Characterisation of oxygen regulation mechanisms in *Rhizobium leguminosarum* for repurposing as tools in the engineering of nitrogen fixation

Paul Johan Rutten

Wolfson College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2021

Declaration

I declare that this thesis was composed by myself and that the work contained herein is my own except where explicitly stated in the text. The work has not been submitted for any degree or professional qualification except as specified.

Parts of this thesis are taken from or based on two publications, both of which I was first author on [1, 2]. These articles, in order of publication, are given as appendices at the end of this thesis. All work and results reported in this thesis were produced by me unless explicitly specified otherwise.

- Paul Johan Rutten

Acknowledgements

At the bottom of this thesis is a list of the people on whose scientific work this DPhil has been built. Without them, this work would not have been possible. However, there is another group of people who are not on that list. Without them, I would not have been able to do this work. They may not be the traditional giants on whose shoulders science is built, but they have been every bit as vital and it is with enormous gratitude that I put them here at the start of this thesis.

My first thanks go to María Rendón, who keeps Plant Sciences sparkling. My Spanish may not have been as good as she had hoped, but our conversations nevertheless made my day. I also want to thank Rob Bryant and James Wisker. Doing low-oxygen work requires many very heavy cylinders of nitrogen gas, and they kept me supplied for four years without ever breaking a sweat.

I have used many pieces of equipment during this DPhil that I had never touched before, and it would not have been possible without others generously giving their time to teach me. Thank you to Pedro Bota, who taught me how to perform GC-MS and never lost patience with me despite the many instrument breakdowns I caused. Thank you to Niloufer Irani, who gave me days of her time to teach me how to do confocal microscopy. One of the images I took thanks to her adorned the cover of the February 2021 issue of PLoS Genetics, and is a testament to her skill and generosity. Thank you to John Baker, who took many photographs of my work. Photographing bacterial colonies is a nightmare, and John still received me with a smile even when I came back for a third day of attempts.

A very special thanks go to Helen Prescott and Lida Chen, whose work and dedication managing our lab space made every experiment so much easier. Managing a lab and keeping everything running smoothly is hard work at the best of times, and even through a global pandemic they made sure we always had what we needed. You are heroes.

I have been fortunate to be surrounded throughout my work by the members of the Poole group. Thank you all for supporting me where you could, and accommodating me where you couldn't. I owe a particular thanks to Tim Haskett and Carmen Sánchez-Cañizares, who supervised me throughout my time in the department. Over the years you have helped me with too many things to list here. I will always strive to do science at the standard you set.

Many others in the department have helped with all the things that go into a DPhil but don't show up in a thesis. Thank you to Gem Toes-Crichton and Nick Kruger, who helped me overcome Oxford University's bureaucracy. Another thank you to Tom King, whose digital efforts allowed me to keep working on my DPhil wherever in the world I was.

It is traditional to thank your DPhil supervisors for all their scientific guidance and advice. I owe everything I have learned during this DPhil about being a scientist to professors Philip Simon Poole and Lee James Sweetlove. However, what I want to thank them for above all else is giving me the freedom to explore myself and my passions outside the lab. Their support has allowed me to do life-changing things that other supervisors would have balked at. My deepest thanks for your forbearance.

I have been fortunate to be fully funded by a BBSRC (now UKRI) scholarship throughout my studies at Oxford. All of my DPhil has therefore been made possible by the support of the British government, and by extension the British public. Should this thesis ever somehow make it out of the academic world and into the hands of one of you, thank you for your incredible generosity giving a Dutch/French citizen this opportunity. I intend to work hard to repay the UK for all it has given me since I arrived almost a decade ago to start my undergraduate degree.

Finally, nothing I have done at Oxford or outside of it would have been possible without my friends and family. You know who you are. Your unwavering support throughout the ups and downs of the last four years have made it all possible. You are a constant reminder of why I set out to do this DPhil in the first place, and whatever I do next, I trust you will keep me pointed in the right direction.

P.S. I was the mysterious cookie leaver.

Colourblind accessibility

The colour scheme used in this thesis has been optimized for compatibility with deuteranopia, protanopia and tritanopia. I would like to acknowledge Luk Cox and Martin Krzywinski, whose colour palettes I used to design the ones found in this thesis.

Abstract

Rhizobia are α- and β-proteobacteria that form a symbiotic partnership with legumes, fixing atmospheric dinitrogen to ammonia and providing it to the plant. Fixation is performed by an oxygen-intolerant nitrogenase enzyme but requires respiration to meet its high energy demands. To overcome this paradox, regulation by oxygen (O_2) in rhizobia is essential for symbiosis and involves multiple O_2 sensing proteins. Interactions between O_2 regulatory systems are common, but their importance was not well understood.

We studied the pea microsymbiont *Rhizobium leguminosarum* biovar *viciae* 3841 (Rlv3841), which employs three systems: hFixL-FxkR-FixK, FnrN and NifA. We found that both the hFixL-FxkR-FixK and FnrN systems are functional, but act at different O_2 concentrations. hFixL-FxkR-FixK is active at a relatively high O_2 concentration (1%). The system induces key symbiosis targets including the high-affinity cbb_3-type terminal oxidase *fixNOQP* and the O_2 sensor *fnrN*. FnrN is largely inactive at 1% O_2 but becomes active inside nodules, where it autoregulates *fnrN* and is critical for full *fixNOQP* expression. Both hFixL-FxkR-FixK and FnrN are required to attain wild-type nitrogen fixation activity. With confocal microscopy, we observed that the two systems act in a hierarchical manner, with hFixL-FxkR-FixK activating at the tip of nodules (in zones I and II), followed by FnrN closer to the nodule core (at the II-III interzone).

The NifA regulator is also O_2 sensitive and of particular interest to engineering efforts because of its central role in the control of nitrogen fixation. Little is known about how rhizobial NifA proteins are controlled at the protein level. Most rhizobial NifA proteins have a GAF domain, but the function of the domain remains unknown. We found that very weak activity could be observed from Rlv3841 NifA when native transcriptional regulation was bypassed. Deleting the GAF domain of Rlv3841 NifA critically impaired its activity. Finally, we engineered NifA and NifV activity in Rlv3841 and were able to detect nitrogen fixation activity in free living conditions. This confirmed the potential of NifA engineering as an avenue to modify native biological nitrogen fixation regulation, albeit substantial work will be needed to improve activity.

The hierarchical arrangement of oxygen regulation in Rlv3841 provides a framework which explains both the multiplicity of oxygen sensors in other rhizobia

and past findings of partial redundancy between them. Our findings demonstrate the complexity of oxygen regulation in nitrogen fixation, and how one of these systems, NifA, could be repurposed to engineer this regulation. A better understanding of oxygen regulation in biological nitrogen fixation could eventually reduce our need for nitrogen fertilizers, substantially improving the carbon footprint and sustainability of modern agriculture.

Contents

List of Figures

Abbreviations

a.u. . . . . . Arbitrary Unit(s)

AAA+ . . . . ATPases Associated with diverse cellular Activities: a domain found in all NifA proteins

ALA . . . . . 5-Aminolevulinic acid

ANOVA . . . . ANalysis Of VAriance

ARA . . . . . Acetylene Reduction Assay, used to measure nitrogenase activity

BNF . . . . . Biological Nitrogen Fixation, the process by which microbes convert atmospheric dinitrogen into ammonia, using an enzyme called the nitrogenase complex

bv. . . . . . Biovar. Also referred to as symbiovar. Variant of a symbiotic diazotroph associated with a certain set of leguminous hosts

cNMP . . . . . Cyclic Nucleotide-Monophosphate, typically used to refer to a cNMP-binding domain in a protein

CRP . . . . . cAMP Receptor Protein

EBP . . . . . Enhancer Binding Proteins, e.g. NifA

EPS . . . . . Exopolysaccharide

FAD . . . . . Flavin Adenine Dinucleotide

Fe protein . . A homodimer protein containing an [Fe_4S_4] iron (Fe) – sulphur (S) cluster. Component of the nitrogenase complex that catalyzes the production of ammonia from atmospheric dinitrogen

FeMoco . . . . Iron-Molybdenum cofactor, found in the FeMo protein of molybdenum-dependent nitrogenases

Fluo . . . . . Fluorescence

FNR . . . . . Oxygen-sensing transcription factor in *E. coli*, named after the Fumarate and Nitrate Reductase genes.

GAF . . . . . A common domain of NifA named after three proteins in which it is found: cGMP-specific phosphodiesterases, Adenylyl cyclases and FhlA

HTH . . . . . Helix-Turn-Helix, a protein domain found in NifA

IDL . . . . . Inter-Domain Linker. Found in the NifA proteins of symbiotic diazotrophs, contains a cysteine-rich motif believed to be involved in oxygen regulation of NifA

IHF . . . . . Integration Host Factor, protein

INseq . . . . . INsertion sequencing

LMB . . . . . Laboratory of Molecular Biology, designation used for strains produced at the John Innes Centre in Norwich. Refers to plasmids when preceded by p, i.e. pLMB

Mo . . . . . Molybdenum element

OD . . . . . Optical Density, measured at 595 or 600 nm in this work (denoted as OD_{595}/OD_{600}). Used as a measure of cell density in liquid media

OPS . . . . . Oxford Plant Sciences, designation used for strains produced at Oxford University. Refers to plasmids when preceded by p, i.e. pOPS

ORF . . . . . Open Reading Frame

ori . . . . . Origin of replication

PAS . . . . . Sensory protein domain. Named after three eukaryotic proteins; Period circadian protein, Aryl hydrocarbon Receptor Nuclear Translocator protein (ARNT) and Single-minded protein. Also referred to as the Per-Arnt-Sim domain

PCR . . . . . Polymerase Chain Reaction

pOGG . . . . . Plasmid Oxford Golden Gate, plasmid assembled using Golden Gate system established in the Poole group

qRT-PCR . . . Quantitative Reverse-Transcription PCR

RBS . . . . . Ribosome Binding Site

Rlv3841 . . . . *Rhizobium leguminosarum* biovar (bv.) *viciae* strain 3841, a well-studied symbiotic diazotroph

RNA-seq . . . RNA sequencing

RU . . . . . Reading University, designation used for strains produced at Reading University. Refers to plasmids when preceded by p, i.e. pRU

SEM . . . . . Standard Error of the Mean

TCS . . . . . Two Component System, type of signal-responsive regulation pathway composed of two proteins, common in bacteria

TSS . . . . . Transcription Start Site

TY . . . . . Tryptone-Yeast growth medium

UAS . . . . . Upstream Activation Sequence, e.g. site bound by NifA to active expression from a promoter

UMS . . . . . Universal Minimal Salts growth medium

WT . . . . . Wild-Type

1

Introduction

Contents

1.1 Oxygen in the *Rhizobium*-legume symbiosis

1.1.1 Symbiosis between rhizobia and legume plants

Rhizobia are soil dwelling α- and β- proteobacteria that form symbiotic partnerships with legume plants in which they fix inert N_2 to biologically accessible NH_3 in return for a carbon and energy source source [3, 4]. The symbiosis has been intensely studied as a potential alternative to nitrogen fertilizers [5–8]. Nitrogen is a limiting factor in the growth of many crop plants and the use of these fertilizers has enabled dramatic improvements in yield [9–11]. However, producing nitrogen fertilizers chemically (via the Haber-Bosch process) consumes 1-2% of the world's total energy supply. Their over-use is also causing global environmental damage due to excess fixed nitrogen and its reactive by-products entering ground water, oceans and the atmosphere [12–15]. Simultaneously, these fertilizers are unused in many areas as their cost makes them unaffordable, creating vast inequalities in crop yields across the world [16]. Their application might be reduced by increased use of legumes and supplementation with artificial rhizobial-crop symbioses or plants engineered to perform nitrogen fixation [17–21]. This would help meet the agricultural demands of a surging world population more sustainably and equitably but will require a comprehensive and detailed understanding of the mechanisms by which nitrogen fixation is regulated [22–24].

Rhizobia taking part in the symbiosis undergo a dramatic change in lifestyle during their move from soil into the symbiotic environment [25, 26]. To initiate symbiosis the bacteria migrate towards plant roots, attach and subsequently enter them [4, 27, 28]. Where rhizobial entry occurs, plants form specialized organs called nodules which create an environment in which nitrogen fixation can occur [29, 30]. Rhizobia within nodules subsequently infect plant cells [31]. Infected plant cells enclose rhizobia into vesicles called symbiosomes [32, 33]. In these intracellular compartments the bacteria differentiate into bacteroids, a highly specialized quasi-organelle form optimized for nitrogen fixation [34, 35]. This shift from their free-living lifestyle in soil to nitrogen fixing bacteroids in nodules requires substantial

changes in bacterial gene expression. These must be carefully coordinated to coincide with the *Rhizobium*'s progress along this transition. A key element of this coordination is oxygen regulated gene expression.

1.1.2 Importance of oxygen in nitrogen fixation

Oxygen regulation is at the heart of the *Rhizobium*-legume symbiosis [36–38]. All rhizobia rely on an enzyme complex known as the nitrogenase to perform nitrogen fixation, and all use a molybdenum-containing form of the complex [39, 40]. Nitrogenase is commonly encoded by a single *nifHDK* operon and contains two components [41, 42]. One of these is the Fe protein, a homodimeric dinitrogenase reductase encoded by *nifH* [43]. The second is the MoFe heterotetramic dinitrogenase protein encoded by *nifDK* [44, 45]. Both proteins contain iron-sulphur clusters required for electron transfer during the enzyme's activity but which are highly intolerant of oxygen [46–49]. The nitrogenase complex is therefore only stable and functional under anoxic conditions. Simultaneously, nitrogen fixation by nitrogenase is a highly energy intensive process [50]. In many species the enzyme relies on an electron bifurcating complex encoded by *fixABCX* to supply it with low-potential electrons [51, 52]. Bacteroids must respire at very high levels to meet this energy demand. Symbiotic nitrogen fixation therefore paradoxically requires a very low oxygen environment permitting nitrogenase function but simultaneously demands a high oxygen supply to meet its energy requirements [53]. Rhizobia and legumes have evolved several strategies to overcome this paradox, many relying on their close cooperation. The oxygen concentration experienced by rhizobia decreases throughout their transition from free-living soil bacteria to nitrogen fixing bacteroids. Free-living rhizobia in soil experience oxygen concentrations up to atmospheric levels, whilst concentrations in nodules are at nanomolar levels [54]. This drop is driven by plant processes that reduce oxygen inside the nodule, ultimately producing near-anoxic conditions in symbiosomes optimal for nitrogen fixation [55].

1.1.3 Regulation of oxygen in plant nodules

The oxygen concentration experienced by bacteroids is determined and regulated by the legume host. Three main legume mechanisms have been identified which adjust nodule oxygen levels and facilitate its supply to bacteroids.

The first of these is a barrier in the cortical layer of nodules that is impermeable to oxygen diffusion [56–58]. Only nanomolar concentrations of oxygen are present beneath the diffusion barrier [59, 60]. Further, the permeability of the layer is variable and controlled by the plant [61–63]. This enables legumes to regulate the influx of oxygen into nodules and maintain a stable internal oxygen concentration despite significant environmental stresses [64–67]. Studies have shown that the barrier is highly responsive, with permeability able to adjust within minutes [68].

A second mechanism regulating nodule oxygen is the production of leghaemoglobin, giving the organs their characteristic red hue [69, 70]. Leghaemoglobins are typically monomeric haemoproteins of 16 kDa similar to human myoglobin but employing a different oxygen binding mechanism [71]. Compared to myoglobin, leghaemoglobins have far higher affinity for oxygen and typically show very fast binding but relatively slow release. The importance of leghaemoglobin has been highlighted by studies linking its expression to that of the nitrogenase complex. In a leghaemoglobin knockout plant, bacteroid nitrogenase expression was eliminated [72]. This connection appears to run in both directions, with a very different pattern of leghaemoglobin expression found in nodules infected by non-fixing rhizobia [73]. In line with these results, a strong correlation between the leghaemoglobin content and nitrogen fixation activity of nodules has been demonstrated [74].

Leghaemoglobin protects nitrogenase by binding free oxygen and enables a high oxygen flux to bacteroids which is required to meet respiratory demands. At the low oxygen concentrations found in nodules there is insufficient free oxygen to support adequate bacteroid respiration. However leghaemoglobin is itself present at a concentration estimated to be six orders of magnitude higher than free oxygen [72]. Most oxygen within nodules is therefore bound by leghaemoglobin and present at a relatively high concentration, with the remainder buffered at nanomolar

levels. This bound oxygen cannot damage the nitrogenase but is accessible to bacteroids and enables sufficient respiration for nitrogen fixation. Leghaemoglobin thus simultaneously limits free oxygen to prevent nitrogenase damage and generates a high concentration of 'safe' oxygen for bacteroids.

Leghaemoglobin binding of oxygen also facilitates its diffusion within the nodule [75, 76]. At the nanomolar concentration present in nodules, unaided diffusion of oxygen would be unable to meet the oxygen flux requirements of respiring bacteroids. Adequate diffusion of oxygen is also essential for its even distribution. Modelling has suggested that in the absence of leghaemoglobin a significant oxygen gradient would exist between the interior and periphery of plant cells infected with bacteroids [77]. This would likely result in bacteroids at plant cell edges receiving excess oxygen whilst those deeper inside would receive too little. Leghaemoglobin ensures both a high oxygen flux and an even distribution by facilitating oxygen diffusion within nodules.

The effects of leghaemoglobin may be fine-tuned by the existence of isoenzymes of the protein. Recent work found two isoenzymes in peas which had different oxygen affinities and spatial expression patterns within nodules [73]. This indicates leghaemoglobins could play a role in creating and shaping the oxygen gradient of indeterminate nodules, but this has received little attention to date.

A third mechanism in nodules appears to specifically control the oxygen levels of infected plant cells [78]. Mitochondria in these cells employ high affinity terminal oxidases and localize to areas at the cell periphery adjacent to intercellular air pockets, where oxygen influx is likely high [79]. This is thought to act as an additional oxygen barrier, with mitochondrial oxygen consumption protecting symbiosomes from the influx of oxygen in these areas. It has also been proposed that this serves to adjust fixation rates in response to oxygen availability. A drop in oxygen influx could reduce mitochondrial ATP production, in turn reducing the rate of nitrogen fixation by bacteroids [80]. This is in line with a proposed model in which oxygen supply is a limiting factor for nitrogen fixation and used by plants to up- or down- regulate the process [81–83].

The combination of these plant mechanisms produces a nodule environment supporting rhizobial nitrogen fixation. Two nodule types exist, called determinate and indeterminate nodules. In determinate nodules, such as those of soybean, all rhizobia reversibly differentiate to fix nitrogen [84–86]. Once the nodule senesces, all rhizobia can return to a free-living lifestyle in soil. In indeterminate nodules, such as those of pea and alfalfa, some rhizobia remain free-living whilst other terminally differentiate into nitrogen fixing bacteroids [87–90]. In these nodules, only the rhizobia which do not differentiate into bacteroids survive nodule senescence. In determinate nodules the plant creates a uniformly low oxygen concentration throughout the nodule [91]. In plants with indeterminate nodules the oxygen diffusion barrier is not present at the nodule apex and oxygen is able to freely enter here, creating a longitudinal oxygen gradient [41, 92]. Undifferentiated bacteria at the tip experience relatively high oxygen concentrations whilst terminally differentiated bacteroids in the central nitrogen fixing zone are in a near anoxic environment. Rhizobial differentiation from the tip to the core is an ongoing process during the growth of indeterminate nodules. The nodule oxygen gradient is used by the bacteria to regulate their differentiation. Four zones, shown in Figure 1.1, have been delineated within indeterminate nodules [93, 94]. Zone I, at the tip of the nodule, has an oxygen concentration similar to the soil and contains undifferentiated bacteria. Zone II holds rhizobia preparing to infect plant cells, contained in so-called infection threads that direct their movement in the nodule [95]. Above Zone III is a key area known as the II-III interzone where bacteria are released from infection threads and infect plant cells [96]. This has been found to coincide with a sharp drop in oxygen concentration and a concurrent induction of several genes important for nitrogen fixation [97–99]. Zone III represents the primary nitrogen fixation zone, whilst Zone IV contains plant cells which are beginning to senesce [100]. Zone V (not shown in Figure 1.1) contains plant cells which have been re-infected by rhizobia from remaining infection threads. Rather than forming a mutualistic relationship, the bacteria in this zone appear to exploit the plant for their own benefit, and the zone is therefore also called the saprophytic zone.

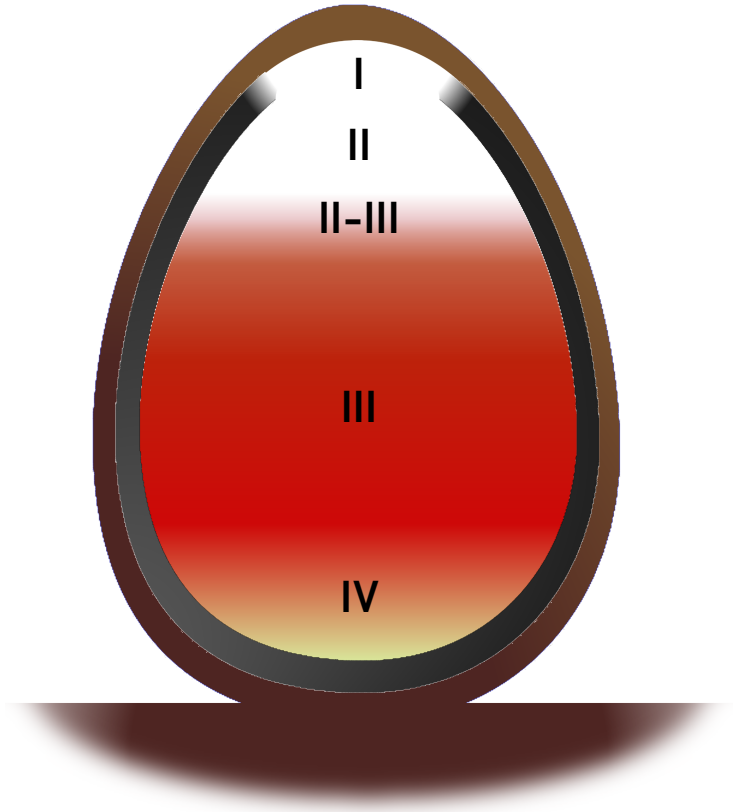


Figure 1.1: Structure of an indeterminate nodule. Legumes such as pea and alfalfa form indeterminate nodules containing free-living rhizobia and terminally differentiated bacteroids. The nodule cortex and root are shown in brown. The oxygen diffusion barrier (black) is absent at the apex of the nodule, creating an oxygen gradient inside. This can be divided into multiple zones. Zone I has free-living rhizobia and oxygen levels similar to surrounding soil. Zone II contains infection threads with bacteria preparing to enter plant cells. At the II-III interzone rhizobia infect plant cells to begin differentiating into bacteroids and oxygen drops sharply. Free oxygen concentration in zone III is at 20-50 nM [56, 59]. This zone houses plant cells infected with nitrogen fixing bacteroids and contains high levels of leghaemoglobin, indicated in red. Zone IV contains senescing cells which are no longer fixing nitrogen. Zone V is not shown, see text for details. Nodule layers and zones are not to scale.

