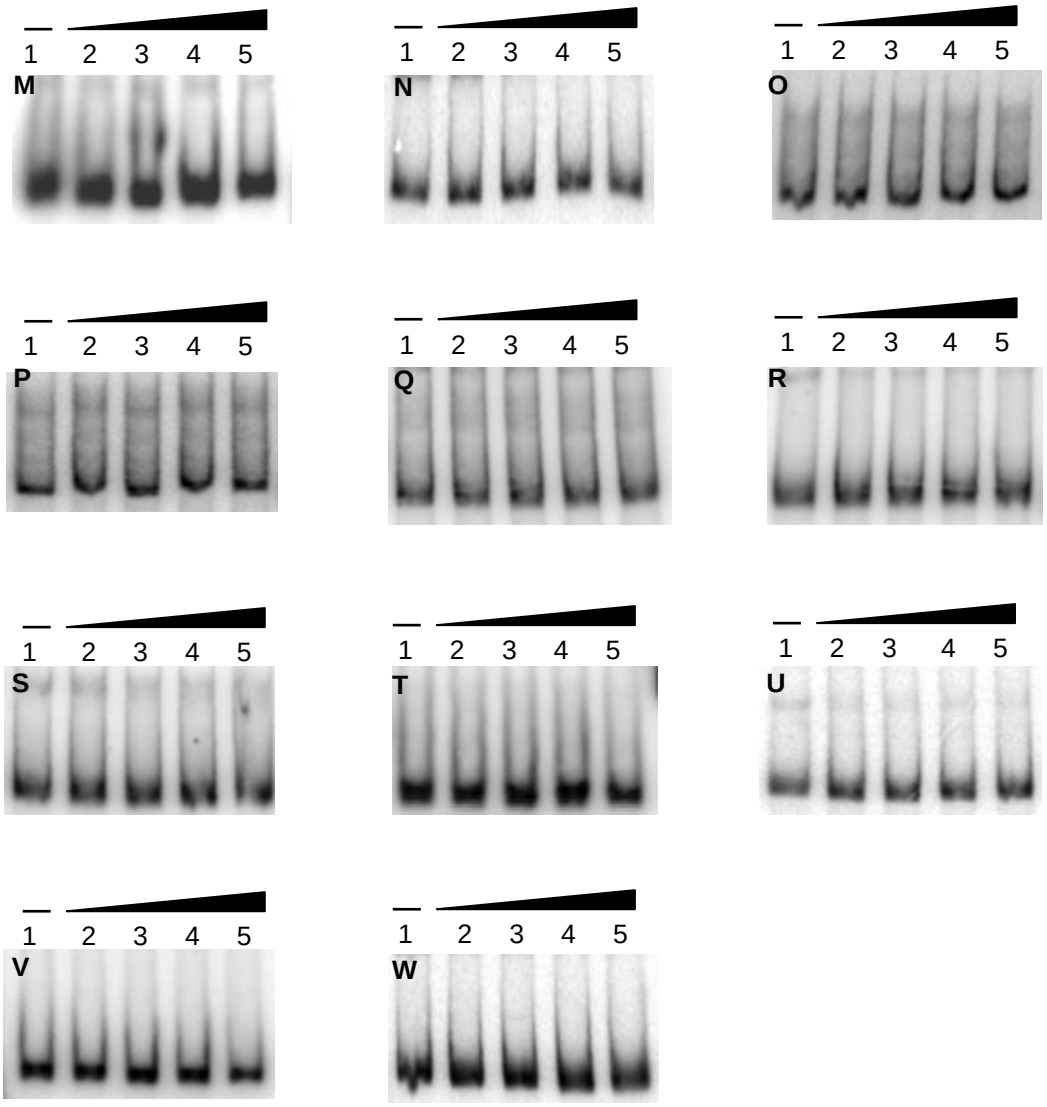


Figure 3.5: EMSA analysis on the newly identified ChIP-on–chip targets and on the positive control region *rex*^{P}. In each case lane 1 contains the probe alone, labelled with γ^{32}P-ATP. Lanes 2 to 5 contain both probe and Rex at increasing concentrations, 25nM (lane 2); 100nM (lane 3); 250nM (lane 4) and 500nM (lane 5). The gels are labelled in order of enrichment in the ChIP-on-chip study and are as follows: **(A)** SCO5207; **(B)** SCO6168; **(C)** SCO3320 (*rex*); **(D)** SCO3615/6; **(E)** SCO3790/1; **(F)** SCO1930; **(G)** SCO3547; **(H)** SCO3101; **(I)** SCO7697; **(J)** SCO5408/9; **(K)** SCO5797; **(L)** SCO5810/1; **(M)** SCO6280; **(N)** SCO3137; **(O)** SCO6917; **(P)** SCO5435/6; **(Q)** SCO6383; **(R)** SCO4461/2; **(S)** SCO6239; **(T)** SCO5032/3 (*oxyR/ahpC*); **(U)** SCO5013; **(V)** SCO2370/1 and **(W)** SCO6218/9. All primers used to generate the EMSA probes are listed in Table 2.3, with the exception of the SCO3320 (*rex*) probe which was generating using primers rexGSrev2 and E68.18ci (Brekasis, 2005).

Band-shifts were obtained for six of the fragments tested, SCO1930; SCO3101; SCO3547; SCO5207; SCO5408/9 and SCO6168. The remaining sites failed to produce band shifts, possibly due to a low affinity for Rex. The EMSAs are shown in order of enrichment in Figure 3.5. Although there is a general trend of highly enriched sites giving positive band shift results, there are also several

examples where high enrichment is not reflected in binding *in vitro* (e.g. SCO3615/6 (18.5-fold) and SCO3790/1(16.6-fold). This suggests that other factors might influence Rex binding *in vivo* (see Discussion).

Section 3.3.3 – The classical ROP site

The discovery of new Rex binding sites allowed the refinement of the consensus sequence of ROP (Table 3.2 and Figure 3.6). The sequences of all identified sites (excluding SCO5207 – explained in section 3.3.4) and their neighbouring nucleotides were compiled into an aligned list and run through a sequence logo generator (WebLogo (Crooks *et al.*, 2004). The logo confirmed that ROP sites have a strong preference for a GTG-n8-CAC sequence and that these nucleotide blocks were most frequently surrounded by AT-rich DNA. The logo also revealed that the consensus ROP site is actually an 18 bp inverted repeat, not 16bp as had previously been reported. There was also a slight preference for AT-rich DNA stretching out across 22bp.

Gene	Sequence	Position
SCO3945 (*cydA*)	ATGTGAACGCGTTCACAA	-101
SCO6168	TTGTGAAAACTTTCACCC	-91
SCO5240 (wblE)	TCGTGAAAGCGTTCACAT	-52
SCO3092 (*ndh*)	TCGTGAAGTTCTTCACAA	-315
	TTGTGAAGGGGCGCACGA	-43
SCO4562 (*nuoA*)	TTGTGACCTGCTTCACAT	-116
SCO3320 (*rex*)	TTGTGCACGCGTTCACAA	-88
SCO5366 (*atpI*)	TTGTGATACGGTTCACGA	-139
SCO4472 (*resA*)	ATGCGAAACTTTTCACAT	-8
SCO1930	TCGTGAAAGCGTGCACAA	+25*
SCO3547	AAGTGAATTCATTCACGA	-215
SCO3101	TTGTGCACCGTCGCACAA	-28
SCO5408/9	TTGTGAACGGAAGCACAA	-32
	TTGTGCTTCCGTTCACAA	-66
SCO4281	TTGTGACTTGAGTCACAA	-148
Consensus	**ttGTGaannnnttCACaa**	

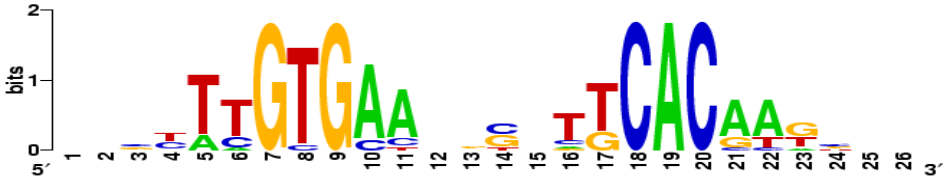


Table 3.2 and Figure 3.6: Sequence logo derived from the genomic regions capable of generating gel shifts in gel retardation studies. The predicted ROP sites in each region and their mid-point positions relative to the translational start sites are indicated in above. *Note that the SCO1930 ROP site sits within the translated region of this gene.

Section 3.3.4 – The Rex operator is 18 bp in length

To further investigate the length of the ROP site a Surface Plasmon Resonance (SPR)-based competition assay was devised for the BIAcore 2000 system (see Figure 3.7 for overview). In brief, the BIAcore system uses a polarised light beam to excite the surface Plasmon of an electrified sensor chip (thin gold plate). The intensity and angle of the reflected light are detected by the instrument. The Plasmon resonance of the chip is altered by changing the composition of the solution in the flow cell, which resides on the opposite face of the sensor surface. Throughout this study the sensor chips used were streptavidin-coated, allowing the attachment of a biotinylated ligand – in the form of 5' biotinylated dsDNA. The interaction of Rex with the attached DNA fragment was then observed as an increase in the response units. This change in response could be altered by the inclusion of NAD^+ or NADH in the injected solution, or by the inclusion of a DNA fragment, in the injected Rex solution. Test fragments that bind to Rex compete with the sensor chip-bound ROP DNA, thereby reducing the signal; fragments that do not bind do not affect the maximal signal (Figure 3.7).

In these experiments the chip-bound ROP site was an 18 bp site (annealed primers NUOROP1 and NUOROP2 (Brekasis, 2005). Experiments were performed using a concentration gradient of competing DNA and compared to both a random DNA control and a protein-only injection per assay. The ROP sites to be tested (16bp vs 18bp ROP^{nuo}) were based on the natural site upstream from the *nuoA-N* operon (ROP^{nuo}). Each site was placed in the same random context, the full sequence of which acted as the random control site (primers Nca and Ncb). The only difference between the two was the presence or absence of an additional thymine at the 5' and 3' ends (based on the native ROP^{nuo}). The results of the competition assays for the 16 bp (primers Nuo-4F and Nuo-4R) and 18 bp sites (primers NuoF and NuoR) are shown in Figure 3.8. The 16 bp site was capable of competing with the biotinylated site on the sensor surface. However this competition only became significantly different, from the random site, at ROP:Rex molar ratios of >5:1 – giving a $\%R_{max}$ of ~30.

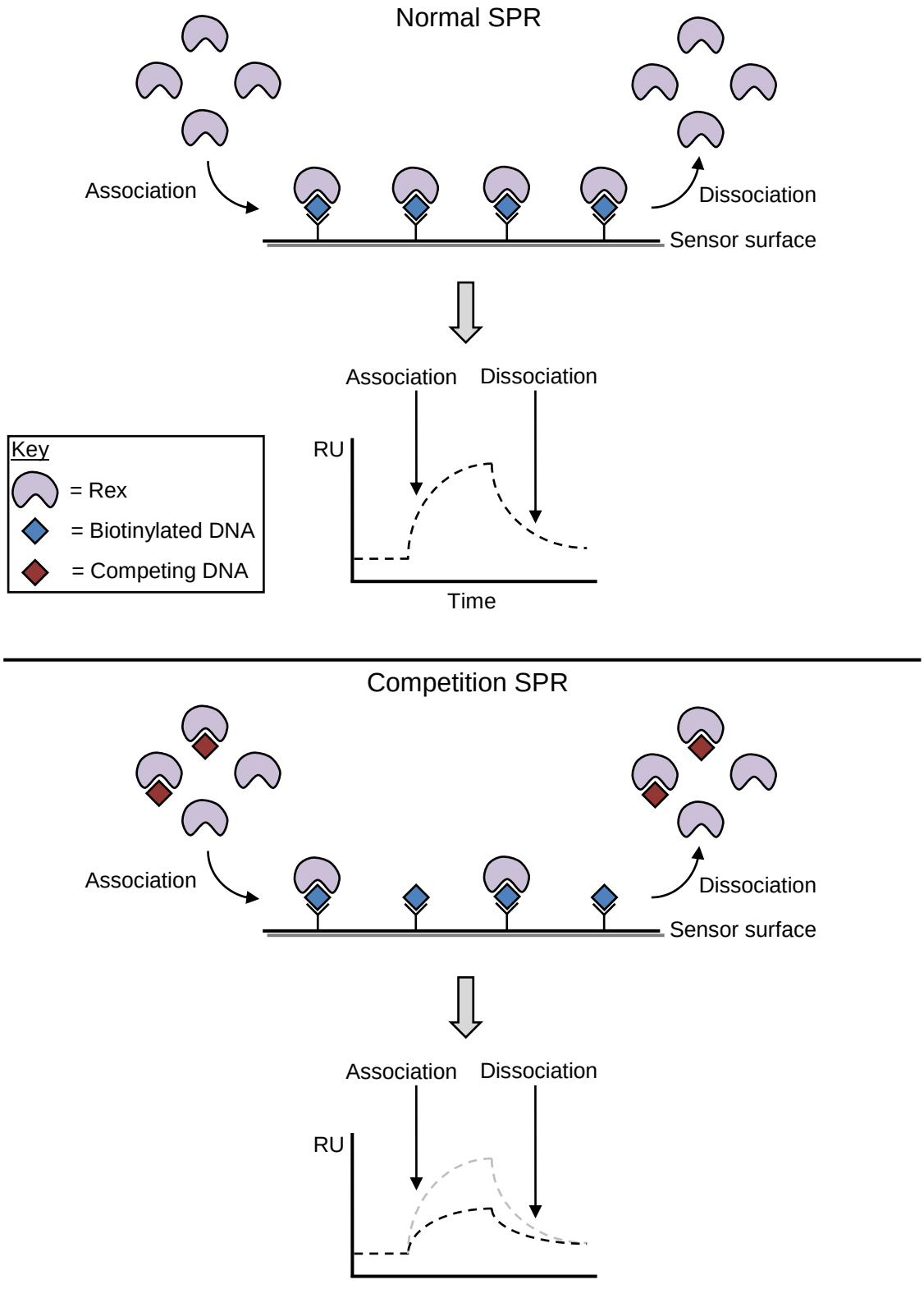


Figure 3.7: Diagrammatic view of the difference between a typical SPR experiment and one were competing DNA was co-injected over the sensor surface. The corresponding drop in RU was proportional to the strength of each competing fragment and the amount of DNA used.

The 18 bp site, however; was able to give the same response at only a 0.5:1 molar ratio. This meant that in order to reduce the concentration of free Rex in the injected sample by 70% you needed 10 times as much of the 16 bp site as the 18 bp, a marked difference for only a one nucleotide extension per half-site. These data confirm that the 5' thymine (position 1/18) is important for Rex binding; the structural basis for this is considered in Section 5.3.1.

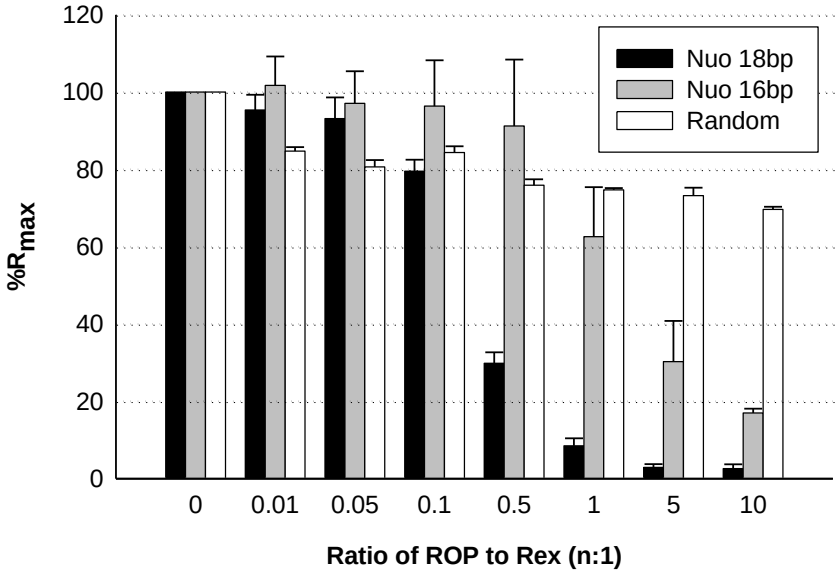


Figure 3.8: SPR competition assay with a 16bp (grey bars) and 18bp (black bars) ROP^{nuo} site placed in a random DNA context, the full sequence of which was used for the random control (white bars). The Rex concentration was fixed at 90nM (dimeric concentration) and the DNA concentration varied between 0nM (ratio of 0:1), 0.9nM (0.01:1), 4.5nM (0.05:1), 9nM (0.1:1), 45nM (0.5:1), 90nM (1:1), 450nM (5:1) and 900nM (10:1). The %Rmax values were calculated using the protein only injections for each run. All injections were done in triplicate.

Section 3.3.5 – Rex appears to interact with half sites

As mentioned previously, most sites that produced gel shifts all shared a common structure; [A/T][A/T]GTG-n8-CAC[A/T][A/T]. However, the SCO5207 upstream region, despite showing high enrichment during ChIP-chip and good gel-shifts (Figure 3.5), was lacking a key element of the consensus ROP site; the sequence (TTGTGAATCCATG**A**ACTA) appeared to be lacking the highly conserved cytosine (shown in bold) . Upon closer inspection this site also appeared to have an additional half-site (TTGTGAA) directly preceding the

ROP^{SCO5207} sequence. This occurrence has previously been observed in the ROP^{cyd} and ROP^{wblE} sites (Brekasis, 2005). However, in these regions there is a one base pair overlap of the last nucleotide of the half-site and first base of the full site. This gives two overlapping but distinct ROP sites, on different faces of the DNA helix, which may explain why the *cydA* and *wblE* promoters are capable of generating a double shift during EMSA analysis (Brekasis, 2005).

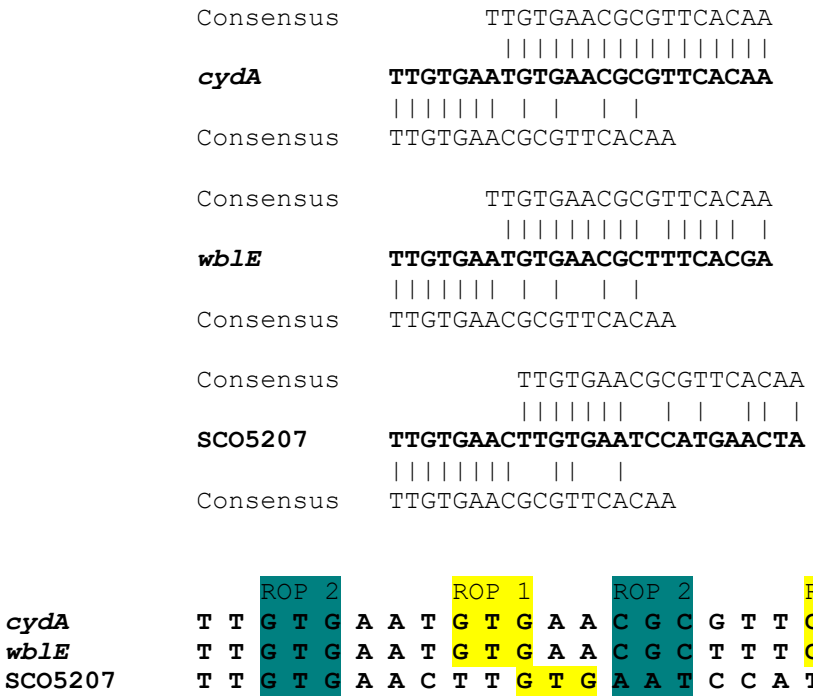


Figure 3.9: Spacing of the *cydA*, *wblE* and SCO5207 ROP sites. The two potential binding sites are compared to an ideal ROP site in each case. The main interaction points of Rex and the DNA (GTG and CAC) are shown in blue (ROP 2) or yellow (ROP1). Note that the *wblE* site is inverse complemented for ease of comparison. The spacing of the SCO5207 site places the two potential binding sites directly adjacent, lacking the 2 bp gap of *cydA* and *wblE*.

The SCO5207 promoter only generates one shift. The SCO5207 data suggested that it might be possible for Rex to bind to ROP sites that lack one or more key elements in one half of the operator. Therefore, having failed to find full ROP sites in the remaining target regions, using the consensus sequence, a search for half sites ([AT][AT]GTG[AT][AT]) was performed instead. This search revealed potential sites in each of the remaining regions that had showed up in the ChIP-on-chip list but had failed to generate gel shifts during EMSA analysis. An alignment and sequence logo was generated for these sites, which revealed

that only the TTGTGAA sequence was well conserved (Table 3.3 and Figure 3.10). This raised the question of whether Rex could interact with the half-site alone.

Gene	Sequence	Position
Half sites:		
SCO7697*	CTTGTGAACGCGTGCACGAG	+13
SCO5797	CAAGTGAACGCGTGAATGTC	-27
SCO5810/1	TTAGTGAACGCTCGCGTTCA	-26
SCO6280	AATGTGAACACACACAGCAC	-92
SCO3137*	ATAGTGAAAAGTTCCAGGTG	-122
SCO6917	TTTGTGAATTGAAACCGCCG	-34
SCO5435/6	ATTGTGAACTAAATGAACGC	-27
SCO6383	TGAGTGAACGTAATCTCGCC	-71
SCO4461/2	TTTGTGCATACTCGGTGGGT	-51
SCO6239	ATAGTGAATGGAGGAGGAAA	-73
SCO5032/3	TTAGTGAAATAGCTACACTC	-95
SCO5013*	ACTGTGATGTAGATGGGGAA	+49
SCO2370/1	ATCGTGTACCTCGCGGGGGC	-185
SCO6218/9	ATTGTGAAGATTGCATGAGA	-172
Consensus	**attGTGaA------c-g---**	
Tandem half-sites:		
SCO5207	CTTGTGAACTTGTGAATCCATGAACTAG	-23
SCO3615/6	CTTGTGAACGTGTGACACACCGCACTTT	-186
SCO3790/1	CTTGTGATCTTGTGCAGGGTCTGGCATG	-182
Consensus	**CTTGTGAaCtTGTGaa-c--cg-aCttg**	

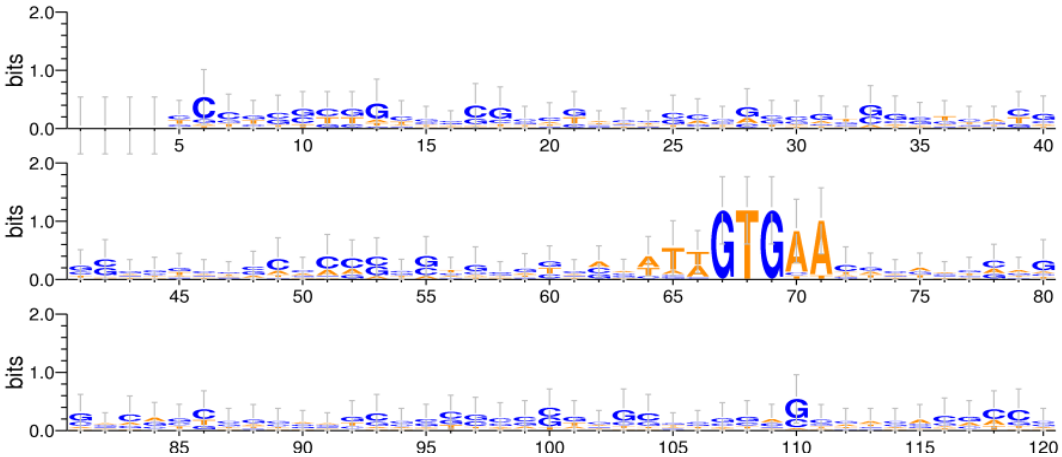


Table 3.3 and Figure 3.10: List of half-site only regions from the ChIP-on-chip targets and corresponding sequence logo for these sites. Only the SCO5207 sequence, in this list, was capable of generating a gel shift by EMSA. Note that where a divergent gene region is listed the position numbering is based on the gene with the lowest SCO number (i.e. the leftmost gene on the *coelicolor* chromosome). *Indicates that the site found was situated within the coding region of a gene.

To test this; an artificial site (primers Full_26_For and Full_26_Rev) was generated in a random context based on our observations of what defines a ROP site. This same sequence was then altered to generate an artificial half-site (primers Half_26_For and Half_26_Rev) by transistionally mutating each of the nucleotides in the second half of the ROP site. These two fragments were then analysed along with the full random sequence (primers Random26_F and Random26_R), in which they sat, using the SPR competition assay (Figure 3.11). The biotinylated fragments used in this case were 16 bp ROP^{nuo} (primers 1a and 1b) and random control (primers Ran_Bio_F and Random_R), in lanes 3 and 2 of the sensor surface, respectively. In this experiment the competing random control fragment had only a slight affect on binding to the sensor surface, decreasing by <5% with a 100-fold molar excess of DNA:Rex. The full artificial ROP site was capable of causing a 50% reduction in binding at a 1:1 ratio of ROP:Rex. The artificial half-site showed a higher level of competition for Rex compared to the random control, although a 100-fold excess DNA was required to achieve 50% inhibition of chip DNA binding.

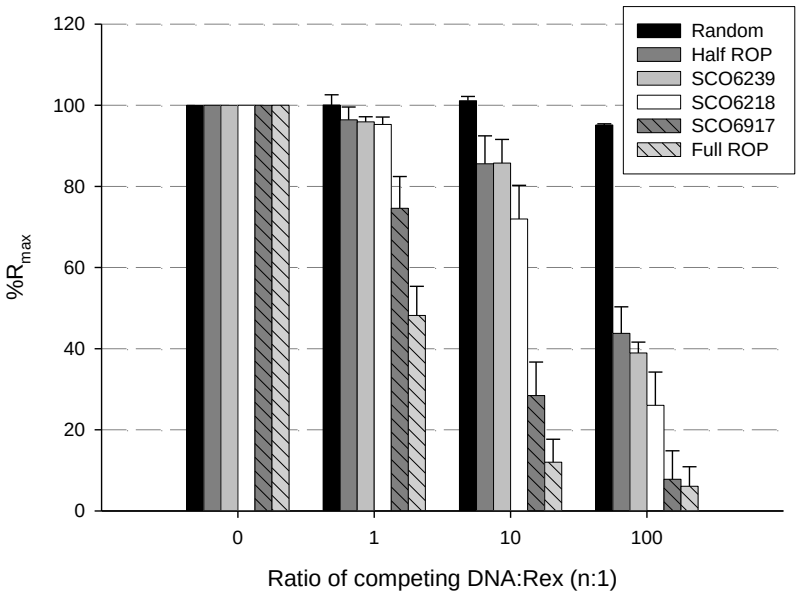


Figure 3.11: SPR competition assay for various ROP sites; native and artificial. The Rex concentration was kept fixed at 90nM (dimeric concentration) and the DNA concentration varied between 0nM (ratio of 0:1), 90nM (1:1), 900nM (10:1) and 9µM (100:1). The %Rmax values were calculated using the protein only injections for each run. All injections were done in triplicate.

A selection of natural putative half-sites were also analysed using this method; SCO6239 (6239_half_For and 6239_half_Rev), SCO6218/9 (6218_half_For and 6218_half_Rev) and SCO6917 (6917_half_For and 6917_half_Rev), all in their native contexts (Figure 3.11). The results showed that all three of the native sites gave a stronger response than the artificial half-site, however only the SCO6917 site gave a response similar to the full ROP site. Analysis of the SCO6917 site revealed that one of the highly conserved 'CAC' residues was present in this fragment but lacking in the others. It is most likely this residue that was responsible for the strength of the response for this fragment. Overall the data indicates that Rex does bind to the natural half-sites, albeit at an affinity that is too low to be detected by EMSA.

Section 3.4 – The Regulon

Previous transcriptome studies on Rex compared RNA harvested from the M145 and S106 (Δ*rex*) strains grown under aerobic conditions (D. Brekasis and M. Paget, personal communication). This identified only a few genes that were >2-fold up-regulated in the Rex null strain, with only two having corresponding ROP sites (*cydA* and *ndh*). Using the ChIP-on-chip method more potential binding sites were identified. However, since most of these genes were not upregulated in the Δ*rex* background, they have not been confirmed as members of the Rex regulon. This section covers the attempts to identify such regulation and also speculates on the possible roles that the genes have in maintaining the redox poise.

Section 3.4.1 – Investigating repression with a *rex*^{G102A} strain

The Rex^{G102A} protein was described previously due to its inability to sense the redox poise (Brekasis and Paget, 2003). This mutation falls on the central glycine of the Rossmann fold (GXGXXG) and physically blocks NADH binding. The result *in vitro* is a protein that can bind to DNA but will no longer dissociate on cue, which could potentially result in a super-repressor *in vivo*.

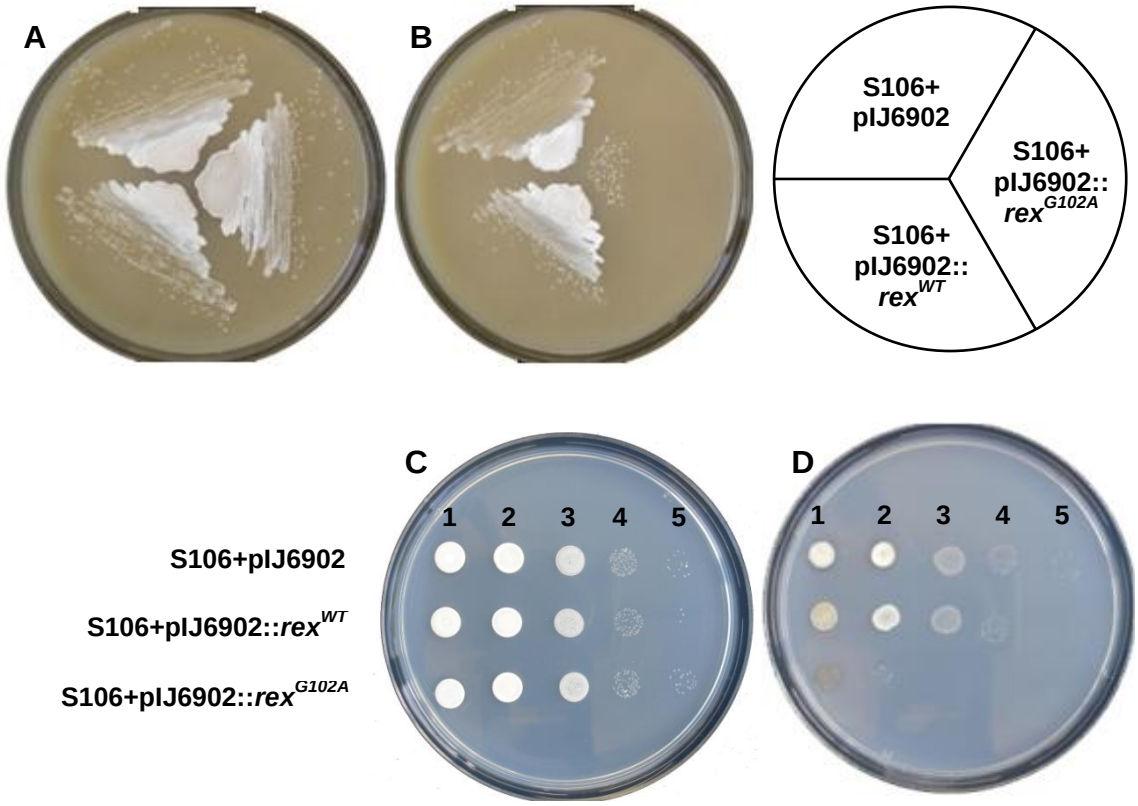


Figure 3.12: Phenotype analysis of the thiostrepton inducible *rex*^{G102A} super-repressor strain. Each strain was either streaked over MS agar (panels A and B) or spotted onto MM agar containing 1% glucose and trace elements (panels C and D), in the presence of 20µg/ml apramycin and 12.5µg/ml thiostrepton (panels B and D), or apramycin alone (panels A and C). For the MM plates 10µl of spore suspension was spotted at each position, from a serial dilution of the spore stock. At each position the following dilution had been made; undiluted stock (1), 10^2-fold (2), 10^4-fold (3), 10^6-fold (4) and 10^8-fold (5). The plates were grown at 30˚C for 2 days prior to photographing the results.

The expression of such a super-repressor might be useful to confirm regulation of target genes by the constitutive down-regulation of promoter activity. In order to test this hypothesis the *rex*^{G102A} fragment was isolated from pSX142::G102A (Brekasis and Paget, 2003) using its flanking *Nde*I and *Bam*HI sites. This was then ligated into *Nde*I/*Bam*HI-cut pIJ6902, along with the *rex*^{WT} fragment from pSX142 to generate vectors pSX407 and pSX407::G102A. The *Streptomyces* plasmid, pIJ6902, is an integrative vector with a multiple cloning site, allowing insertion of an open reading frame under the control of the thiostrepton-inducible *tipA* promoter. These plasmids were then conjugated into the S106 (*Δrex*) strain, along with the vector-only control (pIJ6902). By introducing the *rex* gene into pIJ6902 in this way the negative feedback loop normally observed by

rex was bypassed. The resultant strains were grown on media in the presence or absence of thiostrepton (Figure 3.12). Spores were either streaked directly onto MS agar or diluted and spotted onto MM agar, containing NMMP trace elements and 1% glucose. The vector-only control strain and *rex*^{WT} strain grew well in both the presence and absence of thiostrepton. However, the *rex*^{G102A} strain appeared to only grow in the absence of thiostrepton suggesting that the resulting protein was toxic to the cells. As the same phenotype was not observed for the wild type protein it would appear that it is not simply the level of Rex present in the cell that causes the detrimental effect but is in fact the constitutive repression of one or more target genes. This further implies that one or more Rex target genes are essential for viability – the essentiality of one such gene, *ndh*, is discussed in Chapter 4.

In order to analyse the effect of Rex^{G102A} expression on target gene expression, RNA was isolated following induction and transcript levels determined by qPCR. Cultures were grown in the absence of thiostrepton, to an OD_{450nm} of ~0.7, prior to inducing expression of Rex^{G102A}. Both the *rex*^{WT} and pIJ6902 strains were treated in the same way to ensure that any differences caused by the late addition of the thiostrepton were accounted for in the control strains. A 10ml sample was taken just prior to the addition of 12.5µg/ml thiostrepton and then at 20, 40 and 60 minutes thereafter. All culturing was done in NMMP and was performed in triplicate. The resultant samples were pelleted, immediately resuspended in Kirby mix and then taken through the RNA purification procedure detailed in Section 2.3.3. The concentrations of the purified samples were assessed using the Nanodrop^{TM} and 1µg was taken through an additional DNAse step and subsequent reverse transcriptase reaction (Section 2.3.8). For qPCR analysis 2µl of ~50ng/µl cDNA (assuming an RT efficiency of 100%) were used per 25µl reaction. Standard curves were generated for each test gene using *S. coelicolor* gDNA dilutions, which were then used to convert Ct values to copy numbers for each gene. For cross-comparison between samples all values were to be normalised to an internal reference gene, 16S rRNA (primers 16S_QF/QF - A. Tabib, personal communication). An initial run was then performed for one of the biological replicates on the 16s rRNA and *cydA* genes.

As the 16s rRNA signal was so much higher than that of *cydA* the copy numbers were scaled down by dividing each 100-fold. The *cydA* copy number was then divided by this to cross compare between each sample; the results are shown in Figure 3.13. In the S106 vector only control strain the *cydA* expression appeared to be fairly constant. However, the expression dropped greater than 10-fold in the S106 (pIJ6902::*rex*) and S106 (pIJ6902::*rex*^{G102A}) strains. It appeared that the assay was working therefore the experiment was repeated for *ndh*, *ahpC*, SCO3547 (H^+-translocating pyrophosphatase) and cytochrome c biogenesis operon (SCO4472-4).

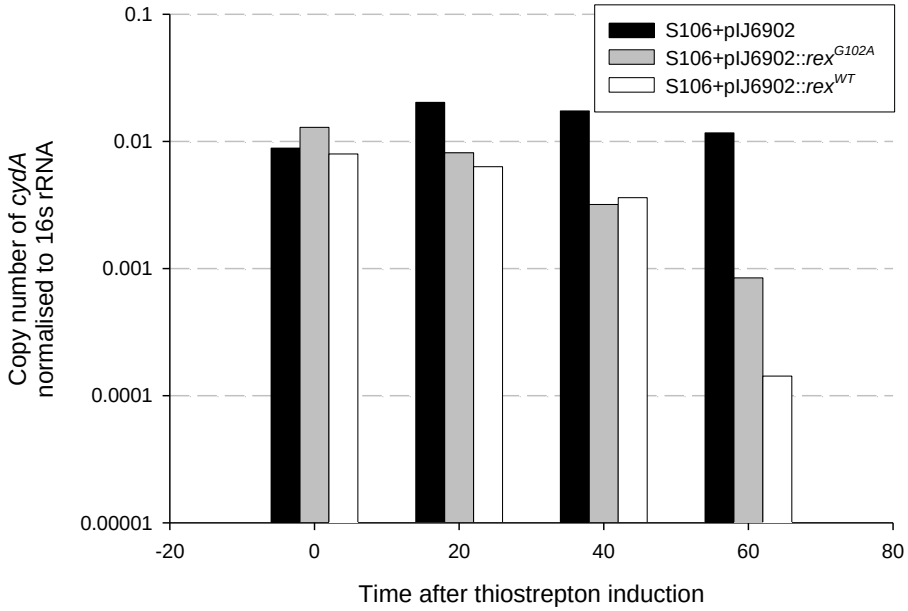


Figure 3.13: RT-qPCR results for *cydA* in the S106+pIJ6902::rex^{G102A} (grey), S106+pIJ6902::rex^{WT} (white), and S106 vector only (black) strains. RNA was harvested before and after induction with 12.5µg/ml thiostrepton at 0, 20, 40 and 60 minutes. The copy numbers were assessed using a gDNA standard curve and were normalised to those of 16s rRNA.

In this case only the 0 and 60 minute time-points were used as this gave the largest difference for the *cydA* gene. The experiment was performed for two biological replicates and is shown in Figure 3.14. The SCO3547 signal was extremely weak suggesting this gene is not expressed under the conditions used, it was therefore excluded from the results. The expression profiles for both the *ndh* and *ahpC* genes appeared to be unaffected by the presence of Rex^{WT} and Rex^{G102A}. There was however a slight decrease in the expression of

SCO4472 when *rex*^{G102A} was induced, compared to the vector only and *rex*^{WT} strains. The significance of this would need to be further verified either by multiple replicates or by *in vitro* transcript mapping. These data suggest that although the expression of a *rex* super-repressor decreases the expression of some target genes, at some promoters the effect is minor and that this approach might not be suitable for the general confirmation of regulation.