1.1.4 Adaptation to nodule oxygen conditions by rhizobia

Free oxygen in much of the nodule is at nanomolar levels and rhizobia have evolved to survive and respire under these conditions during nitrogen fixation. Key to this adaptation is the ability of bacteroids to access leghaemoglobin-bound oxygen [101]. The concentration of this is estimated to be six orders of magnitude higher than that of free oxygen [72]. To access this oxygen, symbiotic bacteria employ a specialized cbb_3-type terminal oxidase encoded by the *fixNOQP* operon in their

respiratory chain which has very high affinity for oxygen [102, 103]. The complex was first identified in *Bradyrhizobium japonicum* and is broadly required for nitrogen fixation in symbiotic rhizobia [104–107]. Assembly and maturation of this symbiotic terminal oxidase requires the *fixGHIS* operon [108, 109]. The *fixNOQP* and *fixGHIS* operons are often similarly regulated and located in proximity to each other. A notable exception is *Azorhizobium caulinodans* in which individual mutants of both operons retained significant nitrogen fixation activity [110, 111]. This suggests the species employs multiple terminal oxidases performing a similar function.

1.1.5 Mechanisms of oxygen regulation in rhizobia

This introduction will cover oxygen regulation of nitrogen fixation in several species to highlight variations as well as common themes across rhizobia. Three oxygen regulation systems are covered in detail: the FixL/hFixL sensors, the FnrN sensor and the NifA sensor. The FixK protein, which does not sense oxygen but serves as an important intermediate for FixL and hFixL regulation, is also covered. The focus will be on strains of five well studied species; *Ensifer meliloti* (previously *Sinorhizobium meliloti*), *Azorhizobium caulinodans*, *Bradyrhizobium japonicum*, *Rhizobium etli* and *Rhizobium leguminosarum* bv. *viciae*. The divergence of rhizobial species predates the evolution of their legume hosts and the nitrogen fixing partnership [112]. Symbiotic capability does not therefore descend from a single ancestor but has instead been horizontally transferred extensively between rhizobial strains and species [113, 114]. Successful transfers have relied on the rapid adaptation of pre-existing bacterial regulatory mechanisms under plant selection to arrive at nodulation and fixation competent rhizobia [115]. It appears pre-existing host regulators were commonly repurposed to control symbiotic nitrogen fixation. As a result, whilst a core set of symbiotic genes are conserved across all symbiotic diazotrophs, the regulation of these genes varies significantly at short evolutionary ranges, at both the species and strain level [116]. It is difficult to determine the cause and significance of these variations; some may result from differing external

pressures, whilst others reflect the diversity of host regulation which existed prior to horizontal transfer of symbiotic functions.

Regulation mechanisms have been ordered according to the oxygen concentration at which they activate. The FixL-FixJ and hybrid FixL-FxkR systems are active at relatively high oxygen concentration and are thought to be one of the first oxygen-sensing mechanisms to act during the symbiotic transition. Both in turn lead to production of the FixK protein which is not oxygen sensitive itself. The oxygen concentration at which FnrN homologs operate is less well understood but appears to fall between that of FixK and NifA, although some overlap is likely. The NifA factor is active only at very low oxygen concentrations and is one of the final regulators of bacteroid differentiation. The extensive interconnection of these regulators will also be discussed.

1.2 The FixL-FixJ and hybrid FixL-FxkR cascades

The FixL-FixJ (FixLJ) regulatory cascade is an oxygen-responsive two-component system (TCS) [38]. TCSs are composed of a sensor-regulator protein pair and are ubiquitously used by bacteria to respond to a range of environmental conditions [117, 118]. Variants of the cascade across several rhizobial species are shown in Figure 1.2. The FixL oxygen concentration sensor activates the FixJ regulator under microaerobic conditions, which in turn induces expression of genes controlled by the cascade. The threshold oxygen concentration at which the cascade becomes active is not precisely known, but activity has been recorded at a concentration as high as 5% [119]. Studies typically observe activity at 1% [1]. The FixLJ system was first discovered in *E. meliloti* as the regulatory mechanism responsible for microaerobic *fixNOQP* expression [120]. FixLJ variants are also present in *A. caulinodans*, *B. japonicum*, *R. etli* and *R. leguminosarum* bv. *viciae* strains [121–123]. Of note, no homolog of the cascade has been found in non-symbiotic diazotrophs, suggesting its role is intimately tied to nitrogen fixation in a symbiotic context.

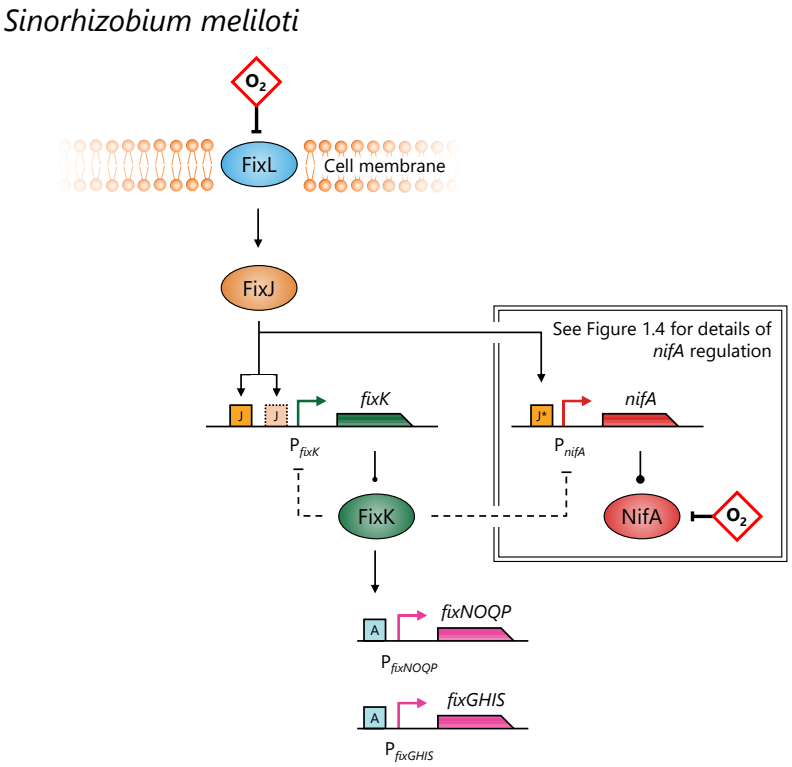
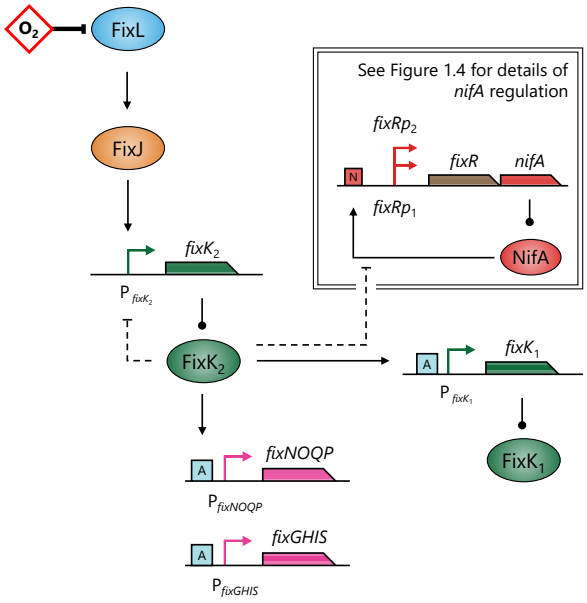


Figure 1.2: Oxygen regulatory networks controlling rhizobial nitrogen fixation genes. Oxygen is indicated in red diamonds. TSS are shown as right-angled arrows. Lines ending in arrows or bars indicate activating and repressing regulation, respectively. Lines end at operators where these are known or predicted from sequence information. Dashed lines represent theorized connections or where the mechanism of action is unclear. Operator sites are shown as boxes. Anaeroboxes are shown in blue and marked with 'A'. FixJ and FxkR (K-box) operators are shown in orange and marked with 'J' and 'K' respectively. *E. meliloti* FixJ binds different operator sequences upstream of *nifA* and *fixK*, and in the latter binds a second low-affinity non-consensus site (faded orange, dotted outline). Pointed rectangles are genes or operons. Lines ending in circles indicate translation. Proteins are shown as ellipses. FixL proteins thought to be membrane associated are indicated. The map shown for *Rhizobium leguminosarum* corresponds to Rlv VF39: Rlv3841 instead has two copies of hFixL (see Figure 4.7 for details) whilst Rlv UPM791 has no functional FixK homolog and two copies of FnrN. Note this figure continues on the next two pages. See text for details of regulatory connections and Figure 1.4 for details of *nifA* regulation.

Bradyrhizobium japonicum



Azorhizobium caulinodans

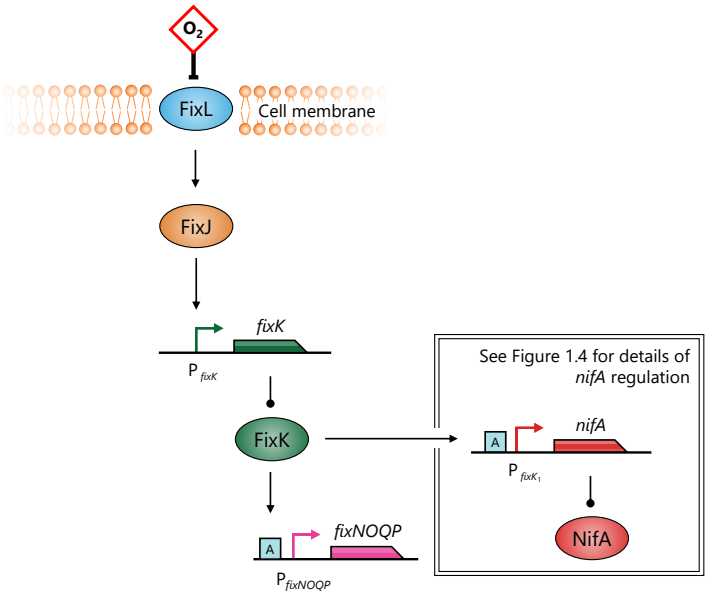
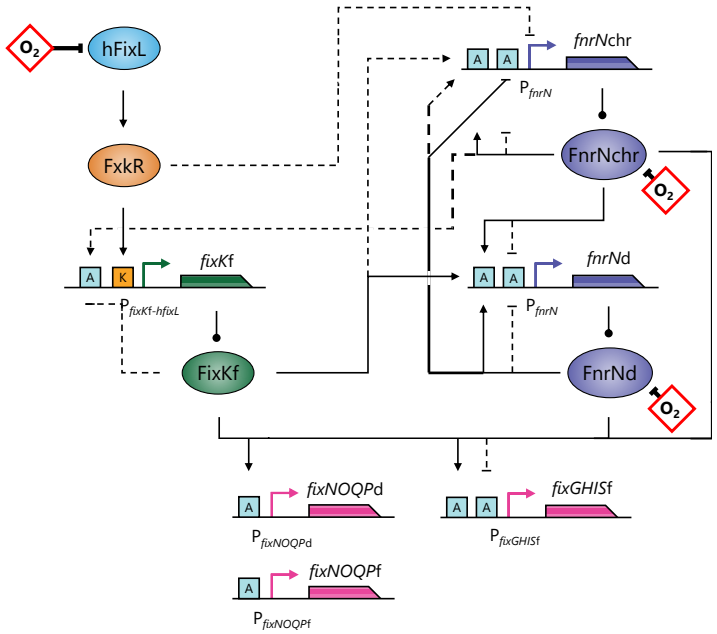


Figure 1.2: Oxygen regulatory networks controlling rhizobial nitrogen fixation genes (continued).

Rhizobium etli CFN42



Rhizobium leguminosarum

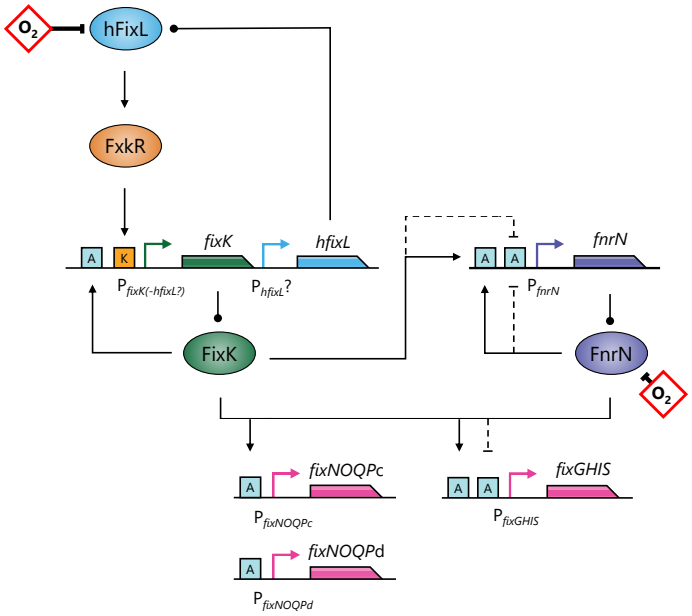


Figure 1.2: Oxygen regulatory networks controlling rhizobial nitrogen fixation genes (continued).

The FixLJ cascade controls only a small number of genes directly, commonly including *fixK* and in certain species *nifA* [38]. However, if indirect targets of the system are included its regulon is one of the largest of any TCS studied to date [124–126]. These indirect targets vary at the species and strain level but typically include the *fixNOQP* and *fixGHIS* operons, controlled via FixK.

Recent findings suggest two forms of the FixL protein exist, each corresponding to a different cascade. The first form of FixL is found in *A. caulinodans*, *E. meliloti* and *B. japonicum* [127]. FixL proteins from these species are approximately 55% homologous and around 55 kDa in size. A second form of FixL appears to be employed by species including *R. leguminosarum* bv. *viciae* VF39 (Rlv VF39) and *R. etli* CFN42 [104, 128]. This has been named hFixL (hybrid FixL). The hFixL variant is larger, at a size of 70 kDa, and shows over 85% homology within these species. It has less than 40% identity to FixL variants from the first group. These two FixL forms appear to act on two different cascades, as evidenced by the presence or absence of a *fixJ* homolog encoding the traditional partner of FixL. Homologs of *fixJ* are present in all species with the first type of FixL but none have been found in Rlv VF39 or *R. etli* CFN42 [129]. A significantly altered system appears to operate in these organisms. hFixL retains its oxygen-sensing role but regulates targets such as *fixK* through a partner called FxkR (*fixK* regulating), discussed in more detail in section 1.2.2. The importance of FixL has been found to correlate with its form. In species with the canonical FixLJ cascade it plays a crucial role, whereas in species using the second form hFixL is largely dispensable. FixL is required for nitrogen fixation by *E. meliloti* and both free-living and symbiotic fixation by *A. caulinodans* [120, 121]. Disruption of *fixL* or *fixJ* in *B. japonicum* led to a 90% drop in nitrogen fixation activity [122]. By contrast, null mutations of their *hfixL* homolog reduced nitrogen fixation activity by only 50% in Rlv VF39 and had a minimal effect on fixation by *R. etli* CFN42 [123, 128, 130].

FixL type varies at the strain level so is not necessarily uniform for a given species. A distribution of both FixL forms exists across *R. etli* strains, with the *hfixL* variant found to be more common [128]. The *R. etli* CNPAF512 strain for instance employs

FixL and FixJ [130]. However, contrary to other species with FixLJ, disabling the cascade in this strain reduced but did not abolish nitrogen fixation [131]. Whilst the pathway still senses oxygen this no longer appears to be done by FixL and the regulatory targets of FixLJ are largely unknown in this strain. Of note, different FixL forms and their respective cascades appear mutually compatible. Sequence information suggests both systems are present in *E. meliloti* SM11, and may also be found in the *E. meliloti* 1021 strain [129, 132]. In the latter strain the hFixL cascade may function outside of symbiosis but this has yet to be further investigated [133].

1.2.1 Structures

FixL structure

FixL is an oxygen sensor which activates the FixJ transcription factor by phosphorylation under microaerobic conditions. Non-hybrid FixL proteins have two main components, shown in Figure 1.3 [134]. At the N-terminus is a sensory Per-Arnt-Sim (PAS) domain which contains an oxygen-binding haem group [135]. The C-terminus contains a histidine kinase (HK) module composed of a dimerization and histidine phosphotransfer domain and a catalytic ATP binding domain. This module is responsible for signal transmission to FixJ by phosphorylation [136]. The FixL proteins of *A. caulinodans* and *E. meliloti* have multiple predicted transmembrane helices and are likely membrane bound or associated, with the PAS domain sensing extracellular oxygen concentration [120, 137]. The first full FixL structure was recently determined using the *B. japonicum* protein and this is instead cytoplasmic [138]. Reported binding affinity (K_d) values for FixL proteins range from 50 µM for the *E. meliloti* protein to 738 µM for the *R. etli* protein [139, 140].

Under microaerobic conditions release of oxygen from the PAS domain causes activation of the FixL kinase function. Activation is mediated by intramolecular signalling through conformational changes in both the sensor PAS domain and the coiled-coil region connecting it to the HK module [138, 141]. Once active the HK autophosphorylates FixL on a conserved histidine residue. The phosphate is then transferred to an aspartate residue on FixJ, switching it to its active conformation

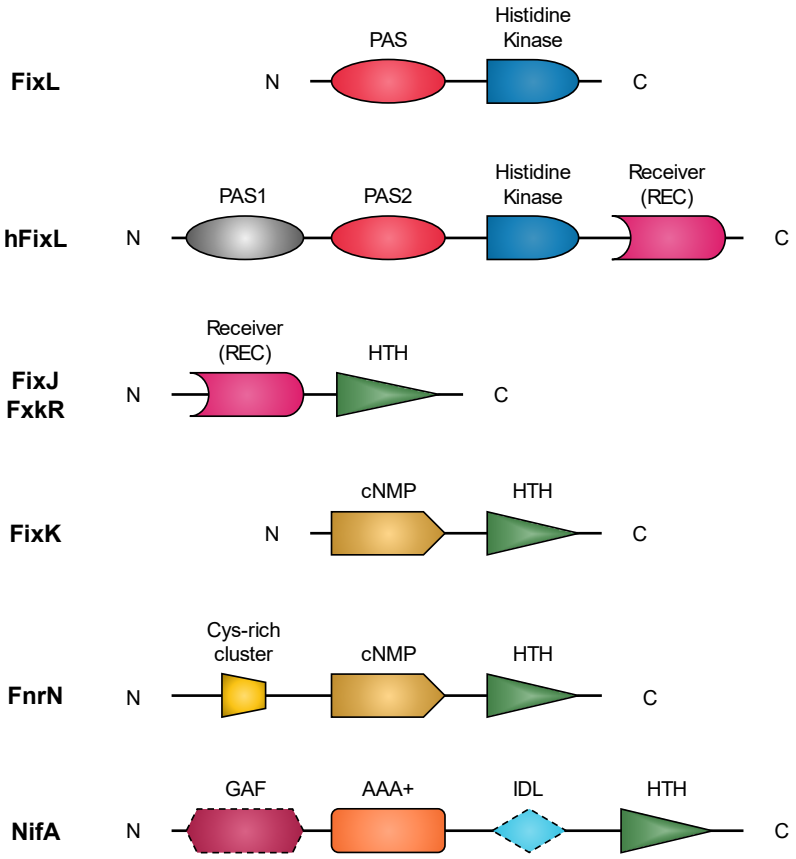


Figure 1.3: Components of rhizobial oxygen regulation proteins. Ellipses are PAS domains: PAS and PAS2 (red) in FixL and hFixL respectively are oxygen sensing, but PAS1 (grey) in hFixL is not. Histidine kinase modules are dark blue, receiver (REC) domains are pink and helix-turn-helix domains (HTH) are shown as green triangles. FixJ and FxkR have the same domains but belong to the NarL/FixJ and OmpR/PhoB families, respectively. cNMP domains are brown and the cysteine-rich cluster at the FnrN N-terminus is shown in yellow. NifA contains a GAF (purple), AAA+ (orange) and HTH domain and an inter-domain linker (IDL, light blue diamond). NifA domains with a dotted outline are not found in some species. Lengths are not to scale. More detailed figures of oxygen regulation in Rlv3841 are given in Chapter 3.

[142, 143]. Oxygen-bound FixL also has phosphatase activity on FixJ, thereby decreasing the cascade's background activity under aerobic conditions [144]. In summary FixL strictly regulates FixJ-dependent transcription in response to a microaerobic oxygen concentration.

FixJ structure

FixJ is a transcription factor that induces expression of FixLJ cascade targets when phosphorylated by FixL under microaerobic conditions [145]. The protein

is approximately 22 kDa and shows sequence conservation of around 50% across *A. caulinodans*, *E. meliloti* and *B. japonicum* [38]. FixJ has features typical of a TCS regulator protein, with an N-terminal receiver domain (REC) and a C-terminal helix-turn-helix (HTH) transcription activating domain [117, 146]. Under aerobic conditions, the non-phosphorylated receiver domain inhibits the function of the transcription activating domain [147]. Under anaerobic conditions FixL phosphorylates the receiver, alleviating this repression. Derepression is due to phosphorylation-induced conformational changes which propagate throughout the tertiary structure of the protein [148, 149]. In *E. meliloti* phosphorylation causes FixJ to adopt an open configuration, thereby relieving steric inhibition of the transcription activating domain by the receiver domain [150]. Simultaneously, phosphorylation leads to a conformational change in the receiver that exposes a dimerization interface in the domain [151]. The resulting dimerization of FixJ has been shown in *E. meliloti* to significantly increase the protein's affinity for its target promoters.

Hybrid FixL structure

Certain rhizobial species employ a different form of FixL, called hFixL (hybrid FixL). These include *R. etli* CFN42, Rlv VF39 and *R. leguminosarum* bv. *viciae* 3841 (Rlv3841), with research to date having focused primarily on the first two species [123, 128, 152]. The hFixL protein retains the ability to sense oxygen but regulates a non-canonical pathway, as these organisms have no FixJ homolog. hFixL combines structural elements of canonical FixL and FixJ; a domain homologous to the FixJ receiver domain is present at the protein's C-terminus [153]. In Rlv VF39, this appended domain was shown to be required for microaerobic induction of genes regulated by hFixL [154]. Alignments show the receiver domain contains an aspartate residue at position D573 that is analogous to the FixJ residue phosphorylated by canonical FixL. In the canonical FixJ protein, phosphorylation of its receiver domain exposes a dimerization interface important for its function. This does not appear to occur in hFixL and there is no evidence to suggest the protein dimerizes [140].

In *R. etli* CFN42, hFixL contains not only this appended receiver domain but also a second PAS (PAS1) domain at its N-terminus alongside the canonical haem-PAS domain (PAS2) [140]. Many of the protein's functions and characteristics appear to be a result of inter-domain interactions. For instance, the protein's kinase activity required the new PAS1 domain. Furthermore, whilst PAS1 has no haem binding capability and no apparent mechanism to sense oxygen directly, the domain does modulate the oxygen affinity of the canonical PAS2. Wild-type (WT) *R. etli* CFN42 hFixL has one of the lowest measured oxygen affinities of any FixL variant, but deletion of PAS1 increased the protein's affinity for oxygen 8-fold [140]. At present it remains unknown how the PAS1 domain decreases the oxygen affinity of hFixL. The effect may be mediated by protein binding to the domain or PAS1 may bind a small molecule to regulate hFixL, amongst other possibilities. The protein's oxygen affinity was also modulated by the appended receiver domain. A D573N mutation at the conserved phosphorylation site in this domain doubled the oxygen affinity of hFixL. These results suggest modulation of hFixL oxygen affinity is an important mechanism regulating the protein's function in *R. etli* CFN42 and involves its PAS1 and receiver domains.

No transmembrane domains are predicted in the hFixL proteins of *R. etli* CFN42 or Rlv VF39, unlike the canonical FixL proteins of *E. meliloti* and *A. caulinodans*. It is therefore believed that hFixL proteins are cytoplasmic and sense intracellular oxygen. This localization is supported by heterologous expression studies in *E. coli* which showed the *R. etli* CFN42 protein was highly soluble even when overexpressed [140]. Of note, *B. japonicum* FixL appears to combine elements from both canonical and hybrid FixL protein types. Like the hFixL proteins of *R. etli* CFN42 and Rlv VF39, *B. japonicum* FixL is cytoplasmic and senses intracellular oxygen concentration [138]. The *B. japonicum* protein also contains a second N-terminal PAS domain that likely regulates the protein's oxygen affinity as shown in *R. etli* CFN42 hFixL. This suggests classifying some FixL proteins as members of the canonical or hybrid group may be impossible, with proteins instead falling along a spectrum between these two forms.

It remains unknown what led to the evolution of hFixL proteins and what their benefits and regulatory implications are. In other organisms hybrid histidine kinase proteins in TCSs allow multiple signals to be integrated and it is suggested they enable more finely tuned regulation [155]. No studies have yet investigated this possibility in rhizobia. Studying organisms such as *E. meliloti* SM11 where both FixL forms act in parallel will be of interest as these may shed light on differences in the role of the two forms and lead to an improved understanding of their respective functions.

1.2.2 Role in oxygen regulation

The FixLJ cascade

The FixLJ cascade is active at a relatively high oxygen concentration and appears to be one of the earliest oxygen regulators of symbiotic establishment. In nodules, activation of the cascade likely occurs in the II-III interzone [94, 98, 104].

The most common target of FixLJ microaerobic induction is *fixK*, as found in *E. meliloti*, *B. japonicum* and *A. caulinodans* [115, 120, 121, 156]. In *E. meliloti*, FixLJ also directly activates *nifA* [124]. The FixJ DNA binding motif and its regulon have historically been extensively investigated in this species. Early attempts to identify a FixJ operator failed to find a binding motif common to both the *fixK* and *nifA* promoters. More recent work has demonstrated that each promoter has a different consensus sequence and their activation appears to proceed by a different mechanism [157].

The first FixJ operator type, found in the *fixK* promoter, has a GTAGTTTCCC consensus sequence and is bound by a FixJ dimer. It shows high cooperativity; binding of a single FixJ protein promotes recruitment of a second monomer to form the active dimer. However, mutation studies demonstrated that this site is not critical for *fixK* induction in *E. meliloti*. Instead, downstream of this operator was found a second site that was critical for *fixK* induction [158]. This second operator has a far lower affinity for FixJ than the first site, and its sequence shows no homology to either the *fixK* or *nifA* operators. A model has been proposed whereby

the first, upstream consensus site recruits FixJ binding to this second, downstream operator. The first site does not appear to be essential and its elimination could be compensated by increased expression of FixJ [145].

A second FixJ operator type is found in the *nifA* promoter of *E. meliloti*. *Ab initio* methods suggest the consensus sequence for this second FixJ operator is a semi-palindromic GTACGTAG motif. This appears to have a lower binding affinity than the consensus sequence of the first *fixK* operator. A sequence with poor homology to this motif is found upstream of *nifA* and was shown to be responsible for its regulation by FixJ [157]. Results indicate this second operator type binds multiple FixJ dimers, unlike the single dimer bound at the *fixK* site. This is supported by the finding in footprinting experiments that a very large region around the site was protected, and that multiple protein-DNA complexes were visible during gel shift titration. However, a FixJ mutant unable to form dimers was still able to induce *nifA* [148]. Induction of *nifA* by FixJ therefore does not appear to require its oligomerization at the promoter.

In line with their different FixJ operator sites and binding characteristics, the mechanism by which FixJ induces expression from the *fixK* and *nifA* promoters also differs. At the *fixK* promoter the FixJ receiver domain plays a role in RNA polymerase recruitment [159]. A mutation of the domain interfering with this recruitment reduced activation of *fixK* 10-fold in *E. meliloti*. By contrast the mutation had no effect on FixJ induction of *nifA*. This suggests FixJ promotes transcription of *nifA* through a different, as yet undetermined, mechanism. Recent genome-wide work in *E. meliloti* has begun to shed light on members of the FixLJ regulon besides the canonical *fixK* and *nifA* targets. This species contains a chromosome and two extrachromosomal replicons pSymA and pSymB [160]. It is believed the FixLJ cascade and its original regulon were initially introduced through horizontal transfer of pSymA [161]. Its regulon appears subsequently to have begun encompassing other targets in the *E. meliloti* genome. In accordance with this theory, putative FixJ operators were found to be very unevenly distributed, with a majority found on pSymA and a much smaller number on the chromosome. All pSymB

sites were within coding regions so none are thought to be functional. However, it has been theorized that some FixJ operators in the coding regions of pSymA and the chromosome do have a functional role. A significant number of these have been found and they may serve as 'reservoirs', partially controlling the intracellular localization of FixJ to facilitate its diffusion to operators in spatially proximal promoters. Analysis of a selection of new putative FixJ targets confirmed that two genes involved in proline metabolism are also regulated by the FixLJ cascade. Several studies have demonstrated the importance of proline metabolism to symbiosis but its role remains poorly understood [162–164]. It is likely that many processes controlled by FixLJ and important in symbiotic regulation remain to be discovered.

The cascade may also serve a function outside of nodules [165]. In *A. caulinodans*, FixLJ activation occurs in free-living cultures even in an environment with atmospheric oxygen concentration [121, 166, 167]. This may be due to the microaerobic internal environment created by this species to enable free-living nitrogen fixation, using multiple terminal oxidases [168]. In *B. japonicum*, which fixes only in symbiosis, free-living cells also showed FixLJ activation of gene targets [119]. Induction was demonstrated when cells were exposed to a 5% O_2 atmosphere and activation gradually increased as oxygen dropped to 0.5%. Few studies have investigated the role of FixLJ outside symbiosis and the biological significance of this remains largely unknown.

The hFixL-FxkR cascade

Several rhizobia employ a hFixL protein which contains an appended domain homologous to the receiver domain of FixJ. These species typically have no FixJ homolog, and instead hFixL oxygen sensing is transmitted to gene induction by the FxkR (*fixK* regulating) protein. This situation has been studied primarily in *R. etli* CFN42 and Rlv VF39 [104, 128]. The *fixL* and *fixJ* genes commonly form an operon located near other nitrogen fixation genes, suggesting they were acquired together in a horizontal gene transfer event. By contrast, *fxkR* in *R. etli* CFN42 and Rlv VF39 is not in the vicinity of its *hfixL* signalling partner, suggesting it was

not acquired in the same transfer event and may be derived from a pre-existing host regulator. Sequence information suggests Rlv3841 also encodes two putative homologs of hFixL and one of FxkR [152]. *E. meliloti* SM11 encodes homologs of both the hybrid and canonical FixL proteins, demonstrating the proteins and their respective pathways are not mutually exclusive [132]. It is unknown at present whether any crosstalk occurs between such parallel systems. Unlike the FixLJ system that is generally critical for nitrogen fixation activity, mutations eliminating hFixL-FxkR regulation are found to have a limited effect on fixation. In *R. etli* CFN42, the pathway appears dispensable [129].

As in the traditional FixLJ cascade, hFixL regulates FxkR at the protein level by phosphorylating it under microaerobic conditions [129]. Transcription of *fxkR* is thought to be constitutive. FxkR has a predicted weight of 27 kDa and a typical response regulator structure with an N-terminal receiver domain and a C-terminal HTH DNA-binding domain. These domains are analogous to those found in FixJ, but the two proteins are not related. FxkR belongs to the OmpR/PhoB family whereas FixJ belongs to the eponymous NarL/FixJ family. Based on studies of other OmpR/PhoB family proteins, it is thought that phosphorylation of the FxkR receiver domain induces a conformational change leading to protein dimerization. This dimer brings together two FxkR HTH domains which bind a DNA target sequence composed of a pair of direct-repeat half sites [169].

As expected given their different families it appears there is limited or no cross-talk between regulation by FxkR and FixJ. An *R. etli* CFN42 *fxkR* mutant could not be complemented by heterologous *E. meliloti* FixJ [129]. It therefore appears hFixL cannot activate the canonical FixJ protein or FixJ is not functional in this strain. In contrast complementation of the *R. etli* CFN42 *fxkR* mutant was possible with Rlv VF39 *fxkR*, demonstrating the similarity of their hFixL-FxkR cascades.

In *R. etli* CFN42, microaerobic activation of *fixK*f is due to hFixL and requires the FxkR intermediary in a manner analogous to the role of FixJ in transmitting oxygen sensing by FixL. Analysis of a library of confirmed and putatively FxkR-regulated genes identified a GTTACA-N_4-GTTACA consensus binding motif, named

the "K-box" [129]. A K-box is found in front of the *fixK*f and *fixK* genes of *R. etli* CFN42 and Rlv VF39 respectively, and both are induced by hFixL-FxkR under microaerobic conditions. By contrast, no K-box element is found upstream of *fixK* genes in *E. meliloti* SM11 [132]. In this strain, the cascade instead functions to repress activation by the parallel FixLJ system. This repression appears to act by protein level inhibition of the TCS, but the mechanism and target of this repression have yet to be determined. It also remains unclear how hFixL senses oxygen in this strain as the protein has no oxygen-binding haem group [130].

In summary, the hFixL-FxkR cascade performs the same function as the canonical FixLJ pathway. However, to date it has always been found with a system providing redundancy (see section 1.4 for details) so is never essential for nitrogen fixation. It appears that when the pathway is present in parallel with the canonical system, hFixL-FxkR interacts with FixLJ and supresses its activation. Of note, sequence analysis putatively identified the presence of hFixL-FxkR variants in non-rhizobial species. This suggests that, unlike FixLJ, the hFixL-FxkR system is used as an oxygen sensing mechanism in a variety of contexts [129].

1.3 The FixK transcription factor

Most of the regulation exerted by both the hFixL-FxkR and canonical FixLJ cascades acts indirectly, through their induction of *fixK* [126, 170]. FixK proteins are transcription factors which show some 35-45% conservation across commonly studied strains. They act as intermediates, regulating gene expression in response to oxygen sensing by the FixL and hFixL proteins. FixK is generally found to be crucial for nitrogen fixation, including in *E. meliloti*, *B. japonicum* and *A. caulinodans* [120, 121, 156]. Where FixK operates in conjunction with an FnrN-like regulator (see section 1.4), it is generally non-essential [123, 171]. The FixK and FnrN regulons overlap significantly and the two proteins often appear to operate in a redundant fashion.

1.3.1 Structure of FixK

FixK is typically 27 kDa and is a member of the CRP/FNR superfamily of transcriptional regulators. This superfamily can be divided into three subgroups of CRP-like, NtcA-like and FNR-like proteins, the last of which FixK falls into [172–174]. FNR-like proteins can be further subdivided into three classes [131]. Class IA proteins such as *E. coli* FNR directly sense oxygen and play an important role in the cellular response to microaerobic conditions [175, 176]. Class IB includes *B. japonicum* $FixK_1$ and the FnrN proteins which are also able to directly sense oxygen and regulate genes accordingly, discussed in more detail in section 1.4 [177]. Class IC is composed primarily of rhizobial FixK proteins including *B. japonicum* $FixK_2$, but these do not respond directly to oxygen at the protein level [38]. *B. japonicum* $FixK_2$ has been shown to respond to reactive oxygen species at the post-translational level, but it is unclear how widespread this regulation is in rhizobia and what role it plays in symbiosis [178].

Two domains are conserved across all members of the CRP/FNR superfamily [179, 180]. At the C-terminus is a HTH domain which binds a DNA motif and interacts with RNA polymerase to regulate transcription [181]. This domain and its function are very well conserved across FNR-like proteins, notably reflected in

the highly similar DNA motif sequence these proteins bind [97, 125, 179]. Many members of class IA, IB and IC FNR-like protein bind the so-called "anaerobox", a highly conserved palindromic TTGA-N_6-TCAA operator sequence. Protein binding to this motif can positively and negatively regulate expression, both upstream and downstream of the anaerobox [182]. At the N-terminus is a cyclic nucleotide-monophosphate-like binding domain (cNMP). However, FNR-like proteins diverge in the function of their N-terminal region. In members of class IA and IB this region contains a cysteine-rich motif which enables the proteins to directly sense oxygen and control gene expression accordingly (see section 1.4). This oxygen sensing motif is absent in the FixK proteins of *E. meliloti*, *A. caulinodans*, *R. etli*, *R. leguminosarum* and $FixK_2$ of *B. japonicum* [38, 156]. Class IC FixK proteins therefore do not respond to oxygen at the protein level and instead transmit oxygen sensing by the FixLJ or hFixL-FxkR cascade to the level of downstream gene expression [123, 183]. No evidence has been found to date that these proteins integrate other signals at the post-transcriptional level [181].

1.3.2 Role of FixK in oxygen regulation

Across rhizobial species three members of the FixK regulon have been well studied [38]. The first are the *fixNOQP* and *fixGHIS* operons, respectively encoding a terminal oxidase with high oxygen affinity required for symbiosis and a complex required for its assembly [102, 109]. Second, in *E. meliloti* FixK has been shown to repress *nifA*, which encodes the central activator of nitrogen fixation (discussed in more detail in section 1.5). Lastly, autoregulation of *fixK* appears to be a very common mechanism in rhizobia.

In species encoding the canonical FixLJ cascade, FixK is generally critical for the microaerobic induction of both *fixNOQP* and *fixGHIS*. In *B. japonicum* and *E. meliloti* both operons have upstream anaeroboxes and are regulated by FixK [101, 108, 181, 184]. In *A. caulinodans fixNOQP* is likewise under FixK control but this species has no homolog of *fixS* and no anaerobox is present upstream of the *fixGHI* operon [110, 111]. In contrast to its regulation in other symbiotic

diazotrophs, expression of *fixGHI* in *A. caulinodans* is not controlled by FixK, does not respond to oxygen and occurs under free-living conditions. These differences probably reflect the organism's ability to perform free-living nitrogen fixation using multiple terminal oxidases [168, 185–187].

In species employing the hFixL-FxkR cascade, FixK appears to be a non-essential regulator of *fixNOQP* [123, 130]. In Rlv VF39, a *fixK* mutant was largely unaffected in microaerobic induction of this operon [104, 154]. *R. etli* CFN42 encodes two copies of *fixK*, named *fixK*d and *fixK*f, respectively located on its pCFN42d and pCFN42f plasmids [128]. Regulation of *fixK*f but not *fixK*d is under the control of the hFixL-FxkR cascade, and only FixKf appears to be important for microaerobic induction. *R. etli* CFN42 also encodes two copies of *fixNOQP*, *fixNOQP*d and *fixNOQP*f, both controlled by FixKf [188]. Expression of *fixNOQP*d is the more important of the two for nitrogen fixation, and requires FixKf. However, although the absence of hFixL suppressed *fixNOQP*f expression, *fixNOQP*d was still expressed at significant levels. Expression of *fixK*f may be induced by a second regulator, potentially the FnrNchr protein discussed in more detail in section 1.4.

Less commonly, FixK also regulates expression of *nifA*, the central activator of nitrogen fixation [38, 41]. In *E. meliloti*, FixK represses expression of *nifA* whilst FixJ activates it [182, 189]. The mechanism of FixK repression is unknown and may be indirect as no anaerobox is present upstream of *nifA*. In combination, these opposing regulatory mechanisms probably balance expression of *nifA*. In *A. caulinodans*, FixK instead activates expression of *nifA* [167, 190]. In *B. japonicum*, *nifA* expression is under indirect $FixK_2$ control [125, 191]. In this species expression of $rpoN_1$, encoding the sigma factor σ^{54} required for NifA-mediated transcriptional activation, is controlled by $FixK_2$ [192]. The FixLJ-$FixK_2$ pathway therefore indirectly regulates the activity of NifA in this species. However, control by the pathway is not complete as a redundant $rpoN_2$ paralog exists which does not appear to be regulated by oxygen. A similar situation exists in *R. etli* CNPAF512 where only one of the *rpoN* paralogs is believed to be regulated by FixK. As expected from post-transcriptional regulation of NifA in this strain, the FixLJ cascade induced

nifH through NifA upregulation but had no effect on *nifA* expression. Because the strain has three copies of *nifH*, abolishing FixLJ regulation reduced but did not eliminate its ability to fix nitrogen. Some rhizobial species thus appear to employ control of RpoN concentration by FixK as a tool to modulate NifA activity.

Autoregulation by FixK is also a common mechanism, regardless of the FixL cascade type. *E. coli* FNR auto-represses its expression by binding to a sequence downstream of its own promoter, presumably sterically hindering transcription by RNA polymerase [193]. In *B. japonicum* $FixK_2$ is negatively auto-regulated [156, 181, 194]. The same effect has been shown in *E. meliloti* [182]. The mechanism of FixK auto-repression in both these species remains poorly understood and two theories have been put forward. The first is a direct repression effect. Two putative FixK binding sites have been found upstream of *E. meliloti fixK* at -487 and -43 relative to the transcription start site (TSS) [158]. It has been suggested that at high concentration FixK binds to both these sites to form a repression loop, resulting in downregulation of *fixK* expression. In a proposed alternative mechanism, repression is an indirect effect [127]. In *E. meliloti*, FixK induces expression of *fixT* [195]. FixT appears to act as an anti-kinase, repressing phosphorylation of FixL [196]. This inhibits the FixLJ cascade and in turn represses *fixK* induction, thus completing the FixK autoregulation loop [197]. However, this system remains poorly understood and it is unclear if the role of FixT is to close this loop or whether the protein acts to integrate another signal into FixLJ-mediated gene expression.

In *R. etli* CFN42, two regulatory inputs have been found at the level of *fixK*f expression. First, in line with the situation in other rhizobia, FixKf auto-represses its expression [128]. Further, a CRP/FNR-type regulator encoded by *stoR*d was recently shown to repress *fixK*f expression [171]. A knockout of this gene appeared to enhance the nitrogen fixation of *R. etli* CFN42 but the function and importance of StoRd remain unclear. A homolog of *stoR*d was found in *E. meliloti* but eliminating this had no effect on nitrogen fixation, suggesting the protein's role varies significantly between species [126]. In Rlv VF39, autoregulation activates rather than represses *fixK* [123]. The *hfixL* and *fixK* genes may form an operon

in this strain suggesting both are auto-activated directly by FixK and indirectly by hFixL. This operon arrangement has yet to be confirmed and the biological significance of such a feedback loop is unknown.

Recent work has begun to explore the wider FixK regulon and found that this extends far beyond the three common targets discussed above. In *B. japonicum*, the direct regulon of $FixK_2$ is reported to contain over 200 members, an order of magnitude more than that of FixJ [170]. A study using shotgun proteomics recently reported over 600 genes were uniquely expressed in *B. japonicum* under microoxic conditions [198]. In *E. meliloti*, FixK was found to have one of the largest regulons of any TCS studied to date [126]. The role of many of these targets and the importance of their regulation by FixK has yet to be established. Notable examples are the *B. japonicum* haem biosynthesis pathway genes *hemA*, *hemB* and *hemN*, all members of the FixK regulon [181, 199]. In rhizobia these genes are typically expressed under symbiotic conditions and *hemA* mutants abolish nitrogen fixation. However, they are not essential in *B. japonicum* which also expresses them under non-symbiotic conditions [200]. In *R. etli* CFN42, FixKf also regulates the response to nitric oxide (NO) [201]. Nitric oxide signalling is implicated in symbiosis, but its role is unclear. As with most other members of its regulon, the importance of FixK haem biosynthesis control is therefore unclear. Further investigations of the FixK regulon are likely to reveal many additional functions regulated in response to oxygen concentration.