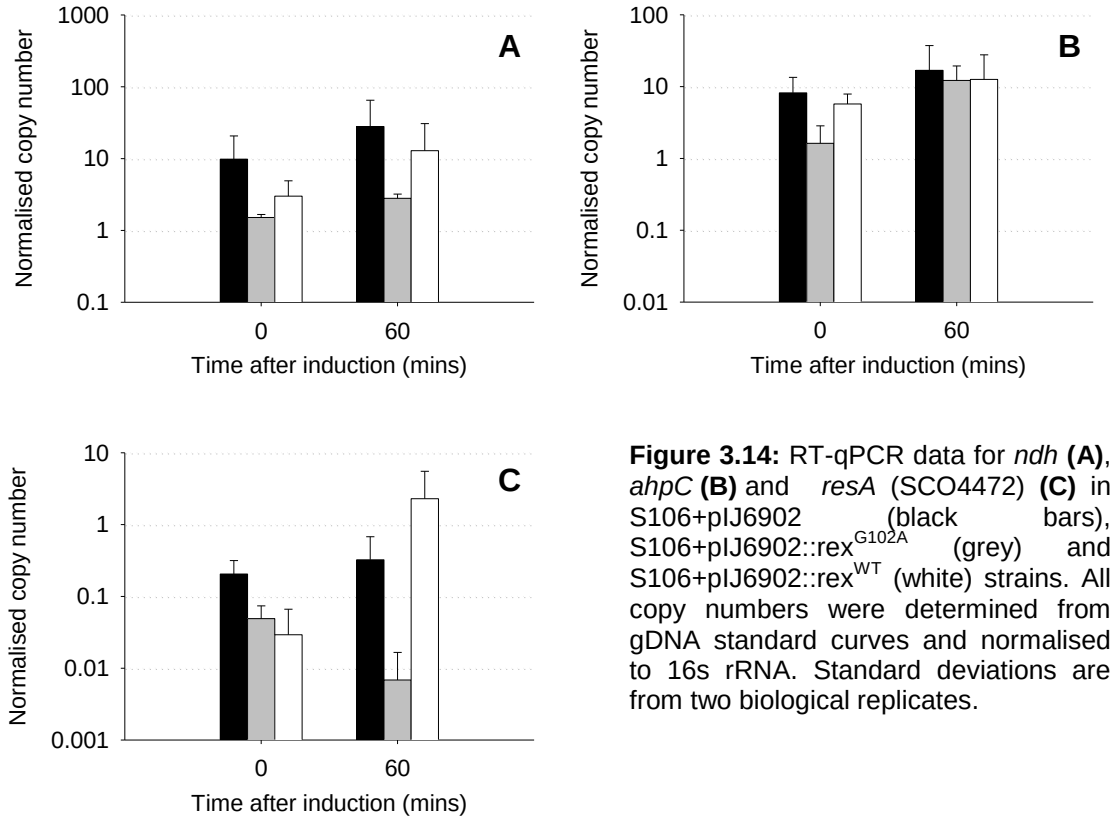


Figure 3.14: RT-qPCR data for *ndh* **(A)**, *ahpC* **(B)** and *resA* (SCO4472) **(C)** in S106+pIJ6902 (black bars), S106+pIJ6902::rex^{G102A} (grey) and S106+pIJ6902::rex^{WT} (white) strains. All copy numbers were determined from gDNA standard curves and normalised to 16s rRNA. Standard deviations are from two biological replicates.

Section 3.4.2 – Potential functions of Rex regulon members

Rex was originally identified through its ability to repress the cytochrome *bd* terminal oxidase operon (Brekasis, 2005, Brekasis and Paget, 2003). It was observed that under aerobic conditions this operon was repressed and that this effect could be bypassed through the deletion of *rex*, which caused constitutive expression (Brekasis, 2005, Brekasis and Paget, 2003). With *S. coelicolor* being an obligate aerobe it would seem that the best course of action during oxygen

limitation is to make use of what little oxygen is still available. Since its initial discovery, other Rex targets were identified that also have a role in the respiratory chain and energy generation: NADH dehydrogenases, *ndh* and *nuo*; cytochrome biogenesis operon (SCO4472-4); and potential ATP synthase gene (*atpI*). The following sections describe some new Rex regulon members with possible roles energy generation and/or redox control, phosphate metabolism, and antibiotic production.

Energy generation and/or redox control

***SCO3547*:** SCO3547 encodes a proton-translocating pyrophosphatase (*hppA*). In the *Streptomyces* Annotation Server (StrepDB) this gene is listed as similar to that of the photosynthetic proteobacterium *Rhodospirillum rubrum*. The proton-translocating pyrophosphatases have the ability to generate a proton-motive force by utilising the energy released from hydrolysing pyrophosphate (Garcia-Contreras *et al.*, 2004). In *Rhodospirillum rubrum* this enzyme appears to be used to generate energy in order to fill the deficit when switching from respiration to photosynthesis, under oxygen limitation (Garcia-Contreras et al., 2004). The *S. coelicolor* protein has also been characterised as having the ability to pump hydrogens, using the energy released from PPi hydrolysis, but the biological relevance of this is not yet known (Hirono and Maeshima, 2009).

***wblE*:** The *wblE* gene was previously identified as a potential regulon member (Brekasis, 2005); however little is known about this gene. It is named due to its similarity to WhiB, which is known to play a key role in differentiation and sporulation as its deletion results in a white colony phenotype (Chater, 1972). The *wblE* gene however does not appear to have such a vital role in sporulation as its deletion had no discernible impact on differentiation (Homerova *et al.*, 2003). This gene is still of interest though considering that its namesake is a potential transcriptional regulator (Davis and Chater, 1992). The closest *wblE* homologue in *M. tuberculosis* is WhiB1, a gene whose expression is regulated by the cAMP-receptor protein (CRP) (Agarwal *et al.*, 2006). The WhiB family of proteins has been shown to contain a 4Fe-4S cluster in both *Streptomyces* and *M. tuberculosis* (Jakimowicz *et al.*, 2005, Singh *et al.*, 2007). The work of Singh *et al.* suggested that this iron-sulphur cluster provides some form of oxygen and

nitric-oxide sensing within the protein (Singh et al., 2007). There are four conserved cysteine residues in each member of the WhiB family to help co-ordinate the cluster; these residues are also present in WblE (Jakimowicz et al., 2005). This may mean that it too may be redox sensitive but responding to an alternative signal than Rex.

***SCO5207*:** The second most highly enriched site in the ChIP-on-chip list was SCO5207. The function of this protein is unknown, however it does appear to consist of two tandem cystathionine β-synthase (CBS) domains, also known as Bateman domains (Bateman, 1997). The CBS domains are renowned for their ability to interact with adenosyl moieties, which allows them to act as 'energy-sensing modules' within multi-domain proteins (Scott *et al.*, 2004). The function of CBS-only proteins is still unclear, however it appears that a CBS domain expressed separately from its usual accompanying subunits can still regulate the function of that protein (Pimkin *et al.*, 2009), so perhaps these domains act as regulators for other proteins in the cell. In ATU1752 from *Agrobacterium tumefaciens* the CBS domain has been shown to co-ordinate an NADH molecule (PDB ID: 3fhm). It therefore seemed plausible that SCO5207 may act as a type of NADH sensing module in the cell. The gene was therefore amplified by PCR and introduced into pET15b for overexpression. A wavelength scan was performed on the resultant protein, looking for the signature 340nm peak of NADH, however only a 280nm peak was obtained (protein) with a slight shoulder at 260nm (data not shown). The 260nm peak could possibly represent NAD^+; this would be confirmed by converting the dinucleotide to NADH via reduction by alcohol dehydrogenase and rescanning for a 340nm peak. This reaction was performed but no 340nm peak was obtained, even when the protein was heat-denatured to ensure release of the bound cofactor (data not shown). If there is indeed a cofactor present in SCO5207 then it is neither NADH nor NAD^+.

***oxyR/ahpC*:** Rex appears to bind in the intergenic region between *oxyR* and *ahpCD*, the peroxide-sensitive transcriptional regulator and alkyl hydroperoxidase genes. OxyR was originally identified to be a redox-sensitive transcriptional activator, using the formation or reduction of a disulphide bridge

to activate or de-activate the protein, respectively (Zheng *et al.*, 1998). The crystal structure of *Escherichia coli* OxyR has since been solved in both its oxidised and reduced forms (Choi *et al.*, 2001). In *S. coelicolor* OxyR also acts as a positive regulator, requiring H_2O_2 activation, and its regulon includes the *ahpCD* operon as well as its own gene (Hahn *et al.*, 2002). Alkyl hydroperoxidases are able to reduce peroxides in the cell by use of internal cysteine residues, which are then re-reduced by other proteins. The *ahpC* gene of *S. aureus* has been deleted with no discernible effect on growth (Cosgrove *et al.*, 2007). However, in combination with a *katA* deletion (encoding a catalase) the cells grew poorly in aerated cultures, with wild-type growth levels recoverable by reduced aeration (Cosgrove *et al.*, 2007). The two genes apparently have differing functions, with *ahpC* deletion causing sensitivity to organic peroxides and *katA* deletion inhibiting the response to H_2O_2; nevertheless both genes appear to be important for the peroxide stress response (Cosgrove *et al.*, 2007). The *S. coelicolor* AhpC has four well conserved cysteine residues. In *Salmonella typhimurium* these residues allow an inter-subunit disulphide bond to form within the protein, during peroxide stress, which is reduced by a combination of NADH and AhpF (Poole, 1996). *S. coelicolor* lacks an annotated *aphF*. *M. tuberculosis* also lacks AhpF; however it seems that in this species AhpC may be reduced by AhpD, the product of the second gene in the *ahpCD* operon (Bryk et al., 2002) so perhaps this is also the case in *S. coelicolor*. The overlap of the -10 and -35 boxes of the divergent *oxyR*/*ahpCD* genes, and the position of the potential ROP site, would suggest that if Rex does indeed regulate these genes then it would regulate both (Figure 3.15).

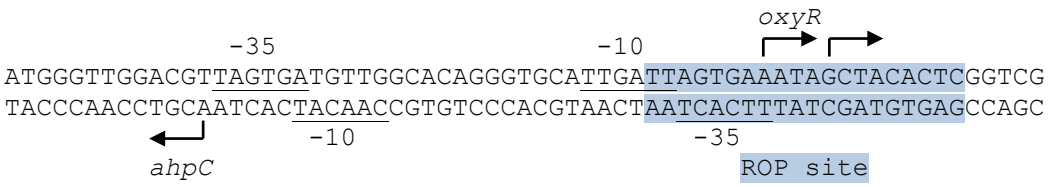


Figure 3.15: The intergenic region between the *oxyR* and *ahpC* genes. The -10 and -35 boxes and transcriptional start sites are indicated (Hahn, *et al*., 2002). The position of the predicted ROP site is highlighted in blue, overlapping the -35 of the *ahpC* and -10 of the *oxyR* gene.

Phosphate metabolism

Within the list of possible targets are three genes that form part of the PhoP regulon; SCO3790, SCO3791 and SCO7697 (Sola-Landa *et al.*, 2008). SCO3790 is a conserved hypothetical protein, which is targeted for export via the twin-arginine secretion system (Widdick *et al.*, 2006). Little is known about the function of SCO3791 and SCO7697, although the latter is noted as a secreted phytase due to similarity to the *B. subtilis* 3-phytase precursor. Continuing the phosphate utilisation theme are other genes such as SCO3547 and SCO6218, both phosphatases. Although the exact function of these genes is not clear it would seem that phosphate uptake and utilisation is somehow linked to the function of Rex. If this was the case then there might be a response to phosphate limitation in the S106 (*Δrex*) or S106 (pSET152::rex^{G102A}) strains. These strains were therefore inoculated onto SMMS agar, along with M145, S106 (pSET152::rex) and S106 (pSET152), under phosphate limited and phosphate- replete conditions (Figure 3.16). There was no obvious growth delay for the S106 (*Δrex*) strains under phosphate limitation, however they did appear to produce more actinorhodin under these conditions. Interestingly the S106 (pSET152::rex^{G102A}) strain had the opposite result, producing more actinorhodin in the presence of phosphate.

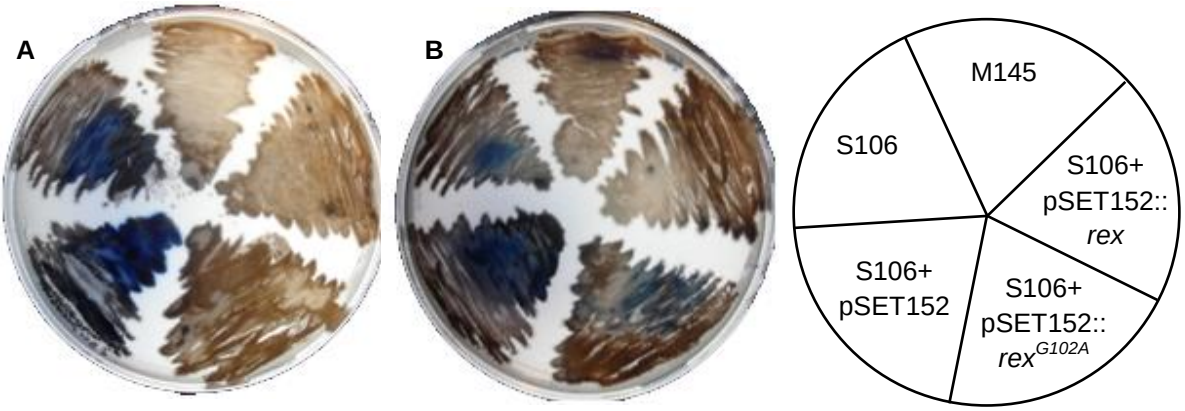


Figure 3.16: Phenotype analysis of the *rex*^{G102A} super-repressor strain and S106 *rex* null strain. Each strain was streaked over SMMS agar containing 1mM **(A)** or 15mM **(B)** phosphate. The plates were grown at 30˚C for 2 days prior to ammonium fuming and photographing the results.

Antibiotic production

SCO6280 was originally referred to as *kasO,* a potential regulator of a type I polyketide synthesis cluster (Takano *et al.*, 2005). This gene has however been subsequently renamed to *cpkO*, cryptic polyketide O (Gottelt *et al.*, 2010, Pawlik *et al.*, 2007). The *cpkO* gene forms part of a larger polyketide cluster, with deletion of *cpkO* having a negative effect on the expression of other members of the cluster (Takano *et al.*, 2005). The *cpkO* gene is itself regulated by the γ-butyrolactone receptor, ScbR, with two distinct binding sites in the *cpkO* promoter region (Takano *et al.*, 2005). The binding of ScbR to these regions was inhibited by γ-butyrolactone molecule SCB1, with addition of SCB1 having a positive effect on expression of *cpkO* (Takano *et al.*, 2005). The predicted ScbR sites in this promoter are centred at ~230bp and ~20bp upstream of the transcriptional start, with the second site overlapping the -35 and -10 regions (Takano *et al.*, 2005). The predicted Rex binding site for *cpkO* would lie in between these two operators, ~55 bp upstream of the transcriptional start site. Given the effect that ScbR has on expression of *cpkO* it would appear that any attempts to study the affect Rex has on this gene would require that the other repressor was first silenced by promoting SCB1 production. More recent work on *cpkO* has focused on promoting expression of the polyketide synthase that it is thought to regulate (Gottelt *et al.*, 2010, Pawlik *et al.*, 2007). A recent study showed that it may in fact be ScbR2, a homologue of ScbR contained within the *cpk* cluster, that regulates expression of polyketide synthesis from *cpk* (Gottelt et al., 2010). It appears that CpkO itself may act as an activator for the cluster, which includes *scbR2*, with ScbR2 then acting to repress expression of both *cpkO* and other cluster members in *S. coelicolor* (Gottelt et al., 2010). This work has also managed to identify a yellow-pigmented product, of this cluster, that has antibacterial properties (Gottelt et al., 2010). It is however not yet clear why Rex would regulate the expression of *cpkO*.

Section 3.5 – Conclusions

The ChIP-on-chip data revealed a large number of potential ROP sites *in vivo*. However, only a handful of the sites appeared to generate stable Rex:ROP complexes *in vitro* (as judged by EMSA). In many cases, putative binding regions that failed to generate EMSA shifts lacked a conventional dyad symmetrical ROP site ([A/T][A/T]GTG-n8-CAC[A/T][A/T]) but did contain a highly conserved half-site (TTGTGAA). By using an SPR competition assay it was possible to observe a weak interaction between the 3 half-sites tested; SCO6239, SCO6218/9 and SCO6917, as well as the artificial half-site. The strength of the interactions with each of these sites appeared to vary, suggesting that there are other determinant factors at these sites that alter the affinity of Rex. Not much can be determined at a sequence level from such a small number of sites; however the region that gave the highest level of competition (SCO6917) had one sequence element in the other half of the operator that the others lacked; the first cytosine of the inverted repeat usually associated with a ROP site (tt**C**ACaa). None of the other second-half sequence elements were present in this region. The SCO6239 region contained the two terminal adenosines and yet this site performed no better than the artificial ROP site in the SPR experiment. Given that these residues are less highly conserved in high affinity binding sites (those that gave gel shifts) than the 'CAC' motif it would appear that these residues only act to aid binding, not prohibit it when absent. Although none of the "half-sites" gave gel shifts with Rex, they were overrepresented in the ChIP-chip data. One possibility is that other components or factors might be present *in vivo* to help stabilise bound Rex that are absent *in vitro*. Factors such as DNA availability, conformation and co-operativity between other DNA-binding proteins have the potential to drastically affect the binding of transcriptional regulators (Minchin and Busby, 2009). Histones are a common occurrence in eukaryotic nucleoid structures and their various modifications are well characterised due to their potential to silence their neighbouring genes. Histone-like proteins are similarly associated with the bacterial nucleoid but their function is less well characterised. HU was originally identified in *E. coli*, and was observed to improve transcription from a λ–DNA template *in vitro*

(Rouviere-Yaniv and Gros, 1975). In *S. coelicolor* two histone-like proteins have been identified, HupA and HupS, the former being expressed in vegetative growth whereas the latter appears to be limited to just aerial hyphae, with a potential role in spore maturation (Salerno *et al.*, 2009). Transcriptional regulator CRP is also renowned for its ability to alter the DNA conformation upon binding, thus altering the way in which that region is presented to other DNA-binding proteins (e.g. RNA polymerase) (Chen *et al.*, 2001).

Potential Rex operator sites occur upstream of a range of genes but the strongest sites have a clear role in respiration, such as *cyd*, *ndh* and *nuo*. The regulation of potential targets has been less apparent than their binding sites, even with the use of the super-repressor Rex^{G102A}. It would seem that whilst Rex may well bind to many loci, it has only a minor role in regulation of the majority of these genes. The Rex regulon has also been studied in *Staphylococcus aureus* and *Bacillus subtilis* (Pagels *et al.*, 2010, Wang *et al.*, 2008). In both of these species lactate dehydrogenase and lactate permease feature prominently, both of which are lacking from the S-Rex regulon but not from the *S. coelicolor* genome. Within StrepDB there is annotated both a putative lactate permease and at least one lactate dehydrogenase. However the *ldh* genes appear to be specific to $_{D}$-lactate, which is not the stereoisomer recognised by the *S. aureus* and *B. subtilis* enzymes. Lactate dehydrogenase is an enzyme capable of converting pyruvate (from glycolysis) to lactate, with the additional benefit of recycling the reduced NADH produced by substrate-level phosphorylation (KEGG enzyme 1.1.1.27). The lactate permease is a symporter; in this context it is used to excrete the fermentative product lactate. However; *S. coelicolor* does not ferment. Borodina *et al.* analysed (*in silico*) the biochemical pathways that exist in *S. coelicolor* and found that the main limiting factor for fermentation was most likely the means by which they transported glucose into the cell, they use a proton symporter (Borodina *et al.*, 2005). They speculate that this method of glucose uptake would never be sufficient during fermentation as it would require that the proton-motive force was preserved for the purposes of providing the substrate for glycolysis (Borodina *et al.*, 2005). They also point out that in species capable of anaerobic growth the glucose symporter is replaced by other means of uptake that do not detract from the

proton-motive force of the respiratory chain (Borodina *et al.*, 2005). Thus fermentation alone may not be the most energy efficient option for *S. coelicolor,* they could however combine pathways for energy generation.

S. coelicolor cannot grow anaerobically but the genome encodes a number of options for surviving prolonged periods of oxygen limitation: (1) a lactate dehydrogenase, as well as phosphate acetyltransferase and acetate kinase, together capable of generating ATP through substrate level phosphorylation, (2) the ResDE and DosRS two-component systems for regulating gene expression under oxygen limitation, and most importantly (3) three copies of the *narGHJI* operon, each encoding a respiratory nitrate reductase (van Keulen *et al.*, 2007). The ResDE system of *B. subtilis* is an anaerobic control system involving the sensor kinase ResE and response regulator ResD (Geng *et al.*, 2007). This system uses both oxygen limitation and nitric oxide as the signal to effectively switch from oxygen to nitrate as the terminal electron acceptor (Geng *et al.*, 2007). The DosRS system of *M. tuberculosis* is akin to the ResDE system in that it too consists of a sensor kinase (DosS) and response regulator (DosR), and it too is induced by both anoxia and nitric oxide, however the DosRS system is a little more complex including both a third component DosT and a third signal carbon monoxide (Kumar *et al.*, 2007). The DosR regulon of *M. tuberculosis* is essential for persistence during latent Mtb infection (Leistikow *et al.*, 2010). The cells respond to the harsh environmental conditions provided by the host by ceasing cell division and maintaining cellular energy levels (Leistikow *et al.*, 2010). There are 48 genes associated with the DosR system, including a nitrate reductase (*narX*) and nitrite extrusion protein (*nark2*) (Voskuil *et al.*, 2003). Respiratory nitrate reductases are able to perform an analogous function to Complex I (NDH-1) but with the concomitant reduction of nitrate. Apparently all three respiratory nitrate reductases are active in *S. coelicolor*, with the ability to reduce nitrate to nitrite (Fischer *et al.*, 2010). This reduction cannot be continued to reduce nitrite to ammonium but the reduction of nitrate itself is potentially enough to reduce the load on the quinone pool, coupling NADH oxidation to nitrate reduction, under oxygen limitation (Fischer *et al.*, 2010). None of the *nar* genes appear in the ChIP-chip list thus Rex does not

appear to directly regulate their expression but it seems likely that these genes also play a role during redox stress.

The appearance of the peroxide response regulator gene; *oxyR,* in the ChIP-chip list was an interesting surprise. As peroxide production is generally associated with hyperoxia, instead of anoxia, it was not immediately apparent why this would be Rex regulated. However, the NDH-2 (*ndh*) of *S. coelicolor* also appears to be Rex-regulated (Brekasis, 2005, Brekasis and Paget, 2003). This enzyme is reportedly the main source of respiratory produced H_2O_2 in *E. coli* (Messner and Imlay, 1999, Seaver and Imlay, 2004). The flavin core of NDH-2 is highly reactive. When NDH-2 cannot immediately pass the electrons, obtained from NADH oxidation, to the quinone pool it will react with oxygen instead to produce peroxide (Messner and Imlay, 1999, Seaver and Imlay, 2004). Thus it would seem that any potential regulation of the peroxide stress response genes *oxyR*/*ahpCD* by Rex may in fact be to counter the effects of *ndh* induction.

Chapter 4

Results II: The NADH Dehydrogenases

"When you're finished changing, you're finished."

Benjamin Franklin (1706-1790)

Section 4.1 – Overview

NADH is ubiquitous in nature, involved in a multitude of cellular activities. This cofactor is particularly important for respiration where it acts as a shuttle to pass high energy electrons, released during substrate oxidation, to the respiratory chain. NAD(H) is constantly cycled between its reduced and oxidised states, a process which must constantly be maintained in order to support life. As both a substrate (NAD^+) and product (NADH) of redox reactions this cofactor has a key role in energy generation by acting as an electron carrier. The electrons that it transfers, in its reduced state, are passed to the quinone pool, a process requiring the membrane-associated NADH ubiquinone oxidoreductases. This chapter covers the investigation of the two main respiratory NADH dehydrogenases in *S. coelicolor*; NADH ubiquinone oxidoreductase (from the *nuo* operon) and NADH dehydrogenase (from *ndh*). Both are possible Rex targets and have the potential to directly influence the redox poise. A greater understanding of the regulation and expression of these genes should increase our understanding of how redox poise is maintained *in vivo*, and provide insights into the role of Rex. Whilst one might expect the purpose of all Rex regulon members is to respond to redox stress and ultimately restore the ratio, none have the potential for such a direct effect on the NAD^+/NADH redox poise as the NADH dehydrogenases. However, the differential roles of different classes of NADH dehydrogenases in *S. coelicolor* are poorly understood and so the relevance of their apparent control by Rex to cellular redox control is not currently understood.

In *S. coelicolor* the *ndh* gene appears to be essential, whereas its proton-pumping counterpart; *nuo*, is not. In fact it appears that expression of *nuo* is limited to just growth on solid media, with no detectable transcripts in liquid cultures. Through use of an inducible disruption strain it has been possible to deplete *ndh* from the cells, resulting in the derepression of the Rex target gene, *cydA*. It would appear that *ndh* has a key role to play in maintaining the redox poise of the cell.

Section 4.2 – The NADH dehydrogenases

Three types of membrane-associated NADH dehydrogenase exist in nature; Type I is proton-translocating, whereas Type II is not. Note that from this point forward the two classes of enzyme will be referred to as NDH-1 (Type I) and NDH-2 (Type II), however the genes from which these proteins are expressed will be referred to as *nuo* and *ndh*, respectively. The third type (NQR) is most similar to NDH-1 as it alters the membrane potential but instead of translocating protons it pumps sodium (Kerscher *et al.*, 2008). Both Type I and II NADH:quinone oxidoreductases appear in the list of ChIP-chip targets, whereas there does not appear to be a version of NQR in the *S. coelicolor* genome. The biological roles of the different types is not that clear (Kerscher et al., 2008). However, in bacteria, there appears to be a link between anaerobic respiration and *nuo* (Tran et al., 1997). NDH-1 is the equivalent of Complex I found in the mitochondrial electron transport chain. In *E. coli* it is this form of NADH dehydrogenase that is expressed under anaerobic growth, coupling the oxidation of NADH with generation of a proton-motive force (Unden and Bongaerts, 1997). In *S. coelicolor* NDH-1 is expressed from the 14 gene *nuoA-N* operon (SCO4562-75) operon. The crystal structure from *Thermus thermophilus* indicates the presence of a prosthetic group in the form of FMN, and also multiple iron-sulphur clusters. This protein complex includes multiple membrane-spanning helices, with the NADH-binding site positioned within the cytosol (Efremov *et al.*, 2010). NDH-2 is simple by comparison, consisting of a single protein expressed from the *ndh* gene (SCO3092). The protein also contains a flavin moiety (Bandeiras *et al.*, 2002). This section covers the gene and protein domain organisation, and what little is known about the potential regulation of these two NADH dehydrogenases.

Section 4.2.1 – The potential function of the NADH dehydrogenases

The type I NADH:quinone oxidoreductase is represented twice in the *S. coelicolor* genome but only once as a complete operon (Table 4.1 and Figure 4.1). The second copy of this operon is lacking genes *nuoD2*, *nuoE2*, *nuoF2*

and *nuoG2*. Furthermore, the third gene in the *nuoA2-N2* operon is only distantly related to *nuoC* (13% identity). The *nuoD2* gene is not contained within the operon; it is instead located ~1.2Mbp upstream of *nuoA2*. The *nuoB2*, *nuoL2* and *nuoM2* genes appear to be quite well conserved among other actinomycetes, however the rest of the *nuoA2-N2* operon appears to be lacking in most species (using the STRING functional protein association networks tool (Jensen *et al.*, 2009)). Interestingly the full length *nuo* operon of *Mycobacterium sp. JLS* is most similar to the *nuoA2-N2* operon of *S. coelicolor*, with its equivalent of *nuoA-N* lacking a large number of genes. This arrangement does not appear to be shared amongst other mycobacterial species. The lack of conservation of a complete *nuoA2-N2* operon in other species would suggest that either the individual components have other functions within the cell or that it is purely a remnant of a duplication event. It is only the start of the *nuoA-N* operon that appears in the ChIP-chip list and only the *nuoA* promoter region that contains a ROP site (Table 4.1 and Figure 4.1). In the recent structure of the *Thermus thermophilus* NDH-1 (PDB ID: 3M9S) (Efremov et al., 2010) seven of the 14 subunits are present; *nuoB, C, D, E, F, G* and *I*. The flavin moiety (FMN) is attached to the equivalent of the *S. coelicolor nuoF* subunit (chain A of PDB), along with an iron-sulphur cluster. There is a Rossmann fold present in this chain, as may be expected given its interaction with FMN, and a characteristic GXGXXG motif (Gly-Arg-Gly-Gly-Ala-Gly). This motif is completely conserved within the equivalent subunit (*nuoF*) of *S. coelicolor* NDH-1. The membrane spanning domain shown in the *T. thermophilus* PDB is actually that of *E. coli* (PDB ID: 3M9C) formed from NuoA, J, K, L, M, and N (Efremov et al., 2010). It is this portion that allows NDH-1 to act as a proton pump.

The type II NADH:quinone oxidoreductases are often referred to as alternative or uncouplers due to their ability to carry out the redox reaction without generating a charge differential across the membrane (Bertsova *et al.*, 1998) (Camougrand *et al.*, 1983). This appears wasteful, because it reduces ATP synthesis. However, although NDH-2 re-oxidises NADH without proton-translocation it still passes electrons to the quinone pool. These quinones can

Gene	Length	Gene	Length	Score (%)
NuoA	119	NuoA2	146	31
NuoB	184	NuoB2	232	39
NuoC	255	NuoC2	453	13
NuoD	440	NuoD2	383	37
NuoE	290	NuoE2	-	-
NuoF	449	NuoF2	-	-
NuoG	843	NuoG2	-	-
NuoH	467	NuoH2	322	37
NuoI	211	NuoI2	197	28
NuoJ	285	NuoJ2	197	29
NuoK	99	NuoK2	130	42
NuoL	654	NuoL2	664	31
NuoM	523	NuoM2	534	29
NuoN	552	NuoN2	515	30

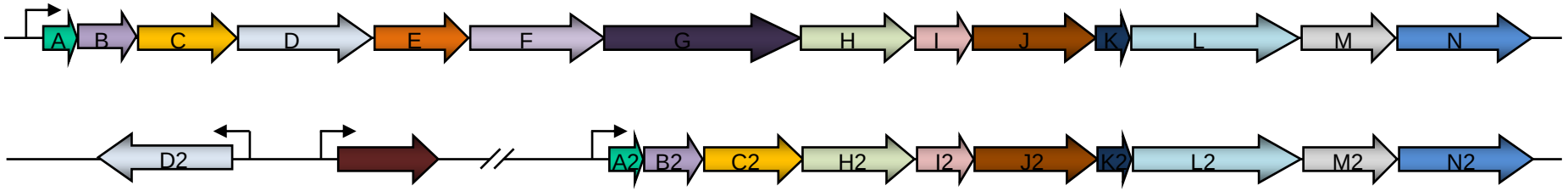
Table 4.1 and Figure 4.1: The two Type I NADH:quinone oxidoreductase operons in *S. coelicolor*. Like-for-like comparison (ClustalW scores) of the genes from each operon is shown in the table and organisation is shown in the figure below. Note that *nuoE2*, *nuoF2* and *nuoG2* are completely absent. *nuoD2* is present in the genome but is separate from the second *nuo* operon and is on the complementary strand. The unlabelled gene (depicted in maroon) is there simply to emphasize that *nuoD2* is completely isolated from the other genes. The operon annotated as *nuoA-N* (rather than *nuoA2-N2*) is the operon represented within the ChIP-chip data. An alignment of the promoter regions is also shown, with the ROP^{nuo} site indicated in bold and underlined.

```
NuoP            GGAGATCACAAAGCTTGTGTAATACCCCGTGTCGCAGATCACAGAGCGTCGGGCATAGGA 60
Nuo2P           -------GTCATACACGT-CAGCGTGCCGTGCGGCGCGACGCCGCGCCATGAG-GTGGAC 51
                       *  *  **  *     *****  **    * * * **   * *  * *

NuoP            TGCGAGGCAGTTGGGCTTGTGACCTGCTTCACATGTTCGCGATCTTCGTCGGGACGGGCG 120
Nuo2P           CGCCGGGATCCACTGAGGAAAACCCCACCCCCGCGTCT-CGAGGGGCGACGGGCTGCTTA 110
                 **  **       *      ***     * *  **   ***    ** ****  *

NuoP            GGGCTCGTGGGGCTG--TTGGGGCGGCTGTGAGTCCAGTGCAACCGCCAGCAGT-----C 173
Nuo2P           CCGTCTCTTTATCAGGCTCTAGACTCCCGTGCGTACGGATCGTCGAACCGCAGCGTTCAC 170
                  *    *    * *  *   * *  * *** ** * *  *  *   * ****      *

NuoP            AGTGCCGACTGAGAGGAGCGAGGAGCG--- 200
Nuo2P           GAGGTTCACGACAGGGAGCGAGGGGCGCAC 200
                  *   **     ********* ***
```



then be recycled by cytochrome terminal oxidases, which in some cases translocate protons. As mentioned previously there appears to be a link between *nuo* and oxygen limitation in *E. coli* (Tran et al., 1997), making NDH-1 a key source of the proton motive force in the absence of oxygen, whereas this role is taken up by the terminal oxidase under aerated growth. As with *nuo*, *ndh* (SCO3092) is not the only putative Type II NADH dehydrogenase in the *S. coelicolor* genome, there are four paralogues with % identities >30%. In a study of the *B. subtilis* NDH-2 it was shown that out of three putative type II NADH dehydrogenases only one (*ndh*) had a detrimental effect on growth when mutated (Gyan *et al.*, 2006). It is interesting that the other genes are not able to compensate for loss of *ndh* in *B. subtilis*; it appears that the same is true of the *S. coelicolor* homologues as *ndh* depletion prevents cell growth (Section 4.3.7). This presents two possibilities: either the homologues are unable to carry out the same function as NDH-2 or they are not expressed, and are therefore unable to compensate. Nevertheless it is only the *ndh* promoter that contains a Rex binding site and thus it is this gene that this study will focus on.

Section 4.3 – Regulation and expression of *ndh*

As mentioned previously, the NADH dehydrogenases potentially have a direct influence on the NAD^+/NADH redox poise. If this is the case then loss of that enzyme activity would cause the ratio to shift, causing Rex to dissociate and de-repress its targets. To investigate the role of *ndh* in maintaining the NAD^+/NADH redox poise, two approaches were taken: (1) the *ndh* gene was disrupted to see the effect on expression of the regulon; (2) the *ndh* gene was overexpressed to confirm that it encoded an enzyme with NADH dehydrogenase activity.

Section 4.3.1 – The *ndh* promoter

The *ndh* gene is one of the few genes to be upregulated in the S106(*Δrex*) strain (Brekasis, 2005). The promoter region also appears to be highly enriched

in the ChIP-on-chip data (Section 3.2.4). Analysis of the *ndh* promoter had identified two potential ROP sites, one centred at -43bp (TTGTGAAGGGGCGCACGA) and one at -315bp (TCGTGAAGTTCTTCACAA). Note that the ROP positions are numbered based on the experimentally determined coincident transcription and translation start site of *ndh*; this is located ~84bp upstream of the current annotated start site (D. Brekasis, personal communication). From looking at the positions of the two ROP sites one would expect that Rex would best exert control by binding to the promoter-proximal site, which potentially overlaps the -35 promoter region. However, sequence analysis revealed that the promoter-distal site conformed better to the consensus ROP site, with a high proportion of conserved sequence elements (Table 4.2). To study the relative importance of the two ROP sites, each was mutated, individually and in combination, and the promoter regions analysed *in vitro* by EMSA and *in vivo* using the reporter plasmid pLST920.

Fragment	Position	Sequence
ROP1	-43 bp	TT**GTG**AAGGGGCG**CAC**GA
ΔROP1	-43 bp	TTACAAAGGGGCGTGTGA
ROP2	-315 bp	TC**GTG**AAGTTCTT**CAC**AA
ΔROP2	-315 bp	TCACAAAGTTCTTTGTAA
Consensus ROP		**ttGTGaannnnttCACaa**

Table 4.2: The sequences of the two ROP sites within the *ndh* promoter, ROP1 (-43 bp) and ROP2 (-315 bp) upstream from the translational start site. Also shown are the sequences of the mutated ROP sites at each position that were used for both EMSA analysis and for the *neo* reporter assay. Only the highly conserved 'GTG' and 'CAC' of each site (shown in bold) was mutated (underlined regions), with each nucleotide having undergone a transition mutation.

Section 4.3.2 – Rex binds preferentially to the upstream *ndh* ROP site

In order to study the *ndh* promoter the entire region was amplified by PCR using primers ndh_414_for and ndh_414_rev to generate a 414 bp fragment. This fragment was ligated into EcoRV-cut pBlueScript II SK+, selecting for colonies that had the desired orientation of the fragment for subsequent isolation with *Kpn*I (pSX414). The resulting vector was used as a template for PCR-mediated mutagenesis of the two ROP sites. The -315 bp site was mutated using primers ndh_SDM1_for and ndh_SDM2_rev (ΔROP2), and the -43 bp site mutated with primers ndh_SDM3_for and ndh_SDM4_rev (ΔROP1), generating vectors

pSX415 and pSX416. Each nucleotide of the conserved 'GTG' and 'CAC' of each ROP was altered by a transition mutation (Table 4.2). A double mutant was then generated using the ndh_SDM1_for and ndh_SDM2_rev primers on the pSX416 vector; the resultant plasmid was pSX417 (ΔROP1+2). Each of the 414bp fragments were subsequently amplified for EMSA analysis using primers ndh_414_for and ndh_414_rev (Figure 4.2).

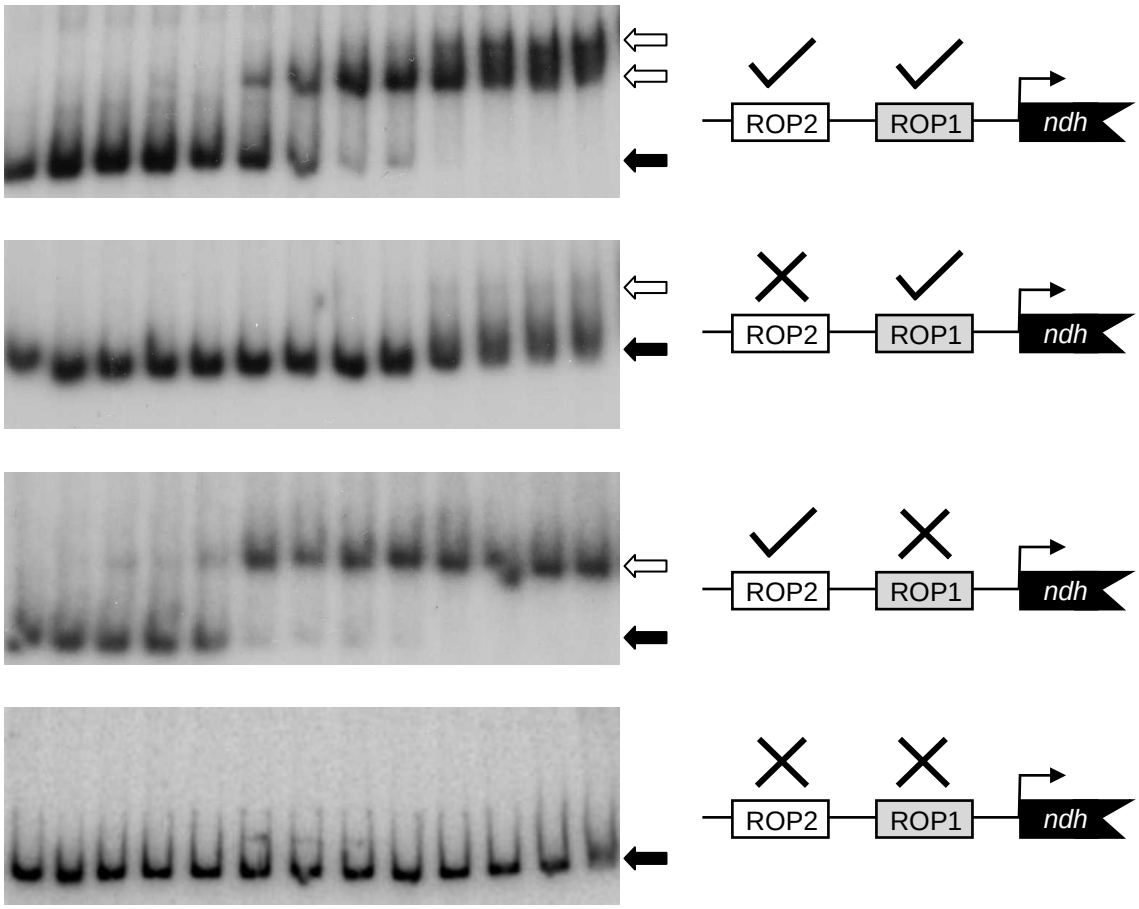


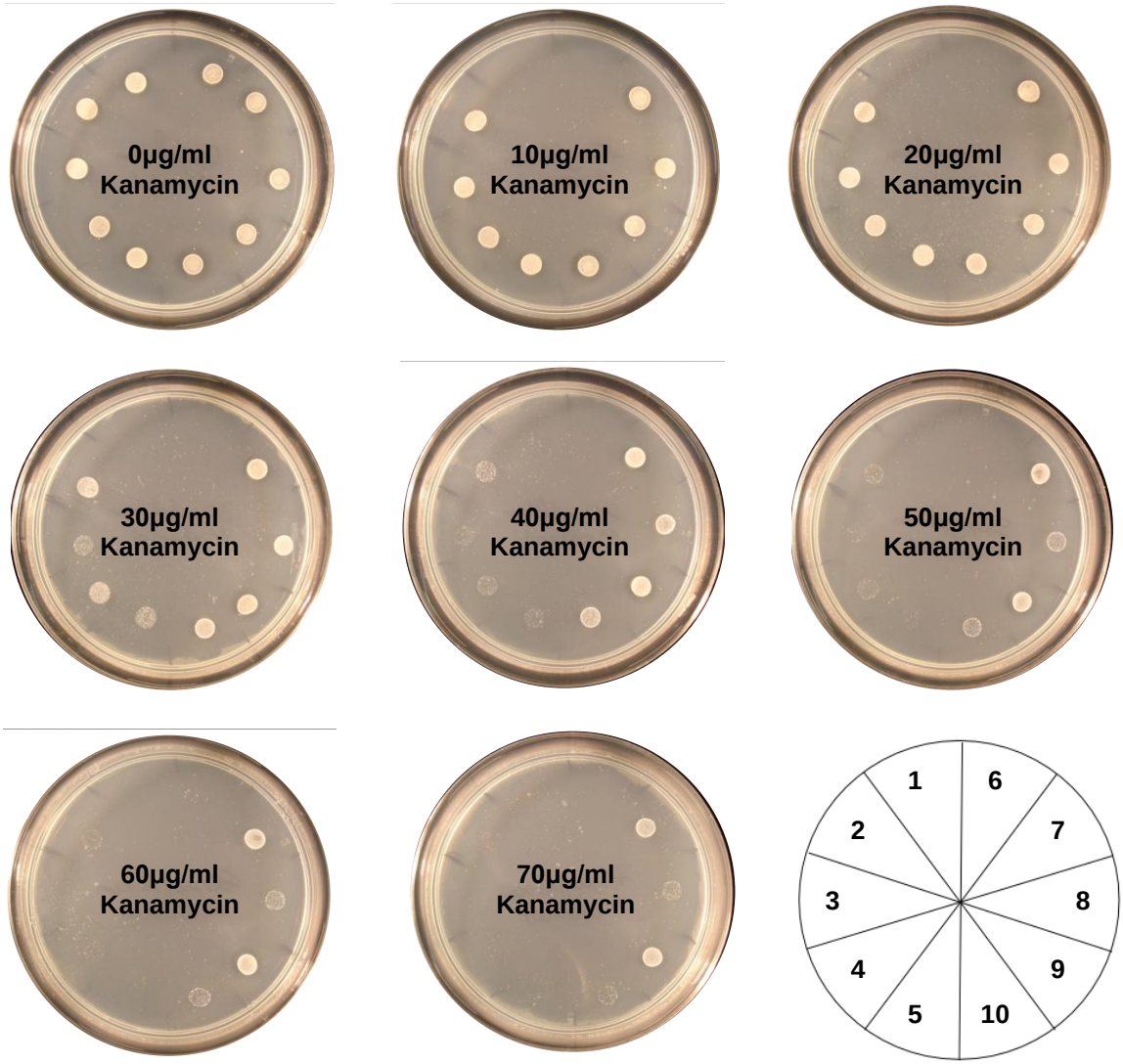
Figure 4.2: EMSA with *S. coelicolor* Rex and the *ndh* (SCO3092) promoter region (414bp). In each case 1ng of [γ-^{32}P]-labelled probe was incubated with Rex for 30 minutes at 25°C prior to running on a 6% non-denaturing polyacrylamide gel. The protein concentration in lane was as follows; 0nM (lane 1), 2.5nM (lane 2), 5nM (lane 3), 7.5nM (lane 4), 10nM (lane 5), 25nM (lane 6), 50nM (lane 7), 75nM (lane 8), 100nM (lane 9), 250nM (lane 10), 500nM (lane 11), 750nM (lane 12) and 1μM (lane 13). The open arrows indicate a gel shift, whereas the closed arrows indicate the unbound probe.

As had previously been observed, the wild type *ndh* promoter fragment was capable of producing two gel shifts, one at a Rex concentration of ~25nM and the other at ~100nM (Brekasis, 2005). As expected the ΔROP1+2 double mutant was incapable of producing a gel shift. The two single ROP^{ndh} mutants

were still both capable of generating gel shifts, however to differing extents. The ΔROP2 fragment produced a single shift at a Rex concentration of ~5nM, the same result with the ΔROP1 fragment required ~250nM of protein. Interestingly the ROP1 site produced a stronger signal in the gel shift when the ROP2 site was intact than it did in the absence of ROP2. This is suggestive of possible co-operativity between the two ROP sites; however this was not confirmed. The results of the EMSA analysis would confirm our initial assessment, that whilst the -43bp site (ROP1) is in a prime location to alter gene expression it is in fact the -315bp site (ROP2) that gives the strongest Rex-ROP complex.

Section 4.3.3 – Reduced expression from the ΔROP2 *ndh* promoter

In an attempt to confirm the findings of the previous subsection *in vivo* the fragments were isolated with *Kpn*I and introduced into the *neo* reporter construct pLST920 (Stratigopoulos *et al.*, 2002), cut with the same enzyme. The pLST920 vector contains a promoterless *neo* gene, which confers kanamycin resistance only upon insertion of an active promoter. The orientation of the inserts was confirmed by restriction analysis and the resultant vectors conjugated into M145 and S106 (*Δrex*). Each of the strains was spotted onto MM agar, containing increasing concentrations of kanamycin; from 0 to 50μg/ml (Figure 4.3). From the EMSA results one would have expected that the M145 (pLST920::wt P^{ndh}) strain would have the lowest level of kanamycin resistance as in this strain Rex would be capable of binding to both ROP sites to inhibit expression of the *neo* reporter gene. One would therefore also expect the M145 (pLST920::ΔROP1+2 P^{ndh}) double mutant to have the highest level of kanamycin resistance. In fact the opposite result was observed; the double ROP mutant had a lower level of kanamycin resistance than the wild type *ndh* promoter construct. In fact it appeared that any attempt to disrupt the -315bp ROP site had a detrimental effect on expression of the reporter gene, which was surprising given its distance from the transcriptional start site. The S106 (*Δrex*) strains were all more resistant to kanamycin but the same trend was followed as detailed above.



Strain	Plasmid	Level of resistance	Plate position
M145	pLST920	-	1
	pLST920::wt P^{ndh}	++	2
	pLST920::ΔROP2 P^{ndh}	+	3
	pLST920::ΔROP1 P^{ndh}	++	4
	pLST920::ΔROP1+2 P^{ndh}	+	5
S106	pLST920	-	6
	pLST920::wt P^{ndh}	+++	7
	pLST920::ΔROP2 P^{ndh}	++	8
	pLST920::ΔROP1 P^{ndh}	+++	9
	pLST920::ΔROP1+2 P^{ndh}	++	10

Figure 4.3: Kanamycin reporter assay to investigate the effect of disruption of each of the *ndh* ROP sites on gene expression. Each strain was spotted onto MM agar containing the indicated amount of Kanamycin. The relative levels of resistance are summarised in the table above.

As the S106 (*Δrex*) strain lacked Rex it would seem that the reduction in gene expression, observed when the promoter-distal ROP site is disrupted, was not due to the regulation by Rex. The most likely cause of this phenotype was

therefore the presence of an activator binding site overlapping the -315bp ROP site. The *ndh* promoter sequence is not that highly conserved in other streptomycetes and yet both ROP sites are (Figure 4.4). Clearly both sites have an important role in the expression of the *ndh* gene.

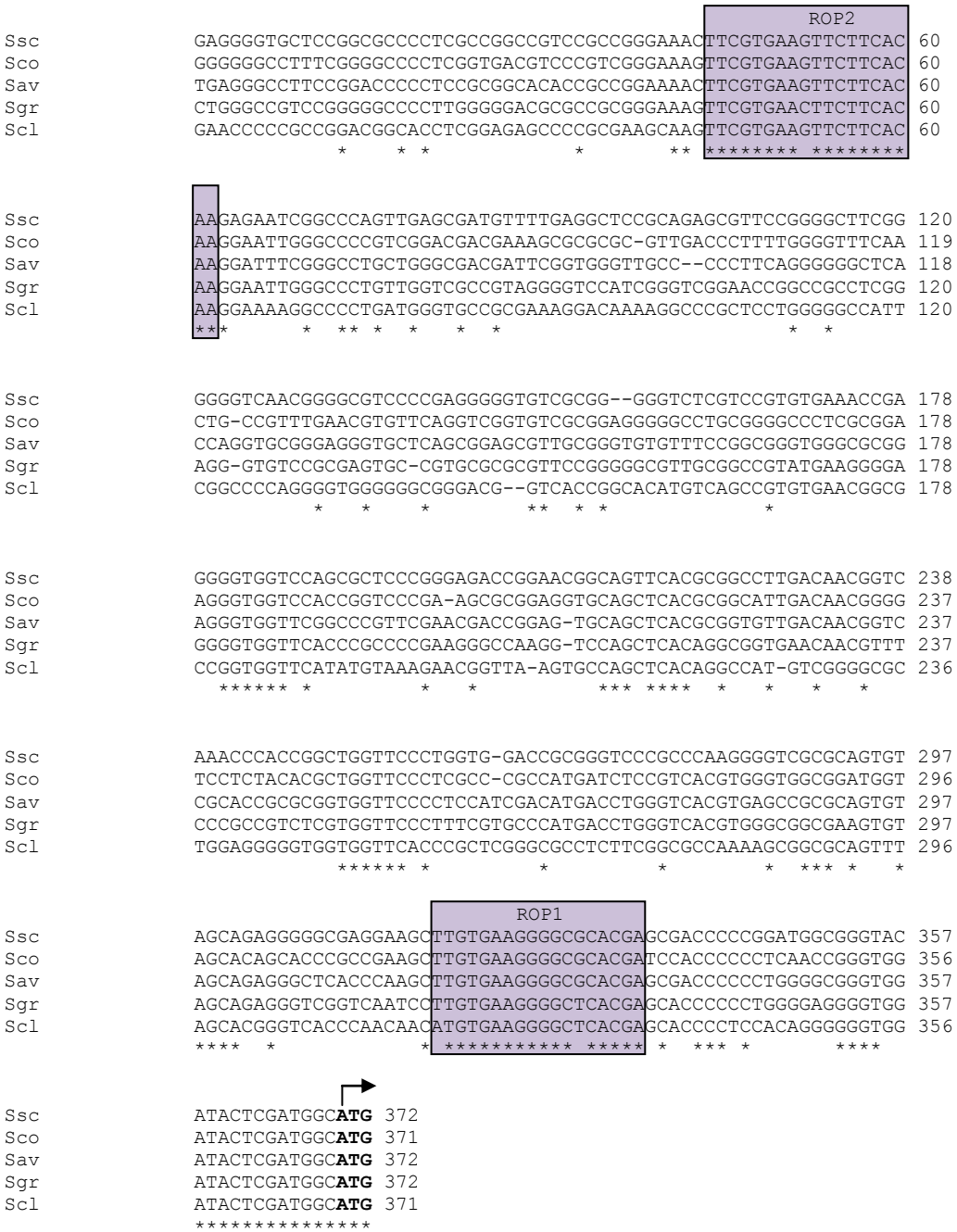


Figure 4.4: The aligned promoters of the *ndh* genes from *Streptomyces scabiei* (Ssc), *Streptomyces coelicolor* (Sco), *Streptomyces avermitilis* (Sav), *Streptomyces griseus* (Sgr) and *Streptomyces clavuligerus* (Scl). The purple boxes indicate the positions of the two ROP sites and the arrow indicates the translational start site.

Section 4.3.4 - Generating an *ndh* mutant using the REDIRECT approach

The REDIRECT© method of gene disruption involves the replacement of a target gene with an antibiotic resistance cassette, using homologous recombination to produce the switch (Figure 2.1) (Gust *et al.*, 2004). This method of gene disruption is commonly used for *Streptomyces* knock-out construction (Kallifidas *et al.*, 2010, Dedrick *et al.*, 2009, Bennett *et al.*, 2007, Hoskisson *et al.*, 2006). The first stage required the isolation of the *S. coelicolor* cosmid that contained the *ndh* gene (StE25). The regions upstream and downstream of *ndh* were then incorporated into the sequence of the 5' ends of primers designed to amplify the apramycin disruption cassette of pIJ773 (primers SCO3092_KO_For and SCO3092_KO_Rev). The primers were designed using the annotated translational start site of *ndh* in order to generate an in-frame deletion; however an alignment of *ndh* homologues suggests that the start site actually resides 84bp upstream of this. Whilst this still results in an in-frame deletion it should be noted that instead of leaving the first 8 amino acids of NDH-2 intact this disruption would actually express the first 36 amino acids of this protein, in addition to the last 10 amino acids incorporated by the position of the reverse primer. The disruption cassette included the apramycin resistance gene: it also contained an origin of transfer and FLP recognition sites, for use later on in this procedure. The disruption cassette, complete with homologous ends, was recombined into the StE25 cosmid using the recombinogenic BW25113 (λRed) strain. The resultant recombinant cosmid was analysed by restriction digest, which confirmed the presence of the apramycin cassette in the cosmid (Figure S2). This was subsequently transformed into the methylation negative/conjugation positive *E.coli* strain ET12567 (pUZ8002), allowing the cosmid to be conjugated into *S. coelicolor* M145. The cosmid itself contained the *neo* gene, conferring kanamycin resistance, in addition to the apramycin resistance conferred by the disruption cassette. This meant that single cross-over strains, where the cosmid had integrated into the genome without replacing the *ndh* gene, would have a Kan^R/Apr^R phenotype. Single cross-over (Kan^R/Apr^R) colonies were isolated and used to screen for double cross-over (*ndh* disruption; Kan^S/Apr^R) strains. Unfortunately after screening more than 3000 colonies no Kan^S/Apr^R strains could be isolated.

Section 4.3.5 – Attempts to induce *ndh* KO with an additional gene copy

The failure to isolate a double-cross over strain suggested that *ndh* might be an essential gene. If the introduction of a second copy of the gene allowed the disruption strategy to work then it would support the idea that the gene was essential; this approach was therefore taken with *ndh*. The entire *ndh* open reading frame and ~105 bp of its upstream region were amplified using primers SCO3092_ROP1 and SCO3092_REV. This was ligated into EcoRV-cut pBlueScript II SK+ generating pSX418. The *ndh* gene was then isolated using EcoRI sites, that had been incorporated within the primer sequences, and was ligated into pHJL401 (Larson and Hershberger, 1986) cut with the same enzyme (pSX419). The *E. coli*/*S. coelicolor* shuttle vector pHJL401 confers resistance to ampicillin and thiostrepton (Larson and Hershberger, 1986), allowing it to be used in conjunction with the Kan^R/Apr^R disruption cassette containing strains. The resultant plasmid was then transformed into *S. coelicolor* protoplasts produced from the single cross-over and M145 strains. The resultant spores were then diluted and used to inoculate agar plates to form single colonies in order to screen for Kan^S/Apr^R double cross-over strains. On the first round of screening for M145 Δ*ndh*::*apr* (pHJL401::ndh) and on the second round of screening for S106 Δ*rex* Δ*ndh*::*apr* (pHJL401::ndh) two Kan^S/Apr^R/$Thio^R$ colonies were isolated. Initial attempts to isolate thiostrepton-sensitive derivatives that had lost the plasmid-borne *ndh* proved unsuccessful. It would appear that the REDIRECT© method had worked but that perhaps *ndh* was an essential gene in *S. coelicolor.*

Section 4.3.6 – Generation of an inducible *ndh* disruption strain

In order to confirm whether or not *ndh* was essential in *S. coelicolor* the gene was to be placed under the control of an inducible promoter. If it were essential then the resultant strains would only be viable in the presence of the inducer, whereas a non-essential gene could be 'switched off' without killing the cells. Ideally this construct would be used to promote the second recombination event, as with the M145 Δ*ndh*::*apr* (pHJL401::ndh) strain, guaranteeing that the only copy of *ndh* in the cells was under the control of the inducer. Several

attempts were therefore made to clone *ndh* into a derivative of pIJ6902 in which the apramycin resistance cassette had been replaced with a strep/spec marker (pSX420); in this plasmid *ndh* would be inducible by thiostrepton and the integrative plasmid could be introduced into M145 *Δndh*::*apr* (pHJL401::ndh) allowing the loss of pHJL401::*ndh*). However, pSX420::ndh could not be constructed due to plasmid instability in *E. coli*, and so an alternative route to a null mutant was sought. The principle behind this method was that a single cross-over recombination event would place the chromosomal copy under the control of an inducible promoter and the resulting second (incomplete) copy of *ndh* would be tagged for degradation (Figure 4.5). The first stage was to generate the degradation tagged *ndh* sequence. The *ndh* gene was isolated using primers SCO3092Complete_For and SCO3092Complete_Rev, introducing *Nde*I and *Eco*RV sites at the start and just before the stop codon, respectively. The fragment was blunt-end ligated into *Eco*RV-cut pBlueScript II SK+ to generate pSX403. Design of the degradation tag was based on the *S. coelicolor* transfer-messenger RNA sequence (tmRNA), incorporating the majority of the proteolysis tag normally encoded by the tmRNA (Tu *et al.*, 1995). The tmRNA (also referred to as 10Sa RNA) of *E. coli* was identified from the presence of a short peptide sequence in the C-terminus of truncated forms of heterologously expressed interleukin-6 (Tu et al., 1995). It was found that this peptide was the product of the *ssrA* gene but was not a transcriptional fusion (Tu et al., 1995). The structure of the tmRNA is thought to be similar in appearance to tRNA, with an alanine anticodon that shifts translation from the original transcript to the tmRNA-encoded transcript (Keiler *et al.*, 1996, Jentsch, 1996, Felden *et al.*, 1997). The tmRNA-encoded peptide acts as a tag to target the truncated protein for degradation by C-terminal specific proteases (Keiler et al., 1996). The amino acid sequence used to target *ndh* for degradation in *S. coelicolor* was RDSSQQAFALAA (from tmRNA website, (Williams and Bartel, 1998)). This was flanked by sticky ends for *Sph*I and *Eco*RI, which were formed by annealing primers Ndh_deg_a and Ndh_deg_b. The annealed fragment was ligated into *Sph*I/*Eco*RI cut pSX403, to form pSX404.

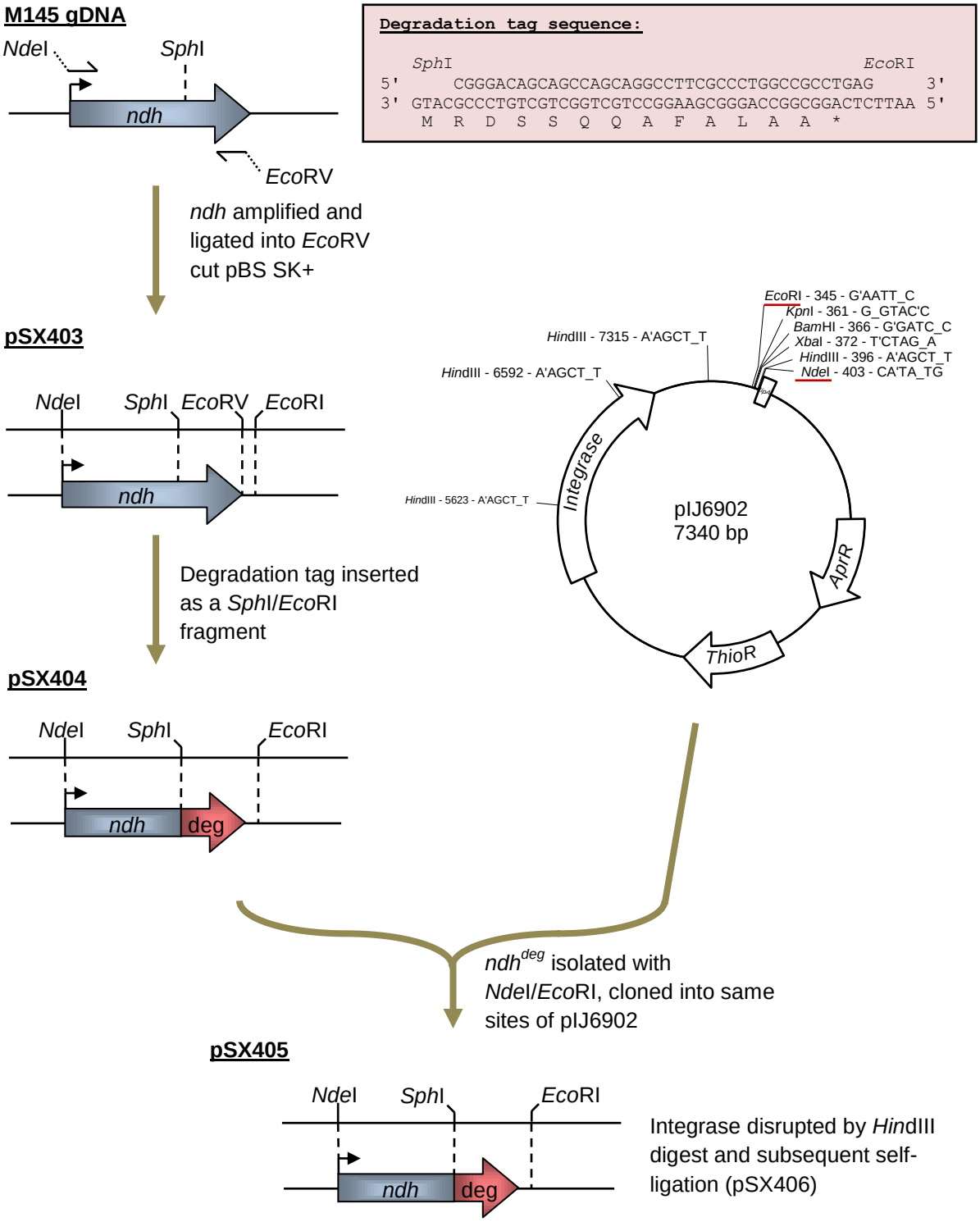
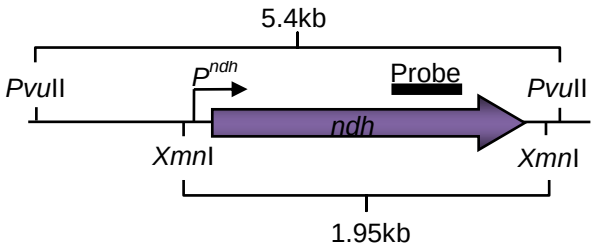
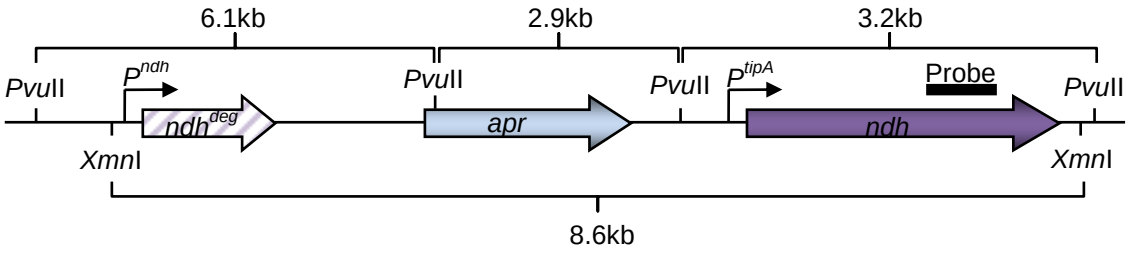


Figure 4.5: Diagrammatic view of the *ndh*^{deg} construction. The sequence of the degradation tag is shown above, with flanking *Sph*I and *Eco*RI sticky ends. The pSX405 construct was converted to a suicide vector by *Hin*dIII digest and subsequent re-ligation, which eliminated the majority of the integrase gene.

Wild type:



Ndh^{deg} disruption:



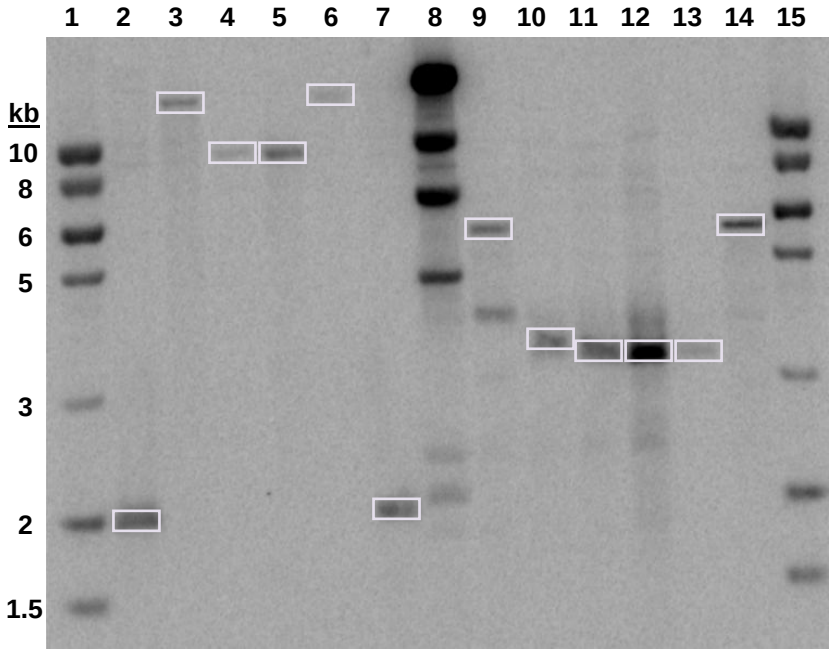


Figure 4.6: Sothern blot of two M145Δndh^{deg} (lanes 3, 4,10 and 11) and S106Δndh^{deg} (lanes 5, 6, 12 and 13) strains compared to the M145 (lanes 1 and 9) and S106 (lanes 7 and 14) strains, containing pIJ6902 alone. The chromosomal DNA in lanes 2 to 7 were cut with *Xmn*I and lanes 9 to 14 were cut with *Pvu*II. Lanes 1 and 15 contained the standard DNA ladder (Hyperladder) and lane 8 contained a *Hin*III-cut λ ladder. The diagrams above show the expected bands for each strain, note that in lanes 3 and 6 the bands were much larger than expected.

In this construct, the degradation tag is fused at amino acid position 350 of NDH-2 (out of 474 residues). The entire *ndh*^{deg} fragment was isolated using the *Nde*I/*Eco*RI sites and introduced into pIJ6902, cut with the same enzymes (pSX405). The final stage was to disrupt the pIJ6902 integrase gene to force a homologous recombination event once introduced into *S. coelicolor*. This was done by a complete digestion with *Hin*dIII and subsequent self-ligation, leaving only the first ~400bp of the integrase gene (pSX406). The final construct was then conjugated into M145, and apramycin resistant colonies isolated. Growth was carried out in the presence of 12.5µg/ml thiostrepton, the inducer. The resultant strains were characterised by Southern blot analysis to confirm the presence of the disruption cassette within the *S. coelicolor* chromosome (Figure 4.6). The results showed that all of the *Δndh*^{deg} strains isolated contained the cassette, however only one M145 and one S106 (*Δrex*) strain contained a single disruption cassette (as indicated by the *Xmn*I digest) – these strains were therefore taken forward for subsequent analysis.

Figure 4.7: Growth of Ndh^{deg} mutants on both MS agar in presence and absence of 10µg/ml thiostrepton. All plates also contained 20µg/ml apramycin. The plates were incubated at 30°C for 4 days.

Section 4.3.7 – Characterising the *ndh*^{deg} disruption strain

As an initial test the strains were streaked onto MS agar in the presence and absence of 10µg/ml thiostepton (Figure 4.7). Whilst the growth of both of the *ndh*^{deg} strains appeared unaffected by the disruption cassette with the inducer (thiostrepton) present, growth in the absence of thiostrepton was severely

hindered. In fact the only observable growth for these strains was a few colonies where the concentrated spore stock had been applied. The same effect was also observed when ~10^8 spores were spread across MM agar lacking inducer and a disc placed centrally containing either 10µl 0.5mg/ml thiostrepton (in DMSO) or 10µl DMSO alone (Figure 4.8). Thiostrepton, but not DMSO, induced growth in the immediate vicinity of the disc. Towards the edge of the agar plates there was no growth, presumably because the concentration of thiostrepton was too low to induce *ndh*. Interestingly, a ring of actinorhodin was observed around the region of growth, which suggests that low levels of *ndh* (and high $NADH/NAD^+$ ratio) might induce the production of this antibiotic.

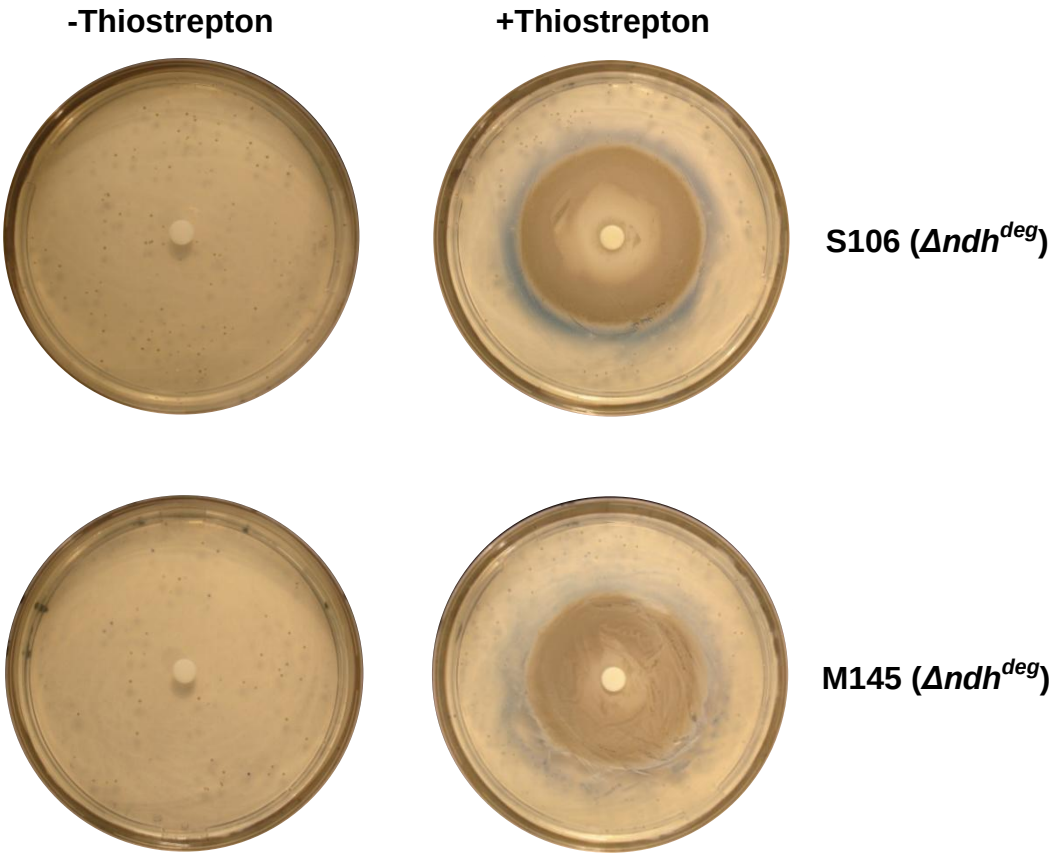


Figure 4.8: Growth of *ndh*^{deg} mutants on MM agar containing 1% glucose, NMMP minor elements and 20µg/ml apramycin. Approximately 10^8 spores, in 100µl sterile water, were spread directly onto the agar. Sterile discs were then placed in the centre of each plate and either 10µl DMSO or 10µl 0.5mg/ml thiostrepton (in DMSO) were pipetted onto the disc. Plates were then incubated at 30ºC for 3 days prior to ammonium fuming and photographing.

In the absence of thiostrepton, some colonies grew, even in the presence of apramycin, which maintains selection for the single cross-over recombination event. These colonies might be suppressor strains, and if this were the case then the suppressor mutation would either be within the disruption vector or elsewhere on the chromosome. A selection of putative suppressors were isolated and grown on non-selective media, to attain strains that had lost the disruption vector. Reintroduction of the original vector into the resultant strains would therefore indicate whether the suppression had come from the vector or from another locus. Screening for loss of the vector was done by replica plating onto DNA media +/- 50µg/ml apramycin. Unfortunately no apramycin sensitive colonies could be isolated therefore it is still unclear as to why these colonies were able to grow in the absence of thiostrepton. One possibility is that the suppressors had mutations that increased basal apramycin resistance thereby removing the selective pressure for the disruption mutation. Regardless, the data indicate that thiostrepton is required for growth of the *ndh*^{deg} strain, suggesting that *ndh* is essential, at least under the growth conditions tested.

Section 4.3.8 – Depletion of NDH-2 induces expression from *cyd*^{P1}

In order to investigate the importance of *ndh* in maintaining the cellular NADH/NAD$^+$ redox poise, the level of NDH-2 was depleted in the M145 *ndh*^{deg} strain by removal of the inducer thiostrepton. The effect of NDH-2 depletion on NADH/NAD$^+$ redox state was analysed indirectly by monitoring the expression of the *cyd*^{P1} promoter. Spores were harvested from MS agar containing 20µg/ml apramycin and 12.5µg/ml thiostrepton. Initial experiments using spores to inoculate NMMP liquid media containing apramycin and thiostrepton were unsuccessful because the cultures grew extremely poorly (data not shown). Therefore mycelial preparations were generated in YEME media (Section 2.2.1) and used to inoculate 60ml NMMP, containing 20µg/ml apramycin and 1µg/ml thiostrepton. Note that the level of thiostrepton used was lowered for this section of the depletion assay in order to limit the amount of carry-over into the depleted samples. Once the cultures had reached an OD_{450nm} of ~0.8-1 a 15ml aliquot was taken and treated with Kirby and phenol/chloroform (as described in

Section 2.3.3). This sample served as the 0 minute time-point. The remaining culture was pelleted at 4,000rpm for 1 minute, washed twice with 50ml of pre-warmed NMMP and finally resuspended in 5ml of the pre-warmed media. This was split evenly between two flasks of pre-warmed NMMP, one containing 20µg/ml apramycin alone and one containing 20µg/ml apramycin and 10µg/ml thiostrepton. The cultures were returned to the incubator, harvesting 15ml samples at 2, 4 and 6 hour intervals. The RNA was purified as detailed in Section 2.3.3 and analysed by S1 nuclease protection assay (Figure 4.10). The S1 nuclease mapping probe was for the *cydA* gene, which has two promoters (indicated by closed and open arrows in Figure 4.10).