1.4 The FnrN transcription factor

The FnrN proteins are transcription factors which directly sense oxygen concentration and regulate rhizobia during symbiosis. *B. japonicum* $FixK_1$ belongs to the same group and will also be discussed in this section [125]. Like FixK, FnrN homologs are FNR-like proteins and typically induce expression of *fixNOQP* and *fixGHIS*. Both FnrN and FixK mediate transcriptional regulation by binding to anaerobox operators [108, 182, 202, 203]. However, whilst FixK proteins are class IC and oxygen insensitive (see section 1.3), FnrN homologs are of class IB and can directly sense oxygen [176, 179]. The presence of FnrN correlates with that of the hFixL-FxkR-FixK pathway, but the protein appears to individually recapitulate the full oxygen sensing and regulating functions of that pathway. Of note, an *E. meliloti fixJ* mutant could be partially complemented by expression of *fnrN*, showing the protein also covers some of the function of the FixLJ cascade [204]. FnrN and FixK usually act in a redundant fashion, with organisms retaining at least some nitrogen fixation activity if either is individually eliminated [123]. This is in contrast to organisms employing only the canonical FixLJ-FixK cascade, where it is usually essential for symbiotic nitrogen fixation [38]. Beyond introducing a degree of redundancy, the function of FnrN remains poorly understood. It may serve to produce more finely-tuned regulation, including by responding to a different oxygen concentration than the hFixL-FxkR TCS. FnrN may also provide more responsive regulation for rhizobia exposed to rapidly fluctuating oxygen concentrations [205, 206].

FnrN was first found in Rlv VF39 during a search for endogenous regulators able to activate heterologous *E. meliloti fixNOQP* expression under microaerobic conditions [202, 204]. Two copies of *fnrN* are also found in *R. leguminosarum* bv. *viciae* UPM791 (Rlv UPM791) and a homolog appears to be present in Rlv3841 [152, 206]. This suggests the protein is broadly conserved across strains of *R. leguminosarum*. Likewise, two FnrN homologs regulate nitrogen fixation in *R. etli* CFN42 and one in *R. etli* CNPAF512, suggesting FnrN proteins are also conserved across *R. etli* strains [131, 188]. $FixK_1$ in *B. japonicum* is also an FNR-like class IB protein but no FnrN homologs have been found in *E. meliloti* or *A. caulinodans* to

date [125]. Like *fxkR*, the lack of co-localization between *fnrN* and other symbiotic genes in Rlv VF39 and Rlv3841 suggests it was originally a non-symbiotic regulator. This is supported by the wide distribution of FNR-like proteins, including in soil bacteria [170, 207–209]. In the non-fixing marine bacterium *Dinoroseobacter shibae*, FNR-like proteins regulate its transition from aerobic to anaerobic growth and induce the expression of a high affinity terminal oxidase [210].

Rlv VF39 employs both FixK, regulated by the hFixL-FxkR pathway, and FnrN. In this strain, mutation of *fnrN* was found to reduce nitrogen fixation activity by some 40% relative to WT [204]. A knockout of *hfixL* or *fixK* respectively resulted in 60% and 80% reduction of fixation activity in this strain [123]. A double *fixK fnrN* mutation eliminated all nitrogen fixation activity. Thus, the proteins are semi-redundant, parallel rather than hierarchical, and collaborate to activate expression of genes required for nitrogen fixation. In Rlv UPM791, there is no FixK homolog but a similar redundancy exists between its two *fnrN* homologs [183]. Individual *fnrN* mutants retained nitrogen fixation activity whilst a double mutant abolished it [97, 211]. *R. etli* CFN42 also shows extensive redundancy, with complete elimination of nitrogen fixation only observed in a triple mutant of its two *fnrN* homologs and *hfixL* [188]. In *R. etli* CNPAF512, the *fnrN* mutant showed a severe reduction in nitrogen fixation, indicating a near-essential role for the gene in this strain [131]. The role of the FixLJ system in this species is presently unknown, and it is possible this has been entirely replaced by FnrN as seen in Rlv UPM791.

A different system operates in *B. japonicum*, which contains both the class IC $FixK_2$ protein and the oxygen-sensing class IB $FixK_1$ protein. In contrast to the situation found in *R. leguminosarum* and *R. etli* strains, $FixK_1$ is not a redundant regulator and is under $FixK_2$ control in a hierarchical cascade [125, 156]. $FixK_2$ is required for nitrogen fixation in *B. japonicum* but $FixK_1$ has no effect on fixation, indicating the two perform very different functions in this species [212, 213]. In summary, FnrN is a key regulator of nitrogen fixation in several species. It is commonly used as a parallel, semi-redundant system for the hFixL-FxkR cascade and in some strains replaces it entirely. However, important variations

still exist in the protein's role across species, as demonstrated by the situation in *B. japonicum* where the class IB $FixK_1$ protein is not required for fixation and is under the control of $FixK_2$.

1.4.1 Structure of FnrN

FnrN proteins and $FixK_1$ in *B. japonicum* are FNR-like class IB transcription factors, capable of sensing and responding to oxygen concentration at the protein level. In a typical FNR-like protein oxygen is sensed by an N-terminal sensor domain [214, 215]. Binding of oxygen triggers a conformational change that results in formation of DNA-binding FNR dimers via an interface at the C-terminal domain [216, 217]. The *E. coli* FNR protein is the best understood example of this model and studies have consistently shown that rhizobial class IB proteins are functionally very similar to this protein [173, 193, 218].

Under anaerobic conditions *E. coli* FNR ligates a $[4Fe\text{-}4S]^{2+}$ iron-sulphur cluster [219, 220]. The cluster is coordinated by four cysteine residues [221, 222]. Three of these are grouped in a so-called cysteine-rich motif (Cys-X_2-Cys-X_5-Cys) at the N-terminus of FNR and the fourth is located at a conserved position in the central part of the protein. The cluster mediates FNR dimerization, the active form of the protein [223]. Exposure to oxygen deactivates FNR by converting the iron-sulphur cluster to $[2Fe\text{-}2S]^{2+}$, and sustained exposure causes the protein to unbind the cluster completely [224, 225]. FNR in *E. coli* therefore rapidly and stringently responds to intracellular oxygen concentration and forms active dimers only under low oxygen conditions. Rhizobial FnrN proteins and *B. japonicum* $FixK_1$ are a different class of FNR-like proteins but employ a very similar oxygen-sensing mechanism. Like *E. coli* FNR, three cysteine residues are grouped in a Cys-$X_{2/3}$-Cys-X_7-Cys motif at the N-terminus of these proteins, with the fourth found in a central position [202, 204, 213]. It is therefore widely assumed rhizobial FnrN proteins respond to oxygen concentration through a mechanism very similar to *E. coli* FNR [202]. This is supported by multiple complementation studies. The Rlv VF39 FnrN protein appears to be able to form heterodimers with *E. coli*

FNR that could bind anaeroboxes and activate microaerobic induction of target genes. Rlv VF39 FnrN could complement an *E. coli fnr* mutant for growth on nitrate and promote anaerobic induction of several FNR targets including *narGHJI*, *nirB* and *fdnGHI* [213]. Complementation of the *E. coli fnr* mutant was also possible with *B. japonicum* $FixK_1$ but not its class IC $FixK_2$ protein. Conversely, *E. coli* FNR complemented an Rlv VF39 *fnrN* mutant for regulation of *fixNOQP*. In summary, *B. japonicum* $FixK_1$ and the rhizobial FnrN proteins show strong similarity to the *E. coli* FNR protein, and all appear to act as integrated oxygen-sensing gene regulation systems.

At their C-terminus, homologs of both class IC FixK and class IB FnrN proteins as well as *E. coli* FNR encode a very highly conserved HTH DNA-binding domain [173, 179, 226]. Under anaerobic conditions a dimerization interface is exposed in this domain and protein dimers are formed which can bind DNA and activate transcription. Key residues involved in DNA binding by this domain are conserved across the three proteins and all bind a near-identical palindromic TTGAT-N_4-ATCAA operator called an anaerobox in the promoters of the genes they regulate [221, 227, 228]. There is therefore an inherent overlap in the gene targets of FNR, FixK and FnrN proteins. In rhizobia, this enables FixK and FnrN to regulate the same genes and thus act in a parallel and often redundant fashion.

1.4.2 Role of FnrN in oxygen regulation

E. coli fnr is expressed regardless of oxygen concentration but is subject to negative autoregulation under microaerobic conditions. In contrast expression of rhizobial *fnrN* homologs appears to occur only under microaerobic conditions, but autoregulation is also found in these species.

Both *fnrN* genes in Rlv UPM791 are positively and negatively auto-regulated under microaerobic conditions [205]. This effect is due to the presence of two anaeroboxes in the promoters of the *fnrN* genes; a high-affinity site in the distal region (-42.5 relative to the TSS) and a low-affinity site in the proximal region (-10). FnrN binding to the distal anaerobox induces *fnrN* expression. At high

FnrN concentration, the protein binds to the low-affinity site and represses *fnrN* transcription. This dual regulatory mechanism is proposed to balance FnrN concentration. Two anaeroboxes are also found in front of the *fnrN* genes of Rlv VF39, Rlv3841 and *R. etli* CFN42, suggesting a similar auto-regulatory balancing mechanism operates in these rhizobia [48, 152, 188].

Rlv VF39 *fnrN* is strongly induced under microaerobic conditions as expected from the presence of anaeroboxes upstream of the gene. The presence of these anaeroboxes suggests FixK or FnrN, or both, act to regulate *fnrN* expression. Conflicting results have been published on the importance of *fnrN* autoregulation in this strain. Several studies reported positive autoregulation of *fnrN* [202, 229]. Another study reported that hFixL was essential for microaerobic *fnrN* induction, suggesting no or limited autoregulation [154]. FixK apparently did not mediate this regulation and no FxkR-binding K-box has been identified in front of *fnrN*. It is therefore unclear how hFixL regulates *fnrN* in this model. The finding that autoregulation and hFixL both play important roles may be reconciled by a model in which the positive autoregulatory feedback loop must be initiated by hFixL-dependent induction. However, this is in apparent contradiction to other results showing hFixL and FnrN are redundant regulators and that a *hfixL* mutant retains nitrogen fixation activity [123, 131]. Further investigations will be required to clarify the mechanisms leading to *fnrN* expression in Rlv VF39.

In *R. etli* CNPAF512, two possible regulatory mechanisms also exist with the potential to regulate microaerobic *fnrN* expression; autoregulation and control by the organism's traditional FixLJ cascade [131]. Two anaeroboxes are present in front of the gene and positive autoregulation of FnrN under microaerobic conditions has been confirmed. To investigate the possibility of *fnrN* control by a FixLJ-FixK system, the gene's native *R. etli* CNPAF512 promoter was introduced into the *E. meliloti* host. However, no regulation by the endogenous *E. meliloti* FixLJ-FixK system was found. Transcription of *fnrN* in *R. etli* CNPAF512 thus appears to rely solely on autoregulation. This implies FixK does not bind the gene's anaeroboxes to regulate it, or this binding is biologically irrelevant.

The CFN42 strain of *R. etli* has two differentially regulated copies of *fnrN*, *fnrN*d and *fnrN*chr [131, 188]. Under microaerobic conditions *fnrN*d is induced through at least three mechanisms; by FnrNchr, by hFixL and by FnrNd autoregulation. Single mutants in *fnrN*chr, *hfixL* or *fnrN*d only partially reduced microaerobic induction of *fnrN*d. A single mutant in *fixK*f did significantly decrease the gene's expression, whilst *fixK*d had no effect on transcription of either *fnrN* homolog. It is believed FixKf plays such a crucial role because it is required not only for induction by the hFixL-FxkR pathway but also for FnrNd production and consequently *fnrN*d autoregulation. This supports the existence of a cooperative induction model, also theorized in Rlv VF39, with both hFixL TCS-based regulation and autoregulation of FnrN. FixKf is likely involved in both mechanisms, explaining its critical role in *fnrN*d expression. By contrast, expression of the second *fnrN* copy *fnrN*chr instead appears to be largely repressed by both FnrNd and the hFixL-FxkR pathway under microaerobic conditions. However, no expression of *fnrN*chr occurred in an *fnrN*d *hfixL* double mutant, suggesting these proteins also have a positive regulatory role. Of note, hFixL-FxkR repression of *fnrN*chr was not relieved by a mutation in *fixK*f suggesting FxkR directly inhibits transcription. However, no consensus K-box is present in the *fnrN*chr promoter, so it is unclear how FxkR regulates the gene's expression. Further, it has been proposed FnrNchr acts to induce *fixK*f, resulting in another regulatory feedback loop that may act to balance the production of FnrNchr or FixKf, or both. This multitude of interconnections between the hFixL-based cascade and the two FnrN homologs results in exquisitely complex regulation that is evidence of oxygen's essential and finely tuned role in controlling nitrogen fixation in *R. etli* CFN42.

Besides autoregulation, genes induced by FnrN commonly include the *fixNOQP* and related *fixGHIS* operons. FnrN and FixK proteins often appear to function in a redundant fashion when they coexist in an organism, as found in Rlv VF39 and *R. etli* CFN42 [104, 177, 188]. The former encodes two copies of the terminal oxidase, *fixNOQP*c and *fixNOQP*d. FnrN is an important regulator of both, with the hFixL-FxkR pathway also shown to play a significant role [154, 202, 229, 230].

Redundancy can also exist in situations where no class IC FixK homolog is present, as is the case in Rlv UPM791 [183]. This organism encodes two copies of *fnrN*, a duplication which appears to compensate for the lack of a FixK homolog. Both *fnrN* homologs are individually sufficient for microaerobic induction of *fixNOQP* [97, 211]. Thus FnrN and NifA together appear sufficient to regulate expression of core nitrogen fixation machinery and associated functions with no TCS involvement identified to date [205, 206, 231].

Even more redundancy is found in *R. etli* CFN42, which encodes two copies of *fnrN* (*fnrN*chr and *fnrN*d) [188]. Single *fnrN* mutants show no effect on *fixNOQP* expression but half WT levels were reported in a double *fnrN* mutant. The remainder of *fixNOQP* induction is due to FixKf regulation under control of the hFixL-FxkR pathway. Of note, the anaeroboxes of all *fnrN* and *fixNOQP* promoters are identical in this strain but differences are nevertheless observed in the regulatory roles of FnrNchr, FnrNd and FixKf. This may be due to differences in their respective anaerobox affinities, a mechanism likely to play a role in other rhizobia as well. The importance of these three regulators also varies temporally. At 32 days post-inoculation a double *fnrN* mutant had a limited effect on nitrogen fixation. At 42 days, the same *fnrN* double mutant drastically reduced activity. In this strain, the FnrN proteins therefore appear to become more important in the later stages of symbiosis. This is in line with *R. etli* CFN42's use of a highly complex oxygen regulation network to enable very finely tuned regulation, both at a spatial and temporal level. In a similar vein, in *R. etli* CNPAF512 *fnrN* is important for *fixNOQP* and *fixGHIS* expression in the early stages of symbiosis but plays a smaller role in the late stage [131].

B. japonicum's class IB FNR-like protein $FixK_1$ operates in a hierarchical cascade, in contrast to the redundancy found in other rhizobia [156]. The $fixK_1$ gene is microaerobically induced under control of the canonical FixLJ cascade via the class IC $FixK_2$ protein [125]. The $FixK_1$ regulon is far smaller than that of $FixK_2$ and microaerobic induction of both *fixNOQP* and *fixGHIS* depends exclusively on the action of $FixK_2$. There is also evidence that $FixK_1$ acts to repress genes

that are activated by NifA, ensuring these remain unexpressed even at low oxygen concentration until the correct symbiotic conditions have been reached. This difference in the role of *B. japonicum* $FixK_1$ and $FixK_2$ may be because the two proteins no longer bind the same anaerobox. Similar differences may exist between FNR-like proteins across rhizobia more generally, but the structural basis for these putative divergences has yet to be investigated.

1.5 The NifA transcription factor

NifA is the central, essential activator of nitrogen fixation across most studied symbiotic and non-symbiotic diazotrophs [41, 232]. In contrast to the FixLJ, hFixL-FxkR and FnrN regulatory systems, which are generally thought to act in the earlier phases of symbiotic establishment, NifA is primarily involved in the final stage of bacteroid differentiation [233]. Key targets of NifA commonly include the nitrogenase (*nif*) genes found across nitrogen fixing species and the *fix* genes specific to symbiotic diazotrophs [52, 234–236]. The protein and its associated regulatory mechanisms have been best studied in non-symbiotic diazotrophs, notably *Klebsiella pneumoniae* [237]. In these species NifA regulation at the transcriptional and post-transcriptional level integrates multiple signals including oxygen concentration, the energy status of the cell and its nitrogen availability [41, 238, 239]. In symbiotic diazotrophs, oxygen is the primary and often apparently sole NifA regulator at both the transcriptional and protein level [38]. Transcriptional regulation of *nifA* is more poorly conserved than that of *fixK* or *fnrN* and differs substantially at the species and strain level. Complementation relying on heterologous expression of *nifA* from its native promoter therefore often fails, but has been demonstrated in a few cases [240–242]. The ability of NifA to regulate targets is much better conserved, so cross-species complementation is common if the protein is successfully expressed. Expression of *B. japonicum*, *E. meliloti* and Rlv UPM791 NifA targets were activated by heterologous *K. pneumoniae* NifA [243–247]. Likewise, *E. meliloti* NifA activated expression of *nif* gene promoters in *K. pneumoniae* and other rhizobia [244, 245].

1.5.1 Structure of NifA

NifA is a σ^{54}-dependent enhancer binding protein (EBP) [248]. The protein's flexibility has to date prevented the resolution of a full crystal structure. At the centre of the protein is a highly conserved AAA+ domain (ATPase associated with diverse cellular activities) [249]. This contains protein binding interfaces enabling NifA oligomerization and activates transcription by driving DNA unwinding and σ^{54} complex formation. It appears to be the only domain essential for transcriptional

activation by NifA [250–252]. Truncation of the other two domains retained activity, albeit with a loss of regulatory control.

The central domain is flanked at the C-terminus by a DNA-binding HTH domain that is also well conserved amongst members of the EBP family [253, 254]. This binds upstream activator sequence (UAS) motifs with a TGT-N_{10}-ACA consensus sequence that is conserved across free living and symbiotic diazotrophs [255, 256]. In contrast to the anaerobox operator bound by FnrN and FixK proteins which is generally found within a hundred base pairs of the TSS, UAS sites can be up to 1 Kb upstream of target genes [257, 258]. NifA forms a loop in the intervening sequence between the UAS and the TSS during its interaction with RNA polymerase to activate transcription [234]. Because of the range at which this mechanism can occur, functional UAS sites need not be within the traditional promoter region of a gene and can be located within the coding sequence of upstream genes [259].

The third domain of NifA is an N-terminal GAF regulatory domain, named after three proteins in which it is found [260]. Although the GAF domain appears to consistently have a regulatory function, its sequence is poorly conserved and the mechanisms and signals it responds to vary significantly across species [261, 262]. In non-symbiotic diazotrophs including *K. pneumoniae* and *Azotobacter vinelandii* the GAF domain plays an important role in NifA repression by NifL binding [240, 252, 263–265]. This interaction between NifA and NifL integrates multiple regulatory signals in *K. pneumoniae*. Expression of *nifL* is activated by NtrC in response to low nitrogen availability [244, 266]. Further, NifL binds an FAD cofactor allowing it to sense oxygen [41, 267]. The oxidation state of FAD responds to oxygen and induces changes in NifL conformation and its sub-cellular localization. These oxygen-regulated effects control both its binding and repression of NifA. NifL regulation of NifA thus combines a nitrogen signal at the level of *nifL* transcription and an oxygen signal at the protein level [268, 269]. This repression is further regulated by the NifA GAF domain which is required for NifL binding [270]. In *A. vinelandii* GAF binds 2-oxoglutarate as a proxy for sensing nitrogen limiting conditions [271].

Binding of 2-oxoglutarate prevents GAF from interacting with NifL, abolishing NifL-mediated oxygen repression. This enables *A. vinelandii* to perform nitrogen fixation under aerobic conditions, in conjunction with a host of mechanisms protecting the nitrogenase from oxygen [272–274]. Derepression of NifA and activation of fixation under nitrogen limiting conditions in *A. vinelandii* thus requires the GAF domain. In *K. pneumoniae*, the GAF domain does not bind 2-oxoglutarate and nitrogen sensing functions primarily through NifA interaction with the GlnK protein instead. There is some evidence that NifL does play a role in nitrogen regulation but it remains unclear what function the GAF domain serves in *K. pneumoniae* [275].

No homolog of NifL is found in rhizobial diazotrophs and the role of the GAF domain in these organisms is unknown [41, 166]. Deletion of the domain in *E. meliloti* and *B. japonicum* had minimal effect on NifA activity and did not influence oxygen regulation of the protein [262, 276]. One study reported a strong increase in *E. meliloti* NifA activity upon deletion of the GAF domain [244]. In the same species partial domain deletions abolished NifA functionality, suggesting these led to protein misfolding or instability [262]. In R. *leguminosarum* bv. *trifolii*, NifA does not have a GAF domain [277]. The GAF domain therefore appears not to be required for NifA activity in rhizobia, but a regulatory function remains likely. It may act to regulate NifA activity in the presence or absence of an as yet unidentified interacting protein or small molecule. *A. caulinodans* likely employs a NifA control system combining elements from free-living and symbiotic diazotrophs as it is able to fix nitrogen under both conditions [38, 239, 278]. Uniquely amongst rhizobia studied to date, its NifA is inactive in conditions of high fixed nitrogen concentration. It has been theorized this effect is due to repression by an as-yet unidentified protein functionally similar to NifL. Deletion of GAF abolishes NifA function in *A. caulinodans*, suggesting the domain is required to relieve this repression, as shown in *K. pneumoniae*. Similarly, in *Azospirillum brasilense* the GAF domain of NifA apparently interacts directly with the P_{II} protein to mediate regulation in response to fixed nitrogen levels [250, 252]. Sequence conservation in the GAF domain of *Herbaspirillum seropedicae* indicates a similar mechanism operates in that species

[279]. *A. caulinodans* NifA may therefore interact with a protein which regulates its activity in response to nitrogen availability, with the GAF domain relieving this inhibition. This putative mechanism is in contrast to most other symbiotic diazotrophs, where nitrogen levels are found not to regulate NifA [41]. As discussed in more detail below, *A. caulinodans* NifA is sensitive to oxygen at the protein level so the GAF domain and NifL homolog likely play no role in oxygen control.