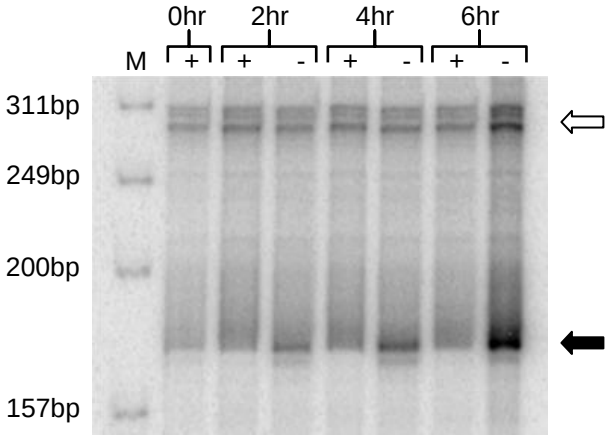


Figure 4.9: S1 Nuclease mapping of the *cyd* promoter region for a thiostrepton depletion study in an *ndh* disruption strain. Induction of the *ndh* gene was achieved by the addition of 10µg/ml thiostrepton (+), in the absence of which only a degradation-targeted version of the gene was present (-). The initial culture was grown in the presence of thiostrepton (0hr), the cell pellet was washed repeatedly and used to inoculate two fresh cultures containing 20µg/ml apramycin and 10µg/ml thiostrepton, or apramycin alone. The cultures were grown for a further 6 hours harvesting at 2 hour intervals. The open arrow indicates expression from the cyd^{P2} promoter, whereas the closed arrow is for the cyd^{P1} (Rex regulated) promoter.

Expression from the P2, Rex independent, promoter remained fairly constant across the time-course as expected. Expression from the cyd^{P1} promoter at the 0 hour time point was quite low and remained at a low level for all of the thiostrepton-replete samples. The thiostrepton-deplete samples however produced a much stronger signal from the cyd^{P1} promoter, which increased over time. As the only difference between these two cultures was the presence or

absence of the inducer, thiostrepton, the observed effect must be due to changes in the expression of *ndh*. This indicates that depletion of NDH-2 in *S. coelicolor* disrupts the ability of Rex to repress its target genes, presumably by increasing the NADH/NAD^+ redox state.

Section 4.4 – Regulation and expression of *nuo*

Nuo is the bacterial equivalent of Complex I in mitochondria, having the ability re-oxidise NADH by passing electrons to the quinone pool and to pump protons across the cell membrane, generating a proton motive force. The *nuo* ROP site was highly enriched in the ChIP-chip data and has been used extensively for binding studies. However the regulation of *nuoA-N* by Rex has not yet been proven. This section aims to characterise the expression profile of *nuo*, as well as its regulation, in order to understand the potential role that it has within the regulon.

Section 4.4.1 - Generating a *nuo*^{deg} mutant

In an attempt to disrupt *nuoA*, the pSX406 suicide vector (see Section 4.3.6) was adapted to control expression of the *nuo* operon in the same manner as used for *ndh*.

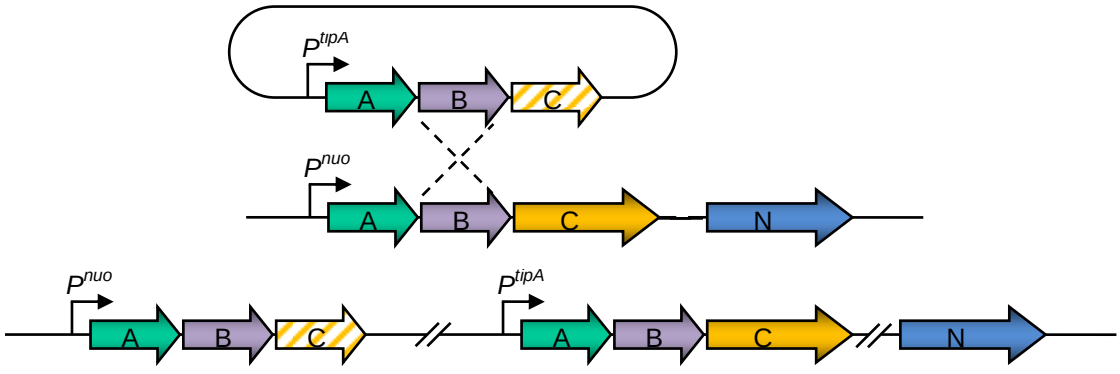


Figure 4.10: Diagrammatic representation of the disruption scheme used to generate the inducible *nuo* knockout *S. coelicolor* strains. In brief a suicide vector, containing part of the *nuo* operon, preceded by the *tipA* promoter, was introduced into the M145 and S106 *S. coelicolor* strains. This placed expression of *nuo* under the control of the *tipA* promoter, which is thiostrepton inducible. Note that the *nuoC* gene on the suicide vector (shaded with stripes) is truncated and contains a degradation signal at its c-terminus. The two preceding genes on this vector, *nuoA* and *nuoB*, are intact.

In order to ensure single cross-over recombinants, at least 1kb of homology was required. However *nuoA* is only 359 bp in length. The region used to generate the disruption cassette was therefore extended to include *nuoA*, *nuoB* and 625bp of *nuoC*. This method had the drawback of introducing two functional copies of both *nuoA* and *nuoB*, which could potentially affect complex assembly (Figure 4.10). The *nuoA*, *nuoB* and *nuoC* segment was amplified using primers NuoKO_F and NuoKO_R. The resultant 1.5kb fragment was ligated into *Eco*RV-cut pBlueScript II SK+ to generate pSX413. This region was isolated as an *Nde*I-*Sph*I fragment and introduced into the same sites of pSX404, linking the degradation tag to the *nuo* fragment. The entire region was isolated with *Nde*I/*Eco*RI and ligated into pSX406 cut with the same enzymes. This vector was then introduced into the M145 and S106 (*Δrex*) strains. The initial exconjugants were noticeably impaired in growth compared to the vector-only control strains (strains containing pIJ6902). At this stage it was not clear whether the delay was due to the presence of the thiostrepton in the overlay, and therefore the induction of the *nuo* operon. The strains were restreaked onto MS agar in the presence and absence of 10µg/ml thiostrepton (Figure 4.11). Unlike the *ndh*^{deg} strains no discernable effect of thiostrepton depletion was observed for the *nuo*^{deg} strains. Both disruption strains however still exhibited a small colony phenotype that had been observed in the initial exconjugants. It would appear that this disruption method had a detrimental effect on the growth of these strains and the process was therefore abandoned.

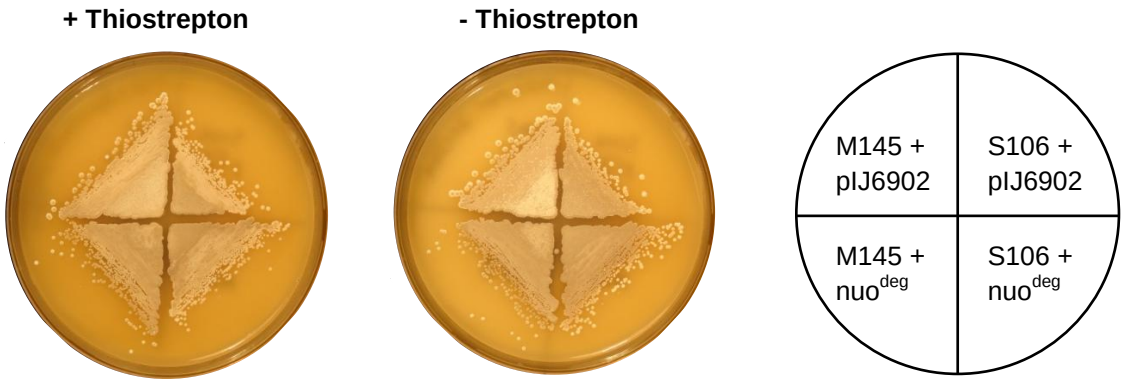


Figure 4.11: Growth of Nuo^{deg} mutants on both MS agar in presence and absence of 10µg/ml thiostrepton. All plates also contained 20µg/ml apramycin. The plates were incubated at 30°C for 4 days.

Section 4.4.2 – The expression of *nuo* under different conditions

From the ChIP-on-chip data it was clear that Rex was able to bind to the *nuoA* promoter region *in vivo*, which confirmed the results from previous binding studies (Brekasis, 2005). However, the regulation of *nuoA-N* has not been confirmed; indeed previous attempts to study *nuoA-N* expression revealed that *nuoA* was not expressed in liquid cultures, which would explain the lack of induction in the S106 (*Δrex*) strain during previous microarray work (D. Brekasis and M. Paget, personal communication). One possible explanation for the lack of *nuoA* expression in liquid media is that it is developmentally regulated. The expression of *nuoA* was therefore investigated on solid agar plates. RNA was harvested (as detailed in section 2.3.3) over a 72 hour time-course, at 24, 36, 48 and 72 hours. Samples were then analysed by S1 nuclease mapping with a *nuoA* probe (from primers NuoAS1a/NuoAS1b (Brekasis, 2005)). The assay gave a strong band at ~160bp (Figure 4.13), with its strength increasing only slightly over the three day time-course. No observable difference between the M145 and S106 (*Δrex*) strains could be distinguished. Thus whilst this experiment confirmed that *nuo* was indeed expressed on solid media, it did not appear to be regulated by Rex. The ROP site within the *nuo* promoter is centred ~20 bp downstream from the predicted transcriptional start site (Figure 4.14). This start site is only approximate, based upon the size of the S1 band and position of the reverse S1 primer but nevertheless the locality of Rex within the transcript would be consistent with its role as a repressor of the *nuo* operon.

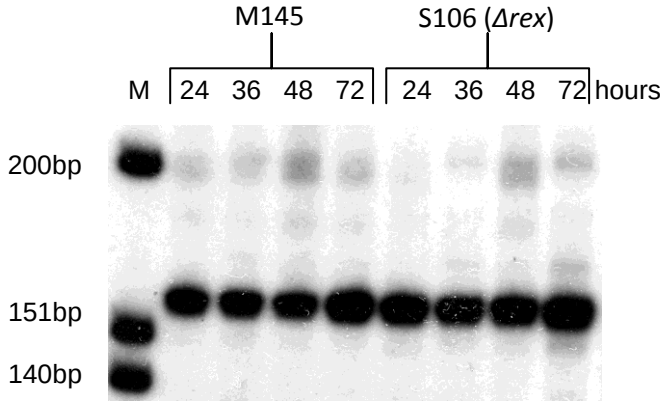


Figure 4.12: S1 nuclease protection assay on the *nuoA* promoter region. RNA was harvested over a 72 hour time-course from *S. coelicolor* cultures grown on MYMTE media.

Section 4.4.3 – Is *nuoA-N* regulated by BldD?

When a gene is restricted to growth under certain conditions, as is the case for *nuo*, it suggests that a regulator is involved. After scanning the promoter region for a possible effector site a potential *bldD* binding site was identified (Figure 4.13).

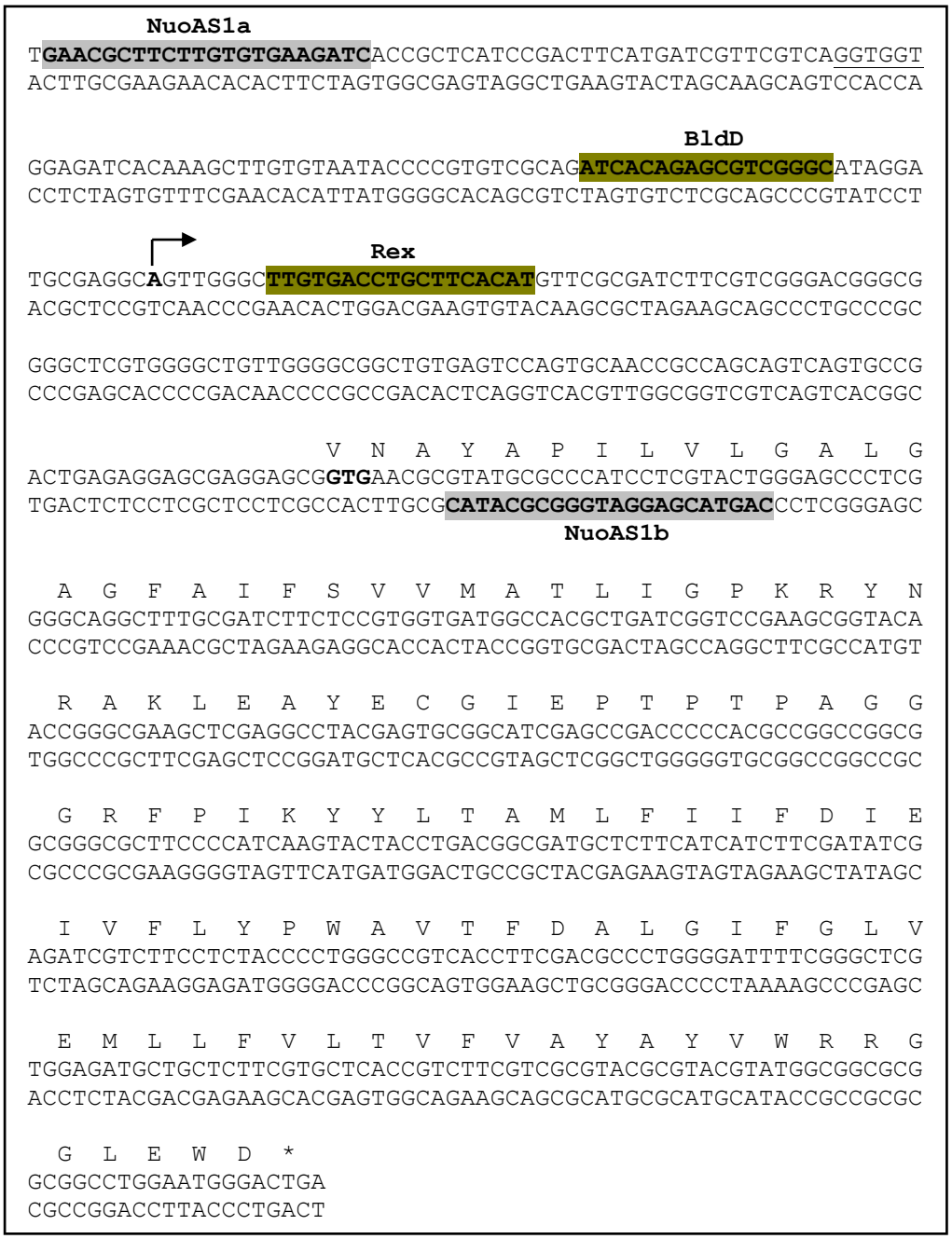


Figure 4.13: The *nuoA* open reading frame and promoter region. The positions of the Rex and BldD binding sites are highlighted in gold. The positions of the two primers used to generate the probe for the S1 nuclease protection assay are highlighted in grey. The approximate position of the transcriptional start site, as assessed by the size of the S1 band (~160bp), is also indicated (by the arrow).

ChIP-on-chip data from another group (C. den-Hengst & M. Buttner, personal communication) supported this. BldD is generally thought to act as a repressor and so it was possible that *bldD* might repress *nuoA-N* during liquid growth. To test this a *bldD* mutant strain 1169 (Merrick, 1976) was utilised in order to bypass possible BldD repression so that the regulation by Rex could be observed. RNA was harvested from liquid culture over a 30 minute oxygen limitation time course and used for an S1 nuclease protection assay, again using the *nuoA* probe. The *cyd* S1 probe was used as the control for the RNA sample (Figure 4.14). The results indicated that whilst the *bldD* mutant responded normally to redox stress, with respect to the induction of *cyd*, *nuo* was still not actively expressed in liquid culture. Rex and BldD may well co-regulate *nuoA-N in vivo,* but apparently they are not the only regulators to do so.

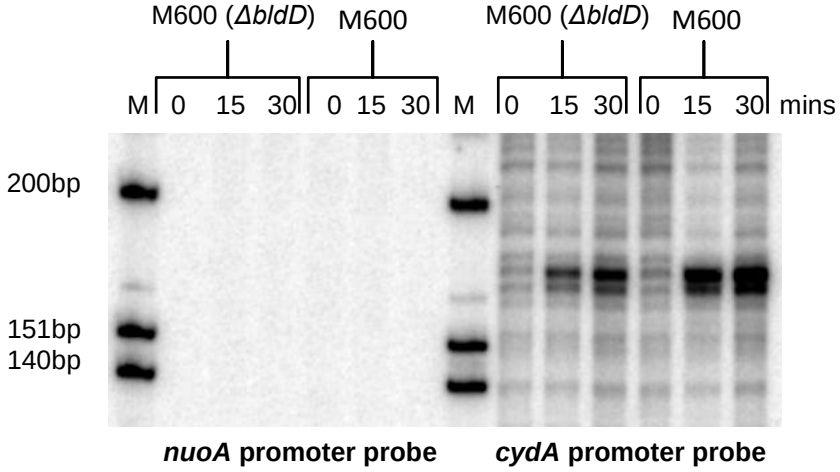


Figure 4.14: S1 nuclease protection assay on the *nuoA* and *cydA* promoter regions. RNA was harvested from *S. coelicolor* cultures grown in NMMP media. The 0 minute time point was harvested under aerobic conditions, whereas the 15 and 30 minute time points were oxygen limited prior to harvesting.

Section 4.4.4 – Reduction in *nuo* expression in the *rex*^{G102A} strain

Finally, the super-repressor mutant, Rex^{G102A} was used in an attempt to confirm regulation of *nuo* by Rex. RNA was harvested from solid media, over a 72 hour time-course, and used for an S1 nuclease protection assay (Figure 4.15). There appeared to be a slight difference in the overall expression levels of *nuoA* in S106 (pSET152) and S106 (pSET152::Rex^{G102A}). The signal for the super-repressor strain was slightly lower than that for the *rex* null strain. However, *nuo*

was still expressed in both strains. This result would suggest that Rex is able to repress expression of *nuo* though the extent to which it does this is not clear given that only the NADH unresponsive *rex*^{G102A} strain showed any discernible difference in expression of *nuo*.

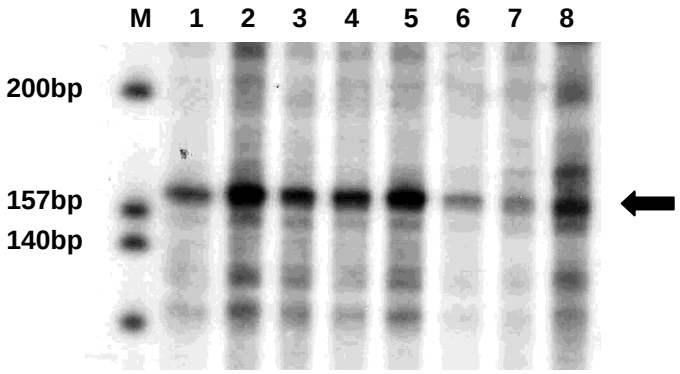


Figure 4.15: S1 Nuclease protection assay of the *nuoA* promoter region in S106+pSET152 (lanes 1-4) and S106+pSET152::rex^{G102A} (lanes 5-8) strains. The cultures were harvested from MYMTE agar over a 72 hour time-course, at 24 (lanes 1+5), 36 (lanes 2+6), 48 (lanes 3+7) and 72 hours (lanes 4+8). The arrow indicates the position of the *nuoA* band.

Section 4.5 – Discussion

The two types of NADH dehydrogenase represented within the ChIP-chip data; *nuo* and *ndh*, are expressed under different conditions. Expression of *nuo* is constitutive on solid media but completely lacking in liquid cultures, whereas *ndh* is apparently essential for growth under all conditions tested. The *ndh* promoter region contains two ROP sites, positioned at -43 bp (ROP1) and -315 bp (ROP2), both of which are capable of producing gel shifts in EMSA analysis. Despite the ROP1 site being positioned for maximal effect on transcription, Rex appears to bind with higher affinity to the ROP2 site. From studies of this promoter *in vivo*, through use of a *neo* reporter construct, it appears that deletion of the ROP2 site actually has a detrimental effect on expression. This result was counterintuitive as one would expect that removal of a repressor binding site would enhance expression of a gene. Thus it seems that the ROP2 site may be recognised by another transcriptional regulator, as well as by Rex. This may explain the presence of a ROP site so far from the transcriptional start site of the gene, if Rex actually functions to block the binding of a transcriptional activator (or chromatin modifier) rather than directly blocking the RNA polymerase. In fact the alignment of the *ndh* promoter regions

of the streptomycetes (Figure 4.4) reveals a dyad symmetrical site (AACTTCGTGAAGTT) overlapping the upstream ROP site, perhaps an operator for an unknown regulator. The *nuo* promoter region is apparently also regulated by more than one factor, having binding sites for both Rex and BldD. The expression profile of *nuo*, being limited to cultures grown on solid media, suggested that it may under the control of a developmental regulator. Indeed, the developmental regulator BldD binds in the *nuoA* promoter region. However, the *bldD* mutant strain 1169 still failed to express *nuo* in liquid cultures. Having failed to observe any obvious differences in the expression of *nuo* in all M145 vs S106 (*Δrex*) cultures it had not been possible to confirm that Rex did indeed regulate *nuo*. In a final attempt to show this the S106 (Δrex^{G102A}) super-repressor strain was utilised. The results suggested a slight reduction in signal for the super-repressor strain, although *nuo* was still expressed. This result did however indicate that Rex is capable of effecting the expression of *nuo* but that the conditions under which Rex represses this promoter are still unknown.

The Type II NADH dehydrogenase of *S. coelicolor*, expressed from *ndh*, is essential. In *M. tuberculosis ndh* is apparently also essential as it is used as a target for a class of anti-tubercular drugs (phenothiazines), utilising the fact that human mitochondria use only NDH-1 and are therefore unhindered by NDH-2 inhibitors (Weinstein *et al.*, 2005). Disruption of this gene in *B. subtilis* caused a marked growth delay but the *ndh* mutant was still viable (Gyan et al., 2006). In *E. coli ndh* disruption is also possible and does not prevent the cells from oxidising NADH (Calhoun and Gennis, 1993). Like *S. coelicolor*, *E. coli* also contains both Type I and Type II NADH dehydrogenase, as well as both cytochrome *bo* and *bd* terminal oxidases. The work by Calhoun *et al.* demonstrated that NADH oxidation is not limited to just NDH-1 or NDH-2 in *E. coli*, but is instead split between the two enzymes (Calhoun *et al.*, 1993). This does not appear to be the case in *S. coelicolor* as *nuo* expression appears to be completely absent in liquid cultures, perhaps causing the dependence upon *ndh* expression. The results of their work also suggest that in *E. coli* the combination of NDH-2 and cytochrome *bd* terminal oxidase, as would be expressed in *S. coelicolor* during oxygen limitation, would result in a marked reduction in the

proton motive force as NDH-2 is incapable of proton translocation and the *bd*-type oxidase is less efficient at it than its *bo*-type counterpart (Calhoun *et al*., 1993). Why then does *S. coelicolor* not induce *nuo* under these conditions to promote generation of a proton motive force? This would suggest that failure to generate a proton motive force was a secondary issue under redox stress, with the primary goal of stress responses being the oxidation of NADH. Studying the enzymatic function of respiratory NADH dehydrogenases is difficult due to both their membrane association and also due to the presence of endogenous NDH enzymes in the expression strain. There were however two such studies in the literature on the NADH dehydrogenases of *E. coli*, in which the rate of NADH oxidation had been determined. The values were given as 25.1 nmol min^{-1} mg^{-1} for NDH-1 (Esterhazy *et al.*, 2008) and 500-600 µmol min^{-1} mg^{-1} for NDH-2 (Jaworowski *et al.*, 1981). This would indicate that whilst NDH-1 is more effective at generating energy, via a proton motive force, NDH-2 is more efficient at oxidising NADH.

Previous binding studies with the *nuoA-N* promoter had revealed a strong ROP site, which was confirmed by the ChIP-chip data. Despite the presence of this site within the *nuoA-N* promoter the regulation of this operon by Rex has not yet been proven. In this study a number of approaches have been taken to study the possible regulation of *nuo*, as well as attempts to isolate and characterise disruption strains in order to identify the role that its product might have in responding to redox stress. A disruption cassette was used to place the operon under the control of a thiostrepton inducible promoter. Unfortunately the targeting cassette itself appeared to have a detrimental effect on growth, as strains were affected in both the presence and absence of thiostrepton. The genes included on the cassette were *nuoA*, *B* and *C*, which have varying purposes and positions within NDH-1. The *nuoA* and *nuoB* genes were intact and therefore duplicated within the genome of the disruption strains. *nuoA* expression appears to be constitutive on solid media (Section 4.4.2). It is possible that constitutive expression of *nuoA* and *nuoB*, from the native *nuo* promoter, results in protein aggregation or toxicity when expressed independently from the rest of the complex. NuoA is thought to form part of the

transmembrane arm of this protein, whereas NuoB and C are located within the enzymatic arm in the cytosol (Efremov et al., 2010). It is however also possible that the disruption cassette hinders expression even in the presence of the inducer. Efforts to confirm that Rex controls *nuo* were hampered by the absence of *nuoA* expression in liquid cultures, the conditions in which Rex regulation is normally investigated. In other species there appears to be a requirement for NDH-1 under anaerobic growth conditions (Tran et al., 1997). It is possible that whilst the aerial hyphae are exposed to atmospheric oxygen concentrations the substrate mycelia, buried within the media and the rest of the culture, are oxygen-limited. If this were the case it would explain why Rex failed to repress expression from the *nuoA* promoter under these conditions. Despite a lack of differential expression of *nuo* in the M145 versus S106 (*Δrex*) strain, the use of the S106 (*Δrex*^{G102A}) strain has again proven useful in emphasizing the repression of Rex. Harvesting of RNA from this strain showed a slight decrease in the expression of *nuo* compared to the S106 (*Δrex*) strain; however the promoter could not be fully silenced by the super-repressor in this study. The super-repressor strain is also under the control of the *rex* promoter, making it autoregulated. This may have limited the availability of Rex^{G102A} *in vivo*, as emphasized by the lack of lethality that is associated with high levels of this protein (Section 3.4.1). It is therefore likely that the low expression levels of *rex*^{G102A} were insufficient to fully silence the expression of *nuo*.

Gyan *et al.* have recently proposed that there is a regulatory feedback loop between the Type 2 NADH dehydrogenase of *B. subtilis* and B-Rex (Gyan *et al.*, 2006). This relationship works because each of its partners is regulated by the presence or absence of NADH. Under aerobic conditions Rex is able to bind to and repress *ndh* but when NADH levels rise, under oxygen limitation, Rex dissociates from the *ndh* promoter. Expression of *ndh* results in a reduction of cellular NADH, lowering the NADH/NAD$^+$ ratio so that Rex can repress expression of *ndh* once more. This regulatory loop means that *ndh* is indirectly sensitive to the redox poise, tailoring *ndh* expression to the availability of the enzyme's substrate. This function is key in maintaining the redox poise of the cell.

Chapter 5

Results III: The Mechanism of Action of Rex

“Thunder is impressive, but it is lightning that does the work.”

Mark Twain (1835-1910)

Section 5.1 – Overview

The action of Rex is best observed through its effect on gene expression, though it is the interplay between the two dinucleotides, NAD^+ and NADH, that modulates its DNA binding activity (Brekasis and Paget, 2003). The physical mechanism by which NADH alters the DNA binding potential of Rex is not fully understood, but can be characterised through a combination of structural analysis, mutagenesis, and biochemical approaches. This chapter will describe a structure-based site-directed mutagenesis analysis of *Streptomyces* Rex. The overall aim of these studies was to understand how the binding of NADH in one part of the protein can ultimately affect the DNA binding potential in a separate domain.

Section 5.2 – Structural overview

Early site-directed mutagenesis studies on Rex relied on sequence-structure predications. For example, the GXGXXG motif, which is characteristic of Rossmann folds, was shown to be essential for NADH binding (Brekasis and Paget, 2003). More recent structural studies on Rex homologues from *Thermus aquaticus* (Sickmier *et al.*, 2005), *Thermus thermophilus* (Nakamura *et al.*, 2007), and *Bacillus subtilus* (Wang *et al.*, 2008), now allow structure-based approaches. At the outset of this project, the only available structure was that of *Thermus aquaticus* Rex (Uniprot accession: Q9X2V5.1), in its NADH-bound state (PDB code: 1XCB; Sickmier *et al.*, 2005). During the course of the project further *T. aquaticus* structures became available including one in the DNA-bound NAD^+-bound state (PDB code: 3IKT) (McLaughlin *et al.*, 2010). As there is no structure available for *S. coelicolor* Rex (S-Rex) all structural analyses in this chapter was based on T-Rex (42% identical), while all mutagenesis was performed on the gene encoding S-Rex protein. Note that because of this the residue numbering can be for either protein but wherever a residue is mentioned the parental protein will be indicated. For ease of comparison a fully numbered alignment of the two proteins is included (Figure

S3), which is colour coded to indicate the similarity of the residues at each position.

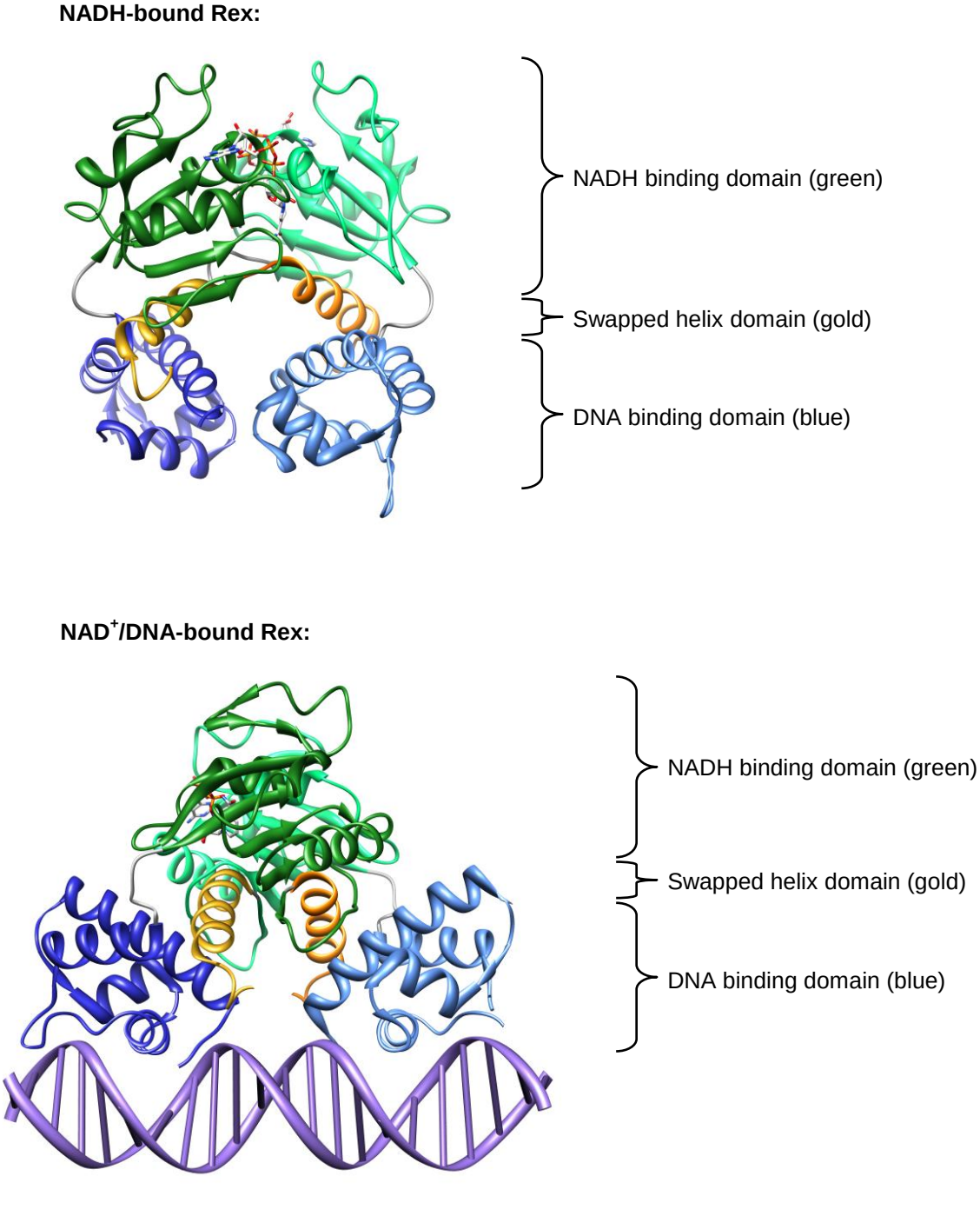


Figure 5.1: Structural overview of NADH- and DNA-bound T-Rex, PDB codes 1XCB and 3IKT, respectively. Chains A and B are coloured by domain, with equivalent domains in each chain shown as different shades of the same colour. The DNA is shown in the lower panel in purple. The NADH or NAD^+ molecules are depicted in stick form on each structure.

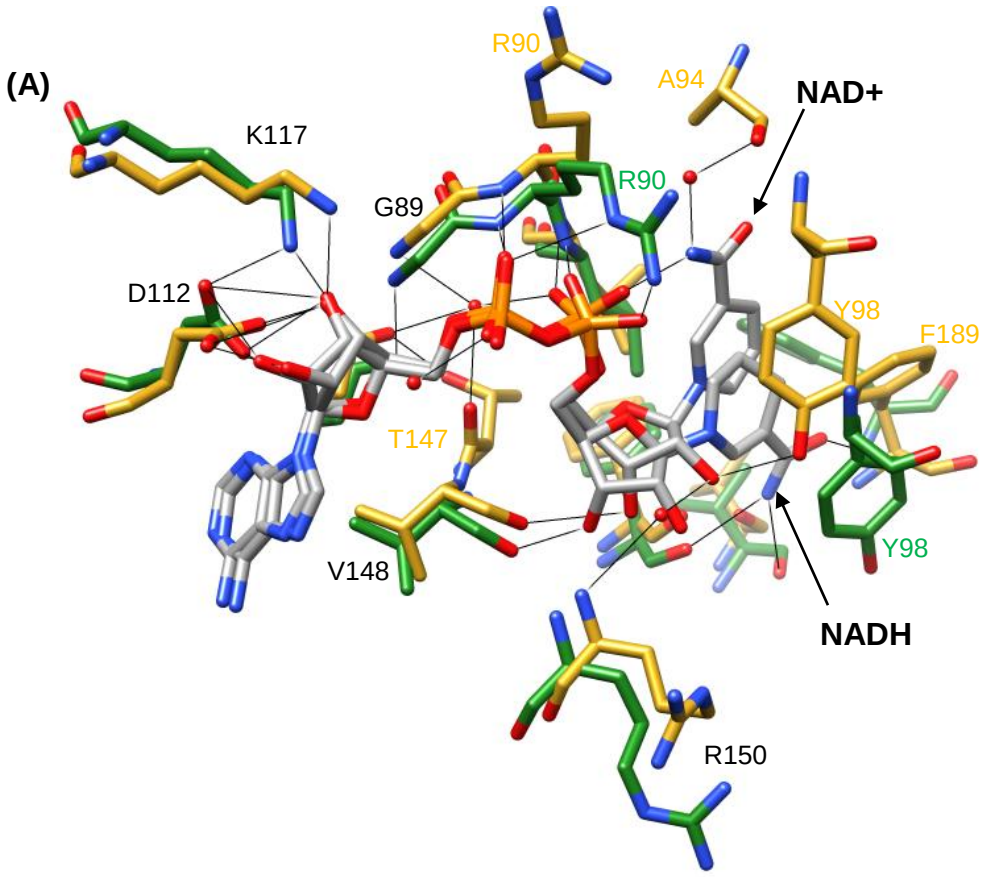
Section 5.2.1 – Structural analysis

All known Rex homologues are dimeric, and have three well-defined domains: the DNA-binding domain, the NADH-binding domain and the swapped helix domain (Figure 5.1). The NADH-binding domain contains the classical Rossmann fold (Rao and Rossmann, 1973), with each domain capable of accommodating one NADH molecule giving a final stoichiometry of 2 NADH molecules per dimer (Sickmier et al., 2005). From biochemical studies we were already aware that NAD^+ must also be capable of binding to Rex in order to act as a competitive inhibitor, a vital role as it allows Rex to sense the ratio of the reduced and oxidised dinucleotides – not just the NADH concentration (Brekasis and Paget, 2003). The structure of T-Rex bound to DNA and NAD^+ confirmed this as it clearly showed a single NAD^+ molecule occupying one of the binding domains (McLaughlin *et al.*, 2010). Comparison of the bound NADH and NAD^+ however revealed that NAD^+ adopts a slightly different conformation within the DNA-bound structure (Figure 5.2). Whilst the adenosyl moiety of NAD^+ and NADH appear to share many of the same contacts within Rex, the nicotinamide ring is flipped back upon itself in the case of NAD^+ when compared to NADH (Figure 5.3). Within the NADH-bound structure there is also a major asymmetry in the orientations of the two F189 residues (Sickmier *et al.*, 2005). This is lacking from the DNA-bound structure, with the F189 adopting a conformation preferable to DNA-binding (McLaughlin *et al.*, 2010). With the F189 residue no longer positioning itself between the two nicotinamide rings the charge on the NAD^+ would not be shielded from the presence of a second, which would result in electrostatic repulsion (McLaughlin *et al.*, 2010). Also with the F189 adopting a different position in the DNA-bound structure there would be a steric clash with the NAD^+ and this residue, if it did not adopt a different conformation than that of NADH (McLaughlin *et al.*, 2010).

The difference between cofactors is important but the most striking difference between the two crystal structures is by far the degree of relative rotation of the subunits. As had previously been reported, the NADH-bound structure was incompatible with DNA binding (Sickmier et al., 2005). In the T-Rex NAD^+-bound structure this steric clash is overcome by a 43º rotation around its axis,

repositioning the recognition helix to interact with the major groove of the DNA (McLaughlin et al., 2010). This is a massive change in conformation triggered by the loss of NADH, a change that NAD^+ fails to induce. The DNA-binding domain itself is a classical winged-helix-turn-helix motif, with the recognition helix positioning itself within the major groove and the wing falling upon the minor groove. Individual interactions will be discussed further on in this chapter.

The final element of Rex is the domain-swapped helix. The role of this domain was the most uncertain as it consisted of a single helix. The positioning of the helix in T-Rex would suggest that it has a role in dimerisation and signal transduction, as it packs between the two functional domains of the opposing chain. However, in *Bacillus subtilis* Rex (B-Rex) this domain does not take up the same conformation (Wang *et al.*, 2008). It still packs against the NADH-binding domain of the opposing chain but the DNA-binding domain appears to be completely isolated (Wang *et al.*, 2008). They do however note that the C-terminal domain of this structure is locked in place by the occurrence of crystal contacts at the N-terminus. This places the swapped helix in a position that is incompatible with NADH occupation of the full dinucleotide binding pocket (Wang *et al.*, 2008).



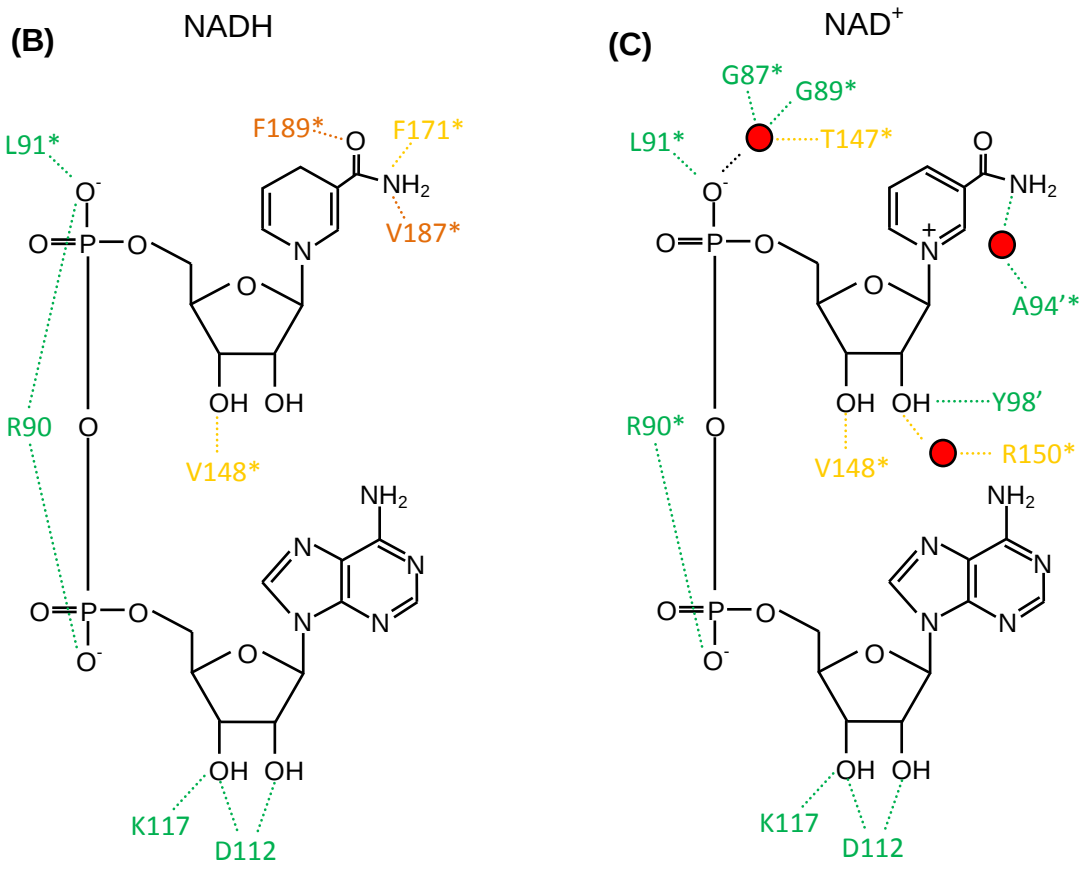
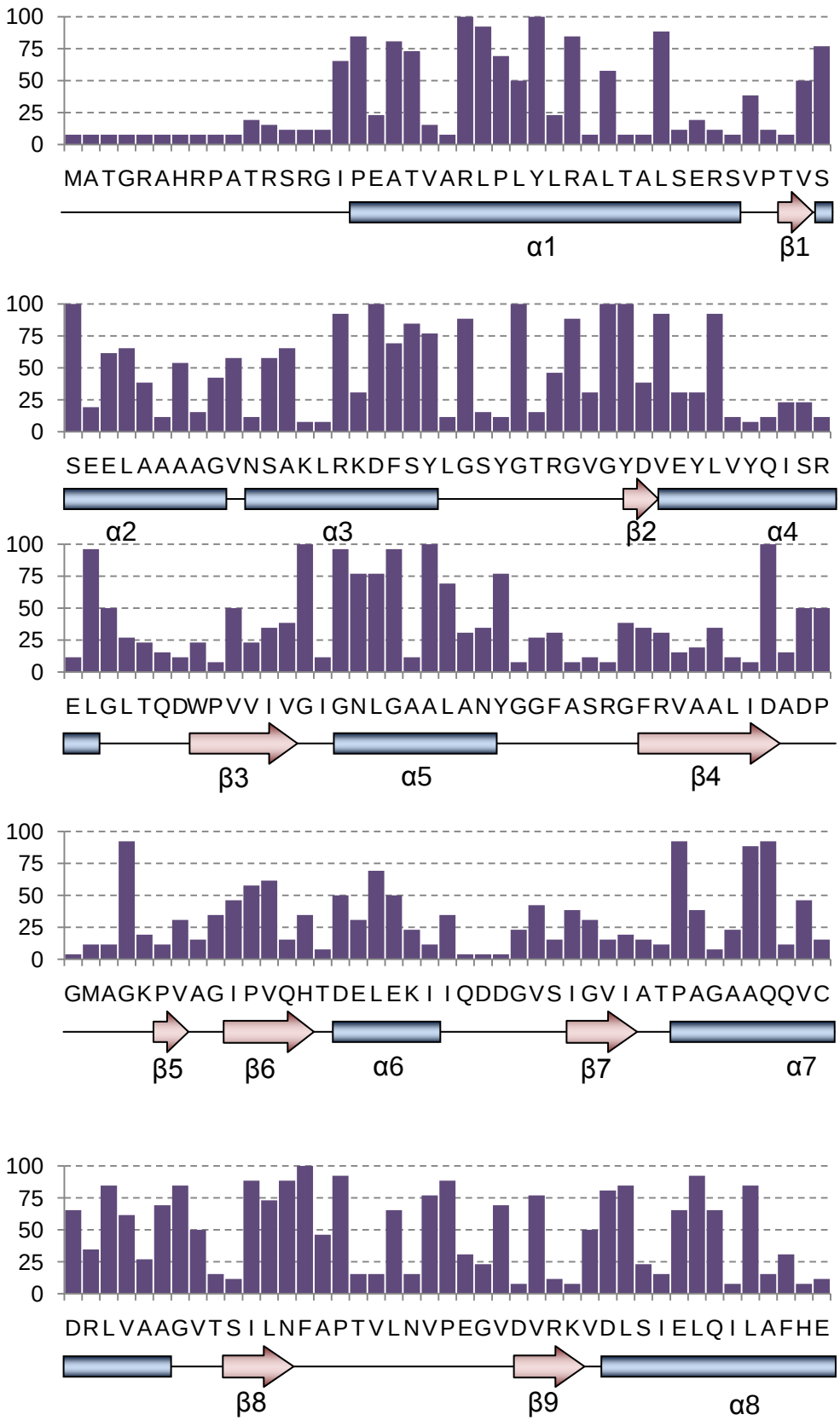


Figure 5.2: Structure of the NADH and NAD^+ binding domains of T-Rex overlayed **(A)**. The interactions are also shown in diagrammatic form for both NADH **(B)** and NAD^+ **(C)**. The dotted lines indicate potential hydrogen bonds on the diagrams, an asterisk indicates that this bond is between the backbone of the residue, and the apostrophe distinguishes between the two chains. Where the bond is relayed via a water molecule the solvent is shown as a red sphere.

Section 5.2.2 – Conservation

In order to identify potentially key residues in the structure and function of Rex, following a multiple alignment, the % conservation at each amino acid position was plotted against amino acid sequence. Of the 258 amino acids of *S. coelicolor* Rex only 11 were 100% conserved among the 26 homologues analysed, which equates to about 5% of Rex (Figure 5.3). Nonetheless, the figure reveals several potentially important residues, and the structures of Rex in the NADH- and DNA-bound states allow hypotheses to be developed. Note that all of the residues discussed below are numbered from the S-Rex protein. The conservation of some amino acids is simple to explain in several cases. For example, R23, S44, G69, G74, and Y75 are 100% conserved and appear to be

involved in DNA binding. R23 and S44 directly interact with the major groove, whereas G69, G74 and Y75 are located at the minor groove in the wing of the recognition helix. The wing potentially co-ordinates the first two nucleotides of each half of the ROP site (TTGTGAA), thus the preference for A and T at these positions (Section 3.3.4) must be caused by the wing. The ROP sequence used in the crystallisation of T-Rex lacked the terminal residue of the S-Rex consensus (CTGTGAA) in each half. In all available T-Rex structures the wing is poorly ordered, which is most likely a consequence of the presence of so many flexible glycines in this region. However, it is possible that inclusion of AT-rich DNA at each end is required to stabilise the wing and hence fix its position within the structure. The GXGXXG of the Rossmann fold is highly conserved although only the first glycine (G100) is 100% conserved. One residue centred at the dimer interface, A107, was completely conserved. This residue appears to mediate a hydrogen bond, via a water molecule, with the amine group of the nicotinamide ring of the NAD^+ molecule. Thus this residue, along with an intramolecular hydrogen bond, appears to stabilise the *syn* orientation of NAD^+. This residue is also responsible for forcing NADH to adopt an *anti*, rather than *syn*, conformation as NADH in the *syn* conformation would clash with the alanine. Thus conservation of this residue appears to be key for maintaining the structural distinction between NAD^+ and NADH. Interestingly, three of the conserved residues appeared to be hydrogen bonding with each other, D61, R23 and Y27, with the aspartate and arginine forming a salt bridge and the tyrosine interacting with the spare oxygen group of the aspartate (Figure 5.4). A similar strategy is used elsewhere in the S-Rex protein between residues (R29, D203 and Y111 – Section 5.4), although these amino acids are not completely conserved. Unlike the R29-D203 salt-bridge, these residues do not alter conformation upon DNA-binding. Based on the relative positions of these residues it would appear that they are key for the interaction between the recognition helix and the stabilisation helix, and therefore the stability of the DNA-binding domain.


100
75
50
25
0
MATGRAHRPATRSRGIPEATVARLPLYLRALTALSERSVPTVS
α1
β1
SEELAAAAGVNSAKLRKDFSYLGSYGTRGVGYDVEYLVYQISR
α2
α3
β2
α4
ELGLTQDWPVVIVGIGNLGAALANYGGFASRGFRVAALIDADP
β3
α5
β4
GMAGKPVAGIPVQHTDELEKIIQDDGVSIGVIATPAGAAQQVC
β5
β6
α6
β7
α7
DRLVAAGVTSILNFAPTVLNVPEGVDVRKVDLSIELQILAFHE
β8
β9
α8

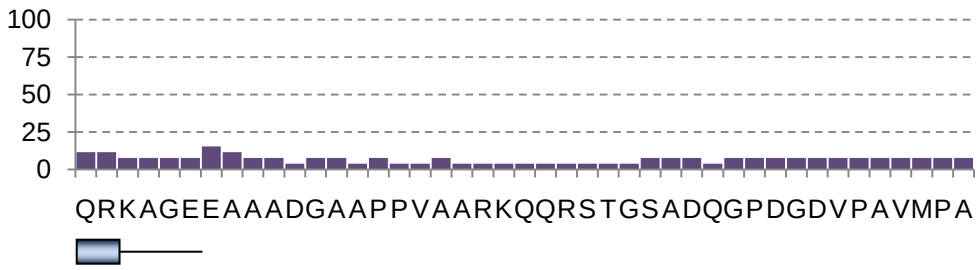


Figure 5.3: Conservation plot of Rex homologues from the following species; *Bacillus anthracis* (YP_081851.1), *Bacillus halodurans* (NP_241417.1), *Bacillus subtilis* (NP_388478.1), *Bacteroides thetaiotaomicron* (NP_812793.1), *Clostridium perfringens* (YP_699563.1), *Clostridium tetani* (NP_782957.1), *Deinococcus radiodurans* (NP_294663.1), *Enterococcus faecalis* (ZP_05597558.1), *Lactococcus lactis* (YP_809112.1), *Lactobacillus plantarum* (NP_784480.1), *Lactobacillus sakei* (YP_394969.1), *Listeria monocytogenes* (ZP_05294788.1), *Oceanobacillus iheyensis* (NP_691573.1), *Porphyromonas gingivalis* (YP_001928136.1), *Rhodopirellula baltica* (NP_864826.1), *Staphylococcus aureus* (ZP_05685717.1), *Streptococcus agalactiae* (ZP_00780168.1), *Streptococcus pneumonia* (NP_358592.1), *Streptococcus pyogenes* (NP_269274.1), *Streptomyces avermitilis* (NP_825915.1), *Streptomyces coelicolor* (NP_627530.1), *Thermoanaerobacter tengcongensis* (NP_622215.1), *Thermotoga maritima* (NP_227984.1) *and Thermus aquaticus* (Q9X2V5.1). Values indicate the % conservation of the S-Rex residue listed at each position in the alignment. The amino acids listed are from the S-Rex sequence and the secondary structural elements are from the T-Rex structure.

Section 5.2.3 – Site-directed mutagenesis scheme

The rational for all mutagenesis in this study followed the same three principles: (i) use the structure of a homologue (T-Rex) to identify regions of interest for any given function, (ii) use the conservation plots of these regions to indicate potential involvement of each residue in the function of this domain, and (iii) use the structure to guide the choice of amino acid substitution and to predict a possible outcome of the mutagenesis. In the case of the DNA binding domain, the rational for targeted disruption was also intimately linked to the conservation of nucleotide residues within the ROP site. By using both pieces of information it was possible to increase the confidence that constructed mutants would have altered DNA binding properties. Please note that in all cases the nucleotides of the ROP site are numbered from 1 to 18 reading 5' to 3' on the coding strand of the gene whose promoter it falls within, and are labelled 1' to 18' reading 5' to 3' on the non-coding strand. Likewise the amino acids on each chain are distinguishable by the presence or absence of an apostrophe after their one-letter code and identifier. Note that in all instances the truncated form of S-Rex (residues 6-233) is used, due to issues with protein cleavage of the full length version (Brekasis, 2005). The functionality of this protein has however been

compared *in vivo* through its ability to complement the S106 (*Δrex*) strain (data not shown).

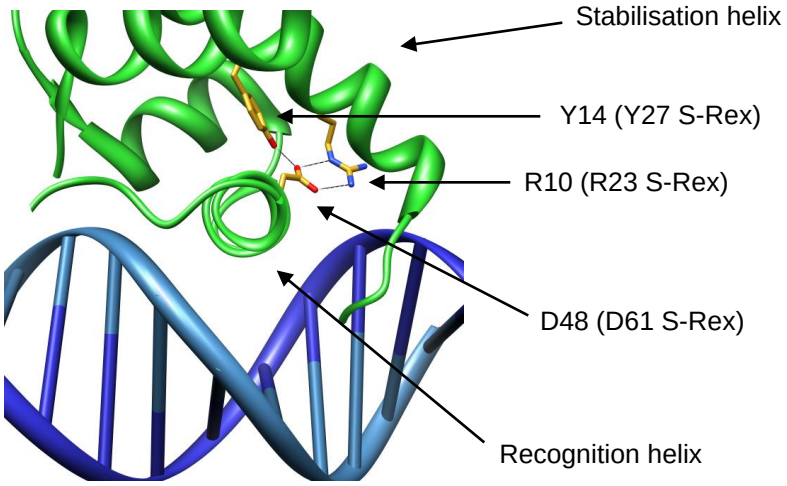


Figure 5.4: Fully conserved salt-bridge stabilises DNA-binding domain of T-Rex. Bridge forming residues R10 and D48, as well as conserved Y14 are indicated above. The recognition and stabilisation helices are also shown.

Section 5.3 – The Rex-ROP complex

From the ChIP-on-chip data it appeared that Rex bound to several loci, but with differing affinities. What distinguishes one ROP site from another and what features of Rex are required for specific binding? This section covers the structural analysis used to identify key Rex-ROP interactions and the subsequent mutagenesis to confirm the importance of selected residues. Note that the wing region within the PDB structure was poorly defined so it was not included in this analysis, despite its high conservation and potential role in DNA binding.

Section 5.3.1 – Interactions with the major groove

As mentioned previously the recognition helix of Rex slots into the major groove of canonical B-form DNA (McLaughlin *et al.*, 2010). In order to identify which residues were required to form this interaction, the structure of DNA-bound T-Rex was analysed (PDB ID: 3IKT) using the protein structure viewer and analysis program UCSF Chimera (Pettersen *et al.*, 2004). The FindHbond tool

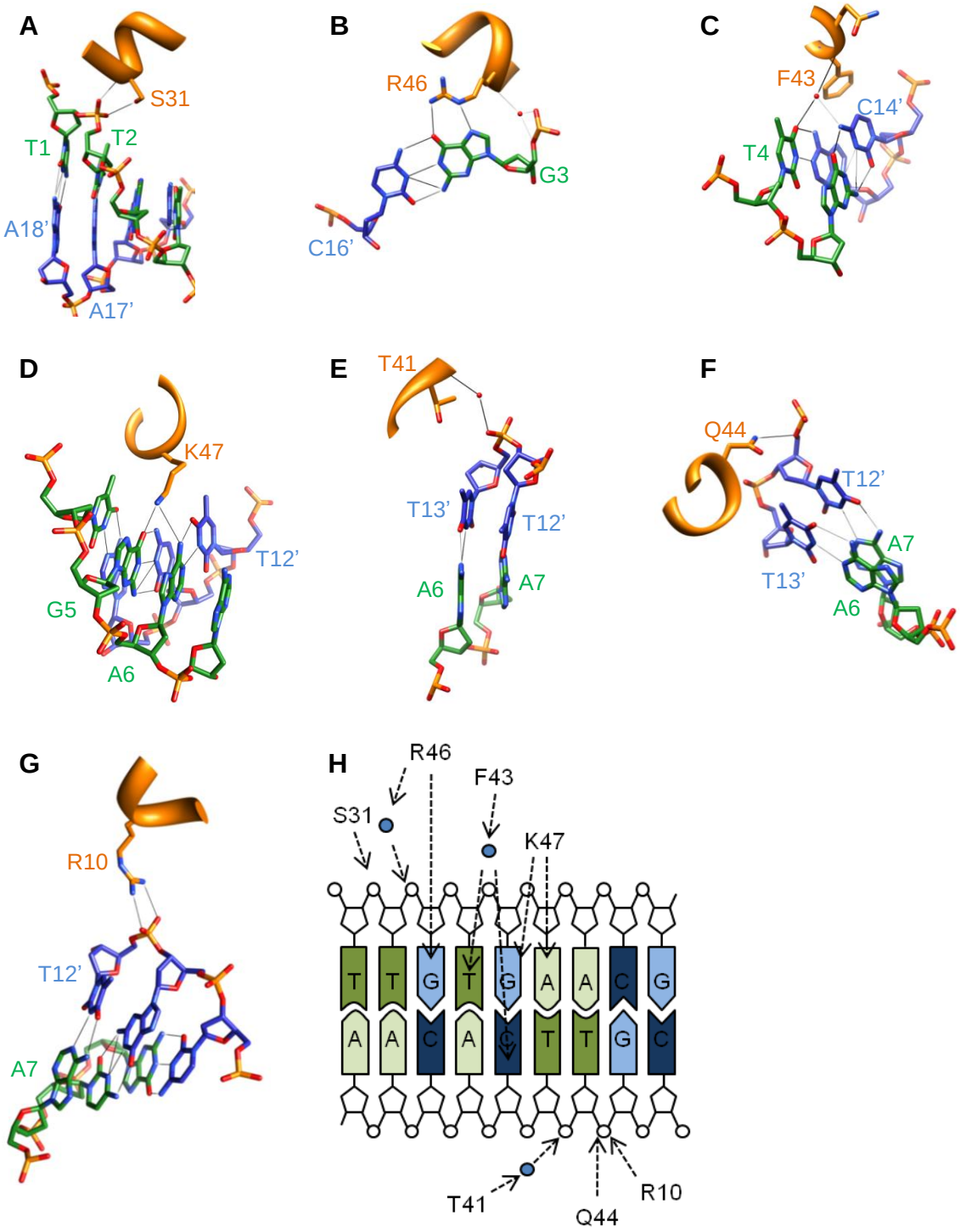


Figure 5.5: Structural overview of the DNA-binding interactions of T-Rex **(A-G)** and diagrammatic summary of these interactions **(H)**. The blue spheres shown in insert H represent water molecules, which appear to bridge a number of the bonds. Note that only one half of the protein is shown in this figure as the interactions are conserved in the opposite half of the structure.

of Chimera was used to predict possible hydrogen bond pairs. These potential H-bonds were then manually filtered for those that fell between the DNA and Rex, and are shown in Figure 5.5. The first bond identified in this way was between S31 of T-Rex (S44 S-Rex), and the phosphate group of Thymine 2 (Figure 5.5A). Guanine 3 appears to be coordinated by R46 (R59 S-Rex), directly via two hydrogen bonds with the side chain, and indirectly via a water molecule (Figure 5.5B). Thymine 4 and Cytosine 14' are each involved in hydrogen bonds between a water molecule and F43 (Figure 5.5C). When F43 was compared between the NADH-bound and DNA-bound structures (Figure 5.6) it appeared that this residue adopted a different conformation on binding to the ROP site. Upon closer inspection it seemed that this may be due to the close proximity of the aromatic ring of F43 to the amine group of Cyt14' when in the NADH-bound conformation. It would seem that when Rex binds to DNA the ring of F43 must adapt its orientation in order to lie parallel with the DNA backbone. Interestingly this residue is replaced by an alanine in S-Rex, which would still allow for the formation of hydrogen bonds with Thy4 and Cyt14', but may change the affinity of Rex for ROP. Lysine 47 (K60 S-Rex) interacts with both Guanine 5 and Adenine 6 of the ROP site (Figure 5.5D) and the phosphate group of Thymine 13' appears to interact with T41 of T-Rex (N54 S-Rex) via a water molecule (Figure 5.5E). Finally the 5' phosphate of Thymine 12' has potential hydrogen bonds with both Q44 (K57 S-Rex) and R10 (R23 S-Rex), Figures 5.5F and G respectively. All of these interactions are summarised in diagrammatic form in Figure 5.5G. In summary, while there are a number of stabilising interactions, the only specific interactions, between amino acid functional groups and nucleotide bases, emanate from Arginine 46 (R59 S-Rex) and Lysine 47 (K60 S-Rex). The arginine residue is very highly conserved among homologues, whereas the lysine residue appears to be replaceable with either an arginine or a glutamine, suggesting that it is the amine group that must be maintained for a functional DNA-binding protein. These residues were therefore targeted for mutagenesis, along with another residue Phenylalanine 43 (detailed above). While this residue was neither highly conserved, nor able to form a specific base interaction it seemed plausible for this residue to be responsible for the marked difference in DNA-binding affinities of T-Rex and S-Rex, Kd 0.1±0.02nM T-Rex, Kd 2.0±0.6nM S-Rex (McLaughlin et al., 2010),

given that it appeared to slot into the major groove within the T-Rex structure but was absent in the S-Rex protein.

Section 5.3.2 – Disrupting Rex's ability to bind to DNA

Site-directed mutagenesis was used to test the importance of three amino acids that were predicted to play key roles in DNA binding: A56, R59 and K60. Each residue was mutated to alanine using the primers A56F_F2 and A56F_R2, R59A_ii_F and R59A_ii_R, and K60A and K60A_R, and vector pSX142 as a template. Each gene was subsequently isolated with *Nde*I/*Bam*HI and cloned into pET15b, cut with the same enzymes, generating pSX143::A56F, pSX143::R59A and pSX143::K60A. These vectors were then used to overexpress each mutant, which were purified via Ni^{2+}-affinity chromatography. Both Rex^{R59A} and Rex^{K60A} failed to give shifts during EMSA analysis (Figure 5.6), indicating that both mutations have a drastic affect on binding – as expected. It had been predicted that an A56F substitution might enhance DNA binding (see Section 5.3.1). However, SPR analysis revealed that Rex^{A56F} had a lower affinity for the DNA than the wild type protein (Figure 5.7). Therefore, this residue is not wholly responsible for the increased DNA-binding affinity of T-Rex compared to S-Rex.

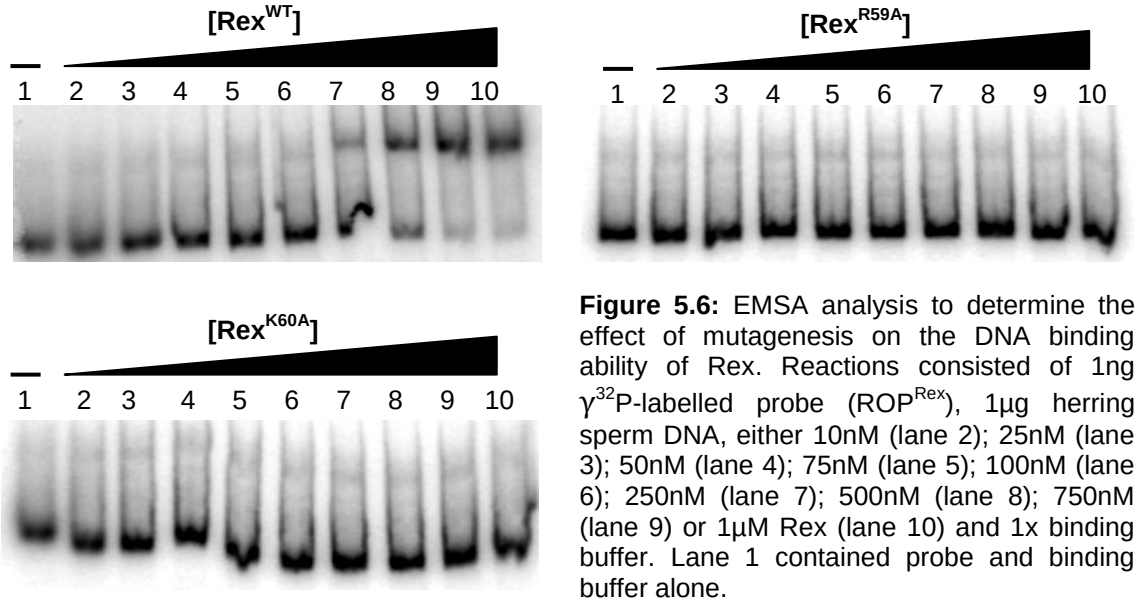


Figure 5.6: EMSA analysis to determine the effect of mutagenesis on the DNA binding ability of Rex. Reactions consisted of 1ng γ^{32}P-labelled probe (ROP^{Rex}), 1µg herring sperm DNA, either 10nM (lane 2); 25nM (lane 3); 50nM (lane 4); 75nM (lane 5); 100nM (lane 6); 250nM (lane 7); 500nM (lane 8); 750nM (lane 9) or 1µM Rex (lane 10) and 1x binding buffer. Lane 1 contained probe and binding buffer alone.

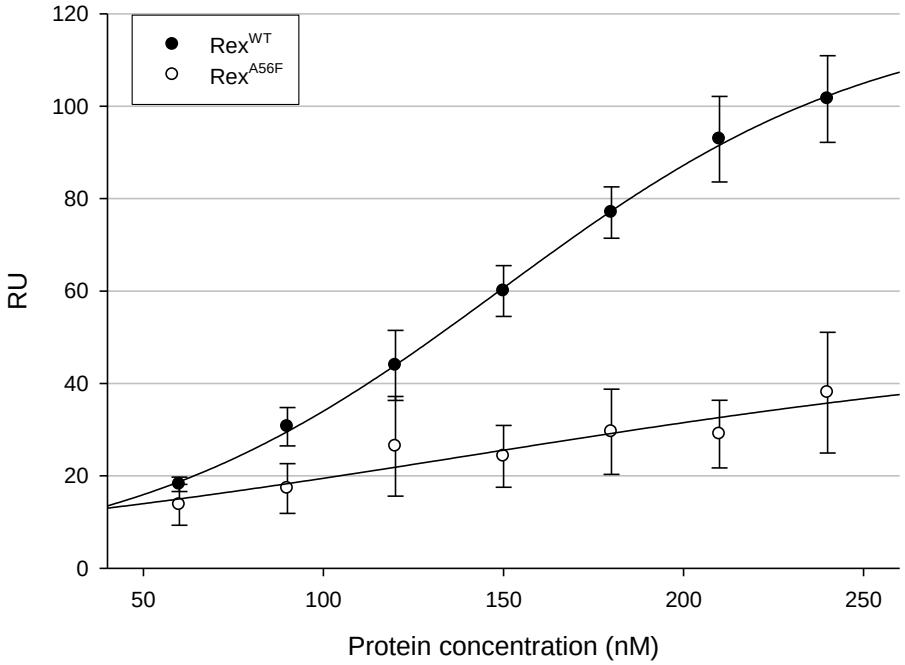


Figure 5.7: SPR analysis of the Rex^{A56F} mutant (open circles), compared to Rex^{WT} (closed circles), over a protein concentration range. All response unit values were background and non-specific binding subtracted.

Section 5.4 – NADH sensing and the signal relay

Having begun this study with the knowledge that an NADH-bound Rex was incompatible with DNA binding we were already aware that a significant structural change must occur upon DNA binding (Sickmier et al., 2005). This section covers the structural analysis and mutagenesis that have helped us to better understand how and why the conformational changes occur.

Section 5.4.1 – The salt-bridge

There are a number of tools available for studying changes in conformation between two structures but sometimes just a simple visual inspection can lead to these discoveries - this was the case for the R29-D203' salt bridge (T-Rex residues R16 and D188'). By applying a surface to the NADH-bound structure in Chimera (Pettersen et al., 2004) it appeared that the space between the DNA-binding domain and the domain swapped helix was extremely solvent accessible (Figure 5.8), with one interaction spanning the void on only one

plane of the protein – a hydrogen bond between R16-D188'. The equivalent residues; R16'-D188 on the other side of the dimer did not interact, thereby generating one example of asymmetry in the NADH-bound structure. The asymmetric unit of the original T-Rex structure contained seven chains, in the form of three dimers and one monomer, all saturated with NADH (Sickmier *et al.*, 2005). In order to ascertain whether this asymmetry was limited to just one of the dimers within the asymmetric unit the others were compared (Figure 5.9).

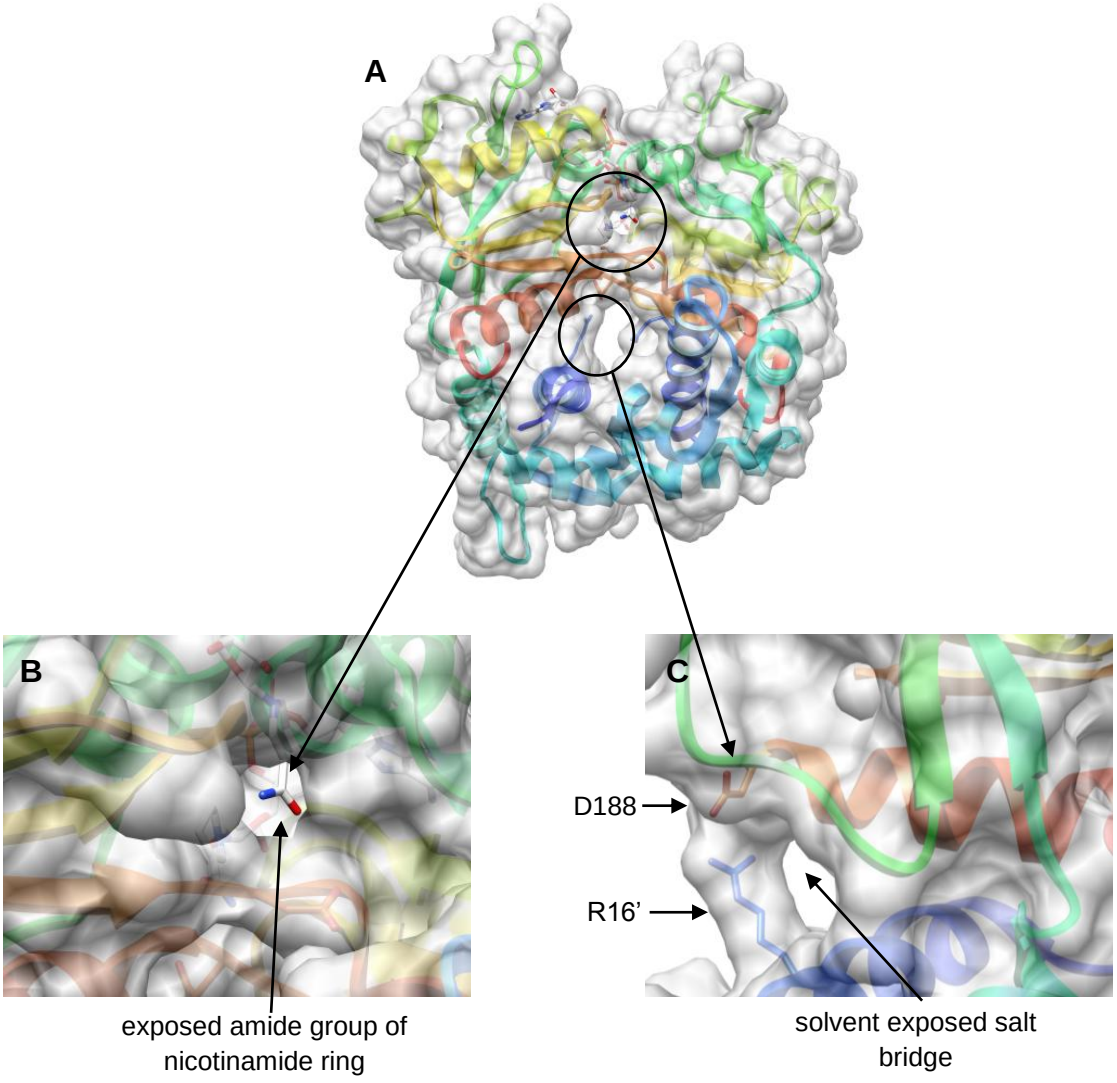


Figure 5.8: Structural overview of NADH-bound T-Rex with surface **(A)**. The solvent accessible amide group of NADH **(B)** and salt bridge **(C)** are shown in greater detail below

This interaction did not appear to be that essential as it was absent in the other two dimers, however there was still a clear difference in the orientations of R16 and R16' in each case. Interestingly this altered conformation corresponded with the F189 asymmetry in each case. Strikingly, in the DNA-bound structure

R16-D188 interactions were present in both sides of the dimer, and replaced the hydrogen bond with a salt bridge. This provides a strong connection between the two subunit chains to potentially stabilise the protein in its DNA-bound form (Figure 5.9). Although this asymmetry was not originally recognised (Sickmier *et al.*, 2005), the pronounced asymmetry of F189 at the dimer interface had been. As mentioned previously F189 has two conformations in the NADH-bound Rex – 'flipped in' and 'flipped out'. This asymmetry is absent in the DNA-bound structure – as is the salt-bridge. With the 'flipped in' F189 always appeared in the same side as the 'broken' salt-bridge (Figure 5.17) it suggests that there might be a connection between the two events. For example NADH binding to one domain might hinder DNA binding in another via the breakage of this salt bridge.

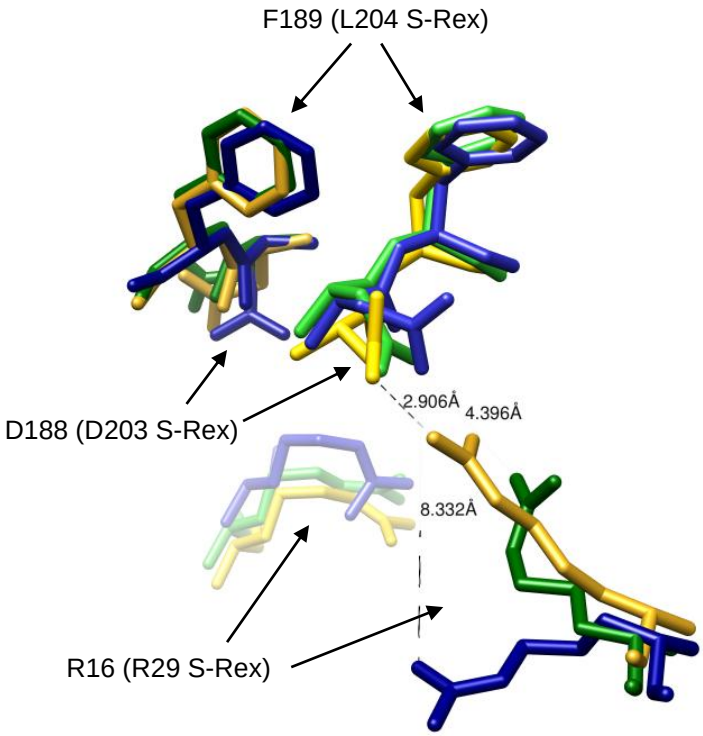


Figure 5.9: Comparison of the three dimers contained within the asymmetric unit of NADH-bound T-Rex (Sickmier *et al.*, 2005). Residues R16 and D188 are shown, as well as the asymmetric F189, 'flipped in' conformation (left-hand side) and 'flipped out' (right-hand side). The three dimmers are coloured gold, green or blue with chains within each dimer coloured with lighter or darker shades of the same colour. The distances between each R16 and D188 is shown on one face of the structure in each case.

The two residues involved in the salt bridge, R16 and D188' T-Rex (R29 and D203 in S-Rex), are highly conserved across the homologues (~80%) thus enforcing our interest in them – both were therefore targeted for mutagenesis. Initially both residues were mutated to alanine by inverse PCR on the pSX142 construct, using primers R29A_F/R and D203A_F/R. The *rex*^{R29A} and *rex*^{D203A} fragments were then isolated by *Nde*I-*Bam*HI digest and introduced into pET15b, cut with the same enzymes. Both proteins were overexpressed and purified via Ni^{2+}-affinity chromatography. Rex^{R29A} expressed well, but Rex^{D203A} produced very little soluble protein. The mutation was therefore replaced with serine - another polar residue, of similar length, which would be incapable of forming the same interaction with R16. The aspartate to serine substitution was introduced in the same manner as for Rex^{D203A}, with primers D203S_F and D203S_R. The abilities of Rex^{R29A}, Rex^{D203A} and Rex^{D203S} to bind to ROP DNA were analysed by electromobility shift assay.