In symbiotic diazotrophs NifL-mediated oxygen control is replaced with direct oxygen sensing by the NifA protein [244, 280]. Oxygen sensitivity at the protein level has been demonstrated for *E. meliloti* and *B. japonicum* NifA. This is mediated by a broadly conserved cysteine-rich motif present in an inter-domain linker (IDL) region between the central AAA+ and HTH domains. All of the conserved cysteine residues in this motif were shown to be essential for the activity of *B. japonicum* NifA [276]. The same conserved residues are present in the *A. caulinodans*, *E. meliloti* and *R. leguminosarum* NifA proteins. Several non-rhizobial species employing NifA homologs, including *A. brasilense* and *H. seropedicae*, also retain the IDL but it is not found in the NifA proteins of *K. pneumoniae* or *A. vinelandii* and appears to be mutually exclusive with the NifL-NifA regulatory system. The cysteine-rich motif is similar to metal-binding domains found in other proteins [281]. It is believed to sense the redox state of the cell as an indirect measure of oxygen concentration [280, 282, 283]. A presently well supported model is that under microaerobic conditions a reduced metal ion, likely Fe^{2+}, is bound by the motif. Binding of this ion at the IDL is likely required for a conformational change that produces active NifA. Metal ion binding also appears required for protein stability. In *E. meliloti* and *B. japonicum*, NifA is rapidly degraded under aerobic conditions. Indirect oxygen control through metal ion binding by the IDL therefore strictly regulates activation by rhizobial NifA proteins.

1.5.2 Role of NifA in oxygen regulation

Rhizobia typically regulate *nifA* transcription in response to oxygen concentration, in contrast to free-living diazotrophs where expression responds to nitrogen levels

[41, 234]. Microaerobic induction is achieved through several mechanisms often acting in combination (Figure 1.4). Positive NifA autoregulation resulting in a feedback loop is a very common motif in rhizobia, likely enabling rapid production of the protein upon transition to nitrogen fixation in nodules. Several species also employ a second layer of oxygen regulation at the transcriptional level, including through the FixLJ cascade.

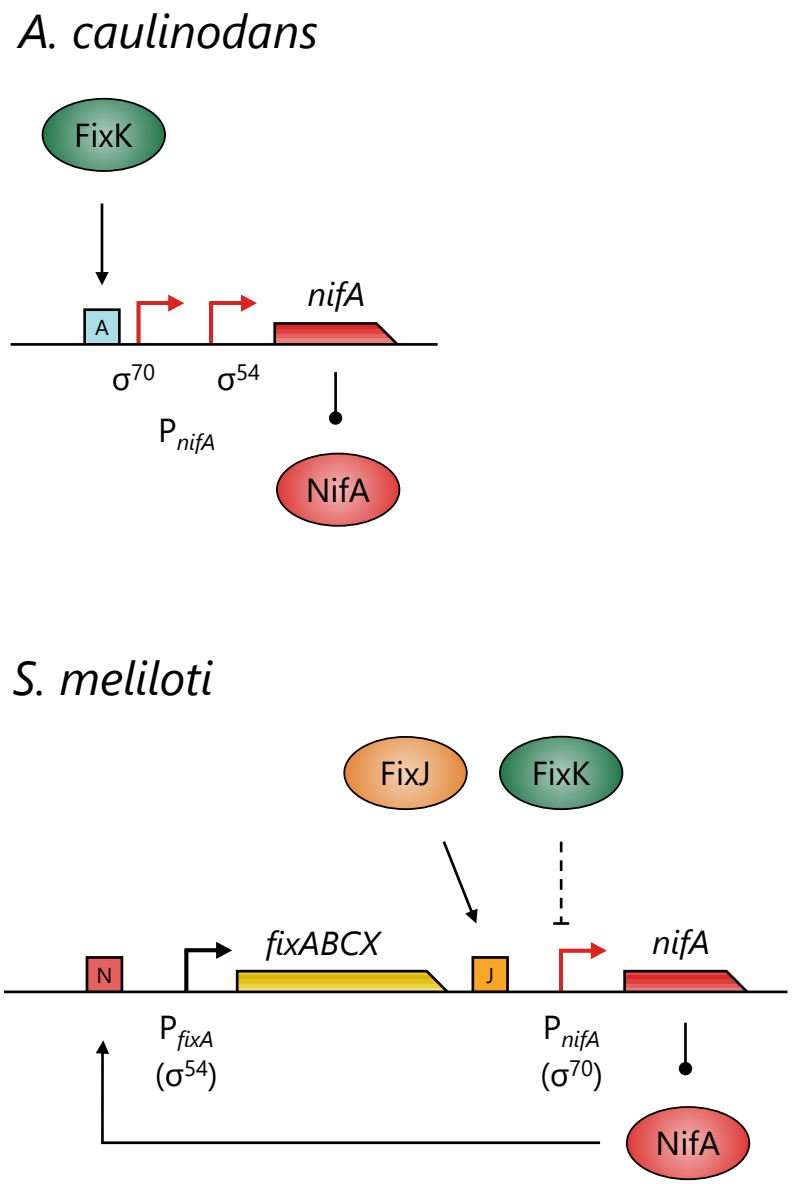


Figure 1.4: Transcriptional regulation of *nifA* across rhizobia. Transcriptional activation and repression are shown as lines ending in arrows and bars, respectively. Squares indicate operators for NifA (N, red), FixJ (J, orange), the anaerobox bound by FixK (A, blue) and the IHF factor operator (I, white). Genes and operons are shown as pointed rectangles. Translation is shown as a line ending in a circle. Proteins are shown as ellipses. The *nifA* gene and its protein are highlighted in red, *fixABCX* in yellow and *fixR* in brown. The FixJ and FixK proteins are given in orange and green, respectively. Proteins which activate transcription but are not directly involved in oxygen regulation are coloured white. Diagrams represent typical systems, but regulation varies at the strain level; see text for details. Figure continues on next page.

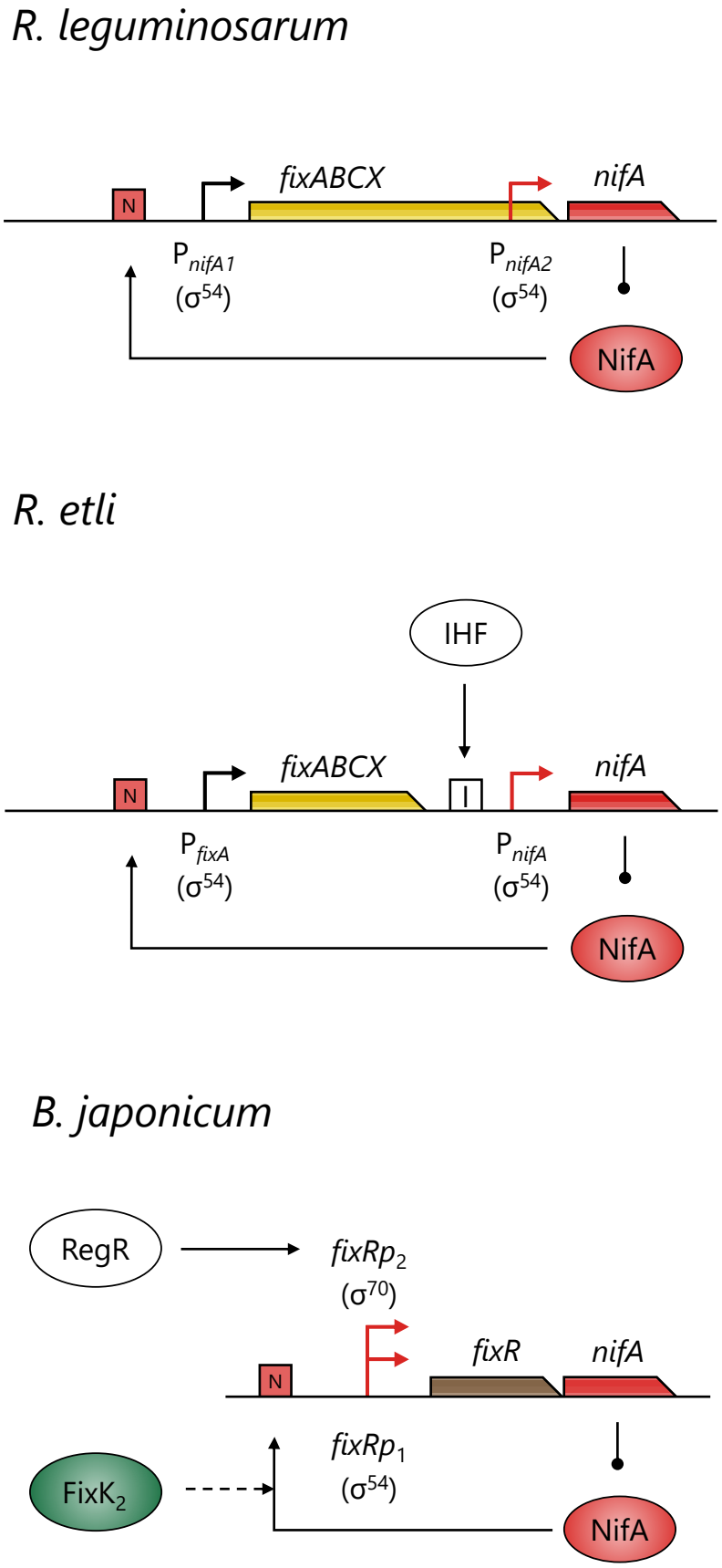


Figure 1.4: Transcriptional regulation of *nifA* across rhizobia (continued).

In *A. caulinodans nifA* is transcribed from two promoters. The first is σ^{54}-dependent, and the second σ^{70}-dependent and regulated by the FixLJ system through FixK [239, 284]. FixK was required under symbiotic and free-living conditions, demonstrating the FixLJ system plays a role in both contexts in *A. caulinodans* [166]. The two promoters appear to be integrated in a single feedback system. Activation of the σ^{70} promoter upregulates activity from the downstream σ^{54} promoter. A putative NifA-binding motif overlaps with the σ^{70} promoter and NifA binding here likely represses transcription from both promoters. This would balance NifA production, not unlike the feedback demonstrated for *fnrN* expression in some species (see section 1.4). The nitrogen status of the cell also plays a role in *nifA* regulation, via the *ntrBC* and *ntrYX* systems [285–287].

In *E. meliloti* a similar dual feedback system operates with two promoters, P_{fixA} (σ^{54}) and P_{nifA} (σ^{70}), regulating *nifA* [288, 289]. The gene is downstream of and co-transcribed with the *fixABCX* operon, itself a target of NifA activation. Thus *nifA* is under P_{fixA} control and auto-regulated, an effect shown to be essential for its expression in nodules [288, 290]. The second promoter P_{nifA} directly upstream of *nifA* is also induced under microaerobic conditions, by FixJ [120, 291]. Simultaneously FixJ induces *fixK*, but FixK represses transcription from P_{nifA}, thereby balancing *nifA* expression when the FixLJ system is active [182]. This system does not appear widespread, and the mechanism of FixK-mediated P_{nifA} repression is unclear. *B. japonicum nifA* is not regulated by FixLJ and likewise there are no FixJ or anaerobox operators upstream of *R. leguminosarum* or *R. etli nifA* [122, 292].

A similar dual control system operates in Rlv UPM791, where the adjacent *fixABCX-nifAB* operons are inside a 10-member co-transcribed gene cluster [234]. Expression of *nifA* occurs only under symbiotic conditions in this strain [293]. Two σ^{54} type promoters regulate expression of the gene, P_{nifA1} transcribing the full cluster and P_{nifA2}, likely located immediately upstream of *fixA*. The promoter regulating the full cluster is auto-activated by NifA and responsible for a majority of *nifA* expression. The P_{nifA2} promoter appears to be an incomplete duplication of the first and does not contain NifA-binding UAS sites. A deletion of P_{nifA1}

expressed *nifA* sufficiently from P_{nifA2} to retain limited nitrogen fixation. Sequence information suggests a similar dual promoter system operates in Rlv3841 [152]. It is possible other regulators act on P_{nifA2}, but their identity has yet to be determined.

In *B. japonicum* and *R. etli* CNPAF512 *nifA* is auto-activated but regulators which do not respond to oxygen also participate in its transcription [289, 294, 295]. Expression of *nifA* in these organisms occurs under aerobic conditions [295, 296]. This suggests the emphasis is on post-transcriptional oxygen regulation of NifA, as found in *A. brasilense*.

Auto-activation in *R. etli* CNPAF512 proceeds from a promoter upstream of *fixABCX* as in Rlv UPM791, accounting for at least half of *nifA* expression [289, 295]. A second promoter exists directly upstream of *nifA* that does not respond to oxygen concentration [297]. It was recently shown that this second promoter is regulated by an IHF-like protein, a mechanism also found in *H. seropedicae* [298, 299]. The implications of this regulation are unclear, and it remains to be determined whether this induces or represses *nifA* transcription.

In *B. japonicum* the gene is part of a *fixR-nifA* operon [283]. Two overlapping promoters regulate expression, $fixRp_1$ and $fixRp_2$ of type σ^{54} and σ^{70} respectively [300]. Under microaerobic conditions $fixRp_1$ is auto-activated by NifA, resulting in five-fold induction [288]. Under aerobic conditions, *nifA* is expressed but the FixLJ-regulated *rpoN* σ^{54} factor required for NifA activity is not [191, 301, 302]. Thus, non-activating binding of NifA to the $fixRp_1$ promoter instead results in auto-repression when the right oxygen conditions are not present. It is likely this mechanism also represses other members of the NifA regulon. The second, housekeeping $fixRp_2$ promoter controlling *nifA* expression is regulated in part by the RegS-RegR (RegSR) TCS [303, 304]. A mutation of *regR* was found to reduce nitrogen fixation to 2% of WT levels, demonstrating the importance of this second promoter. The RegSR system has since been shown to regulate expression of other targets which play important roles in antibiotic resistance, symbiotic host specificity and denitrification [305, 306]. RegR activation of *nifA* expression is independent of oxygen, suggesting the gene can be expressed under aerobic conditions [307, 308].

RegS is instead believed to sense redox conditions at the cell surface in a manner analogous to the well-studied *R. capsulatus* RegBA system. Of note, a *regS* mutant only minimally reduced nitrogen fixation activity. This is in line with denitrification studies which suggest RegR is also regulated by an as yet unidentified alternative sensor kinase [283]. RegR control of *nifA* therefore likely integrates at least one additional control signal that has yet to be identified.

Beyond its well-studied *nif* and *fix* gene targets, active NifA controls a far larger regulon and many members of this have yet to be investigated [309]. It is theorized that control of some NifA targets evolved through "regulatory noise", wherein horizontally transferred genes adapt to their host through changes in their operator sequences [247, 310]. In some rhizobia NifA activates expression of hydrogenase genes allowing cells to use the hydrogen produced by nitrogen fixation as an energy source [247, 293, 311, 312]. In *E. meliloti* and *B. japonicum* NifA also regulates a ferredoxin critical for nitrogen fixation which may function as a direct electron donor to nitrogenase [259, 282]. The NifA regulon in *B. japonicum* appears so broad that it has been suggested the protein should be thought of as a general regulator of anaerobic processes rather than one specific for nitrogen fixation. This may explain why at least part of its expression, through the σ^{70} $fixRp_2$ promoter, apparently escapes oxygen regulation. A recent study of *B. japonicum* DOA9, which like *A. caulinodans* is able to fix under free-living and symbiotic conditions, found that this strain employed two *nifA* homologs [313]. Both were individually sufficient for successful symbiotic nitrogen fixation, but one was specifically essential for free-living fixation. Like the functional similarity of hFixL-FxkR and FnrN, reiteration of NifA may be another rhizobial strategy to improve the robustness of oxygen regulation and create a more finely-tuned system.

As found with FixJ, the NifA regulon is probably augmented by its control of transcription factors. One of these may include the FixLJ cascade itself. In *E. meliloti*, a *nifA* mutant appeared to show increased expression of *fixLJ* [314]. NifA may therefore indirectly suppress FixLJ targets during the final stage of the symbiosis. Controlling additional regulatory mechanisms would enable the protein

to indirectly influence a large pool of genes beyond its direct targets. The full role of NifA thus appears to extend well beyond what is currently known.

1.6 Engineering nitrogen fixation for agriculture

Because of its high cost and environmental impact, there is considerable interest in reducing the use of nitrogen fertilizer in agriculture [15, 315, 316]. Most major crop plants do not form symbioses with biological nitrogen fixers [317]. Thus, extending Biological Nitrogen Fixation (BNF) into crops could partially replace the use of nitrogen fertilizer [318]. This possibility has been the subject of study for several decades [5–7, 17].

Two broad approaches to extend BNF to non-legumes have emerged [18, 23]. The first approach is plant-centric, and aims to engineer nitrogenase activity directly into crops without the need for bacterial symbiosis [19, 21, 319]. Studies in unicellular eukaryotes have demonstrated the validity of this approach [320–322]. Work in plants has focused on expressing nitrogenase components inside mitochondria or chloroplasts, as both could create the appropriate low-oxygen environment necessary for nitrogenase activity [323–325]. To date, most components of the nitrogenase complex and its associated machinery have been expressed individually or in combination in plant mitochondria [326]. However, *in planta* nitrogenase activity has yet to be achieved [327].

The second approach to extending BNF is bacteria-centric, with the aim of engineering new plant-bacteria symbioses into crop plants [22, 328, 329]. This could be done by engineering an existing rhizobium or modifying a natively non-fixing bacterium [330]. Although centred on bacteria, a key part of this approach will likely be modifying crop plants to develop nodules [331]. Significant changes would also need to be made to the native mechanisms that regulate BNF in bacteria [332, 333]. There has been some success in the past in partially or wholly refactoring existing regulation [328, 334, 335]. More recently, the transfer of nitrogen fixation clusters between organisms, retaining most native regulation, has also shown considerable promise [336]. As the central activator of nitrogen fixation, engineered control of NifA could be exploited to modify the regulation of diazotrophs [18]. Using NifA would avoid the need to create an entirely new regulation system but depends on a good understanding of the protein. NifA regulation in non-symbiotic diazotrophs

has been relatively well studied [41]. However, much remains unknown about how NifA is controlled in symbiotic diazotrophs such as *Rhizobium* species [1]. A better understanding of symbiotic NifA regulation would thus substantially facilitate future efforts to engineer nitrogen fixation in natively fixing bacteria. It may also be critical for attempts to transfer fixation machinery into organisms which cannot natively perform BNF.

1.7 Aims of this project

Rhizobia undergo a complex lifestyle transition in their partnership with legumes which requires finely tuned regulation. Nitrogen fixation is energy intensive but the nitrogenase complex must operate under microaerobic conditions, creating a conflicting demand for oxygen. These paradoxical requirements have driven a host of evolutionary adaptations in legumes and rhizobia to establish a successful symbiosis. As rhizobia differentiate from a free-living lifestyle in soil to nitrogen fixing bacteroids in nodules, the oxygen concentration drops several orders of magnitude. Multiple oxygen-sensing systems, including the FixLJ and hFixL-FxkR cascades, FnrN and NifA enable rhizobia to respond appropriately to oxygen concentration during symbiosis. Oxygen regulation systems show considerable inter- and intra- species variation, due in part to the spread of symbiotic nitrogen fixation through horizontal gene transfer (HGT) of symbiosis islands [337]. Several regulatory themes are nevertheless broadly conserved but remain poorly understood. Some interaction between oxygen regulation systems appears to be universal in rhizobia. There is evidence for both hierarchical and parallel arrangements of oxygen regulation systems. Parallel arrangements can produce redundancy, with multiple oxygen regulation systems regulating the same target. Redundancy is also achieved by the reiteration of regulators within an organism, with two (or more) homologs playing subtly different roles. Hierarchical arrangements have also been found but have received less attention. In *R. etli* CFN42, regulation by hFixL-FxkR is important at the start of symbiosis but FnrN becomes critical in later stages [188]. How such hierarchical arrangements are created, and what advantage they confer, is poorly understood.

In Chapters 3 and 4, our first aim was to understand how two widespread systems, the hFixL-FxkR-FixK and FnrN pathway, interact. We chose to study the model organism Rlv3841, which encodes both systems [338]. We investigated whether the two could interact, with hFixL-FxkR-FixK inducing *fnrN*, as suggested by sequence data [152]. Past work in rhizobia has typically found that the FixLJ-FixK/hFixL-FxkR-FixK cascades are active at a relatively high oxygen concentration [132,

139]. The oxygen range at which FnrN is active was not known. Because of the oxygen gradient inside nodules, understanding the oxygen sensitivity of a system helps pinpoint its role during symbiosis. We therefore also set out to establish the oxygen sensitivity of FnrN to understand what role it plays in the context of hFixL-FxkR-FixK and NifA regulation. One possibility was that FnrN bridged the gap between these systems.

Our second aim in Chapters 3 and 4 was to understand the role of autoregulation in oxygen regulation. Autoregulation is common in oxygen regulation systems, acting both to balance protein production and to enable a rapid increase in transcription through a self-amplifying loop [41]. This suggests rapidly producing these proteins is important to establish symbiosis, although the reason for this remains unclear. Sequence data from Rlv3841 indicated that *fnrN* was auto-regulated, as has been suggested in the past with other organisms [131, 152, 205]. We aimed to determine if a similar mechanism was taking place in Rlv3841 and understand how this interacted with potential hFixL-FxkR-FixK mediated *fnrN* induction.

In Chapter 5, we focused on NifA, another key oxygen regulation system that regulates the final stages of bacteroid differentiation [249]. Much remains unknown about how rhizobial NifA proteins are controlled. Our first aim was to better understand the regulation of the NifA protein in Rlv3841 ($NifA_{Rlv3841}$). We investigated whether bypassing native transcriptional regulation could enable activity in free-living Rlv3841. We also investigated the role of the GAF domain, found in almost all rhizobial NifA proteins but of unknown function in these proteins [238, 277, 339]. It was thought to integrate one or more signals mediated by protein interactions or small molecule binding. Identifying these could provide important insights into other mechanisms regulating symbiotic establishment, and how these interact with oxygen regulation. We deleted the NifA GAF domain in Rlv3841 ($NifA_{Rlv3841}$ ΔGAF) and studied how this impacted on the activity and regulation of $NifA_{Rlv3841}$.