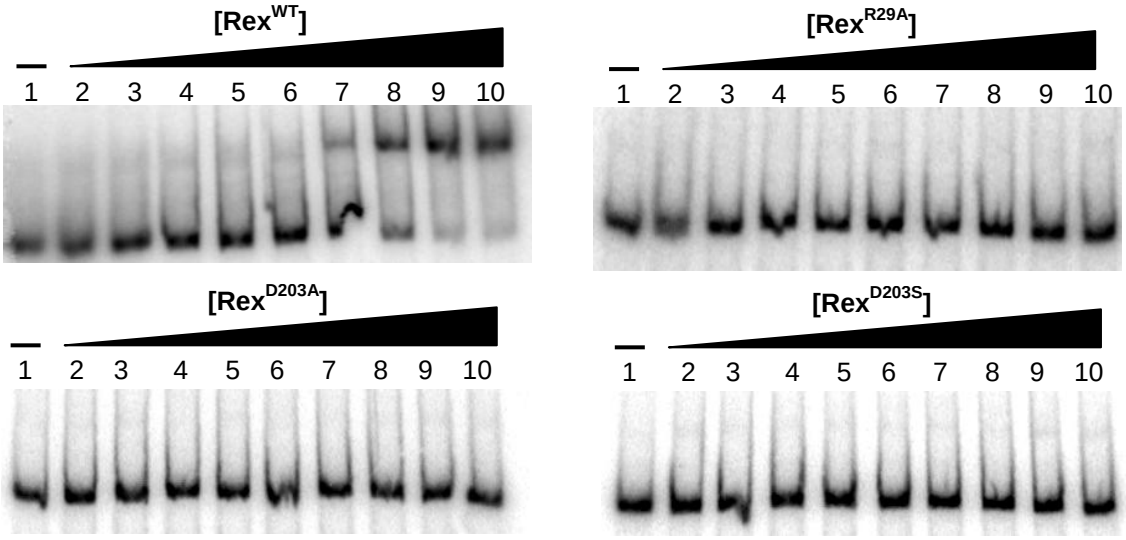


Figure 5.10: EMSA analysis to determine the effect of mutagenesis on the DNA binding ability of Rex. Reactions consisted of 1ng γ^{32}P-labelled probe (ROP^{Rex}), 1µg herring sperm DNA, either 10nM (lane 2); 25nM (lane 3); 50nM (lane 4); 75nM (lane 5); 100nM (lane 6); 250nM (lane 7); 500nM (lane 8); 750nM (lane 9) or 1µM Rex (lane 10) and 1x binding buffer. Lane 1 contained probe and binding buffer alone.

Strikingly, even though each residue is far from the DNA, each failed to bind to the operator even at concentrations of up 1µM protein (Figure 5.10). The Rex^{R29A} mutant was produced in sufficient quantity to undergo size exclusion gel filtration, which confirmed that the protein was still dimeric and indicated that

it co-purified with NADH. The Rex^{D203S} mutant also appeared to be folded, as determined by CD, but did not co-purify with NADH (C. Kielkopf and K. McLaughlin, personal communication). This result would suggest that the salt-bridge does play an important role in stabilising the DNA-bound form of Rex. As a further attempt to study the salt bridge a charge inversion was attempted. R29 was mutated to aspartate and D203 was mutated to arginine in the same protein, to generate $Rex^{R29D:D203R}$. In theory this mutant should still be capable of forming the salt bridge. The two mutations were introduced independently with primers R29D_F/R and D203R_F/R. The two fragments were then adjoined using an internal *Eco*NI site. The resultant protein was analysed by EMSA but unfortunately did not give a gel shift (Figure 5.11). This combined with our observations for the single mutations would suggest that although the salt-bridge is required to stabilise the DNA-bound form of Rex there is an additional function carried out by D203.

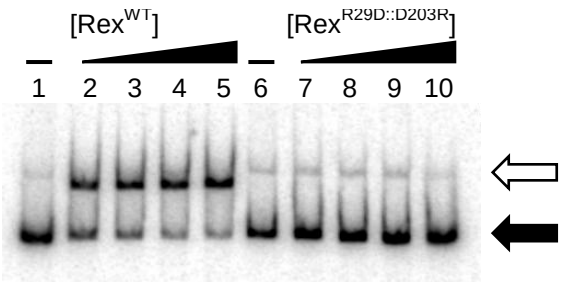


Figure 5.11: EMSAs with the ROP^{Rex} $\gamma^{32}P$ labelled fragment. The probe was mixed with protein at various concentrations, 50nM (lanes 2 and 7), 100nM (3 and 8), 250nM (4 and 9) and 500nM (5 and 10). Lanes 1 and 6 contained the DNA probe alone.

Section 5.4.2 – The sensory triad

The chemical differences between NADH and NAD^+ are slight and yet in Rex one completely inhibits DNA binding whereas the other does not (Brekasis, 2005, Brekasis and Paget, 2003). As mentioned previously, comparison of the NADH- and NAD^+-bound structures revealed a marked difference in the conformations of these two cofactors (Figure 5.2) and also a difference in stoichiometry (McLaughlin et al., 2010). The main difference between the conformation of NAD^+ and NADH in the structures appeared to lie in the position of the nicotinamide ring. In the NADH-bound form the rings are buried

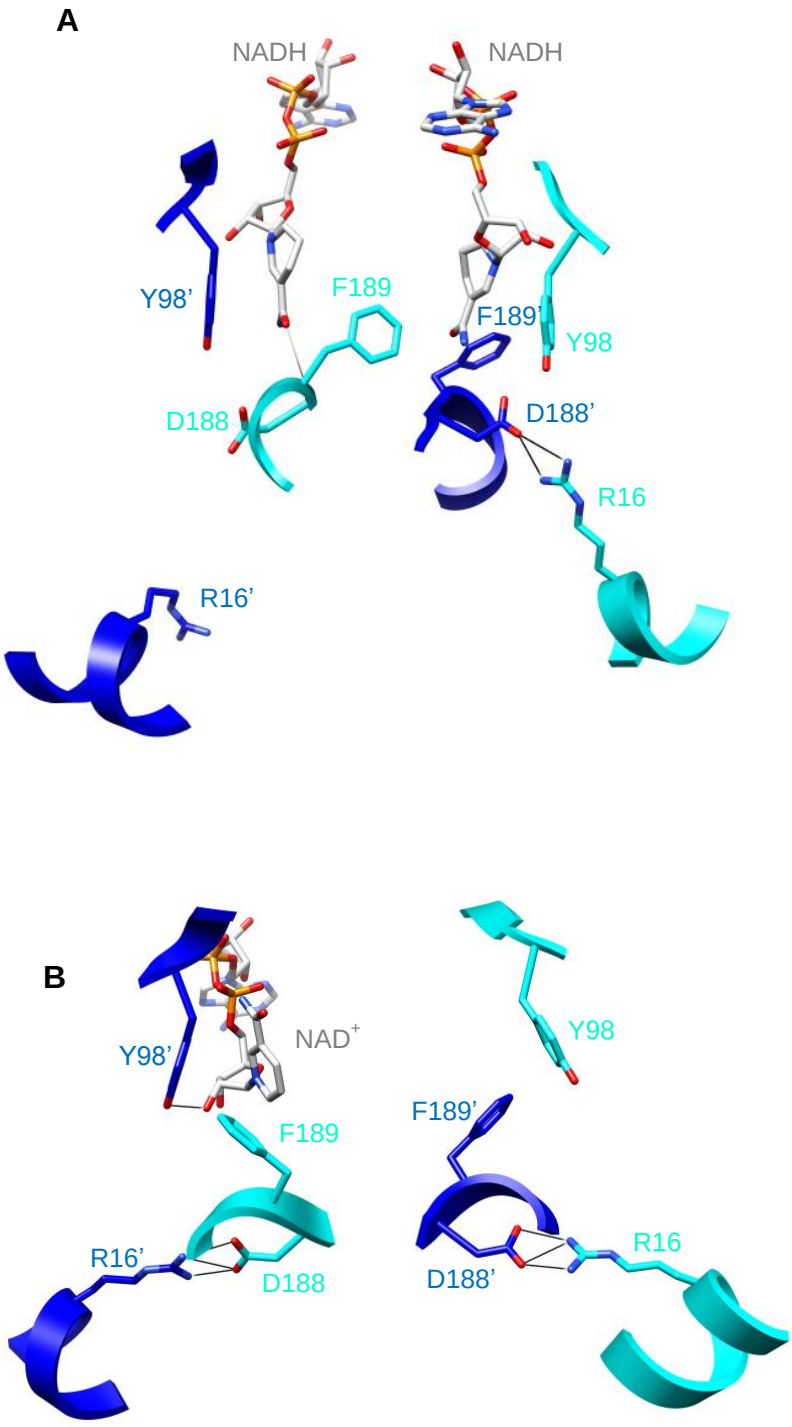


Figure 5.12: Structural view of the sensory triad and salt-bridge of T-Rex, in both its NADH-bound **(A)** and DNA-bound forms **(B)**. In each case one chain is shown in turquoise and the other in blue. The NAD^+ and NADH molecules are depicted in grey.

within the protein and packed against the asymmetrical 'flipped in' F189. In the DNA/NAD^+-bound structure the nicotinamide ring points upwards towards the adenine moiety (*syn*), occupying a position that potentially hinders binding of a second NAD^+ molecule. In the NADH-bound Rex each NADH molecule appears to close a solvent accessible channel (Figure 5.8), that would otherwise lead to the hydrophobic F189, with the binding of both forcing the F189 into its alternative conformation to better accommodate the two cofactors. This asymmetry was proposed to play a key role in the redox sensing mechanism of Rex. However, surprisingly, a F189A T-Rex mutant apparently retained its ability to bind ROP DNA and respond to NADH (data not shown). This suggests that this residue is not a critical component of the redox sensing mechanism and implies that other residues play crucial signalling roles. As mentioned, a correlation between the conformation of D188 (D203) in one chain and the 'flipped in' F189 in the other was identified – this suggested a link between NADH binding, F189 altering its conformation, and the ability of D188 to form its salt bridge. On closer inspection of this region, it appeared that there may be another residue involved – Y98 (Y111), due to its close proximity to both the F189 and D188 residues (Figure 5.12). The F189 equivalent in S-Rex, L204, had previously been substituted with an alanine – a mutation that drastically affected the multimeric state of the protein (D. Brekasis, personal communication). Mutations of D188 (D203) proved to be similarly problematic (Section 5.4.1), and so particular attention was paid to Y98 (Y111). The alanine-substituted Rex^{Y111A} had already been successfully generated and appeared to be both dimeric and stable (D. Brekasis, personal communication). To complement this mutation a more conservative substitution, to phenylalanine, was generated which would potentially still allow NADH binding but lacking the charge that would normally provide an interaction with D203. Another mutation was also generated replacing the tyrosine with an arginine; an arginine might still be able to interact with D203 but no longer able to interact with NADH. It was predicted that this would result in an NADH unresponsive protein. The two mutations were introduced using primers Y111F_F, Y111F_R, Y111R_F and Y111R_R on the pSX142 construct. The resultant fragments were then isolated using *Nde*I/*Bam*HI and ligated into pET15b for overexpression. The Rex^{Y111R} protein did not express well but sufficient quantities were obtained for EMSA

analysis. Apparently this protein was folded, as judged by CD (McLaughlin et al., 2010), but did not co-purify with NADH and did not bind to ROP (Figure 5.13). Rex^{Y111F} and Rex^{Y111A} however produced a shift akin to that of the wild type protein (Figure 5.13).

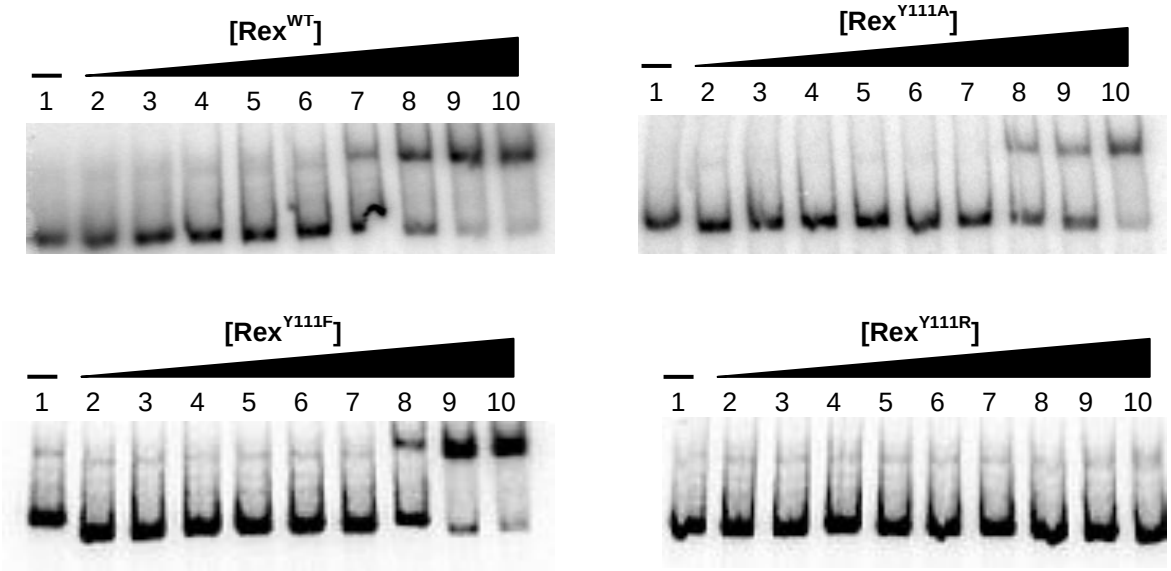


Figure 5.13: EMSA analysis to determine the effect of mutagenesis on the DNA binding ability of Rex. Reactions consisted of 1ng γ^{32}P-labelled probe (ROP^{Rex}), 1µg herring sperm DNA, either 10nM (lane 2); 25nM (lane 3); 50nM (lane 4); 75nM (lane 5); 100nM (lane 6); 250nM (lane 7); 500nM (lane 8); 750nM (lane 9) or 1µM Rex (lane 10) and 1x binding buffer. Lane 1 contained probe and binding buffer alone.

In order to assess whether these mutations had indeed affected the NADH sensitivity of Rex an equilibrium-based SPR method was devised. From our previous observations of the consensus ROP site it appeared that residues and positions 1 and 18 played an important role in stabilising DNA-Rex complexes (Section 3.3.3). It was therefore considered that a 16bp ROP site would generate a specific interaction with Rex but not a stable one. In an SPR assay, this effectively increases the off-rate, allowing the system to reach equilibrium during an injection – crucially without saturating the sensor surface. For each injection the RU value was recorded at the point of equilibrium. For each Rex construct the RU of a protein only injection was taken as the maximum response possible. All injections containing NADH were then compared to this as a percentage of the maximum response ($\%R_{max}$). This method was applied to Rex^{WT}, Rex^{Y111A} and Rex^{Y111F} with NADH concentrations ranging from 180nM to 50µM for Rex^{WT} and Rex^{Y111F}, and from 90nM to 150µM for Rex^{Y111A}. The

discrepancy between the initial added NADH concentrations was because as-purified Rex^{Y111A} was only half saturated with NADH, whereas the Rex^{Y111F} and wt proteins contained 1:1 stoichiometric NADH. In each case the protein concentration was fixed at 180nM and was injected at a flow rate of 30µl/min for 4 minutes. The resulting curves are shown in Figure 5.14.

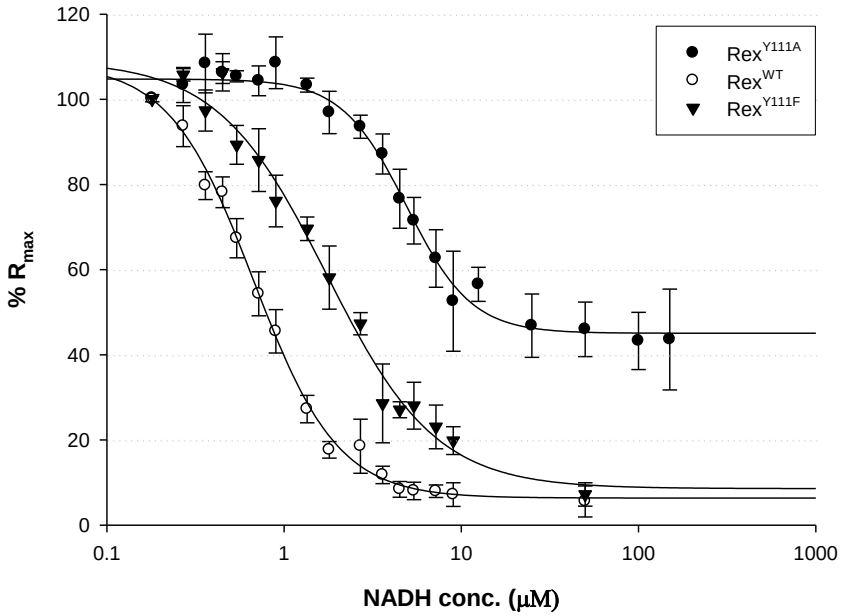


Figure 5.14: SPR data for Rex^{WT}, Rex^{Y111F} and Rex^{Y111A} mutants. The RU values were obtained for each across an NADH concentration range, in triplicate, and compared to a protein only control. The protein concentration was fixed at 180nM and each cycle consisted of a 120µl injection at a rate of 30µl/min, followed by 30µl of 1M NaCl. The minimum NADH concentration of each curve was determined by the amount of NADH that had co-purified with each protein; 180nM for both Rex^{WT} and Rex^{Y111F}, and 90nM for Rex^{Y111A}.

The curve shapes are near identical for Rex^{WT} and Rex^{Y111F}. However, the phenylalanine mutation shifted the curve to the right – indicating that more NADH was required to give the same response. The IC_{50} was read as the mid-point between the start and end plateaus and gave values of 0.79±0.07 and 2.46±0.27 µM NADH for Rex^{WT} and, Rex^{Y111F}, respectively. In the case of Rex^{Y111A}, although NADH inhibited Rex-ROP interactions, there remained a population of protein that would not fully dissociate. This suggested that there might be a mixed population of protein conformers, with some binding to DNA in a non-responsive form. Nonetheless, the IC_{50} for Rex^{Y111A} was predicted to be 4.57±0.71. Y111 plays an important role in redox sensing. Any attempts to

disrupt this residue resulted in a reduced response to NADH (Rex^{Y111A} and Rex^{Y111F}), or an inability to bind DNA at all (Rex^{Y111R}). The repositioning of this residue, upon NADH binding, places the hydroxyl group in proximity to D203, forming a hydrogen bond, breaking the salt-bridge and destabilising the DNA-bound form. Conversely upon DNA binding Y111 forms a hydrogen bond with NAD^+, freeing D203 to form a stabilising salt bridge with R23. Thus this residue has roles in both forms of Rex.

Section 5.5 – A single chain Rex

The structure of Rex bound to NAD^+ (and DNA) or NADH revealed differences in stoichiometry and in the position of the nicotinamide ring. A crucial question is whether the binding of one or two NADH molecules is required to dissociate Rex. However, since Rex is dimeric, mutations affect the equivalent position in both subunits which impedes the construction of a mutant that can only bind a single NADH. One way to solve this problem is to generate a single-chain dimer that allows the specific mutation of one “subunit”. Such an approach was used to test the importance of stoichiometry in tetracycline binding to TetR (Krueger *et al.*, 2003). A single chain mutant would also allow more detailed investigation of Rex-ROP interactions at half-sites.

Section 5.5.1 – Rex^{SC} design and execution

There are two ways to generate a mixed species homodimer: express both copies separately then selectively purify the heterodimer; or translationally fuse the two copies together. In this study only the latter approach was taken. In each case homologous recombination between closely related sequences is a potential problem that can result in plasmid instability. To overcome this, a second copy of *rex* was redesigned using synonymous codons, which should limit recombination, while maintaining optimum codon usage for gene expression in *E. coli*. The sequence of *rex* was therefore altered using the bioinformatic tool Silent Wizard (Damerell, 2007). This tool was run with the

```
                M  A  H  R  P  A  T  R  S  R  G  I  P  E  A  T  V  A  R  L
Rex_Synth       ATGGCGCATCGCCCTGCCACGCGTTCGCGCGGCATCCCGGAAGCGACGGTGGCGCGCCTG 60
Rex             ATGGCACACCGACCGGCGACCCGCAGCCGAGGGATTCCCGAGGCCACCGTCGCCAGGCTT 60
                ***** ** ** ** ** ** **    ** ** ** ** ** ** ** ** **  * **

                P  L  Y  L  R  A  L  T  A  L  S  E  R  S  V  P  T  V  S  S
Rex_Synth       CCTTTATATCTGCGTGCGTTAACGGCCTTAAGCGAACGTAGCGTTCCGACCGTGAGCAGC 120
Rex             CCGCTGTACCTCCGCGCACTGACCGCGCTGTCCGAGCGCTCGGTGCCCACGGTCTCCTCC 120
                **  * ** ** ** **  * ** **  *   *** **    ** ** ** **   *  *

                E  E  L  A  A  A  A  G  V  N  S  A  K  L  R  K  D  F  S  Y
Rex_Synth       GAAGAATTAGCCGCGGCGGCCGGCGTGAATAGCGCCAAATTACGTAAAGATTTTAGCTAT 180
Rex             GAGGAGCTGGCGGCCGCCGCGGGGGTCAACTCCGCGAAGCTGCGCAAGGACTTCTCCTAC 180
                ** **  * ** ** ** ** ** ** **   *** **  * ** ** ** **   ***

                L  G  S  Y  G  T  R  G  V  G  Y  D  V  E  Y  L  V  Y  Q  I
Rex_Synth       CTGGGTAGCTATGGCACGCGTGGCGTGGGTTATGATGTGGAATACCTGGTGTATCAAATT 240
Rex             CTCGGCTCCTACGGGACCCGCGGTGTCGGCTACGACGTCGAGTATCTCGTCTACCAGATC 240
                ** **   *** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **

                S  R  E  L  G  L  T  Q  D  W  P  V  V  I  V  G  I  G  N  L
Rex_Synth       AGCCGTGAGCTGGGTCTGACGCAAGATTGGCCTGTGGTTATTGTGGGCATTGGTAATCTG 300
Rex             TCGCGCGAACTCGGCCTCACCCAGGACTGGCCGGTTGTGATCGTCGGTATCGGCAACCTC 300
                   ** ** ** ** ** ** ** ** ** ***** ** ** ** ** ** ** ** ** **

                G  A  A  L  A  N  Y  G  G  F  A  S  R  G  F  R  V  A  A  L
Rex_Synth       GGCGCGGCCCTGGCGAATTATGGCGGCTTTGCGAGCCGTGGCTTTCGTGTGGCGGCCCTG 360
Rex             GGTGCCGCGCTCGCCAACTACGGTGGTTTCGCCTCCCGCGGGTTCCGCGTCGCCGCGCTC 360
                ** ** ** ** ** ** ** ** ** ** ** **   *** ** ** ** ** ** ** **

                I  D  A  D  P  G  M  A  G  K  P  V  A  G  I  P  V  Q  H  T
Rex_Synth       ATTGATGCGGACCCTGGCATGGCGGGCAAACCGGTGGCGGGTATTCCTGTTCAACATACG 420
Rex             ATCGACGCCGATCCGGGAATGGCCGGAAAGCCCGTCGCCGGCATCCCGGTGCAGCACACC 420
                ** ** ** ** ** ** ***** ** ** ** ** ** ** ** ** ** ** ** **

                D  E  L  E  K  I  I  Q  D  D  G  V  S  I  G  V  I  A  T  P
Rex_Synth       GATGAATTAGAAAAAATTATTCAAGATGATGGCGTGAGCATTGGCGTTATTGCCACGCCG 480
Rex             GACGAGCTGGAGAAGATCATCCAGGACGACGGTGTCTCGATCGGTGTGATCGCGACCCCC 480
                 ** **  * ** ** ** ** ** ** ** ** ** **    ** ** ** ** ** ** **

                A  G  A  A  Q  Q  V  C  D  R  L  V  A  A  G  V  T  S  I  L
Rex_Synth       GCGGGTGCGGCGCAACAAGTGTGTGATCGTCTGGTTGCGGCGGGCGTGACGAGCATTTTA 540
Rex             GCCGGCGCCGCCCAGCAGGTCTGCGACCGCCTCGTGGCCGCCGGTGTCACCTCCATCCTG 540
                ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **   ***  *

                N  F  A  P  T  V  L  N  V  P  E  G  V  D  V  R  K  V  D  L
Rex_Synth       AATTTTGCCCCTACGGTTTTAAATGTGCCGGAAGGTGTGGATGTTCGTAAAGTGGATCTG 600
Rex             AACTTCGCGCCGACCGTGCTGAACGTCCCCGAGGGCGTCGACGTGCGCAAGGTCGACCTC 600
                ** ** ** ** ** **  * ** ** ** ** ** ** ** ** ** ** ** ** **

                S  I  E  L  Q  I  L  A  F  H  E  Q  R  K  A  G  E  E  A  A
Rex_Synth       AGCATTGAATTACAAATTCTGGCGTTTCATGAACAACGTAAAGCCGGTGAAGAAGCGGCC 660
Rex             TCCATCGAGCTGCAGATCCTCGCCTTCCACGAGCAGCGCAAGGCGGGCGAGGAGGCCGCG 660
                  *** **  * ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **

                A  D  G  A  A  P  P  V  A
Rex_Synth       GCGGATGGTGCGGCGCCTCCGGTGGCG 687
Rex             GCCGACGGCGCCGCACCGCCCGTCGCC 687
                 ** ** ** ** ** ** ** ** **
```

```
   L  S  G  G  G  G  S  G  G  G  G  H
AAGCTTTCGGGCGGTGGCGGTTCCGGTGGCGGTGGCCATATG
TTCGAAAGCCCGCCACCGCCAAGGCCACCGCCACCGGTATAC
HindIII                              NdeI
```

$(SG_4)_2$ linker sequence, flanked by sticky ends for *Hin*dIII and *Nde*I

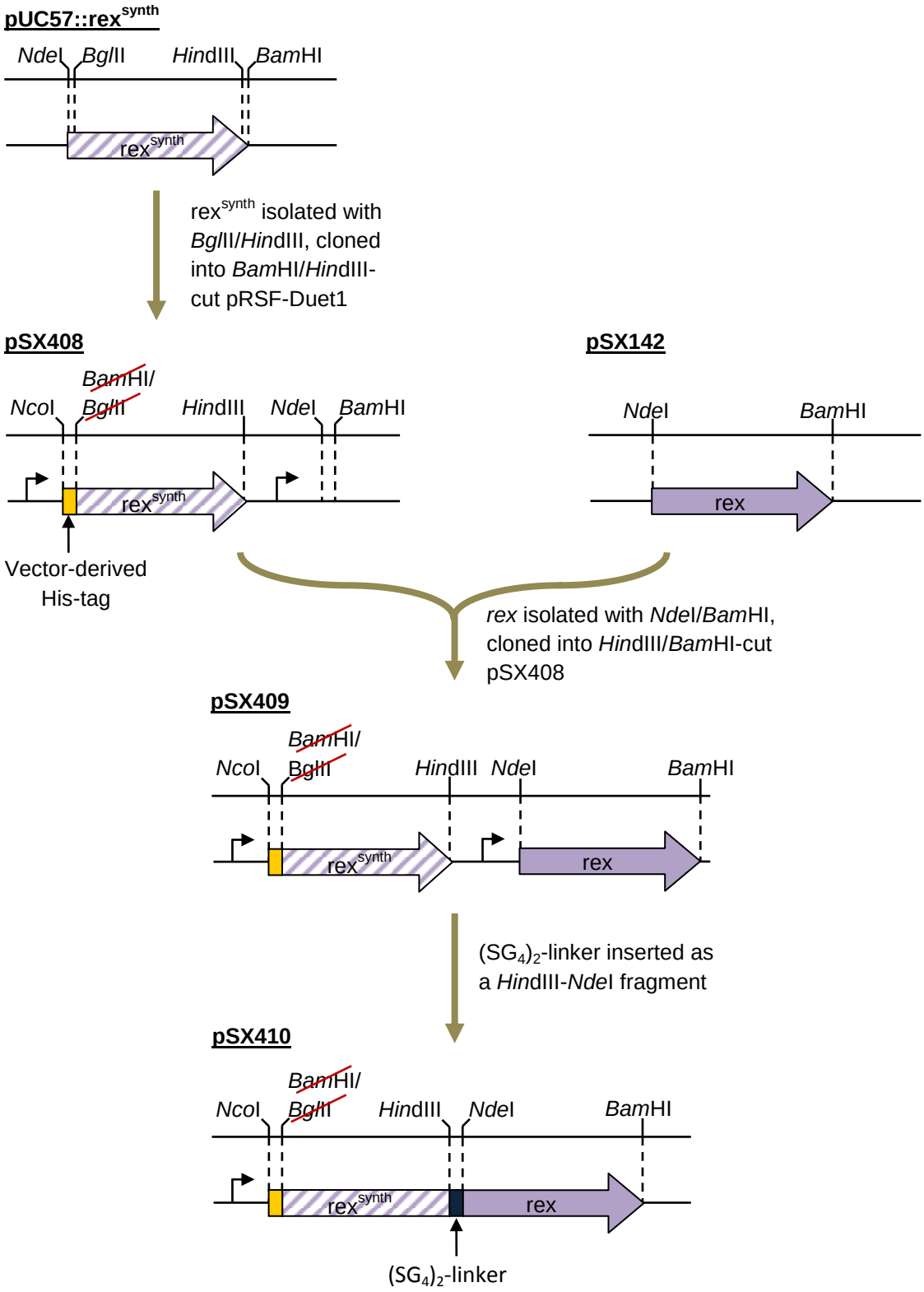


Figure 5.15: Diagrammatic view of the Rex^{SC} construction. The full sequence of the redesigned *rex* gene (Rex^{synth}) is shown on the preceding page as an alignment against the original sequence. The sequence of the linker peptide, which joins the two chains, is also shown.

option to convert the sequence to one as different as possible with a high expression level, using the *E. coli* K12 codon usage table (Figure 5.15). The resulting sequence was then synthesized by GenScript (New Jersey, USA) and placed into pUC57. The synthetic *rex* sequence (*rex*^{synth}) was then isolated as a *Bgl*II/*Hin*dIII fragment and introduced into *Bam*HI/*Hin*dIII cut pRSF-Duet1, to generate pSX408. The original *rex* fragment was subsequently isolated as an *Nde*I/*Bam*HI fragment from pSX142 and introduced into the equivalent sites in pSX408, to create pSX409. Finally, the linker peptide fragment was introduced as a *Hin*dIII/*Nde*I fragment, yielding pSX410. The linker chosen was based on that of Krueger *et al.*, as they had used the same amino acid repeat pattern (Ser-Gly-Gly-Gly-Gly) to successfully generate a TetR fusion (Krueger *et al.*, 2003). As the gap between the C- and N-termini of Rex (maximum distance of ~16.6Å) appeared to be much shorter than TetR (~59.1Å, based on 1QPI biological assembly) only two repeats were used $(SG_4)_2$, instead of the five used for TetR. The full sequence of the linker peptide, and of *rex*^{synth}, is shown in Figure 5.15, along with an overview of the above protocol for the generation of the single-chain Rex (Rex^{SC}). Overexpression of the fusion gene produced a single soluble protein of ~50 kDa (data not shown). An initial test was performed on the protein to check that it was still capable of binding to DNA and responding to NADH (Figure 5.16).

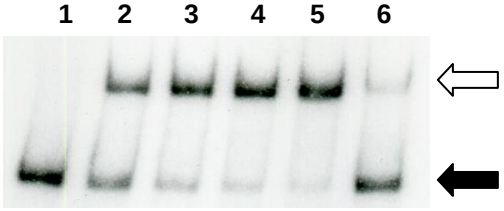


Figure 5.16: EMSA analysis of the Rex^{SC} protein. 1ng of γ^{32}P-labelled ROP^{rex} was used in each reaction (lane 1), along with 1µg of herring sperm DNA. The amount of protein used was 25nM (lane 2), 50nM (3), 100nM (4) and 250nM (5). Lane 6 contained 250nM Rex^{SC} and 100µM NADH. All reactions were incubated at 25°C for 20 minutes prior to loading onto a 6% polyacrylamide gel. The closed arrow shows the position of the probe, whereas the open arrow indicates the shift.

The fusion protein produced a gel shift in EMSA analysis and this was almost entirely lost upon the inclusion of 100µM NADH in the reaction mixture. It would appear that the functionality of Rex was maintained in Rex^{SC}. The *rex*^{SC}

fragment was subsequently transferred into pET15b using the *Nco*I/*Bam*HI sites, yielding pSX411, which was more in line with all previous Rex overexpression work.

Section 5.5.2 – Rex^{SC} mutagenesis

Can a single NADH can trigger Rex dissociation? In order to assess whether a single NADH molecule could trigger Rex dissociation, the G102A super-repressor mutation was introduced into the Rex^{SC} construct as an *Nde*I-*Bam*HI fragment from pSX142::G102A (pSX411::G102A). For mutagenesis of the synthetic region, in order to create double site controls, the region had to be isolated from pET15b::Rex^{SC} using *Eco*O109I and ligated into pBlueScript II SK^{+} cut with the same enzyme, to generate pSX412. This was then used as a template for PCR, with primers G102A_SC_F and G102A_SC_R. The resulting fragment was isolated using *Nco*I/*Nde*I and introduced into the same sites of pSX411::G102A. The single site mutant was annotated as Rex$^{SC::G102A}$ and double site mutant as Rex$^{SC::G102A::G102A}$. The resultant proteins were expressed and purified, including the Rex^{SC} wild type protein. The three proteins were then analysed by SPR in order to assess their NADH-responsiveness (Figure 5.17).

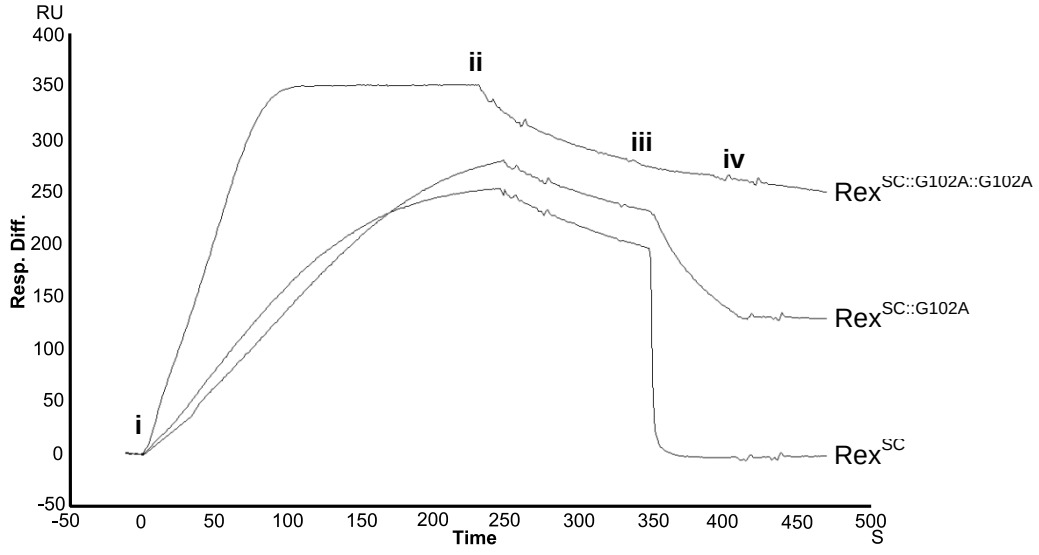


Figure 5.17: SPR data for the Rex^{SC} mutants: WT; G102A and G102A::G102A. The flow rate was kept constant at 30µl/min, with an injection volume of 120µl 180nM Rex, 60µl of 5µM NADH and 30µl 2M $MgCl_2$. The points labelled i and ii indicate the beginning and end of the protein injections, respectively. Points iii and iv indicate the start and end of the NADH injections. Note that the axes labels have been redrawn for clarity but are still to the same scale.

The result was a rapid drop in the response up to the point where the injection ended. However, whereas Rex^{SC} was completely dissociated by NADH $Rex^{SC::G102A}$ was not. This suggests that the $Rex^{SC::G102A}$ is still capable of responding to NADH but with reduced affinity for the inhibitor.