Our second aim in Chapter 5 was to investigate the potential for NifA engineering as a tool to modify the regulation of BNF. Past work had shown that NifA is widely

cross-compatible, with NifA from one organism able to active genes in another [243–247]. We studied the activity of *A. caulinodans* ORS571 NifA in Rlv3841 ($NifA_{ORS571}$), and vice versa the activity of $NifA_{Rlv3841}$ in *A. caulinodans*. Finally, we set out to create strains of Rlv3841 able to fix nitrogen in free-living conditions by engineering both NifV (homocitrate synthase) and NifA activity.

2

Materials and Methods

Contents

2.1 Media, antibiotics and other chemicals

2.1.1 Media

E. coli strains were grown at 37 °C in Luria Bertani (LB) broth (10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl) and shaken at 200 RPM [340]. For solid media, 1.4% w/v agar was added.

Rhizobium and *Azorhizobium* strains were grown in Tryptone-Yeast (TY) media (5 g L^{-1} tryptone, 3 g L^{-1} yeast extract, 6 mM NaCl) unless otherwise specified, at 28 °C and 37 °C respectively [341]. For solid media, 1.75% w/v agar was added. Liquid cultures were shaken at 200 RPM unless otherwise specified. In some experiments, strains were grown in Universal Minimal Salts (UMS) media with glucose at 10 mM, and with or without ammonium chloride as specified [342]. UMS was made with 0.5 mM K_2HPO_4, 0.5 g L^{-1} $MgSO_4.7H_2O$, 0.2 g L^{-1} NaCl and 4.19 g L^{-1} MOPS, then adjusted to pH 7.0. After autoclaving, the following sterile solutions were added: 1 mL of 12 g L^{-1} $FeSO_4.7H_2O$, 1 mL of 75 g L^{-1}

Antibiotic	*E. coli*	Rlv3841	*A. caulinodans* ORS571
Ampicillin (Amp)	100	N/A	100
Gentamicin (Gent)	10	20	20
Kanamycin (Kan)	20	50	20
Neomycin (Neo)	20	80/250*	N/A
Nitrofurantoin (Nit)	20	20	N/A
Streptomycin (Str)	25	500	N/A
Spectinomycin (Spec)	50	100	N/A
Tetracycline (Tet)	2	2	N/A

Table 2.1: Working concentrations of antibiotics used in this work. All concentrations are given in µg mL^{-1}. A higher concentration of neomycin was used for interposon mutant selection, indicated by *.

$CaCl_2.2H_2O$, 1 µL of a 1000X vitamin stock solution (0.375 g L^{-1} EDTA-Na_2, 0.16 g L^{-1} $ZnSO_4.7H_2O$, 0.2 g L^{-1}, $NaMoO_4$, 0.25 g L^{-1} H_3BO_3, 0.2 g L^{-1} $MnSO_4.4H_2O$, 0.02 g L^{-1} $CuSO_4.5H_2O$, 1 g L^{-1} $CoCl_2.6H_2O$) and 1 mL of a solution containing 1 g L^{-1} thiamine hydrochloride, 2 g L^{-1} D-Pantothenic acid calcium salt and 100 mg L^{-1} biotin. Also added was a carbon source: 10 mM glucose and 20 mM succinate for *Rhizobium* and *Azorhizobium* strains, respectively. Ammonium chloride was added, if at all, as specified on a per-experiment basis.

Long-term storage of strains was done in a 15% glycerol solution, snap-frozen with liquid nitrogen and stored at -80 °C.

2.1.2 Antibiotics and other chemicals

Antibiotics were added to cultures as appropriate at the concentrations given in Table 2.1.

Azorhizobium caulinodans ORS571 is auxotrophic for nicotinate so 30 µM was added when this strain was grown in UMS. Strains of *E. coli* ST-18 were grown in media supplemented with 50 µg mL^{-1} 5-aminolevulinic acid (ALA). For blue-white

screening of transformations into *E. coli*, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to LB-agar at 40 µg mL^{-1}.

2.2 Bacterial strains, plasmids and primers

2.2.1 Strains

Individual tables are sorted by alphanumeric order of strain name.

Escherichia coli

Name	Relevant characteristics	Source
DH5α	F supE44 lacU169 hsdR17 recA1 endA1 gyrA96 thi-1 relA1 (80lacZM15)	Hanahan 1983 [343]
EC100D pir+	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (Str^R) nupG pir+(DHFR)	Lucigen (Epicentre)
ST18	S17-1 hemA thi pro hsdR-M- chromosomal integrated [RP4-2 Tc::Mu:Kmr::Tn7, Tra+ Trir Str^R	Thoma and Schobert 2009 [344]
OneShot® PIR1	F- lac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA(MluI)::pir-116	ThermoFisher Scientific (Invitrogen)
OneShot® TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (Str^R) endA1 nupG	ThermoFisher Scientific (Invitrogen)

***Rhizobium leguminosarum* bv. 3841**

Name	Relevant characteristics	Source
D5250	WT + pIJ11282 (P_{neo}:*luxCDABE* in pIJ11268 backbone, lux positive control)	Frederix et al. 2014 [345]
LMB403	*hfixL*$_c$:pK19 single crossover	This work
LMB495	*hfixL*$_9$::ΩSpec	This work
LMB496	*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19	This work
LMB542	WT + pIJ11268 (promoterless *luxCDABE*, lux negative control)	Frederix et al. 2014 [345]
LMB648	*fnrN*::ΩTet	This work
LMB673	*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19 *fnrN*::ΩTet	This work
OPS0376	WT (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0136 ($P_{fixK_{9a}}$:*luxCDABE* in pIJ11268 backbone)	This work
OPS0528	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0136 ($P_{fixK_{9a}}$:*luxCDABE* in pIJ11268 backbone)	This work
OPS1267	WT + pOPS0978 ($P_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1268	WT + pOPS0979 (P_{nifH}:*syfp2* in pOPS0786 backbone)	This work
OPS1269	WT + pOPS0980 (P_{fnrN} in pOPS0786 backbone)	This work
OPS1274	LMB648 (*fnrN*::ΩTet) + pOPS0977 ($P_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1275	LMB648 (*fnrN*::ΩTet) + pOPS0978 ($P_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1277	LMB648 (*fnrN*::ΩTet) + pOPS0980 (P_{fnrN} in pOPS0786 backbone)	This work
OPS1278	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0977 ($P_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1279	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0978 ($P_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1281	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0980 (P_{fnrN} in pOPS0786 backbone)	This work
OPS1287	WT + pOPS0977 ($P_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1294	WT + pOPS0785 (J23106:*syfp2* in pOPS0786 backbone, *syfp2* positive control)	This work
OPS1295	WT + pOPS0786 (Promoterless *syfp2* in pOPS0786 backbone, *syfp2* negative control)	This work
OPS1563	LMB403 (*hfixL*$_c$:pK19 single crossover) + pOPS0980 (P_{fnrN} in pOPS0786 backbone)	This work
OPS1565	LMB495 (*hfixL*$_9$::ΩSpec) + pOPS0980 (P_{fnrN} in pOPS0786 backbone)	This work

OPS1573	LMB403 (*hfixL*$_c$:pK19 single crossover) + pOPS0977 (P$_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1574	LMB403 (*hfixL*$_c$:pK19 single crossover) + pOPS0978 (P$_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1575	LMB495 (*hfixL*$_9$::ΩSpec) + pOPS0977 (P$_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1576	LMB495 (*hfixL*$_9$::ΩSpec) + pOPS0978 (P$_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1808	Δ*fxkR*$_9$	This work
OPS1811	OPS1808 (Δ*fxkR*$_9$) + pOPS0977 (P$_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS1812	OPS1808 (Δ*fxkR*$_9$) + pOPS0978 (P$_{fixNOQP_9}$:*syfp2* in pOPS0786 backbone)	This work
OPS1813	OPS1808 (Δ*fxkR*$_9$) + pOPS0980 (P$_{fnrN}$ in pOPS0786 backbone)	This work
OPS2260	LMB648 (*fnrN*::ΩTet) + pOPS1510 (P$_{lac}$:*fnrN* in pOGG280 backbone, genomically integrated by Tn7)	This work
OPS2428	WT + pOPS1593 (P$_{fixNOQP_9}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2429	WT + pOPS1594 (P$_{fnrN}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2431	LMB648 (*fnrN*::ΩTet) + pOPS1593 (P$_{fixNOQP_9}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2432	LMB648 (*fnrN*::ΩTet) + pOPS1594 (P$_{fnrN}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2434	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1593 (P$_{fixNOQP_9}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2435	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1594 (P$_{fnrN}$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This work
OPS2468	WT + pOPS1644 (P$_{fixNOQP_{10}}$:*syfp2* in pJP2 backbone)	This work
OPS2469	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1644 (P$_{fixNOQP_{10}}$:*syfp2* in pJP2 backbone)	This work
OPS2470	WT + pOPS1607 (J23106:*syfp2* in pOGG276 backbone, *syfp2* positive control, genomically integrated by Tn7)	This work
Rlv3841	Wild type *R. leguminosarum* bv. *viciae* 3841, Str^R derivative of strain Rlv300	Johnston and Beringer 1975 [338]
OPS1707	OPS1737 (Δ*nifA*) + pLMB134 (P$_{tauA}$:*sfgfp* in pLMB51 backbone)	This work
OPS1737	Δ*nifA*	This work

OPS1778	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone)	This work
OPS1779	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone) pOPS1009 (P_{tauA}:$nifA_{Rlv3841}$ in pLMB51 backbone)	This work
OPS1824	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS1009 (P_{tauA}:$nifA_{Rlv3841}$ in pLMB51 backbone)	This work
OPS1825	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS1104 (P_{tauA}:$nifA_{Rlv3841}$ GAF in pLMB51 backbone)	This work
OPS1826	OPS1737 (Δ*nifA*) + pOPS0983 (P_{tauA}:$nifA_{ORS571}$ in pLMB509 backbone)	This work
OPS2080	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone)	This work
OPS2081	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS1009 (P_{tauA}:$nifA_{Rlv3841}$ in pLMB51 backbone)	This work
OPS2082	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS1104 (P_{tauA}:$nifA_{Rlv3841}$ GAF in pLMB51 backbone)	This work
OPS2104	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone) pOPS0983 (P_{tauA}:$nifA_{ORS571}$ in pLMB509 backbone)	This work
OPS2233	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS0983 (P_{tauA}:$nifA_{ORS571}$ in pLMB509 backbone)	This work
OPS2261	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone) pOPS1104 (P_{tauA}:$nifA_{Rlv3841}$ GAF in pLMB51 backbone)	This work
OPS2279	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS1591 (P_{tauA}:$nifA_{Rlv3841}$ in pLMB509 backbone)	This work
OPS2398	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone) pOPS1591 (P_{tauA}:$nifA_{Rlv3841}$ in pLMB509 backbone)	This work
OPS2399	OPS1737 (Δ*nifA*) + pOPS1177 (Rlv3841 P_{fixA}:*sfgfp* in pME6041 backbone) pOPS0982 (P_{tauA}:$nifA_{Rlv3841}$ GAF in pLMB509 backbone)	This work
OPS2401	OPS1737 (Δ*nifA*) + pOPS1178 (Rlv3841 P_{nifH}:*sfgfp* in pME6041 backbone) pOPS0982 (P_{tauA}:$nifA_{Rlv3841}$ GAF in pLMB509 backbone)	This work
OPS2467	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1630 (*hfixL*$_9$ complementation plasmid)	This work

OPS2500	*fixK*$_{9a}$ *fixK*$_{9b}$:pRU877 single crossover	This work
OPS2512	OPS2500 (*fixK*$_{9a}$ *fixK*$_{9b}$:pRU877 single crossover) + pOPS0977 (P$_{fixNOQP_{10}}$:*syfp2* in pOPS0786 backbone)	This work
OPS2513	OPS2500 (*fixK*$_{9a}$ *fixK*$_{9b}$:pRU877 single crossover) + pOPS0978 (P$_{fixNOQP_{9}}$:*syfp2* in pOPS0786 backbone)	This work
OPS2515	OPS2500 (*fixK*$_{9a}$ *fixK*$_{9b}$:pRU877 single crossover) + pOPS0980 (P$_{fnrN}$ in pOPS0786 backbone)	This work
OPS2604	OPS1737 (Δ*nifA*) + pLMB504 (*nifV*$_{A.\ vinelandii}$ in pJP2neo backbone)	This work
OPS2605	OPS1737 (Δ*nifA*) + pLMB504 (*nifV*$_{A.\ vinelandii}$ in pJP2neo backbone) pOPS1591 (P$_{tauA}$:*nifA*$_{Rlv3841}$ in pLMB509 backbone)	This work
OPS2606	OPS1737 (Δ*nifA*) + pLMB504 (*nifV*$_{A.\ vinelandii}$ in pJP2neo backbone) pOPS0983 (P$_{tauA}$:*nifA*$_{ORS571}$ in pLMB509 backbone)	This work
OPS2607	OPS1737 (Δ*nifA*) + pOPS1719 (P$_{lac}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7)	This work
OPS2608	OPS1737 (Δ*nifA*) + pOPS1720 (P$_{A1lacO1}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7)	This work
OPS2609	OPS1737 (Δ*nifA*) + pOPS1719 (P$_{lac}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7) pOPS1591 (P$_{tauA}$:*nifA*$_{Rlv3841}$ in pLMB509 backbone)	This work
OPS2610	OPS1737 (Δ*nifA*) + pOPS1720 (P$_{A1lacO1}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7) pOPS1591 (P$_{tauA}$:*nifA*$_{Rlv3841}$ in pLMB509 backbone)	This work
OPS2615	OPS1737 (Δ*nifA*) + pOPS1719 (P$_{lac}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7) pOPS0983 (P$_{tauA}$:*nifA*$_{ORS571}$ in pLMB509 backbone)	This work
OPS2616	OPS1737 (Δ*nifA*) + pOPS1720 (P$_{A1lacO1}$:*nifV*$_{ORS571}$ in pOGG281 backbone, genomically integrated by Tn7) pOPS0983 (P$_{tauA}$:*nifA*$_{ORS571}$ in pLMB509 backbone)	This work
LMB374	*fixK*$_{9b}$:pRU877 single crossover	Graham Hood (unpublished)

Azorhizobium caulinodans ORS571

Name	Relevant characteristics	Source
ORS571	Wild type *A. caulinodans* ORS571	Dreyfus et al. 1981 [185]
OPS1626	*nifA* + pOPS1122 ($nifA_{\text{ORS571}}$-$rpoN_{\text{ORS571}}$ under P_{mocB} control in pOPS0889 backbone)	Marta Mendes (unpublished)
Δ*nifA*	*A. caulinodans* ORS571 with its *nifA* gene deleted. Note this strain does not have a Poole group name (i.e. OPS code)	Ryu et al. 2020 [336]
OPS1292	*nifA* + pOPS0999 (*A. caulinodans* P_{fixA}:*sfgfp* in pOGG026 backbone)	This work
OPS1293	*nifA* + pOPS0981 (*A. caulinodans* P_{nifH}:*sfgfp* in pOGG026 backbone)	This work
OPS1296	WT + pOPS0999 (*A. caulinodans* P_{fixA}:*sfgfp* in pOGG026 backbone)	This work
OPS1297	WT + pOPS0981 (*A. caulinodans* P_{nifH}:*sfgfp* in pOGG026 backbone)	This work
OPS1621	WT + pLMB134 (P_{tauA}:*sfgfp* in pLMB51 backbone)	This work
OPS1777	*nifA* + pLMB719 (P_{tauA}:*mcherry* in pLMB509 backbone)	This work
OPS1796	*nifA* + pLMB426 (P_{tac}:*mcherry* in pRU1097 backbone)	This work
OPS2232	*nifA* + pOPS0999 (*A. caulinodans* P_{fixA}:*sfgfp* in pOGG026 backbone) pOPS0983 (P_{tauA}:$nifA_{\text{ORS571}}$ in pLMB509 backbone)	This work
OPS2105	*nifA* + pOPS0981 (*A. caulinodans* P_{nifH}:*sfgfp* in pOGG026 backbone) pOPS0983 (P_{tauA}:$nifA_{\text{ORS571}}$ in pLMB509 backbone)	This work
OPS2643	*nifA* + pOPS1724 (P_{mocB}:$nifA_{\text{Rlv3841}}$ in pOPS0889 backbone)	This work
OPS2644	*nifA* + pOPS1725 (P_{mocB}:$nifA_{\text{Rlv3841}}$ GAF in pOPS0889 backbone)	This work

2.2.2 Plasmids

Name	Relevant characteristics	Source
pHP45ΩSpc	pBR322 derivative vector carrying Ω interposon spectinomycin resistance cassette, pHP45 replicon; Amp^R, Spc^R	Fellay et al. 1987 [346]
pHP45ΩTet	pBR322 derivative vector carrying Ω interposon tetracycline resistance cassette, pHP45 replicon; Amp^R, Tet^R	Fellay et al. 1987 [346]
pIJ11268	Broad host range vector based on pJP2 containing promoterless *luxCDABE* operon, used as negative control for Lux assay; Tet^R	Frederix et al. 2014 [345]
pIJ11282	pIJ11268 with neomycin promoter cloned in front of *luxCDABE*, used as positive control for Lux assay; Tet^R	Frederix et al. 2014 [345]
pJET1.2/blunt	*E. coli* vector for cloning PCR products; Amp^R	ThermoFisher Scientific
pJP2	Broad-host-range *gusA* transcriptional promoter probe vector; Tet^R Amp^R	Prell et al. 2012 [347]
pJQ200SK	Suicide vector, pACYC derivative, p15A origin of replication, *lacZ sacB traJ*; $Gent^R$	Quandt and Hynes 1993 [348]
pK19mob	Mobilizable *E. coli* vector for integration mutagenesis (*oriV*), pMB1 replication, RP4 mob; Kan^R	Schafer 1994 [349]
pK19mobSacB	Mobilizable *E. coli* vector for integration mutagenesis (*oriV*, *sacB*), pMB1 replication, RP4 mob; Kan^R	Kirchner and Tauch 2003 [350]
pLMB441	Internal fragment of $hfixL_c$ amplified from Rlv3841 with primers pr0988/0989, cloned into pK19mob digested with XbaI.	Graham Hood [2]
pLMB581	$hfixL_9$ amplified from Rlv3841 with pr1270/1271 cloned into pJET1.2/blunt	Graham Hood [2]
pLMB585	$hfixL_9$ digested out of pLMB581 with XbaI/XhoI cloned into pJQ200SK, digested with XbaI/XhoI.	Graham Hood [2]
pLMB590	ΩSpc from SmaI digested pHP45ΩSpc cloned into pLMB585 digested with StuI (blunted); $Gent^R$ Spc^R	Graham Hood [2]
pLMB732	Rlv3841 *fnrN* amplified with primers pr1381/1382 cloned into pJQ200SK at XbaI/XhoI site.	Graham Hood [2]
pLMB733	ΩTet from EcoRI digested pHP45ΩTet cloned into pLMB732 digested with MfeI; $Gent^R$ Tet^R	Graham Hood [2]

pOGG276	Mobilizable vector for hosting sequences to be genomically inserted via mini-Tn7. R6Kγ; Gent^{R} (genomic insert) Amp^{R} (backbone)	Beatriz Jorrín (unpublished)
pOGG280	Mobilizable vector for hosting sequences to be genomically inserted via mini-Tn7. R6Kγ; Kan^{R} (genomic insert) Amp^{R} (backbone)	Beatriz Jorrín (unpublished)
pOPS0136	$P_{fixK_{9a}}$ amplified from Rlv3841 with oxp0287/0288 cloned into pIJ11268 digested with BamHI/KpnI	Lucie McMurtry [2]
pOPS0785	Reporter plasmid backbone with an MCS for fusing promoters to *mruby3*. Constitutive *syfp2* expression, used as positive control for *in vitro* fluorescence assays. Contains *parABCDE* stability system. Gent^{R}.	Joshua Roworth [2]
pOPS0786	Reporter plasmid backbone with an MCS for fusing promoters to *syfp2*. Constitutive *mruby3* expression. Contains *parABCDE* stability system. Used as negative control for *in vitro* fluorescence assays. Gent^{R}.	Joshua Roworth [2]
pOPS0977	$P_{fixNOQP_{10}}$ amplified from Rlv3841 with primers oxp3039/3040 cloned into pOPS0786 digested with KpnI.	This work
pOPS0978	$P_{fixNOQP_{9}}$ amplified from Rlv3841 with primers oxp3041/3042 cloned into pOPS0786 digested with KpnI.	This work
pOPS0979	P_{nifH} amplified from Rlv3841 with primers oxp3043/3044 cloned into pOPS0786 digested with KpnI.	This work
pOPS0980	P_{fnrN} amplified from Rlv3841 with primers oxp3045/3046 cloned into pOPS0786 digested with KpnI.	This work
pOPS1199	Fragments upstream (oxp2874/2875) and downstream (oxp2876/oxp2877) of $fxkR_9$ amplified from Rlv3841 and cloned into pK19mobSacB digested with PstI and EcoRI, used to make the markerless mutant.	This work
pOPS1510	Rlv3841 *fnrN* gene amplified with oxp4115/4116 cloned into pOGG280 digested with BsaI.	This work
pOPS1593	$P_{fixNOQP_{9}}$ fused to *syfp2* amplified from pOPS0978 with primers oxp4354/4355 cloned into pOGG276 digested with XbaI.	This work