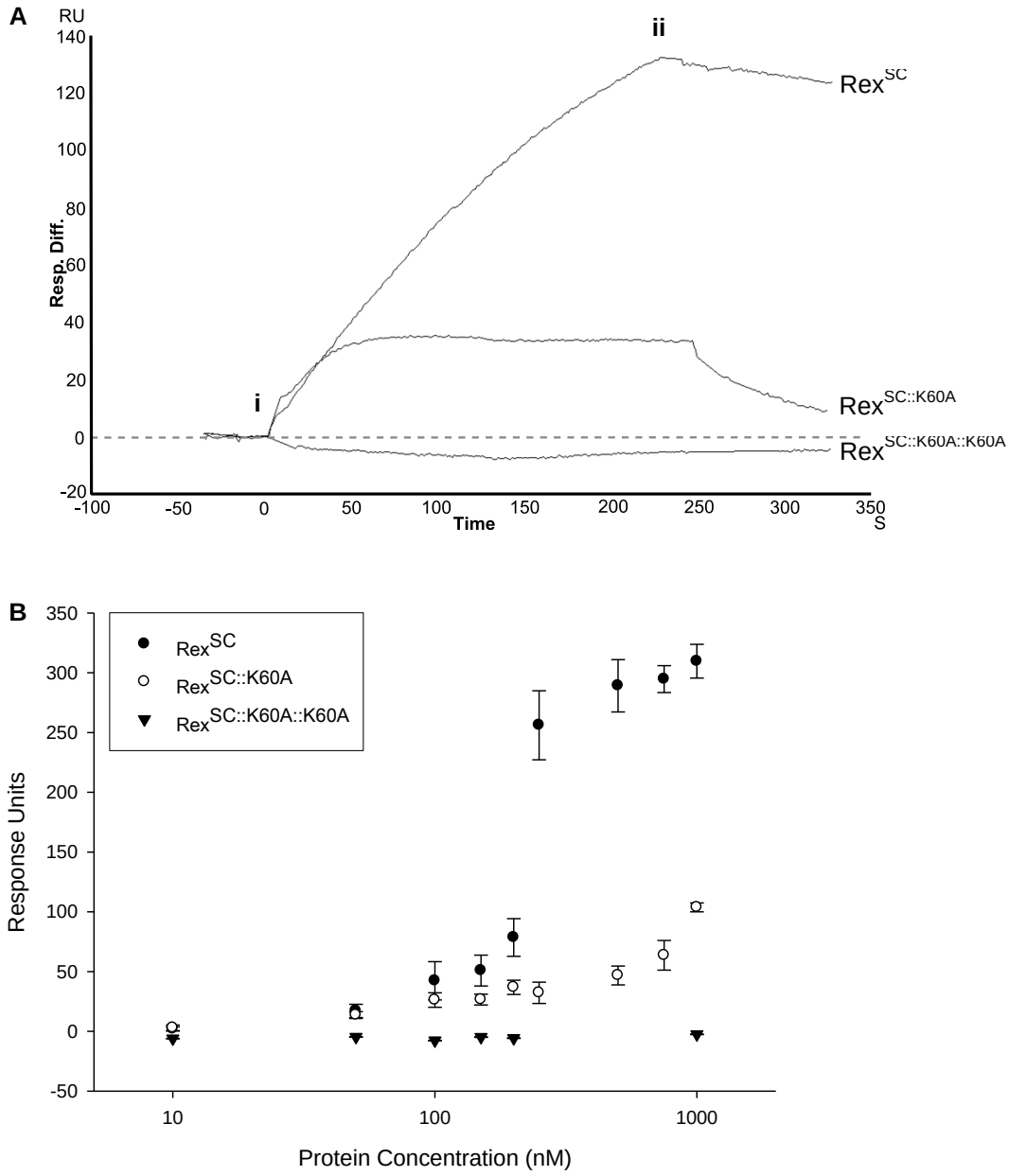


Figure 5.18: SPR data for the Rex^{SC} mutants: WT; K60A and K60A::K60A. **(A)** Raw data showing the response curves for each protein. **(B)** Binding curves showing the change in response units over a protein concentration in each case. The flow rate was kept constant at 30µl/min, with an injection volume of 120µl of Rex and 30µl 2M $MgCl_2$. The Rex concentrations were **(A)** 180nM and **(B)** 10, 50, 100, 150, 200, 250, 500, 750 and 1000nM. The points labelled i and ii indicate the beginning and end of the protein injections, respectively. Note that the axes labels have been redrawn for clarity but are still to the same scale.

Can Rex recognise DNA via a single DNA-binding domain? The identification of the Rex regulon by ChIP-chip revealed several native binding sites that appeared to be comprised of only a half-site. This suggested that Rex might be able to bind such sites by invoking only key protein-DNA interactions in one major groove. To further test this it was decided to remove a key protein-DNA interaction in one of the recognition helices. As the K60 residue of Rex was shown to be essential for DNA binding (Section 5.3.2) it was used to disrupt DNA binding in one half of the Rex "dimer". The mutagenesis scheme was as for $Rex^{SC::G102A}$ and $Rex^{SC::G102A::G102A}$, using primers K60A_SC_F and K60A_SC_R. The resultant proteins were again analysed by SPR (Figure 5.18). As had previously been observed the Rex^{SC} protein was capable of binding to ROP DNA, and consistent with the Rex^{K60A} mutant the single chain K60A double mutant was also incapable of binding to DNA. The SPR responses for the $Rex^{SC::K60A}$ mutant were much lower than those of wild type Rex^{SC}, however the protein did appear to bind specifically to ROP^{nuo}. This result would suggest that although two functional DNA-binding domains are required for stable complex formation, a protein with only one active DNA-binding domain can still interact with a ROP operator. It was noted that none of the Rex^{SC} proteins co-purified with any discernable NADH; the proteins all nevertheless bound specifically to ROP instead of the randomised DNA on a preceding lane of the sensor chip. Unfortunately all of the single-chain proteins were slightly unstable, possibly due to the lack of NADH or an affect of the fusion itself, as a noticeable decline was observed in the response maximum over time. Due to this it has not been possible to determine K_ds or IC_{50}s for these proteins. This issue would need to be addressed before proceeding any further with this system.

Section 5.6 – Conclusions

Rex consists of two DNA-binding domains, which co-ordinate the dyad symmetrical repeat (TTGTGAA-n4-TTCACAA). By a combination of structural analysis and mutagenesis this study has demonstrated the importance of two positively charged residues, R46 and K47 (T-Rex), in forming

a tight Rex-ROP complex. However, Rex has also been shown to interact with regions lacking a full operator site, containing TTGTGAA alone (Section 3.3.5). Through use of a single-chain Rex construct, the ability of a single DNA-binding domain to form a specific interaction with ROP DNA was shown. However, this interaction was much weaker than its wt counterpart, suggesting that a single DNA-binding domain can interact with a target site but that both are required for tight binding. NADH-bound Rex adopts a conformation that is incompatible with DNA binding (Sickmier *et al.*, 2005). Comparison with the DNA-bound structure reveals a 43° rotation, centred at residue D188 (T-Rex), which re-positions the DNA-binding domains into adjacent major grooves of the DNA. The structure of an NADH-binding mutant, Rex^{R90D}, has also been determined (McLaughlin *et al.*, 2010). The relative Cα positions of this and DNA-bound T-Rex^{WT} were identical, with the apo-Rex locked in the DNA-bound conformation. This would implicate NADH, not DNA, as the modulator of the structural alterations to Rex.

The relationship between NADH dissociation and DNA-binding has been a major theme of this work. By studying the difference in the NADH- and DNA-bound structures, a model whereby NADH binding triggers a massive conformational change is proposed (Figure 5.19). Several factors are involved centring around a sensory triad at the heart of the protein, involving residues R16, Y98 and D188 (T-Rex). The mechanism relies on the ability of D188 to switch bond partners between R16 (DNA-bound) and Y98 (NADH-bound). DNA-/NAD^+-bound Rex contains a R16-D188' salt bridge on each side of the protein, stabilising the Rex-ROP complex. Within this structure one NAD^+ molecule occupies a single NAD(H) binding site, and adopts a *syn* conformation. This is a conformation NADH is not able to adopt due to a steric clash with A94; the same residue is also responsible for the lack of *anti* NAD^+. The *syn* conformation of NAD^+ is stabilised by an intra-molecular hydrogen bond. The absence of a second NAD^+ within DNA-bound Rex is most likely due to the electrostatic repulsion that would occur between the two cofactors. Conversely the binding of two *anti* NADH molecules is possible and requires the repositioning of F189 between the two cofactors ('flipped in'). Upon NADH dissociation, the phenylalanine is permitted to 'flip out' allowing the structure to rotate on its axis. This movement repositions Y98 and D188, breaking the

hydrogen bond and freeing D188 to form a salt bridge with R16', which has drawn closer by the movements. This bond forms on each face and locks Rex into its DNA-bound state. Discovery of the sensory triad of Rex has provided a unique insight into the redox sensing mechanism of the Rex-family repressors.

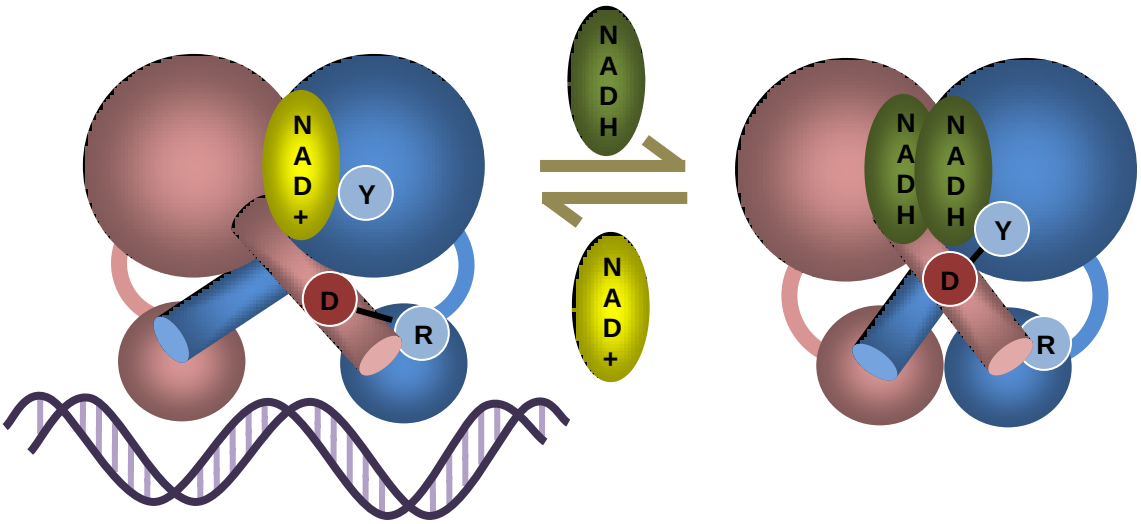


Figure 5.19: A model for redox sensing in Rex highlighting the importance of sensory triad residues: R16; Y98 and D188. The repositioning and alteration in bond partners during the switch from DNA- to NADH-bound forms is indicated above. Figure adapted from (McLaughlin *et al.*, 2010).

Studies on the Rex^{SC} NADH-binding mutants have proven interesting, despite requiring further work. As all of the single chain proteins lack NADH one would expect that they must adopt a DNA-bound conformation when purified. They all bind specifically to ROP DNA and have comparable natural off-rates, suggesting stability on the DNA is unaltered. Upon NADH injection both the Rex^{SC} wt and $Rex^{SC::G102A}$ proteins will dissociate, however rapid dissociation appears to require the binding of two NADH molecules (wt). Nonetheless, the single site mutant will respond to NADH but is much less sensitive. This suggests that the presence of one NADH within the protein is enough to trigger the structural changes associated with full NADH-binding. For example, a single binding event might break one of the salt-bridges forming an unstable intermediate, which cannot bind a second NADH. This would present two possible options: (1) succumb to the reduced stability and dissociate from the DNA, or (2) await NADH-dissociation and re-form the salt-bridge. Clearly, further studies are required; for example it is likely that full NADH-induced dissociation is possible for the single binding site mutant at higher NADH

concentrations, as this would presumably increase the on-rate of NADH onto the protein thus making option two unfeasible.

Salt-bridges are a common occurrence in transcription factor structures (e.g. PhoP, SlyA, DesT, etc). However inter-chain salt-bridges, like the ones in Rex, that stabilise the DNA-bound form alone are rare. There was however a recently identified example; RAG1. RAG1 is a eukaryotic protein involved in V(D)J recombination during lymphocyte development (Yin *et al.*, 2009). RAG1 binds to two DNA molecules and, through homo-dimerisation, 'tethers' them together (Yin *et al.*, 2009). This dimerisation requires one of its α-helices to 'kink' in the middle, positioning one end in the major groove of the DNA and the other located to interact with the other subunit (Yin *et al.*, 2009). Stabilisation of this kink requires the formation of a salt-bridge between an arginine on one subunit and a glutamate of the other, deletion of which had a severe effect on DNA-binding (Yin *et al.*, 2009).

The triad components are highly conserved across the Rex homologues (R16 ~ 85%, Y98 ~ 77% and D188 ~ 81%). However, only three homologues have had their crystal structures solved: *T. aquaticus*, *T. thermophilus* and *B. subtilis* (Nakamura *et al.*, 2007, Wang *et al.*, 2008). The structure of Tth-Rex is extremely similar to T-Rex, as is expected from two species of the same genus, perhaps especially so given that these are both thermophiles and so any divergence in the sequences could potentially have drastic effects on protein stability at their native temperatures. The elements of the B-Rex protein are conserved but the overall conformation is quite different due to a crystal packing issue. Unfortunately T-Rex is the only homologue to have both its NADH- and DNA-bound structures determined thus it is not yet possible to compare the salt-bridge function in other species. Its conservation would however suggest it is also functionally important within these proteins.

Chapter 6

General Discussion

"From error to error, one discovers the entire truth."

Sigmund Freud (1856-1939)

Section 6.1 – Overview

Bacterial respiratory chains can vary greatly, incorporating components akin to those of eukaryotic mitochondria or plant photosystems, and can be adapted to use whatever substrates and terminal electron acceptors that are available to them. For a strict aerobe such as *S. coelicolor*, anaerobic growth is not an option, posing quite a challenge for a soil-dwelling microbe faced with frequent bouts of oxygen limitation (van Keulen *et al.*, 2003). In other bacteria oxygen deprivation is a key signal, direct or indirect, to switch from aerobic to anaerobic pathways – a much less energy efficient process (Poole and Cook, 2000, Unden and Bongaerts, 1997). In *S. coelicolor* it is a signal to adapt the aerobic respiratory chain to continue energy generation at lower oxygen concentrations. One example of this is induction of the cytochrome *bd* terminal oxidase, which is thought to have a higher affinity for oxygen, allowing aerobic respiration to continue at lower oxygen tensions (Poole and Cook, 2000). As these alternative pathways are generally less efficient at energy generation it would not make sense for the organisms to induce them under aerated conditions; thus tight control is important for maximising ATP yield. In order to switch these systems on and off, as required, Gram positive bacteria, including *S. coelicolor* have evolved a novel transcriptional repressor; Rex (Brekasis, 2005, Brekasis and Paget, 2003). Rex is able to directly sense a shift in the NADH/NAD^+ redox poise that is caused by oxygen limitation (Brekasis, 2005, Brekasis and Paget, 2003). Binding to operators to repress during aeration and dissociating when NADH levels rise (Figure 6.1). Other bacterial redox sensors exist, e.g. ArcAB, ResDE, Fnr, but Rex appears to be unique in its method of redox sensing. Previous work had focused on identifying and characterising this protein as a redox-sensitive repressor (Brekasis, 2005, Brekasis and Paget, 2003, Sickmier *et al.*, 2005). The current study has focused on characterising the Rex regulon and dissecting the mechanism through which it conveys the binding of NADH in one domain to a distinct DNA-binding domain.

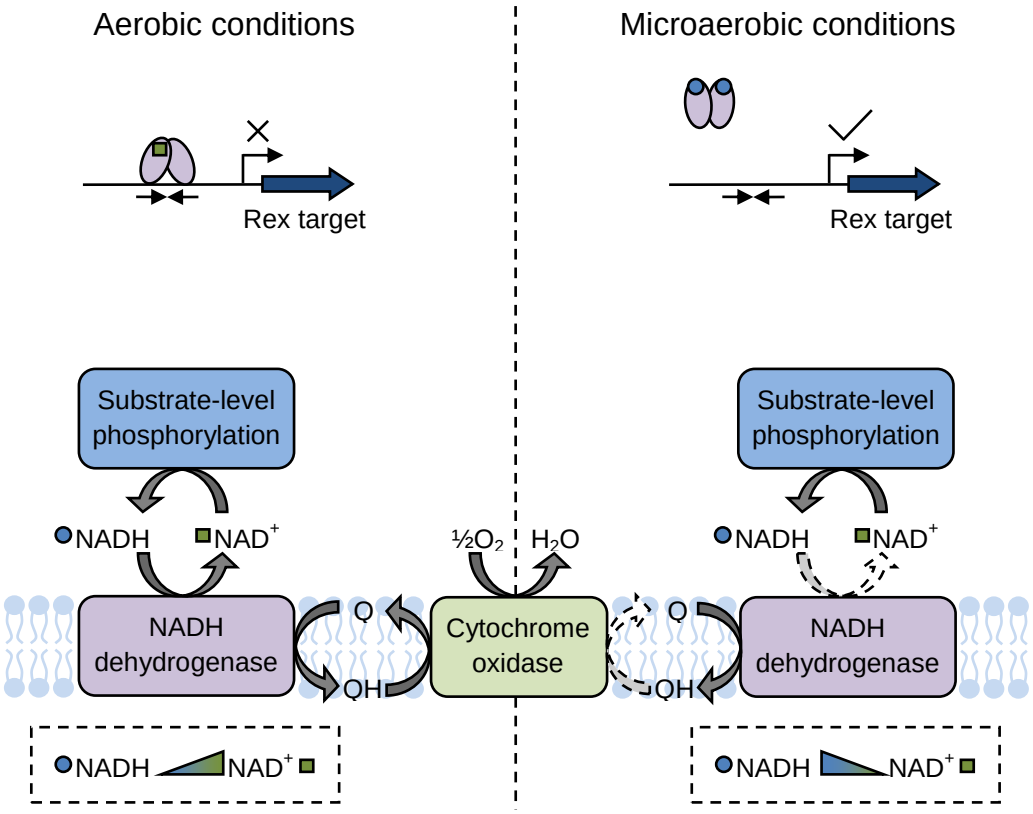


Figure 6.1: Overview of the changes in redox poise, and the resultant change in the regulation by Rex, that occur when cells switch between aerobic and microaerobic growth.

Section 6.1.1 – The Rex operator site

During this study ChIP-on-chip was used to identify 22 new Rex binding sites *in vivo*. The binding to six of these sites was confirmed *in vitro* by EMSA analysis; in the promoters of genes: SCO1930, SCO3101, SCO3547, SCO5207 and SCO6168, and in the region between divergent genes: SCO5408 and SCO5409. By devising a more sensitive binding analysis method, using SPR, a further 3 sites were tested and shown to interact with Rex; SCO6218/9, SCO6239 and SCO6917. Analysis of the sequences of all of ChIP-chip sites revealed two classes of ROP site; strong and weak binders. Furthermore this revealed that the S-Rex recognition site was specific to an 18 bp operator site, not 16 bp as had previously been reported. The strong sites were defined by the presence of both the ‘GTG’ and ‘CAC’ elements of the consensus ROP site (ttGTGaa-n4-ttCACaa) and by their ability to withstand a gel shift when bound to

Rex. Weak sites were lacking all or most of these conserved elements, but interestingly only in one half of the ROP site. This finding was further supported by use of a single-site DNA-binding mutant. Rex can still interact with specific DNA in this form but is unable to form a stable Rex-ROP complex. The repression method of these weaker sites remains unclear as one would expect that the binding would be too weak for simple promoter occlusion. However it is possible that the binding is stabilised by either another protein or simply by the DNA topology *in vivo*, which would explain the enrichment observed for these sites during ChIP-on-chip, but failure to bind *in vitro*.

Rex is apparently not unique in its ability to recognise half-sites, CtrA is a cell-cyle regulator of *Caulobacter crescentus* that is also able to do so (Spencer *et al.*, 2009). CtrA is activated by phosphorylation, which increases its affinity for a 15 bp site (Spencer *et al.*, 2009). However, CtrA can also recognise half-sites but with weaker affinity than the full binding site (Spencer *et al.*, 2009). The affinity for the half-site is unaltered by phosphorylation of CtrA, and the affinity for the full site is comparable when CtrA is in this state, allowing occupation of the weaker sites (Spencer *et al.*, 2009). Interestingly other sequence elements surrounding the half-site appear to stimulate binding; these elements are however not required by the full CtrA site (Spencer *et al.*, 2009).

It is not yet clear what the function of these sites is due to the weakness of the interaction. It would however be interesting to determine the redox sensitivity of these sites as this may indicate their purpose, i.e. if they are more sensitive to NADH then they may allow for a more rapid derepression of the targets and if they were less sensitive then they would allow for differential derepression of target genes.

Section 6.1.2 – Homeostatic redox control

Rex is not an essential gene in *S. coelicolor* but overexpression of the *rex*^{G102A} super-repressor was lethal. The cause of the toxicity was most likely the inability to rapidly de-repress one or more target genes. Given the essential nature of *ndh* this would seem an obvious source of the Rex^{G102A} toxicity. However this was not supported by the RT-qPCR data. The position of the tightest ROP site

in the *ndh* promoter (~300 bp upstream) would also suggest that this was not the target responsible. Another possibility is the *rex-hemACD* operon. Previously it was shown that mutations in *rex* have a polar effect on the downstream *hemACD* genes resulting in very poor growth (Brekasis and Paget 2003). Rex also binds upstream of the essential ATP synthase operon although it is not clear how tightly this operon is controlled by Rex. As the inhibitory effect of the super-repressor was only observed on solid media it is also possible that the effect is only caused under these conditions. As *nuo* (NDH-1) appears to be constitutively active during growth on solid media it would seem possible that may be responsible, however the growth inhibition for the *nuo* disruption strain (*nuo*^{deg}) was not as severe as observed for the super-repressor. Another gene potentially linked to growth on solid media is *wblE*, however this gene has successfully been deleted by others without any noticeable effects on growth (Homerova *et al.*, 2003). One possible culprit is *resA*, of the cytochrome biogenesis operon. This gene gave a high enrichment by ChIP-on-chip (20.84-fold), and also produces a gel-shift during EMSA analysis (D. Brekasis, personal communication). Interestingly the cytochrome biogenesis operon (*resABC*) of *B. subtilis* appears to be essential (Sun *et al.*, 1996).

The *ndh* gene is not just regulated by Rex; it also plays an important role in the regulation of Rex. By depleting *ndh* from the cells the ability of Rex to repress target gene; *cydA*, was disrupted. Thus by depriving the cells of NDH-2 the NADH/NAD$^+$ ratio was altered enough to cause the dissociation of Rex. When both Rex and NDH-2 are present in the cell they will act upon each other, with Rex regulating *ndh* expression and NDH-2 maintaining the redox poise at a level at which Rex can repress. This importance of this feedback loop has already been shown in *B. subtilis* (Gyan *et al.*, 2006) and has also been demonstrated in *S. coelicolor* through the work in this study. Interestingly both *B. subtilis* and *S. aureus* appear to lack Rex operators upstream of their *rex* genes, meaning that in these systems *rex* does not appear to be auto-regulated (Gyan *et al.*, 2006, Larsson *et al.*, 2005, Pagels *et al.*, 2010). Systems that employ negative auto-regulation have been shown to reach steady-state much faster than non-autoregulated systems, due to the production being linked to the

product (Rosenfeld *et al.*, 2002). This potentially speeds up the reaction time in response to changes in the redox poise within the *S. coelicolor* system.

It has been reported previously that Rex acts as a redox sensor, not just an NADH sensor, by detecting differences in the NADH/NAD^+ redox poise (Brekasis and Paget, 2003, Sickmier *et al.*, 2005). Interestingly it appears that NADH is the only cofactor capable of modulating the structure of Rex, NAD^+ has no stabilising effect on DNA-binding (McLaughlin *et al.*, 2010). Thus it would seem that NAD^+ functions as a competitive inhibitor by blocking the NADH-binding site.

Section 6.1.3 – A conserved salt-bridge is essential for Rex functionality

Rex is a dimeric protein containing three structural domains: (1) DNA-binding, (2) NAD(H)-binding and (3) domain-swapped helix (McLaughlin *et al.*, 2010, Sickmier *et al.*, 2005). The binding domains for DNA and NADH are physically separated by the third domain from the opposing chain, and yet NADH binding is able to destabilise DNA-binding. Based on the location of the domain swapped helix this had seemed an obvious means of signal relay between domains, an assumption that proved to be right. A residue at the base of this helix; D188 (T-Rex nomenclature), has a vital role in sensing and responding to redox stress (McLaughlin *et al.*, 2010). The D188, along with R16 and Y98 (T-Rex), form a sensory triad at the centre of Rex. By altering the conformations of each of these residues, a salt-bridge between D188 and R16, of the DNA-binding domain, is formed or broken to stabilise and destabilise the Rex-ROP complex, respectively. The sensory triad residues adopt different conformations depending on the occupancy of the NADH binding site. This discovery has unveiled a mechanistic link between the two functions of Rex; DNA-binding and redox-sensing, furthering our understanding of how this novel repressor operates *in vivo*.

Section 6.1.4 – The regulon

Expansion of the regulon by ChIP-on-chip proved successful; however very little is known of the majority of the newly identified targets. In other systems Rex

has been shown to regulate genes involved in fermentative pathways, such as lactate dehydrogenase and lactate permease (Pagels *et al.*, 2010, Wang *et al.*, 2008). However, the genes while present in the *S. coelicolor* genome are not present in the Rex regulon. In the *B. subtilis* system, both *ndh* and *cydABCD* are also Rex regulated (Gyan *et al.*, 2006, Larsson *et al.*, 2005). The *S. aureus* Rex regulon however appears to be particularly geared towards anaerobic growth, in addition to *ldh* genes it contains *nirC*, *nirR*, *narG* (nitrate/nitrite reduction), as well as alcohol dehydrogenases; *adh1* and *adhE* (Pagels *et al.*, 2010). *S. coelicolor* does not appear to be capable of anaerobic growth, thus perhaps explaining the lack of these genes in the Rex regulon. It is however able to survive long bouts of oxygen limitation (van Keulen *et al.*, 2007). Interestingly within the list of Rex targets were the genes encoding the oxidative stress response regulator, *oxyR*, and the alkyl hydroperoxidase, *ahpC*. These genes appear to be lacking from the regulons of both SA-Rex and B-Rex. It is proposed that NDH-2 is a major source of peroxide in *E. coli* (Messner and Imlay, 1999, Seaver and Imlay, 2004). Therefore OxyR may have a protective role during redox stress due to the derepression and subsequent expression of *ndh*. The induction of *oxyR* in response to Rex de-repression is however yet to be confirmed.

Section 6.2 – Rex – what's next?

The regulon has been expanded to include 22 new bindings sites, however the regulation of these sites has not been confirmed. The strength of binding to each of these sites varies, as does the distance from the annotated start sites of the target genes. Despite having a good knowledge of the mechanism of action of Rex it is still not known how Rex alters gene expression. It is assumed that it blocks polymerase progression but this has not been confirmed. It is also possible that Rex precludes the binding of other regulators involved in other stress responses.

Section 6.2.1 – The regulon and its regulation by Rex

There are a number of approaches that could be used to study the expression of genes, e.g. microarrays or RT-qPCR using cDNA, and S1 nuclease protection assays using RNA. The limitation of these methods for use with the Rex system is that they require the target gene to be expressed in the absence of Rex, which as has previously been observed it not always the case. One approach attempted in this study was the use of the super-repressor strain S106 (*rex*^{G102A}), to emphasize any repression by Rex; however high level expression of this protein was toxic to the cells. The protein does not hinder growth when under the control of its own promoter, thus it may still be a useful construct that simply requires optimisation. For example the use of this strain in combination with other stress inducers may be enough to stimulate expression of a number of targets, which would be hindered by the Rex^{G102A} protein. It is already possible to speculate about the possible inducers of some targets, for example: the use of peroxide to induce *oxyR*/*ahpC*, or the use of phosphate limitation to induce Pho-regulon members; SCO3790, SCO3791 and SCO7697. Even general alterations to the carbon or nitrogen sources may impact expression by altering the metabolic pathways used. From what is known of oxygen limitation sensors one would also expect that nitric oxide or carbon monoxide treatment, or even different means of oxygen limitation such as displacement with argon gas, may impact the transcription profiles of Rex-targeted genes, allowing their regulation to be investigated.

As for how Rex regulates these genes – ideally this requires a combination of DNAse footprinting to map the binding site, and S1 nuclease protection assay (or RNA-seq) to map the transcriptional start sites of the genes. If Rex binds within a transcript then it likely represses by blocking RNAP progression, if overlapping potential -10 and -35 boxes then it is promoter occlusion but if the ROP site lies further upstream then another means of repression is being adopted. Therefore by mapping the ROP site and the transcriptional start site one can begin to investigate how Rex alters the gene expression of its target. From what has been observed of expression of both type I and II NADH dehydrogenases (*ndh* and *nuo*) Rex is not always the sole regulator of a gene thus another means of studying the regulation would be to map the positions of other regulators. By identifying these sites and potentially even characterising

the strength of binding of each one could build a picture of when and why a gene is induced *in vivo*. Other sampling methods could also be adopted as the majority of RNA harvesting has been done from late-exponential liquid cultures, not necessarily providing a full story of the role of Rex in cells. The solid media approach worked well for studying *nuo* expression but ideally harvesting would be done for each layer within the cultures (substrate hyphae and aerial mycelia) in order to account for the differences in oxygen availability at each layer. With the emergence of more sensitive transcriptomic analysis methods such as RNA-seq/RT-qPCR this is becoming more feasible as lower sample quantities are required. Having already characterised the way in which the Rex protein functions the next step is to map the biological role of Rex under all growth stages and culture conditions.

References

Agarwal, N., T. R. Raghunand & W. R. Bishai, (2006) Regulation of the expression of whiB1 in Mycobacterium tuberculosis: role of cAMP receptor protein. *Microbiology* **152**: 2749-2756.

Alberic, P., A. Vennink, S. Cornu, H. Bourennane & A. Bruand, (2009) A snapshot of soil water composition as an indicator of contrasted redox environments in a hedged farmland plot. *Sci Total Environ* **407**: 5719-5725.

Alexeeva, S., K. J. Hellingwerf & M. J. Teixeira de Mattos, (2003) Requirement of ArcA for Redox Regulation in Escherichia coli under Microaerobic but Not Anaerobic or Aerobic Conditions. *J. Bacteriol.* **185**: 204-209.

Andersson, S. G., O. Karlberg, B. Canback & C. G. Kurland, (2003) On the origin of mitochondria: a genomics perspective. *Philos Trans R Soc Lond B Biol Sci* **358**: 165-177; discussion 177-169.

Bandeiras, T. M., C. Salgueiro, A. Kletzin, C. M. Gomes & M. Teixeira, (2002) Acidianus ambivalens type-II NADH dehydrogenase: genetic characterisation and identification of the flavin moiety as FMN. *FEBS Lett* **531**: 273-277.

Bao, K. & S. N. Cohen, (2001) Terminal proteins essential for the replication of linear plasmids and chromosomes in Streptomyces. *Genes Dev* **15**: 1518-1527.

Baruah, A., B. Lindsey, Y. Zhu & M. M. Nakano, (2004) Mutational Analysis of the Signal-Sensing Domain of ResE Histidine Kinase from Bacillus subtilis. *J. Bacteriol.* **186**: 1694-1704.

Bateman, A., (1997) The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends in Biochemical Sciences* **22**: 12-13.

Becker, S., G. Holighaus, T. Gabrielczyk & G. Unden, (1996) O2 as the regulatory signal for FNR-dependent gene regulation in Escherichia coli. *J Bacteriol* **178**: 4515-4521.

Bekker, M., S. Alexeeva, W. Laan, G. Sawers, J. Teixeira de Mattos & K. Hellingwerf, (2010) The ArcBA two-component system of Escherichia coli is regulated by the redox state of both the ubiquinone and the menaquinone pool. *J Bacteriol* **192**: 746-754.

Bennett, J. A., R. M. Aimino & J. R. McCormick, (2007) Streptomyces coelicolor genes ftsL and divIC play a role in cell division but are dispensable for colony formation. *J Bacteriol* **189**: 8982-8992.

Bentley, S. D., S. Brown, L. D. Murphy, D. E. Harris, M. A. Quail, J. Parkhill, B. G. Barrell, J. R. McCormick, R. I. Santamaria, R. Losick, M. Yamasaki, H. Kinashi, C. W. Chen, G. Chandra, D. Jakimowicz, H. M. Kieser, T. Kieser & K. F. Chater, (2004) SCP1, a 356,023 bp linear plasmid adapted to the ecology and developmental biology of its host, Streptomyces coelicolor A3(2). *Mol Microbiol* **51**: 1615-1628.

Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A.

Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill & D. A. Hopwood, (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **417**: 141-147.

Bertsova, Y. V., A. V. Bogachev & V. P. Skulachev, (1998) Two NADH:ubiquinone oxidoreductases of Azotobacter vinelandii and their role in the respiratory protection. *Biochim Biophys Acta* **1363**: 125-133.

Bierman, M., R. Logan, K. O'Brien, E. T. Seno, R. N. Rao & B. E. Schoner, (1992) Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **116**: 43-49.

Borodina, I., P. Krabben & J. Nielsen, (2005) Genome-scale analysis of Streptomyces coelicolor A3(2) metabolism. *Genome Res* **15**: 820-829.

Brekasis, D., (2005) Identification and characterisation of Rex, a novel sensor of the NADH/NAD^+ redox poise in *Streptomyces coelicolor*. DPhil Thesis, School of Life Sciences, University of Sussex, Falmer.

Brekasis, D. & M. S. B. Paget, (2003) A novel sensor of NADH/NAD+ redox poise in Streptomyces coelicolor A3(2). *EMBO J.* **22**: 4856–4865.

Brown, G. C., (2001) Regulation of mitochondrial respiration by nitric oxide inhibition of cytochrome c oxidase. *Biochim Biophys Acta* **1504**: 46-57.

Bryk, R., C. D. Lima, H. Erdjument-Bromage, P. Tempst & C. Nathan, (2002) Metabolic enzymes of mycobacteria linked to antioxidant defense by a thioredoxin-like protein. *Science* **295**: 1073-1077.

Bucca, G., E. Laing, V. Mersinias, N. Allenby, D. Hurd, J. Holdstock, V. Brenner, M. Harrison & C. P. Smith, (2009) Development and application of versatile high density microarrays for genome-wide analysis of Streptomyces coelicolor: characterization of the HspR regulon. *Genome Biol* **10**: R5.

Calhoun, M. W. & R. B. Gennis, (1993) Demonstration of separate genetic loci encoding distinct membrane-bound respiratory NADH dehydrogenases in Escherichia coli. *J Bacteriol* **175**: 3013-3019.

Calhoun, M. W., K. L. Oden, R. B. Gennis, M. J. de Mattos & O. M. Neijssel, (1993) Energetic efficiency of Escherichia coli: effects of mutations in components of the aerobic respiratory chain. *J. Bacteriol.* **175**: 3020-3025.

Camougrand, N. M., R. B. Caubet & M. G. Guerin, (1983) Evidence for an alternative and non-phosphorylating pathway for NADH reoxidation in a yeast strain resistant to glucose repression. *Eur J Biochem* **135**: 367-371.

Challis, G. L. & D. A. Hopwood, (2003) Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. *Proc Natl Acad Sci U S A* **100 Suppl 2**: 14555-14561.

Chater, K. & R. Losick, (1996) The mycelial life-style of *Streptomyces coelicolor* A3(2) and its relatives. In: Bacteria as Multicellular Organisms. J. Shapiro & M. Dworkin (eds). New York: Oxford University Press, pp. 149-182.

Chater, K. F., (1972) A Morphological and Genetic Mapping Study of White Colony Mutants of Streptomyces coelicolor. *J Gen Microbiol* **72**: 9-28.

Chater, K. F., (1998) Taking a genetic scalpel to the Streptomyces colony. *Microbiology* **144**: 1465-1478.

Chen, S., J. Vojtechovsky, G. N. Parkinson, R. H. Ebright & H. M. Berman, (2001) Indirect readout of DNA sequence at the primary-kink site in the CAP-DNA complex: DNA binding specificity based on energetics of DNA kinking. *J Mol Biol* **314**: 63-74.

Cho, H. Y., H. J. Cho, Y. M. Kim, J. I. Oh & B. S. Kang, (2009) Structural insight into the heme-based redox sensing by DosS from Mycobacterium tuberculosis. *J Biol Chem* **284**: 13057-13067.

Choi, H., S. Kim, P. Mukhopadhyay, S. Cho, J. Woo, G. Storz & S. E. Ryu, (2001) Structural basis of the redox switch in the OxyR transcription factor. *Cell* **105**: 103-113.

Chong, P. P., S. M. Podmore, H. M. Kieser, M. Redenbach, K. Turgay, M. Marahiel, D. A. Hopwood & C. P. Smith, (1998) Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of Streptomyces coelicolor A3(2). *Microbiology* **144 (Pt 1)**: 193-199.

Claessen, D., R. Rink, W. de Jong, J. Siebring, P. de Vreugd, F. G. Boersma, L. Dijkhuizen & H. A. Wosten, (2003) A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in Streptomyces coelicolor by forming amyloid-like fibrils. *Genes Dev* **17**: 1714-1726.

Colangeli, R., A. Haq, V. L. Arcus, E. Summers, R. S. Magliozzo, A. McBride, A. K. Mitra, M. Radjainia, A. Khajo, W. R. Jacobs, Jr., P. Salgame & D. Alland, (2009) The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc Natl Acad Sci U S A* **106**: 4414-4418.

Cosgrove, K., G. Coutts, I. M. Jonsson, A. Tarkowski, J. F. Kokai-Kun, J. J. Mond & S. J. Foster, (2007) Catalase (KatA) and alkyl hydroperoxide reductase (AhpC) have compensatory roles in peroxide stress resistance and are required for survival, persistence, and nasal colonization in Staphylococcus aureus. *J Bacteriol* **189**: 1025-1035.

Crack, J. C., N. E. Le Brun, A. J. Thomson, J. Green & A. J. Jervis, (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.

Crooks, G. E., G. Hon, J. M. Chandonia & S. E. Brenner, (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.

Damerell, D., (2007) Silent Wizard. http://silentwizard.nixbioinf.org/

Datsenko, K. A. & B. L. Wanner, (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.

Davis, N. K. & K. F. Chater, (1990) Spore colour in Streptomyces coelicolor A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Mol Microbiol* **4**: 1679-1691.

Davis, N. K. & K. F. Chater, (1992) The Streptomyces coelicolor whiB gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet* **232**: 351-358.

Dedrick, R. M., H. Wildschutte & J. R. McCormick, (2009) Genetic interactions of smc, ftsK, and parB genes in Streptomyces coelicolor and their developmental genome segregation phenotypes. *J Bacteriol* **191**: 320-332.

Efremov, R. G., R. Baradaran & L. A. Sazanov, (2010) The architecture of respiratory complex I. *Nature* **465**: 441-445.

Efromovich, S., D. Grainger, D. Bodenmiller, S. Spiro & K. P. Robert, (2008) Genome-Wide Identification of Binding Sites for the Nitric Oxide-Sensitive Transcriptional Regulator NsrR. In: Methods in Enzymology. Academic Press, pp. 211-233.

Elliot, M. A., M. J. Bibb, M. J. Buttner & B. K. Leskiw, (2001) BldD is a direct regulator of key developmental genes in Streptomyces coelicolor A3(2). *Molecular Microbiology* **40**: 257-269.

Elliot, M. A., T. R. Locke, C. M. Galibois & B. K. Leskiw, (2003) BldD from Streptomyces coelicolor is a non-essential global regulator that binds its own promoter as a dimer. *FEMS Microbiol Lett* **225**: 35-40.

Esbelin, J., J. Armengaud, A. Zigha & C. Duport, (2009) ResDE-dependent regulation of enterotoxin gene expression in Bacillus cereus: evidence for multiple modes of binding for ResD and interaction with Fnr. *J Bacteriol* **191**: 4419-4426.

Esterhazy, D., M. S. King, G. Yakovlev & J. Hirst, (2008) Production of reactive oxygen species by complex I (NADH:ubiquinone oxidoreductase) from Escherichia coli and comparison to the enzyme from mitochondria. *Biochemistry* **47**: 3964-3971.

Felden, B., H. Himeno, A. Muto, J. P. McCutcheon, J. F. Atkins & R. F. Gesteland, (1997) Probing the structure of the Escherichia coli 10Sa RNA (tmRNA). *RNA* **3**: 89-103.

Fischer, M., J. Alderson, G. van Keulen, J. White & G. Sawers, (2010) The obligate aerobe Streptomyces coelicolor A3(2) synthesizes three active respiratory nitrate reductases. *Microbiology*.

Flärdh, K. & M. J. Buttner, (2009) Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* **7**: 36-49.

Garcia-Contreras, R., H. Celis & I. Romero, (2004) Importance of Rhodospirillum rubrum H(+)-pyrophosphatase under low-energy conditions. *J Bacteriol* **186**: 6651-6655.

Geng, H., P. Zuber & M. M. Nakano, (2007) Regulation of respiratory genes by ResD-ResE signal transduction system in Bacillus subtilis. *Methods Enzymol* **422**: 448-464.

Georgellis, D., A. S. Lynch & E. C. Lin, (1997) In vitro phosphorylation study of the arc two-component signal transduction system of Escherichia coli. *J Bacteriol* **179**: 5429-5435.

Gottelt, M., S. Kol, J. P. Gomez-Escribano, M. Bibb & E. Takano, (2010) Deletion of a regulatory gene within the cpk gene cluster reveals novel antibacterial activity in Streptomyces coelicolor A3(2). *Microbiology* **156**: 2343-2353.

Green, J. & M. S. Paget, (2004) Bacterial redox sensors. *Nat Rev Microbiol* **2**: 954-966.

Gust, B., G. Chandra, D. Jakimowicz, T. Yuqing, C. J. Bruton & K. F. Chater, (2004) Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Adv Appl Microbiol* **54**: 107-128.

Gust, B., Kieser, T., Chater, KF, (2002) REDIRECT© technology: PCR-targeting system in Streptomyces coelicolor.

Gyan, S., Y. Shiohira, I. Sato, M. Takeuchi & T. Sato, (2006) Regulatory loop between redox sensing of the NADH/NAD(+) ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in Bacillus subtilis. *J Bacteriol* **188**: 7062-7071.

Hahn, J. S., S. Y. Oh & J. H. Roe, (2002) Role of OxyR as a peroxide-sensing positive regulator in Streptomyces coelicolor A3(2). *J Bacteriol* **184**: 5214-5222.

Hesketh, A., H. Kock, S. Mootien & M. Bibb, (2009) The role of absC, a novel regulatory gene for secondary metabolism, in zinc-dependent antibiotic production in Streptomyces coelicolor A3(2). *Mol Microbiol* **74**: 1427-1444.

Hirono, M. & M. Maeshima, (2009) Functional enhancement by single-residue substitution of Streptomyces coelicolor A3(2) H+-translocating pyrophosphatase. *J Biochem* **146**: 617-621.

Homerova, D., J. Sevcikova & J. Kormanec, (2003) Characterization of the Streptomyces coelicolor A3(2) wblE gene, encoding a homologue of the sporulation transcription factor. *Folia Microbiol (Praha)* **48**: 489-495.

Hoogerheide, J. C., (1975) Studies on the energy metabolism during anaerobic fermentation of glucose by baker's yeast. *Radiat Environ Biophys* **11**: 295-307.

Hopwood, D., K. F. Chater & M. Bibb, (1995). In: Genetics and Biochemistry of Antibiotic Production. L. Vining & C. Stuttard (eds). Newton, Massachusetts: Butterworth-Heinemann, pp. 65-102.

Hoskisson, P. A., S. Rigali, K. Fowler, K. C. Findlay & M. J. Buttner, (2006) DevA, a GntR-like transcriptional regulator required for development in Streptomyces coelicolor. *J Bacteriol* **188**: 5014-5023.

Iuchi, S. & E. C. Lin, (1988) arcA (dye), a global regulatory gene in Escherichia coli mediating repression of enzymes in aerobic pathways. *Proc Natl Acad Sci U S A* **85**: 1888-1892.

Iuchi, S., Z. Matsuda, T. Fujiwara & E. C. Lin, (1990) The arcB gene of Escherichia coli encodes a sensor-regulator protein for anaerobic repression of the arc modulon. *Mol Microbiol* **4**: 715-727.

Jakimowicz, P., M. R. Cheesman, W. R. Bishai, K. F. Chater, A. J. Thomson & M. J. Buttner, (2005) Evidence That the Streptomyces Developmental Protein WhiD, a Member of the WhiB Family, Binds a [4Fe-4S] Cluster. *Journal of Biological Chemistry* **280**: 8309-8315.

Jaurin, B. & S. N. Cohen, (1985) Streptomyces contain Escherichia coli-type A + T-rich promoters having novel structural features. *Gene* **39**: 191-201.

Jaworowski, A., H. D. Campbell, M. I. Poulis & I. G. Young, (1981) Genetic identification and purification of the respiratory NADH dehydrogenase of Escherichia coli. *Biochemistry* **20**: 2041-2047.

Jensen, L. J., M. Kuhn, M. Stark, S. Chaffron, C. Creevey, J. Muller, T. Doerks, P. Julien, A. Roth, M. Simonovic, P. Bork & C. von Mering, (2009) STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* **37**: D412-416.

Jentsch, S., (1996) When proteins receive deadly messages at birth. *Science* **271**: 955-956.

Junemann, S., (1997) Cytochrome bd terminal oxidase. *Biochim Biophys Acta* **1321**: 107-127.

Kallifidas, D., B. Pascoe, G. A. Owen, C. M. Strain-Damerell, H. J. Hong & M. S. Paget, (2010) The zinc-responsive regulator Zur controls expression of the coelibactin gene cluster in Streptomyces coelicolor. *J Bacteriol* **192**: 608-611.

Keiler, K. C., P. R. Waller & R. T. Sauer, (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**: 990-993.

Kerscher, S., S. Drose, V. Zickermann & U. Brandt, (2008) The three families of respiratory NADH dehydrogenases. *Results Probl Cell Differ* **45**: 185-222.

Kieser, T., M. Bibb, M. Buttner, K. Chater & D. Hopwood, (2000) *Practical Streptomyces Genetics (2nd ed.)*. John Innes Foundation, Norwich, England.

Kim, I. K., C. J. Lee, M. K. Kim, J. M. Kim, J. H. Kim, H. S. Yim, S. S. Cha & S. O. Kang, (2006) Crystal structure of the DNA-binding domain of BldD, a central regulator of aerial mycelium formation in Streptomyces coelicolor A3(2). *Mol Microbiol* **60**: 1179-1193.

Kita, K., K. Konishi & Y. Anraku, (1984) Terminal oxidases of Escherichia coli aerobic respiratory chain. II. Purification and properties of cytochrome b558-d complex from cells grown with limited oxygen and evidence of branched electron-carrying systems. *J Biol Chem* **259**: 3375-3381.

Kois, A., M. Swiatek, D. Jakimowicz & J. Zakrzewska-Czerwinska, (2009) SMC protein-dependent chromosome condensation during aerial hyphal development in Streptomyces. *J Bacteriol* **191**: 310-319.

Krueger, C., C. Berens, A. Schmidt, D. Schnappinger & W. Hillen, (2003) Single-chain Tet transregulators. *Nucleic Acids Res* **31**: 3050-3056.

Kumar, A., J. C. Toledo, R. P. Patel, J. R. Lancaster, Jr. & A. J. Steyn, (2007) Mycobacterium tuberculosis DosS is a redox sensor and DosT is a hypoxia sensor. *Proc Natl Acad Sci U S A* **104**: 11568-11573.

Kwon, O., D. Georgellis & E. C. Lin, (2000) Phosphorelay as the sole physiological route of signal transmission by the arc two-component system of Escherichia coli. *J Bacteriol* **182**: 3858-3862.

Larson, J. L. & C. L. Hershberger, (1986) The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. *Plasmid* **15**: 199-209.

Larsson, J. T., A. Rogstam & C. von Wachenfeldt, (2005) Coordinated patterns of cytochrome bd and lactate dehydrogenase expression in Bacillus subtilis. *Microbiology* **151**: 3323-3335.

Lazazzera, B. A., H. Beinert, N. Khoroshilova, M. C. Kennedy & P. J. Kiley, (1996) DNA binding and dimerization of the Fe-S-containing FNR protein from Escherichia coli are regulated by oxygen. *J Biol Chem* **271**: 2762-2768.

Leistikow, R. L., R. A. Morton, I. L. Bartek, I. Frimpong, K. Wagner & M. I. Voskuil, (2010) The Mycobacterium tuberculosis DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J Bacteriol* **192**: 1662-1670.

Leskiw, B. K., R. Mah, E. J. Lawlor & K. F. Chater, (1993) Accumulation of bldA-specified tRNA is temporally regulated in Streptomyces coelicolor A3(2). *J Bacteriol* **175**: 1995-2005.

Loui, C., A. C. Chang & S. Lu, (2009) Role of the ArcAB two-component system in the resistance of Escherichia coli to reactive oxygen stress. *BMC Microbiol* **9**: 183.

Luijsterburg, M. S., M. F. White, R. van Driel & R. T. Dame, (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* **43**: 393-418.

MacMicking, J., Q. W. Xie & C. Nathan, (1997) Nitric oxide and macrophage function. *Annu Rev Immunol* **15**: 323-350.

MacNeil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons & T. MacNeil, (1992) Analysis of Streptomyces avermitilis genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**: 61-68.

Malpica, R., B. Franco, C. Rodriguez, O. Kwon & D. Georgellis, (2004) Identification of a quinone-sensitive redox switch in the ArcB sensor kinase. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 13318-13323.

McLaughlin, K. J., C. M. Strain-Damerell, K. Xie, D. Brekasis, A. S. Soares, M. S. Paget & C. L. Kielkopf, (2010) Structural Basis for NADH/NAD(+) Redox Sensing by a Rex Family Repressor. *Mol Cell* **38**: 563-575.

Merrick, M. J., (1976) A morphological and genetic mapping study of bald colony mutants of Streptomyces coelicolor. *J Gen Microbiol* **96**: 299-315.

Messner, K. R. & J. A. Imlay, (1999) The identification of primary sites of superoxide and hydrogen peroxide formation in the aerobic respiratory chain and sulfite reductase complex of Escherichia coli. *J Biol Chem* **274**: 10119-10128.

Mettert, E. L. & P. J. Kiley, (2005) ClpXP-dependent Proteolysis of FNR upon Loss of its O2-sensing [4Fe-4S] Cluster. *Journal of Molecular Biology* **354**: 220-232.

Mettert, E. L., F. W. Outten, B. Wanta & P. J. Kiley, (2008) The impact of O(2) on the Fe-S cluster biogenesis requirements of Escherichia coli FNR. *J Mol Biol* **384**: 798-811.

Minchin, S. D. & S. J. Busby, (2009) Analysis of mechanisms of activation and repression at bacterial promoters. *Methods* **47**: 6-12.

Nakamura, A., A. Sosa, H. Komori, A. Kita & K. Miki, (2007) Crystal structure of TTHA1657 (AT-rich DNA-binding protein; p25) from Thermus thermophilus HB8 at 2.16 A resolution. *Proteins* **66**: 755-759.

Negre, N., S. Lavrov, J. Hennetin, M. Bellis & G. Cavalli, (2006) Mapping the distribution of chromatin proteins by ChIP on chip. *Methods Enzymol* **410**: 316-341.

Pagels, M., S. Fuchs, J. Pane-Farre, C. Kohler, L. Menschner, M. Hecker, P. J. McNamarra, M. C. Bauer, C. von Wachenfeldt, M. Liebeke, M. Lalk, G. Sander, C. von Eiff, R. A. Proctor & S. Engelmann, (2010) Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in Staphylococcus aureus. *Mol Microbiol* **76**: 1142-1161.

Pascoe, B., (2009) Functionl analysis of Zur, the zinc uptake regulator of streptomyces coelicolor A3(2). DPhil Thesis, School of Life Sciences, University of Sussex, Falmer.

Pawlik, K., M. Kotowska, K. F. Chater, K. Kuczek & E. Takano, (2007) A cryptic type I polyketide synthase (cpk) gene cluster in Streptomyces coelicolor A3(2). *Arch Microbiol* **187**: 87-99.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng & T. E. Ferrin, (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.

Pillai, S. & S. P. Chellappan, (2009) ChIP on chip assays: genome-wide analysis of transcription factor binding and histone modifications. *Methods Mol Biol* **523**: 341-366.

Pimkin, M., J. Pimkina & G. D. Markham, (2009) A regulatory role of the Bateman domain of IMP dehydrogenase in adenylate nucleotide biosynthesis. *J Biol Chem* **284**: 7960-7969.

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

Poole, L. B., (1996) Flavin-dependent alkyl hydroperoxide reductase from Salmonella typhimurium. 2. Cystine disulfides involved in catalysis of peroxide reduction. *Biochemistry* **35**: 65-75.

Poole, R. K. & G. M. Cook, (2000) Redundancy of aerobic respiratory chains in bacteria? Routes, reasons and regulation. *Adv Microb Physiol* **43**: 165-224.

Portnoy, V. A., D. A. Scott, N. E. Lewis, Y. Tarasova, A. L. Osterman & B. O. Palsson, (2010) Deletion of cytochrome oxidases and quinol monooxygenase blocks aerobic-anaerobic shift in Escherichia coli K-12 MG1655. *Appl Environ Microbiol*. **76**: 6529-40.

Pul, U., R. Wurm & R. Wagner, (2007) The role of LRP and H-NS in transcription regulation: involvement of synergism, allostery and macromolecular crowding. *J Mol Biol* **366**: 900-915.

Rao, S. T. & M. G. Rossmann, (1973) Comparison of super-secondary structures in proteins. *J Mol Biol* **76**: 241-256.

Reents, H., R. Munch, T. Dammeyer, D. Jahn & E. Hartig, (2006) The Fnr regulon of Bacillus subtilis. *J Bacteriol* **188**: 1103-1112.

Rosenfeld, N., M. B. Elowitz & U. Alon, (2002) Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* **323**: 785-793.

Rouviere-Yaniv, J. & F. Gros, (1975) Characterization of a novel, low-molecular-weight DNA-binding protein from Escherichia coli. *Proc Natl Acad Sci U S A* **72**: 3428-3432.

Rustad, T. R., A. M. Sherrid, K. J. Minch & D. R. Sherman, (2009) Hypoxia: a window into Mycobacterium tuberculosis latency. *Cell Microbiol* **11**: 1151-1159.

Sala, C., D. C. Grainger & S. T. Cole, (2009) Dissecting regulatory networks in host-pathogen interaction using chIP-on-chip technology. *Cell Host Microbe* **5**: 430-437.

Salerno, P., J. Larsson, G. Bucca, E. Laing, C. P. Smith & K. Flardh, (2009) One of the Two Genes Encoding Nucleoid-Associated HU Proteins in Streptomyces coelicolor Is Developmentally Regulated and Specifically Involved in Spore Maturation. *J. Bacteriol.* **191**: 6489-6500.

Sardiwal, S., S. L. Kendall, F. Movahedzadeh, S. C. G. Rison, N. G. Stoker & S. Djordjevic, (2005) A GAF Domain in the Hypoxia/NO-inducible Mycobacterium tuberculosis DosS Protein Binds Haem. *Journal of Molecular Biology* **353**: 929-936.

Scott, J. W., S. A. Hawley, K. A. Green, M. Anis, G. Stewart, G. A. Scullion, D. G. Norman & D. G. Hardie, (2004) CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. *J Clin Invest* **113**: 274-284.

Seaver, L. C. & J. A. Imlay, (2004) Are respiratory enzymes the primary sources of intracellular hydrogen peroxide? *J Biol Chem* **279**: 48742-48750.

Sickmier, E. A., D. Brekasis, S. Paranawithana, J. B. Bonanno, M. S. B. Paget, S. K. Burley & C. L. Kielkopf, (2005) X-Ray Structure of a Rex-Family Repressor/NADH Complex Insights into the Mechanism of Redox Sensing. *Structure* **13**: 43-54.

Singh, A., L. Guidry, K. V. Narasimhulu, D. Mai, J. Trombley, K. E. Redding, G. I. Giles, J. R. Lancaster & A. J. C. Steyn, (2007) Mycobacterium tuberculosis WhiB3 responds to O2 and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proceedings of the National Academy of Sciences* **104**: 11562-11567.

Sola-Landa, A., A. Rodriguez-Garcia, A. K. Apel & J. F. Martin, (2008) Target genes and structure of the direct repeats in the DNA-binding sequences of the response regulator PhoP in Streptomyces coelicolor. *Nucl. Acids Res.* **36**: 1358-1368.

Spencer, W., R. Siam, M. C. Ouimet, D. P. Bastedo & G. T. Marczynski, (2009) CtrA, a global response regulator, uses a distinct second category of weak DNA binding sites for cell cycle transcription control in Caulobacter crescentus. *J Bacteriol* **191**: 5458-5470.

Stratigopoulos, G., A. R. Gandecha & E. Cundliffe, (2002) Regulation of tylosin production and morphological differentiation in Streptomyces fradiae by TylP, a deduced gamma-butyrolactone receptor. *Mol Microbiol* **45**: 735-744.

Sun, G., E. Sharkova, R. Chesnut, S. Birkey, M. Duggan, A. Sorokin, P. Pujic, S. Ehrlich & F. Hulett, (1996) Regulators of aerobic and anaerobic respiration in Bacillus subtilis. *J. Bacteriol.* **178**: 1374-1385.

Takano, E., (2006) Gamma-butyrolactones: Streptomyces signalling molecules regulating antibiotic production and differentiation. *Curr Opin Microbiol* **9**: 287-294.

Takano, E., H. Kinoshita, V. Mersinias, G. Bucca, G. Hotchkiss, T. Nihira, C. P. Smith, M. Bibb, W. Wohlleben & K. Chater, (2005) A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of Streptomyces coelicolor. *Mol Microbiol* **56**: 465-479.

Thomas-Chollier, M., O. Sand, J. V. Turatsinze, R. Janky, M. Defrance, E. Vervisch, S. Brohee & J. van Helden, (2008) RSAT: regulatory sequence analysis tools. *Nucleic Acids Res* **36**: W119-127.

Tolla, D. A. & M. A. Savageau, (2010) Regulation of aerobic-to-anaerobic transitions by the FNR cycle in Escherichia coli. *J Mol Biol* **397**: 893-905.

Tran, Q. H., J. Bongaerts, D. Vlad & G. Unden, (1997) Requirement for the proton-pumping NADH dehydrogenase I of Escherichia coli in respiration of NADH to fumarate and its bioenergetic implications. *Eur J Biochem* **244**: 155-160.

Tseng, C. P., J. Albrecht & R. P. Gunsalus, (1996) Effect of microaerophilic cell growth conditions on expression of the aerobic (cyoABCDE and cydAB) and anaerobic (narGHJI, frdABCD, and dmsABC) respiratory pathway genes in Escherichia coli. *J Bacteriol* **178**: 1094-1098.

Tu, G.-F., G. E. Reid, J.-G. Zhang, R. L. Moritz & R. J. Simpson, (1995) C-terminal Extension of Truncated Recombinant Proteins in Escherichia coli with a 10Sa RNA Decapeptide. *Journal of Biological Chemistry* **270**: 9322-9326.

Unden, G. & J. Bongaerts, (1997) Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. *Biochim Biophys Acta* **1320**: 217-234.

van Keulen, G., J. Alderson, J. White & R. G. Sawers, (2007) The obligate aerobic actinomycete Streptomyces coelicolor A3(2) survives extended periods of anaerobic stress. *Environ Microbiol* **9**: 3143-3149.

van Keulen, G., H. M. Jonkers, D. Claessen, L. Dijkhuizen & H. A. Wosten, (2003) Differentiation and anaerobiosis in standing liquid cultures of Streptomyces coelicolor. *J Bacteriol* **185**: 1455-1458.

Voskuil, M. I., D. Schnappinger, K. C. Visconti, M. I. Harrell, G. M. Dolganov, D. R. Sherman & G. K. Schoolnik, (2003) Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program. *J Exp Med* **198**: 705-713.

Wang, E., M. C. Bauer, A. Rogstam, S. Linse, D. T. Logan & C. von Wachenfeldt, (2008) Structure and functional properties of the Bacillus subtilis transcriptional repressor Rex. *Mol Microbiol* **69**: 466-478.

Weinstein, E. A., T. Yano, L. S. Li, D. Avarbock, A. Avarbock, D. Helm, A. A. McColm, K. Duncan, J. T. Lonsdale & H. Rubin, (2005) Inhibitors of type II NADH:menaquinone oxidoreductase represent a class of antitubercular drugs. *Proc Natl Acad Sci U S A* **102**: 4548-4553.

Widdick, D. A., K. Dilks, G. Chandra, A. Bottrill, M. Naldrett, M. Pohlschroder & T. Palmer, (2006) The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. *Proc Natl Acad Sci U S A* **103**: 17927-17932.

Williams, K. P. & D. P. Bartel, (1998) The tmRNA Website. *Nucleic Acids Res* **26**: 163-165.

Wright, L. F. & D. A. Hopwood, (1976) Actinorhodin is a chromosomally-determined antibiotic in Streptomyces coelicolar A3(2). *J Gen Microbiol* **96**: 289-297.

Yin, F. F., S. Bailey, C. A. Innis, M. Ciubotaru, S. Kamtekar, T. A. Steitz & D. G. Schatz, (2009) Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. *Nat Struct Mol Biol* **16**: 499-508.

Zheng, M., F. Aslund & G. Storz, (1998) Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**: 1718-1721.

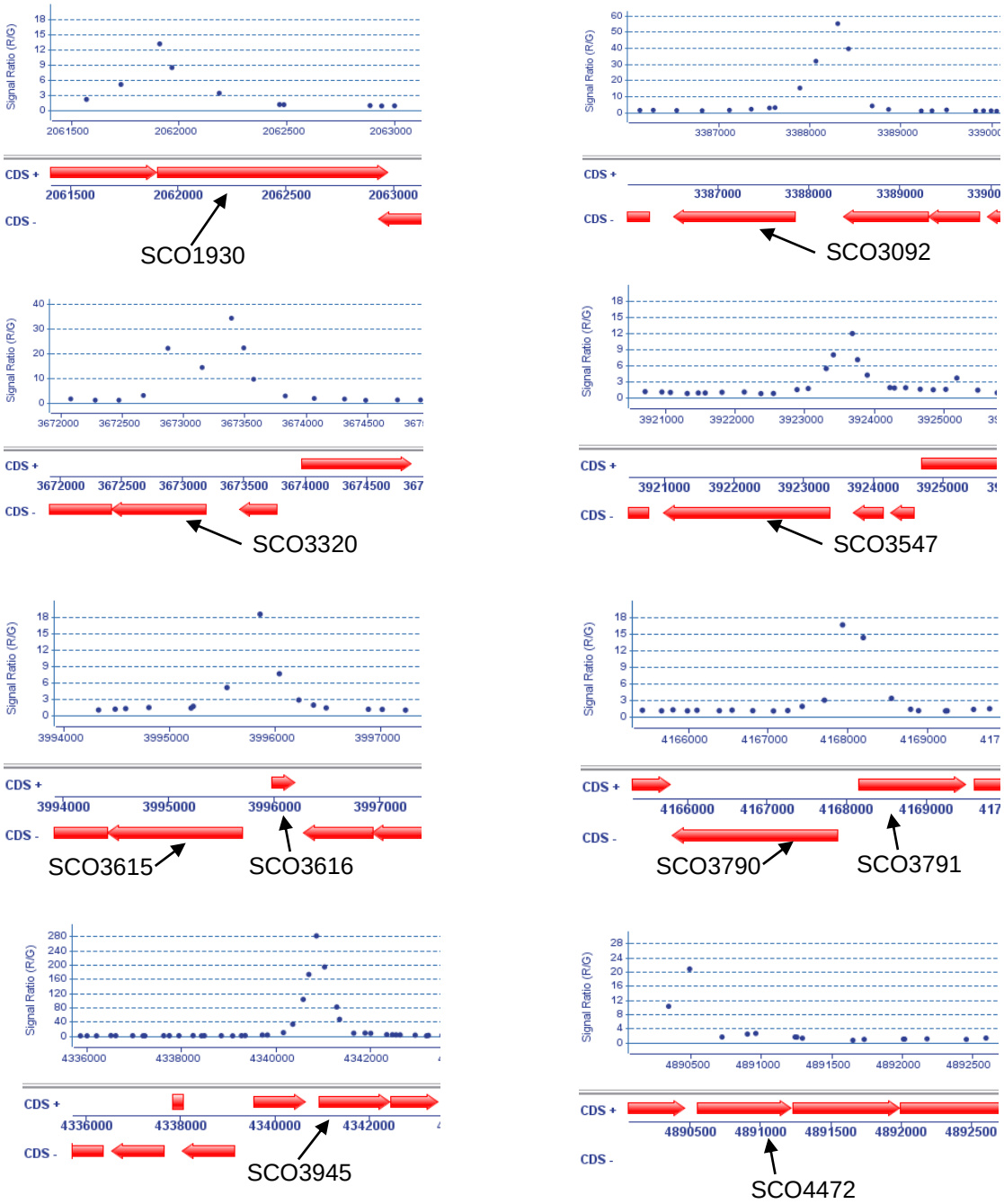
Appendix

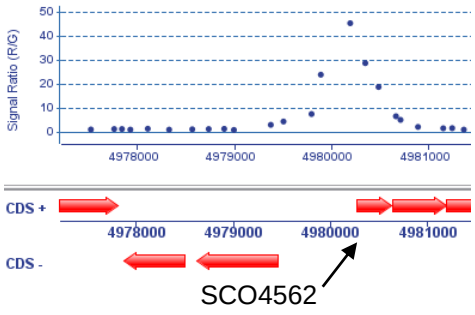
Supplementary Materials

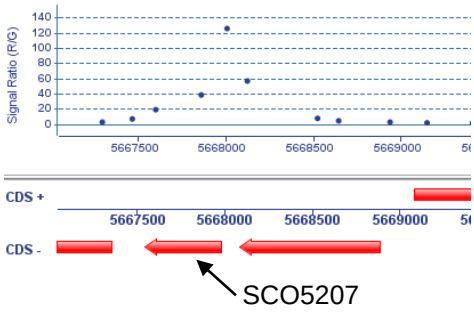
Table S1: The composition of samples used in each hybridisation chamber of the ChIP-on-chip microarray slide. For each strain 2 samples were labelled with Cy3 and two with Cy5 – constituting a dye match. The samples in chamber 4 were a mixture of biological replicates in order to generate a large enough test amount (~150ng).

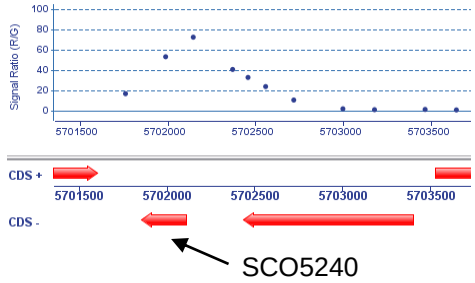
Chamber	Strain	Replicate	Label
1	S106 (pSX402)	1	Cy5
	S106 (pSET152)	2	Cy3
2	S106 (pSX402)	1	Cy3
	S106 (pSET152)	2	Cy5
3	S106 (pSX402)	2	Cy5
	S106 (pSET152)	1	Cy3
4	S106 (pSX402)	1,2,3	Cy3
	S106 (pSET152)	2,3	Cy5

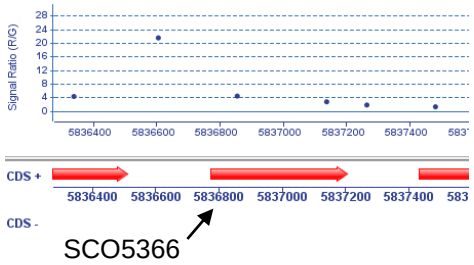
Figure S1: Genomic location and signal enrichment for the Rex targets that were identified by ChIP-on-chip. Regions were viewed using Chip Browser and the target gene in each case is indicated above.

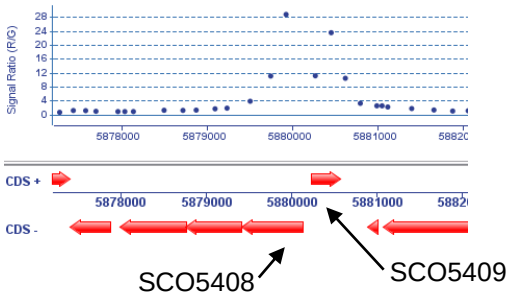


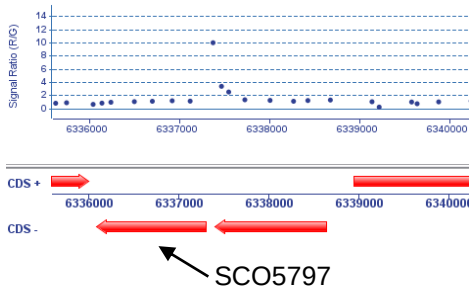

Signal Ratio (R/G)
CDS +
CDS -
SCO4562

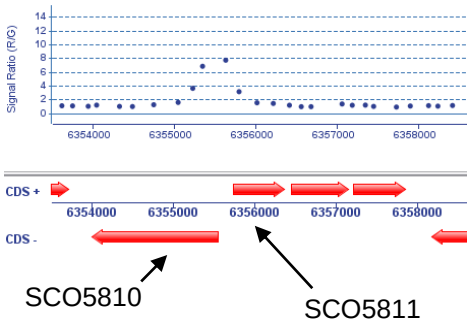

Signal Ratio (R/G)
CDS +
CDS -
SCO5207

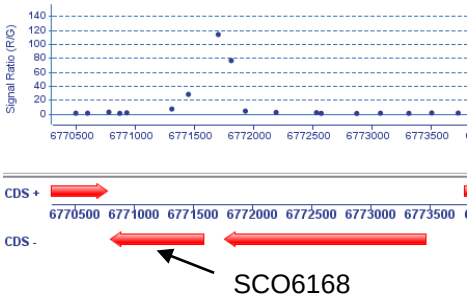

Signal Ratio (R/G)
CDS +
CDS -
SCO5240

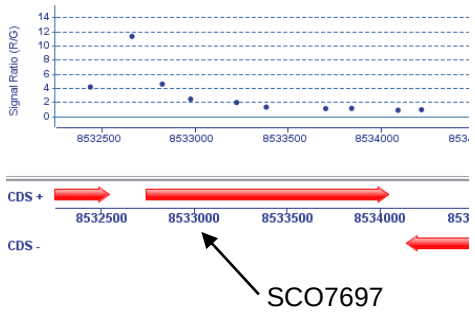

Signal Ratio (R/G)
CDS +
CDS -
SCO5366


Signal Ratio (R/G)
CDS +
CDS -
SCO5408
SCO5409


Signal Ratio (R/G)
CDS +
CDS -
SCO5797


Signal Ratio (R/G)
CDS +
CDS -
SCO5810
SCO5811


Signal Ratio (R/G)
CDS +
CDS -
SCO6168


Signal Ratio (R/G)
CDS +
CDS -
SCO7697

Previously identified sites	Newly identified sites
SCO3945 (*cydA*)	SCO1930
SCO3092 (*ndh*)	SCO2370/1 (*aceE2*)
SCO4562 (*nuoA*)	SCO3101
SCO5366 (*atpI*)	SCO3137 (*galE1*)
SCO5240 (*wblE*)	SCO3547 (*hppA*)
SCO3320 (*rex*)	SCO3615/6* (*ask*)
SCO4280/1*	SCO3790/1*
SCO4472 (*resA*)	SCO4461/2*
	SCO5013
	SCO5032/3* (*ahpC/oxyR*)
	SCO5207
	SCO5408/9*
	SCO5435/6* (*dcuS/dctA*)
	SCO5797
	SCO5810/1*
	SCO6168
	SCO6218/9*
	SCO6239
	SCO6280 (*cpkO*)
	SCO6383
	SCO6917
	SCO7697 (*phyC*)

Table S2: Comparison of previously identified ROP sites with newly identified sites, as identified by ChIP-on-chip. The * indicates a site present in the intergenic region between two divergent genes.

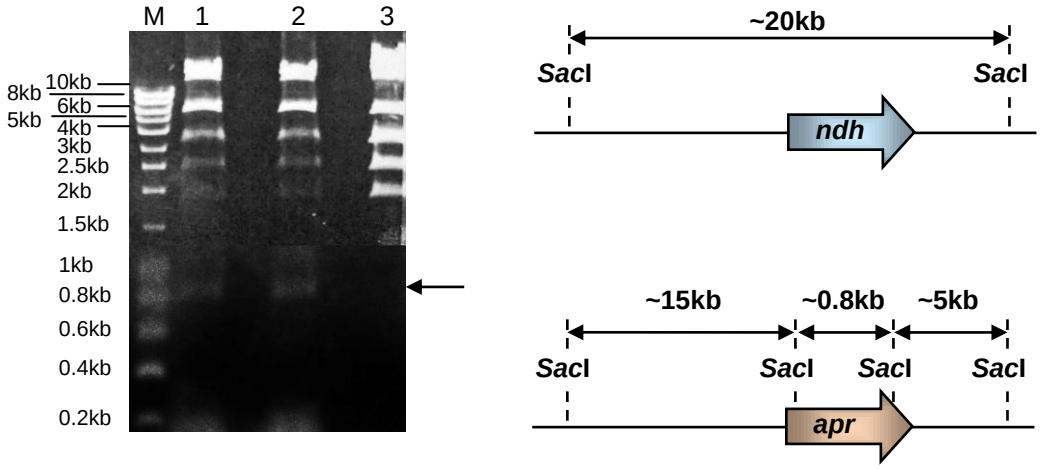


Figure S2: Agarose gel electrophoresis of *Sac*I-cut *S. coelicolor* cosmid StE25 (lane 3) and *ndh* disruption cosmid (lanes 1 and 2). Samples were digested for 3 hours and analysed on a 1% agarose-TBE gel alongside a DNA-ladder (HyperLadder I - Bioline). The diagram on the right-hand side shows the approximate positions of the *Sac*I restriction sites in each cosmid. The presence of the apramycin resistance cassette in the recombinant cosmid is indicated by the 800bp fragment in lanes 1 and 2, generated by internal sites within the apramycin gene.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
S-Rex	M	A	T	G	R	A	H	R	P	A	T	R	S	R	G	I	P	E	A	T	V	A	R	L	P	L	Y
T-Rex	-	-	-	-	-	-	-	-	-	-	-	-	-	M	K	V	P	E	A	A	I	S	R	L	I	T	Y
	-	-	-	-	-	-	-	-	-	-	-	-	-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
S-Rex	L	R	A	L	T	A	L	S	E	R	S	V	P	T	V	S	S	E	E	L	A	A	A	A	G	V	N
T-Rex	L	R	I	L	E	E	L	E	A	Q	G	V	H	R	T	S	S	E	Q	L	G	E	L	A	Q	V	T
	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41

	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81
S-Rex	S	A	K	L	R	K	D	F	S	Y	L	G	S	Y	G	T	R	G	V	G	Y	D	V	E	Y	L	V
T-Rex	A	F	Q	V	R	K	D	L	S	Y	F	G	S	Y	G	T	R	G	V	G	Y	T	V	P	V	L	K
	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68

	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108
S-Rex	Y	Q	I	S	R	E	L	G	L	T	Q	D	W	P	V	V	I	V	G	I	G	N	L	G	A	A	L
T-Rex	R	E	L	R	H	I	L	G	L	N	R	K	W	G	L	C	I	V	G	M	G	R	L	G	S	A	L
	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95

	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
S-Rex	A	N	Y	G	G	F	A	S	R	G	F	R	V	A	A	L	I	D	A	D	P	G	M	A	G	K	P
T-Rex	A	D	Y	P	G	F	G	-	E	S	F	E	L	R	G	F	F	D	V	D	P	E	K	V	G	R	P
	96	97	98	99	100	101	102	-	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121

	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162
S-Rex	V	A	G	I	P	V	Q	H	T	D	E	L	E	K	I	I	Q	D	D	G	V	S	I	G	V	I	A
T-Rex	V	R	G	G	V	I	E	H	V	D	L	L	P	Q	R	V	P	G	-	R	I	E	I	A	L	L	T
	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	-	140	141	142	143	144	145	146	147

	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189
S-Rex	T	P	A	G	A	A	Q	Q	V	C	D	R	L	V	A	A	G	V	T	S	I	L	N	F	A	P	T
T-Rex	V	P	R	E	A	A	Q	K	A	A	D	L	L	V	A	A	G	I	K	G	I	L	N	F	A	P	V
	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174

	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216
S-Rex	V	L	N	V	P	E	G	V	D	V	R	K	V	D	L	S	I	E	L	Q	I	L	A	F	H	E	Q
T-Rex	V	L	E	V	P	K	E	V	A	V	E	N	V	D	F	L	A	G	L	T	R	L	S	F	A	I	L
	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201

	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243
S-Rex	R	K	A	G	E	E	A	A	A	D	G	A	A	P	P	V	A	A	R	K	Q	Q	R	S	T	G	S
T-Rex	N	P	K	W	R	E	E	M	M	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228

	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258
S-Rex	A	D	Q	G	P	D	G	D	V	P	A	V	M	P	A
T-Rex	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243

= NO MATCH
= WEAKLY CONSERVED (.)
= STRONGLY CONSERVED (:)
= MATCH (*)

Figure S3: Alignment of *S. coelicolor* Rex versus *T. aquaticus* Rex, the structural homologue used as a template for mutagenesis design in this study. The residue numbers are shown for both proteins at each position and their conservation is indicated by the shading of each square, with the darkest squares being completely conserved and the lighter squares being similar but not identical.

Related Publication

McLaughlin KJ*, Strain-Damerell CM*, Xie K, Brekasis D, Soares AS, Paget MSB and Kielkopf CL (2010) Structural Basis for NADH/NAD^+ Redox Sensing by a Rex Family Repressor. Mol Cell 38(4): 563-575.
(* These authors contributed equally to this work)