pOPS1594	P_{fnrN} fused to *syfp2* amplified from pOPS0978 with primers oxp4354/4355 cloned into pOGG276 digested with XbaI.	This work
pOPS1607	Reporter plasmid assembled into the pOGG276 backbone by Golden Gate assembly: J23106 promoter, RBStd, *syfp2*, DT16 terminator.	This work
pOPS1644	pJP2 digested with HindIII and XbaI to remove *gfp* reporter, replaced with Rlv3841 $P_{fixNOQP_{10}}$ fused to *syfp2*, amplified from pOPS0977 with oxp4550/4551; Tet^R Amp^R	This work
pTNS3	*E. coli* vector expressing *tnsABCD* from PI and lac promoters. Enables Tn7 insertions at the *glmS* site; Amp^R	Choi et al. 2008 [351]
pRK2013	Conjugation helper plasmid; mob^+, Kan^R	Ditta et al. 1980 [352]
pOPS1630	Complementation plasmid for Rlv3841 $hfixL_9$. Open reading frame of $hfixL_9$ amplified with primers oxp4552 (including its native RBS) and oxp4553. Cloned into pOGG250 digested with HindIII and XbaI.	This work
pOPS1645	Fragments upstream (oxp2870/2871) and downstream (oxp2872/2873) of $fixK_{9a}$ amplified from Rlv3841 and cloned into pJQ200SK digested with PstI and EcoRI, used to make the double *fixK* mutant.	This work
pOPS1009	Rlv3841 *nifA* amplified with oxp2410/2412, cloned into pLMB51 digested with BamHI and XbaI.	This work
pOPS1104	Rlv3841 *nifA* amplified without its GAF domain with oxp2411/2412, cloned into pLMB51 digested with BamHI and XbaI.	This work
pOPS0982	Rlv3841 *nifA* amplified without its GAF domain with oxp3030/3029, cloned into pLMB509 digested with NdeI.	This work
pOPS0983	*A. caulinodans nifA* amplified with oxp3026/3027, cloned into pLMB509 digested with NdeI.	This work
pOPS1591	Rlv3841 *nifA* amplified with oxp3028/3029, cloned into pLMB509 digested with NdeI.	This work
pLMB504	*A. vinelandii nifV* amplified with pr1059/1060, cloned into pJP2neo backbone digested with HindIII and XbaI.	Jason Terpolilli (unpublished)

pOPS1719	*A. caulinodans nifV* amplified with oxp4476/4479, assembled by Golden Gate cloning under control of P_{lac} into a pOGG281 backbone.	This work
pOPS1720	*A. caulinodans nifV* amplified with oxp4476/4479, assembled by Golden Gate cloning under control of $P_{A1lacO1}$ into a pOGG281 backbone.	This work
pOPS1177	Rlv3841 P_{fixA} amplified with oxp3103/3104 controlling *sfgfp*, cloned into pME6041 digested with PstI.	This work
pOPS1178	Rlv3841 P_{nifH} amplified with oxp2998/2999 controlling *sfgfp*, cloned into pME6041 digested with PstI.	This work
pOPS0999	*A. caulinodans* P_{fixA} amplified with oxp3031/3032 controlling *sfgfp* (Poole group part pOGG37) and terminated with DT16 (Poole group part pOGG157), cloned by Golden Gate into the pOGG026 backbone.	This work
pOPS0981	*A. caulinodans* P_{fixA} amplified with oxp3033/3034 controlling *sfgfp* (Poole group part pOGG37) and terminated with DT16 (Poole group part pOGG157), cloned by Golden Gate into the pOGG026 backbone.	This work
pLMB719	P_{tauA}:*mcherry*, cloned into pLMB509 backbone.	Robert Green (unpublished)
pLMB134	P_{tauA}:*gfp-mut3*, cloned into pLMB51 backbone.	Jürgen Prell (unpublished)
pLMB426	P_{tac}:*mcherry* amplified from pLMB420 cloned into pRU109 digested with ClaI and HindIII.	Ramakrishnan Karunakaran (unpublished)
pOPS1167	Promoter J23106:*sfgfp*, terminated by the pharmacia terminator, cloned into the pOGG024 backbone.	Ramakrishnan Karunakaran (unpublished)
pOPS0468	Promoter J23106:*sfgfp*, terminated by the DT16 terminator, cloned into the pOGG024 backbone.	Beatriz Jorrín (unpublished)
pOPS1106	Fragments upstream (oxp2431/2432) and downstream (oxp2433/2434) of *nifA* amplified from Rlv3841 and cloned into pJQ200SK digested with PstI and EcoRI, used to make the *nifA* mutant.	This work

pOPS1724	P_{mocB}:$nifA_{Rlv3841}$ amplified with oxp4828/4829, cloned by Golden Gate into the pOPS0889 backbone.	This work
pOPS1725	P_{mocB}:$nifA_{Rlv3841}$ amplified without its GAF domain with oxp4827/4829, cloned by Golden Gate into the pOPS0889 backbone.	This work
pOPS1122	P_{mocB}:$nifA_{ORS571}$-$rpoN_{ORS571}$, cloned by Golden Gate into the pOPS0889 backbone.	Patrick Green (unpublished)
pLMB185	$fixK_{9b}$ amplified from Rlv3841 with pr0516/0517 cloned into pRU877	Graham Hood (unpublished)
pRU877	*gusA* in the pK19mob backbone. Kan^R	Lodwig et al. [353]
pOGG026	Broad host range Golden Gate level 1 cloning backbone containing an RK2 origin of replication and the *parABCDE* stability system. Kan^R	Geddes et al. 2018 [354]
pME6041	Cloning vector with a p15A origin of replication and a pVS1 origin of replication and stability system. Kan^R	Heeb et al. [355]
pLMB51	Low copy taurine inducible *gusA* expression vector with *parABCDE* stability system. Amp^R Tet^R	Tett et al. [356]
pOGG281	Level 1 Golden Gate cloning backbone with R6K origin of replication, ready for Tn7 genomic integration. SpecR	Beatriz Jorrín (unpublished)
pLMB509	High copy taurine inducible *gfp-mut3.1* expression vector. $Gent^R$	Tett et al. [356]
pOPS0889	Rhizopine inducible P_{mocB} expression vector with a pBBR1 origin of replication, carrying the *intBC* transporters and *mocBR*. Kan^R	Barney Geddes (unpublished)

2.2.3 Primers

Name	Description	Sequence (5'-3')
oxp0283	Forward mapping primer for pOPS0786-based reporter plasmids	AGCGTTCTGAACAAATCC
oxp1331	Reverse mapping primer for pOPS0786-based reporter plasmids	TTTTGAAGACAAAAGCTTATTATTTATACAGCTCATCCATACCCAG
oxp0287	Forward primer for amplification of Rlv3841 $P_{fixK_{9a}}$ for cloning into pIJ11268	TTTTGGTACCGATGTCGTCCCCAGTG
oxp0288	Reverse primer for amplification of Rlv3841 $P_{fixK_{9a}}$ for cloning into pIJ11268	AAAAGGATCCTGGAACGCCTCTGC
oxp2327	Forward mapping primer for Tn7 integrations into Rlv3841	GATGATCTTCTCGCTGCCGA
oxp2328	Reverse mapping primer for Tn7 integrations into Rlv3841	GCTCTGGCCAATGAGGTTCT
oxp2874	Forward for amplicon upstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	GTCGACTCTAGAGGATCCCCTTCGGGATCATTGGCGCTG
oxp2875	Reverse for amplicon upstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	CGGTGAAGACGTAGCAGTACTCGTCCTCGAAATAGCGCGTCAG
oxp2876	Forward for amplicon downstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	GCTATTTCGAGGACGAGTACTGCTACGTCTTCACCGCCAG
oxp2877	Reverse for amplicon downstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	TGAATTCGAGCTCGGTACCCTCTTCGGACAGCACATTGAG
oxp3039	Reverse primer for amplification of $P_{fixNOQP_{10}}$ from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGATGTCGTCCCCAGTACGCC
oxp3040	Forward primer for amplification of $P_{fixNOQP_{10}}$ from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTTACGGCGGCCGCGACAGC
oxp3041	Forward primer for amplification of $P_{fixNOQP_9}$ from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGATGTCGTCCCCAGTGCG
oxp3042	Reverse primer for amplification of $P_{fixNOQP_9}$ from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTGGAACGCCTCTGCGTCAC
oxp3043	Forward primer for amplification of P_{nifH} from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGTTTGGCGTTCCTTCATGTGTTC

oxp3044	Reverse primer for amplification of P_{nifH} from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTC-GATGCTGACCGCCTGATC
oxp3045	Reverse primer for amplification of P_{fnrN} from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTG-GTCCTGATCC-CTTTTGAAATCCT
oxp3046	Forward primer for amplification of P_{fnrN} from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAG-GCGCTGTACCTCATGAAAT
oxp3115	Forward mapping primer for $fxkR_9$ mutagenesis	GGTCGTTGTCTCCAGGCGCG
oxp3156	Reverse mapping primer for $fxkR_9$ mutagenesis	TGCGCAGTGGTTGGC-TAGGC
oxp4115	Forward primer for amplification of Rlv3841 *fnrN* gene for cloning into pOGG280	TAATGCCGAATTCGGATC-CCGCGCTGTACCTCAT-GAAATG
oxp4116	Reverse primer for amplification of Rlv3841 *fnrN* gene for cloning into pOGG280	CTATCAACAGGAGTC-CAAGTATGCGCTGATCATC-CGCTC
oxp4354	Forward primer for amplification of Rlv3841 promoters fused to *syfp2* in pOPS0786	AATTCGGATCCGGAGTCG-GTCACATGTGCATC
oxp4355	Reverse primer for amplification of Rlv3841 promoters fused to *syfp2* in pOPS0786	AGGAGTCCAA-GAGCGGGTC-GAAAAAAAAAGCCCG
oxp4550	Forward primer for amplification of $P_{fixNOQP_{10}}$ fused to *syfp2* from pOPS0977	GTCCGGGTACCATGGATC-CATTACGGCGGCCGC-GACAG
oxp4551	Reverse primer for amplification of PfixNOQP10 fused to *syfp2* from pOPS0977	CGGACCATGAT-TACCTCAGTGGTC-GAAAAAAAAAGCCCG-CACTGTC
pK19A	Mapping primer for pK19mob mutagenesis	ATCAGATCTTGATCCC-CTGC
pr0482	Forward mapping primer for $hfixL_c$ mutagenesis, binds in genome	AGTTCGATGTTCGTATCC-GAAC
pr0988	Forward primer for amplification of Rlv3841 $hfixL_c$ internal fragment for cloning into pK19mob	GCAGGTCGACTCTAGATG-GAAGAGCTTCGGACCGAA
pr0989	Reverse primer for amplification of Rlv3841 $hfixL_c$ internal fragment for cloning into pK19mob	CCGGGGATCCTCTA-GAATATCTCGATCGTCA-GACGG

pr1270	Forward primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pJQ200SK	CTCGAGGCTACATCGAC-CACTATCTC
pr1271	Reverse primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pJQ200SK	TCTAGAACACGGGCGT-CATCTTCGAC
pr1272	Forward mapping primer for *hfixL*$_9$ mutagenesis	CGGAAGAGCTTCCAC-GATGA
pr1273	Reverse mapping primer for *hfixL*$_9$ mutagenesis	GCCGTCCGCACCT-GTCGTTC
pr1381	Forward primer for amplification of Rlv3841 *fnrN* gene for cloning into pJQ200SK	GCCTAAAGCGCGTCTG-GTTC
pr1382	Reverse primer for amplification of Rlv3841 *fnrN* gene for cloning into pJQ200SK	AATAAGCCTGCGGCG-CATCC
pr1432	Forward mapping primer for *fnrN* mutagenesis	CTGGGCCATGGTCTC-GATCA
pr1433	Reverse mapping primer for *fnrN* mutagenesis	CATAATCTCGGCAC-CATGGC
oxp4552	Forward primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pOGG250	CTATGACCATGATTACGC-CAAAGCTTTGCAACGAAGC-CGCAGGA
oxp4553	Reverse primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pOGG250	GCTCGGTACC-CGGGGATCCTTCTA-GATCACTTCGCGTCGA-GAAGTTC
oxp3200	Forward mapping primer for *fixK*$_{9a}$ mutagenesis	TACATCATGC-CCAGGAAGCC
oxp3201	Reverse mapping primer for *fixK*$_{9a}$ mutagenesis	ACCTGATATCTTCGCC-GAGC
oxp2870	Forward primer for amplicon upstream of *fixK*$_{9a}$, for cloning into pJQ200SK and markerless mutant generation	TTTTTCTGCAGTTGGAGC-CAAGAAACGAGGCC
oxp2871	Reverse primer for amplicon upstream of *fixK*$_{9a}$, for cloning into pJQ200SK and markerless mutant generation	CGAGATTGACGATTC-GAATTCTGTTCCT-CATTGTCGGTCGGC
oxp2872	Forward primer for amplicon downstream of *fixK*$_{9a}$, for cloning into pJQ200SK and markerless mutant generation	CGACAATGAGGAACA-GAATTCGAATCGT-CAATCTCGCCTGCC

oxp2873	Reverse primer for amplicon downstream of *fixK*$_{9a}$, for cloning into pJQ200SK and markerless mutant generation	TTTTTCTGCAGTGACGGTCAGGAGGTTGTTGAAG
oxp2410	Forward primer for amplification of *nifA*$_{Rlv3841}$ for cloning into pLMB51	TTGGATCCATGATTAAACCAGAGGCGCGGC
oxp2412	Reverse primer for amplification of *nifA*$_{Rlv3841}$ with or without the GAF domain for cloning into pLMB51	AAACTAGTTCACTCCTTCTTCACATCGATACGAAACC
oxp2411	Forward primer for amplification of *nifA*$_{Rlv3841}$ without the GAF domain for cloning into pLMB51	TTGGATCCGAATTGTGCAGCGATGGCTGCT
oxp3030	Forward primer for amplification of *nifA*$_{Rlv3841}$ without the GAF domain for cloning into pLMB509	AGGAGGAAGAACATATGATGGAATTGTGCAGCGATGGCTGCT
oxp3029	Reverse primer for amplification of *nifA*$_{Rlv3841}$ with or without the GAF domain for cloning into pLMB509	TGGTGATGATGCATATGTCACTCCTTCTTCACATCGATACGAAACC
oxp3026	Forward primer for amplification of *nifA*$_{ORS571}$ for cloning into pLMB509	AGGAGGAAGAACATATGTCAGAAGCGCTTGATGTCG
oxp3027	Reverse primer for amplification of *nifA*$_{ORS571}$ for cloning into pLMB509	TGGTGATGATGCATATGATGCCAATGACCGACGCC
oxp3028	Forward primer for amplification of Rlv3841 *nifA*$_{Rlv3841}$ for cloning into pLMB509	AGGAGGAAGAACATATGATGATTAAACCAGAGGCGCGGC
oxp4476	Forward primer for amplification of *nifV*$_{ORS571}$, for cloning into pOGG281	CACTCTGTGGTCTCATACTACATCCCACGTGACCTCATTGTGTTCCG
oxp4479	Reverse primer for amplification of *nifV*$_{ORS571}$, for cloning into pOGG281	CACTTCGTGGTCTCAAAGCTCAGATTGCGCCTTCCACGGCG
oxp3104	Reverse primer amplifying Rlv3841 P$_{fixA}$ for cloning into pME6041	CGAATTCGAAGCTTCGAATTCCCTGGCAGTTTATGG
oxp3103	Forward primer amplifying Rlv3841 P$_{fixA}$ for cloning into pME6041	TGACGTCGACGCGTCGGAGGTGATCGGGGCGGATTGC
oxp2998	Forward primer amplifying Rlv3841 P$_{nifH}$ for cloning into pME6041	TGACGTCGACGCGTCTCGATGCTGACCGCCTGATC
oxp2999	Reverse primer amplifying Rlv3841 P$_{nifH}$ for cloning into pME6041	CGAATTCGAAGCTTCAGCGGAATTCCCTGGCAG

oxp3031	Forward primer amplifying *A. caulinodans* P_{fixA} for cloning into pOGG026	TTTTGAAGACAAG-GAGGGTATCGCTCAC-CGTTCTCAAGTGA
oxp3032	Reverse primer amplifying *A. caulinodans* P_{fixA} for cloning into pOGG026	TTTTGAAGACAA-CATTTGCGTTCGCTCC-CGGAA
oxp3033	Forward primer amplifying *A. caulinodans* P_{nifH} for cloning into pOGG026	CACTCTGTG-GTCTCAGGAGTGATC-GATGCTGACCGCCTGATC
oxp3034	Reverse primer amplifying *A. caulinodans* P_{nifH} for cloning into pOGG026	CACTCTGTGGTCTCA-CATTGTCTTGAATTCCTTC-GAACCGTTGCC
509F	Forward mapping primer for confirmation of pLMB51 and pLMB509 assemblies	CGCCCAACTGGACTCATC-TAACTTC
509R	Reverse mapping primer for confirmation of pLMB509 assemblies	CGCAGTCGGCCTATTGGT-TAAA
oxp0092	Reverse mapping primer for confirmation of pLMB51 assemblies	CCACAGTTTTCGC-GATCCAG
oxp2636	Reverse mapping primer for confirmation of pME6041 assemblies, binds downstream of MCS	GTAACATCAGAGATTTTGA-GACAC
oxp1951	Forward mapping primer for confirmation of pOGG026 assemblies	AGAGCGTTCACCGA-CAAACA
oxp1952	Reverse mapping primer for confirmation of pOGG026 assemblies	TCAACAGGAGTCCAA-GAGCG
oxp3062	Forward mapping primer for confirmation of pOGG276, pOGG280 and pOGG281 assemblies	GAGCGCTTTTGAAGC-TAATTCGA
oxp3063	Reverse mapping primer for confirmation of pOGG276, pOGG280 and pOGG281 assemblies	TCACTTATCTGGTTGGC-CTGC
oxp2667	Forward mapping primer for $nifA_{\mathrm{Rlv3841}}$ mutagenesis	CGAAGGCTTG-GAAATCAGGGGC
oxp2668	Reverse mapping primer for $nifA_{\mathrm{Rlv3841}}$ mutagenesis	ACTCTTTGGC-GACGGGTGGG
oxp2431	Forward primer for amplicon upstream of $nifA_{\mathrm{Rlv3841}}$, for cloning into pJQ200SK and markerless mutant generation	TTTTTCTGCAGGATTTTG-GTGAGGATGCCGC
oxp2432	Reverse primer for amplicon upstream of $nifA_{\mathrm{Rlv3841}}$, for cloning into pJQ200SK and markerless mutant generation	AAGGCTGCGATGAAC-GAATTCCTACGCCG-GTTTCGTATCGA

oxp2433	Forward primer for amplicon downstream of $nifA_{\mathrm{Rlv3841}}$, for cloning into pJQ200SK and markerless mutant generation	ACGAAACCGGCGTAGGAATTCGTTCATCGCAGCCTTCAGGA
oxp2434	Reverse primer for amplicon downstream of $nifA_{\mathrm{Rlv3841}}$, for cloning into pJQ200SK and markerless mutant generation	TTTTTCTGCAGGAGGGATCAAACCTCGCGAT
oxp3155	Forward mapping primer for $fxkR_9$ mutagenesis	GGTCGTTGTCTCCAGGCGCG
oxp3156	Reverse mapping primer for $fxkR_9$ mutagenesis	TGCGCAGTGGTTGGCTAGGC
pr0516	Forward primer amplifying $fixK_{9\mathrm{b}}$ for cloning into pRU877 backbone	CTTCTCGAGCTCTAGAAGCAGCCAGTCGAACATCTG
pr0517	Reverse primer amplifying $fixK_{9\mathrm{b}}$ for cloning into pRU877 backbone	ATTACCTCAGTCTAGAATCATTCGCGACACAGTTTC
pr0518	Mapping primer for $fixK_{9\mathrm{b}}$ mutagenesis	TGCTTTAGGCGTTCTGGCTT
pr0095	Mapping primer for pRU877 mutagenesis	TGCATCGGCGAACTGATCGTTA

2.3 Molecular techniques

2.3.1 DNA isolation

Plasmid DNA was extracted from cultures using the Monarch Plasmid Purification Kit (New England Biolabs). Genomic DNA was extracted from Rlv3841 and *A. caulinodans* cultures with a DNeasy Blood and Tissue Kit (Qiagen). DNA fragments were purified from gel electrophoresis using the GENEJet Gel Extraction Kit (ThermoFisher Scientific). All extractions and purifications were performed according to instructions of their respective kit manufacturers.

2.3.2 Primer design, DNA synthesis and sequencing

PCR primers, including those for all cloning reactions, were designed in Benchling and Geneious R9. Oligonucleotides were synthesized by Eurofins MWG Operon, and larger DNA fragment were synthesized by the Invitrogen GeneArt service (ThermoFisher Scientific). Sanger sequencing was carried out by Eurofins MWG Operon. All plasmids assembled for this project were sequence-confirmed prior to stocking and experimental use.

2.3.3 DNA amplification by polymerase chain reaction (PCR)

PCRs for cloning were carried out with Q5 High-Fidelity DNA Polymerase (New England Biolabs) in a total volume of 50 µL, according to the manufacturer's instructions. Colony PCRs for screening were performed using OneTaq DNA Polymerase (New England Biolabs). A small amount of each colony to be screened was added to a total colony PCR reaction volume of 30 µL. Both cloning and colony PCR used thermocycler conditions set according to the manufacturer's instructions for their respective polymerases, with T_m determined by Benchling and the New England Biolabs Tm Calculator.

2.3.4 DNA gel electrophoresis

Electrophoresis was carried out in 1% agarose (Sigma Aldrich) in TAE buffer (400 mM Tris-Acetate, 1 mM EDTA). DNA was stained with SYBR® Safe (Invitrogen)

at a concentration of 1:10000 in the TAE buffer. The ladder marker used was the 1 Kb Plus DNA Ladder from Invitrogen. Gels were run at 100 V for 60 minutes. For fragments smaller than 200 bp, electrophoresis was carried out in 1.7% agarose with the 100 bp Plus DNA ladder from Invitrogen. Gels were visualized using the GelDoc EZ System (BioRad).

2.3.5 DNA restriction digestion and ligation

Restriction enzymes (New England Biolabs and ThermoFisher Scientific) were used according to the manufacturer's instructions. Reactions were typically performed overnight. For heat-sensitive endonucleases, reactions were stopped by denaturation at 65 °C or 80 °C for 10 minutes. DNA ligations were run overnight with T4 DNA Ligase (Thermo Scientific) according to the manufacturer's instructions.

2.3.6 Golden Gate cloning

Based on past work by Engler et al. and Weber et al., a standardized and modular Golden Gate system has been established in the Rhizosphere group and was used in this work [354, 357–359]. A library of standardized vector backbones has been developed as part of this system, and several are used extensively in this thesis (see section 2.2.2 for details). Golden Gate cloning reactions were carried out with materials supplied by New England Biolabs. Each reaction contained 1 µL of the desired Type IIS enzyme (BpiI or BsaI), 1 µL of T4 DNA ligase (10,000 U), 1.5 µL of BSA (2 mg mL^{-1}), 1.5 µL of T4 DNA ligase buffer (ThermoFisher Scientific) and 40 pmol of the required DNA fragments and vector backbone. Reactions were made up to 15 µL with distilled H_2O and thermocycler parameters were set per the recommendations of New England Biolabs.

2.3.7 Homology-based cloning

Homology-based cloning reactions were carried out using a BD In-Fusion® Kit (Takara Bio). Reactions and their thermocycling conditions followed the manu-

facturer's protocols. Primers were designed with 15 bp overhangs corresponding to the destination vector or adjacent fragments.

2.3.8 Transformations

Chemically competent *E. coli* were used for transformations. In most instances, DH5α cells were used. For assemblies with particularly low transformation efficiency, TOP10 One-Shot® (ThermoFisher) cells were used. For R6K vectors, cells with a *pir*$^+$ (producing a low copy number) or *pir*-116 (producing a high copy number) genotype were used. Most transformations were performed into *E. coli pir* strains propagated in-house, with commercial cells from Lucigen (*pir*$^+$) or ThermoFisher Scientific (*pir*-116). To facilitate the creation of some mutants, ALA auxotrophic *E. coli* ST18 (propagated in-house) cells were used. Cells were made chemically competent by growing them in LB for 3 hours until they reached an OD_{600} of 0.3-0.4, then incubated on ice for 10 minutes. Cultures were washed by centrifuging them twice at 4,000 RPM at 4 °C for 10 minutes. Between centrifugations, cells were resuspended in 10 mL ice cold 0.1 M $CaCl_2$ in a 15% v/v glycerol solution and incubated on ice for 30 minutes. After washing, aliquots of 50-200 µL were snap frozen in liquid nitrogen and stored at -80 °C until use.

Transformations were carried out by thawing aliquots on ice, adding 1-5 µL of purified plasmid or assembly mixture and incubating for 30 minutes. Cells were then heat-shocked at 42 °C for 45 seconds, incubated on ice for 5 minutes and 150 µL of Super Optimal broth with Catabolite repression (SOC) media added. Following addition of SOC, cells were shaken at 200 RPM, 37 °C for 1 hour. The transformation mixture was then spread onto LB-agar plates containing the appropriate antibiotics and incubated at 37 °C overnight.

2.3.9 Conjugations and Tn7 genomic integration

Triparental conjugations were performed based on a protocol developed by Buchanan-Wollaston (1979) [360]. Strains into which DNA was to be conjugated (hosts) were grown on TY slopes for 3 days prior to conjugation. In parallel, *E. coli* cultures

carrying the plasmid(s) of interest (donors) and a strain carrying the pRK2013 helper plasmid were cultured in liquid overnight with the appropriate antibiotics before conjugation. On the day of conjugation, 50 µL of each *E. coli* overnight culture was used to inoculate a new 10 mL LB liquid culture containing the appropriate antibiotics and grown for 4 hours. Simultaneously, *Rhizobium* and/or *Azorhizobium* strains receiving DNA were resuspended from slopes and washed three times to remove antibiotic traces. These cells were then resuspended in 1-3 mL of fresh TY media without antibiotics. Conjugation mixtures were set up with 400 µL of each *E. coli* donor strain, 400 µL of the host strain, and 200 µL of the *E. coli* culture containing pRK2013. This mixture was centrifuged for 5 minutes at 3,500 RCF. The supernatant was discarded and the pellet resuspended in 30 µL of TY, then spotted onto sterile nitrocellulose filters placed on a TY-agar plate without antibiotics. For conjugations into *A. caulinodans*, nitrocellulose filters were not used, and resuspended pellets were instead plated onto TY directly (also referred to as "patch mating"). All conjugation plates were incubated overnight at 28 °C or 37 °C as appropriate for the host strain. The next day, bacteria growing on the filter were streaked onto plates carrying the appropriate antibiotics. Genomic integrations were done using the mini-Tn7 system developed by Choi and colleagues, with vectors created by Beatriz Jorrín in our group [351, 361]. See section 2.2.2 for vector details.

2.4 Mutant generation

Several strains used in this thesis were produced by Graham Hood (see individual mutants for details). These were all re-verified before being used for experimental work.

2.4.1 Rlv3841 *hfixL*$_c$ (RL1879) mutant, LMB403

The LMB403 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. To generate the mutant, a 1 Kb internal fragment of *hfixL*$_c$ was PCR amplified from Rlv3841 with primers pr0988/0989, adding XbaI sites at the 5' and 3' ends. This fragment was cloned into pK19mob digested with XbaI, using BD In-Fusion cloning, to produce plasmid pLMB441. Triparental filter conjugation of pLMB441 into Rlv3841 WT was then performed using kanamycin selection. Colonies were screened by colony PCR using primers pr0482 and pK19A, which bind upstream of *hfixL*$_c$ and inside the integrated pK19 backbone respectively. This produced mutant strain LMB403.

2.4.2 Rlv3841 *hfixL*$_9$ (pRL90020) mutant, LMB495

The LMB495 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. A 1 Kb region containing *hfixL*$_9$ was PCR amplified from Rlv3841 with primers pr1270/1271. This region was subcloned into pJET1.2/ blunt to produce plasmid pLMB581. Plasmid pLMB581 was then digested with XbaI/XhoI and the *hfixL*$_9$ region cloned into pJQ200SK using BD In-Fusion, digested with the same enzymes, to give plasmid pLMB585. A spectinomycin resistance cassette was digested out of the pHP45ΩSpc plasmid with SmaI and cloned into pLMB585 at a unique StuI site blunted using the Klenow fragment to give plasmid pLMB590. Triparental filter conjugation of pLMB590 into Rlv3841 WT was then performed using spectinomycin selection. Colonies were screened by colony PCR using primers pr1272/1273. This gave mutant strain LMB495.

2.4.3 Rlv3841 double *hfixL*$_{c}$ *hfixL*$_{9}$ mutant, LMB496

The LMB496 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. An Rlv3841 mutant in both hfixL genes was generated by triparental filter conjugation of pLMB441 (see section 2.3.9) into strain LMB495, producing double mutant strain LMB496.

2.4.4 Rlv3841 *fnrN* (RL2818) mutant, LMB648

The LMB648 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. A 2.5 Kb region containing *fnrN* was PCR amplified from Rlv3841 with primers pr1381/1382. This fragment was digested with XbaI/XhoI and cloned using BD In-Fusion into pJQ200SK linearized with digestion by the same enzymes to make plasmid pLMB732. A tetracycline resistance cassette was then digested out of the pHP45ΩTet plasmid with EcoRI and cloned into pLMB732 at a unique MfeI site to give plasmid pLMB733. Triparental filter conjugation of pLMB733 into Rlv3841 WT was then performed using tetracycline selection. Colonies were screened by colony PCR using primers pr1432/1433. This gave mutant strain LMB648.

2.4.5 Rlv3841 triple *hfixL*$_{c}$ *hfixL*$_{9}$ *fnrN* mutant, LMB673

The LMB673 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. This triple mutant, in both hfixL genes and the *fnrN* gene, was generated by transducing *fnrN*::ΩTet from LMB648 into LMB496 to produce strain LMB673.

2.4.6 Rlv3841 *fixK*$_{9b}$ (pRL90025) mutant, LMB374

The LMB374 mutant was produced by Graham Hood as part of his thesis work before this project was begun [362]. To generate the mutant, a 1 Kb internal fragment of *fixK*$_{9b}$ was PCR amplified from Rlv3841 with primers pr0516/0517, adding XbaI sites at the 5' and 3' ends. This fragment was cloned into pRU877 digested with XbaI, using BD In-Fusion cloning, to produce plasmid pLMB185.

Triparental filter conjugation of pLMB185 into Rlv3841 WT was then performed using kanamycin selection. Colonies were screened by colony PCR using primers pr0518 and pr0095, which bind upstream of *fixK*$_{9b}$ and inside the integrated pRU877 backbone respectively. This produced mutant strain LMB374.

2.4.7 Rlv3841 double *fixK*$_{9a}$ *fixK*$_{9b}$ (pRL90019 pRL90025) mutant, OPS2500

Two 1 Kb regions, one upstream and one downstream of *fixK*$_{9a}$, were PCR amplified from Rlv3841 with primer pairs oxp2870/2871 and oxp2872/2873 respectively. These were cloned with BD In-Fusion into pJQ200SK digested with PstI and EcoRI to produce plasmid pOPS1645. Triparental filter conjugation of pOPS1645 into LMB374 was then performed using kanamycin and gentamicin selection followed by kanamycin and sucrose selection as previously described [348]. Surviving colonies were screened for loss of gentamicin resistance, kanamycin resistance and using colony PCR with primers oxp3200/3201 to isolate mutant strain OPS2500.

2.4.8 Rlv3841 *fxkR*$_9$ (pRL90026) mutant, OPS1808

Two 1 Kb regions, one upstream and one downstream of *fxkR*$_9$, were PCR amplified from Rlv3841 with primer pairs oxp2874/2875 and oxp2876/2877 respectively. These were cloned with BD In-Fusion into pK19mobSacB digested with PstI and EcoRI to produce plasmid pOPS1199. Triparental filter conjugation of pOPS1199 into Rlv3841 WT was then performed using kanamycin selection. Colonies were screened by colony PCR using primers oxp3155 and pK19A. Colonies with correct integration were subsequently subjected to sucrose selection to remove plasmid pK19mobSacB. Colonies were then screened for loss of kanamycin resistance and using colony PCR with primers oxp3155/3156 to isolate mutant strain OPS1808.

2.4.9 Rlv3841 *nifA* (pRL100196) mutant, OPS1737

Two 1 Kb regions, one upstream and one downstream of *nifA*, were PCR amplified from Rlv3841 with primer pairs oxp2431/2432 and oxp2433/2434 respectively.

These were cloned with BD In-Fusion into pJQ200SK digested with PstI and EcoRI to produce plasmid pOPS1106. Triparental filter conjugation of pOPS1106 into Rlv3841 WT was then performed using gentamicin selection followed by sucrose selection. Surviving colonies were screened for loss of gentamicin resistance and using colony PCR with primers oxp2667/2668 to isolate mutant strain OPS1737.

2.5 Plant experiments

2.5.1 Growth of *P. sativum*

Strains to be inoculated on plants were grown on TY slopes with the appropriate antibiotics for 3 days. Slopes were then washed with 5 mL of sterile distilled water and the resuspended culture diluted 1:100 in 10 mL sterile distilled water. Each seed was inoculated with 1 mL of this diluted culture ($\backsim 1 \times 10^7$ cells).

Before sowing, *Pisum sativum* cv. Avola seeds were surface sterilized by washing with 95% ethanol for 60 seconds, 2% sodium hypochlorite for 5 minutes and five washes with sterile distilled water. Plants were grown in 1 L beakers filled with medium-grade vermiculite and 400 mL of nitrogen-free nutrient solution (4 mM Na_2HPO_4, 3.7 mM K_2PO_4, 1 mM $CaCl_2$, 800 µM $MgSO_4$, 100 µM KCl, 35 µM H_3BO_3, 10 µM Fe-EDTA, 9 µM $MnCl_2$, 0.8 µM $ZnCl_2$, 0.5 µM Na_2MoO_4 and 0.3 µM $CuSO_4$ [363]. All pots were autoclaved before sowing.

Two seeds were planted for each pot. Plants were grown in a growth room (16h light/8h dark) and all pots were covered with cling film to reduce contamination. The cling film was cut after 2-3 days to allow seedlings to grow. After 3-5 days, the slower growing seed was taken out of each pot. A week before plants were harvested, each was watered with 100 mL of distilled water. Plants were harvested after 21 days.

2.5.2 Acetylene reduction assays

Activity from the nitrogenase enzyme, both in symbiotic and free-living diazotrophs, was measured by acetylene reduction assays (ARA). ARAs exploit the promiscuous conversion of acetylene to ethylene by nitrogenase, which involves breaking a triple bond similar to that found in N_2 [364]. To perform the assay on *P. sativum* plants in symbiosis with Rlv3841 strains, plants were placed in 250 mL Duran bottles immediately after harvesting with some wet paper to prevent drying out. Bottles were sealed with airtight neoprene lids throughout the assay. First, 8 mL of atmosphere from inside the bottle was removed with a syringe. This was replaced with 6.4 mL of acetylene to create a 2% acetylene concentration inside the bottle.

Bottles were then incubated for 1 hour, after which a 1 mL sample of the internal atmosphere was removed with a syringe. Samples were run on a Clarus® 480 gas chromatographer (Perkin-Elmer) to determine relative acetylene and ethylene concentrations. The ratio of these concentrations was used to calculate conversion activity, normalised by the total mass of nodules per plant to calculate nitrogenase activity. Nodules were weighed after picking them off plants subsequent to the assay.

2.5.3 Isolation of bacteroids

Bacteroids were isolated from 21-day old root nodules following a differential spin protocol adapted from Tsukada et al. 2009 [365]. Approximately 100 mg of nodules was picked per plant. Nodules were immersed in 1 mL of sterile isolation buffer (1 M K_2HPO_4, 1 M KH_2PO_4, 300 mM sucrose, 2 mM $MgCl_2$) and macerated. The mixture was spun down at 200 RCF for 5 minutes to remove plant debris. The supernatant was transferred to a fresh tube and spun down at 3,500 RCF for 5 minutes. The supernatant from this second spin was discarded and the pelleted fraction, containing the isolated bacteroids, was resuspended in 500 µL isolation buffer and used for microtiter plate measurements (see section 2.6).

2.6 Assays

2.6.1 Low-throughput fluorescence and growth assays on free-living rhizobial strains

Rlv3841 and *A. caulinodans* strains were first grown on TY slopes with appropriate antibiotics for three days. Cells were resuspended and washed three times in UMS without nitrogen or carbon sources by centrifugation at 5,000 RCF for 10 minutes. *A. caulinodans* cells were immediately used to inoculate plates and begin experimental measurements after washing. For Rlv3841 strains, washed cells were first used to inoculate 10 mL liquid UMS cultures to OD_{600} 0.01 and grown overnight without antibiotics, followed by use in experimental work. OD_{600} readings on cultures not in microtiter plates were taken with a GENESYS spectrophotometer (ThermoFisher Scientific). To begin a low-throughput experiment, cultures of both organisms were diluted to OD_{600} 0.1 in 400 µL UMS per well in a 24-well microtiter plate (4titude), with carbon and nitrogen sources as specified on a per-experiment basis. Induction with taurine was done by adding 0-10 µL of a 400 mM sterile taurine stock solution to each well, as required by the experiment. A gas-permeable membrane (4titude) was applied to microtiter plates. Plates were incubated in a FLUOstar Omega plate reader equipped with an Atmospheric Control Unit (both produced by BMG) to adjust O_2 concentration in the 1-21% range as required, and CO_2 concentration to 0.1%. Temperature was set to 28 °C or 37 °C as appropriate for the strains being studied. Fluorescence and OD_{595} readings were taken every 30 minutes and plates shaken at 700 RPM in double orbital mode between readings. Activity from reporters was determined from readings taken at 18 hrs post-inoculation, when all cultures had reached stationary phase.

Fluorescence measurements were filter based and taken using the bottom optic with a gain of 2,000 and orbital averaging (53 readings, 6 mm radius). The excitation and emission wavelengths used for each fluorophore are given in Table 2.6.

Fluorophore	Excitation (nm)	Emission (nm)
sYFP2	520	540
sfGFP/GFP-mut3	485	520
mCherry	560	590

Table 2.6: Fluorophore wavelengths. Excitation and emission wavelengths used for plate-based fluorophore measurements in this study.

2.6.2 High-throughput fluorescence and growth assays on free-living rhizobial strains

Incubating microtiter plates in a plate reader limited experiments to a single, 24-well plate, restricting throughput. To improve throughput, we incubated microtiter plates in an O_2 cabinet, enabling temperature and O_2 concentration control for multiple microtiter plates simultaneously. This was only used for Rlv3841 strains. Cultures were initially grown on slopes, washed and grown overnight as described in section 2.6.1. Identical plates and per-well volumes were used. Nitrogen gas was flushed through the cabinet to attain the required O_2 concentration. A heat bar was used to regulate internal temperature to 28 °C throughout incubation. After inoculation, plates were shaken on a platform shaker at 250 RPM inside the O_2 cabinet. A single endpoint reading was taken after 18 hours, by taking plates out of the O_2 cabinet inserting them one at a time into a FLUOstar Omega plate reader. Readings were taken with the same parameters used for low-throughput experiments.

2.6.3 Fluorescence and luminescence assays on isolated bacteroids

To perform assays, bacteroids were first isolated as described in section 2.5.3. Fluorescence readings were taken using a FLUOstar Omega plate reader (BMG). In a 96-well microtiter plate, 200 µL of each isolated bacteroid sample in isolation buffer was added per well. Fluorescence and OD_{595} readings were then taken as described in section 2.6.1 for free-living rhizobia.

Luminescence measurements were made on a GloMax plate reader (Promega) using the manufacturer's protocol. In experiments where luminescence readings were taken, OD_{600} was also measured on the GloMax instrument, simultaneously. Luminescence was used for the *fixK*$_{9a}$ reporter as this construct was already available from previous unpublished work in our group [366].

2.6.4 Photography

Root nodule photos were taken with an EOS 1100D camera (Canon) after all vermiculite had been cleaned away. Roots were placed on a black background and the camera's macro photography mode used with automatic focus. Flash was disabled. Images were adjusted to the same size after photos were taken using a scale next to the roots.

Colony morphology photos were taken with the assistance of John Baker from the Department of Plant Sciences. Single colonies were picked and streaked on square petri dishes in TY with antibiotics, then grown for 4 days at 28 °C or 37 °C as appropriate. Images were taken with a Z6 camera (Nikon). The camera was positioned at the same distance relative to each plate during photography, and all camera settings including white balance were manually controlled to be identical between photos. Focal length was 105 mm, aperture was f/32, exposure time was 1/3rd of a second and ISO speed was 250. Plates were lit with identical indirect lighting sources for all photos and no flash was used.

2.6.5 Confocal microscopy of nodules

Plants were inoculated with marked strains and grown as described in section 2.5.1. After 21 days, nodules were picked and immersed in water then cut in half longitudinally with a razor blade (Wilkinson Sword). Confocal images were taken with an LSM 880 confocal laser-scanning microscope equipped with the Axio Imager.Z2 (Zeiss), using the manufacturer's ZEN Black software. A Plan-Apochromat 10×/0.45 M27 objective (Zeiss) was used. Excitation was at 514 nm with an Argon laser and emission measurements filtered to a range of 519-572 nm.

Acquisitions were tile scans with 2×3 tiles per image. 31 Z-stack slices were taken for each tile, separated by a height of 10 µm. Finally, a maximum-intensity orthogonal projection was created for each imaged nodule with the ZEN Blue software (Zeiss).

2.6.6 Acetylene reduction assays on *A. caulinodans* cultures

A. caulinodans strains were first grown on TY slopes and washed as described in section 2.6.1. Washed cultures were used to inoculate 10 mL UMS cultures in universals to OD_{600} 0.4. Liquid cultures contained 20 mM succinate, no ammonium chloride, 10 µM scyllo-inosamine as appropriate and antibiotics. These liquid cultures were first incubated for 2 hours in a 3% O_2 atmosphere inside an O_2 cabinet at 28 °C, shaking at 200 RPM. Universals were not closed to allow O_2 to equilibrate. Universals were then sealed with airtight neoprene lids and transferred to a shaking incubator, where they were incubated for a further 2 hours at 37 °C at 200 RPM. To begin the acetylene reduction assay, 3.25 mL of air was removed from each sealed universal with a syringe and the same amount of acetylene was injected to create a 10% headspace acetylene concentration. Cultures were incubated in a shaker at 37 °C at 200 RPM for a further 2 hrs, after which a 1 mL sample of the headspace was extracted with a syringe and analysed for acetylene and ethylene concentration as described in section 2.5.2. When ARAs were completed, an OD_{600} reading was taken from each culture and used to normalise nitrogenase activity.

2.6.7 Gas chromatography – mass spectrometry (GC-MS) assays of homocitrate

Homocitrate standards were produced from pure homocitrate powder (Sigma-Aldrich). A 1 mg mL^{-1} solution in distilled water was first made, from which 0.1 mg mL^{-1} and 0.01 mg mL^{-1} stocks were made by dilution. These stocks were aliquoted into 1.5 mL Eppendorf tubes to create standard curve GC-MS samples.

A. caulinodans cultures to be assayed for homocitrate concentration were first grown and made to fix nitrogen as detailed in section 2.6.8, and an OD_{600} reading was taken for each culture. Following confirmation of nitrogenase activity, cultures

were spun down at 17,000 RCF for 5 minutes and immediately placed on ice. Rlv3841 strains were first grown for three days on TY slopes with appropriate antibiotics, then washed and used to inoculate 50 mL TY liquid cultures in flasks to OD_{600} 0.3 without antibiotics and with 10 mM IPTG where *nifV* expression was LacI controlled. These liquid cultures were grown overnight, then spun down at 4,000 RCF for 20 minutes at 4 °C and immediately placed on ice.

For both *A. caulinodans* and Rlv3841 strains, extraction was performed by resuspending cell pellets in 500 µL of a 60% methanol solution pre-cooled to -20 °C. Resuspended cells were vortexed briefly then stored at -20 °C for 20 minutes, and this cycle repeated four times. Following extraction, cell debris was pelleted by centrifugation at 17,000 RCF for 10 minutes. Supernatant aliquots of 100 µL were used to create GC-MS samples.

50 µL of 0.01 mg mL^{-1} ribitol was added to all samples to act as a standard for GC-MS sensitivity. Samples were subsequently snap frozen with liquid nitrogen and freeze-dried to remove water. Derivatization was carried out on freeze-dried samples by first adding 40 µL of 20 mg mL^{-1} methoxyamine in pyridine (both supplied by Sigma-Aldrich) and incubating samples for 2 hrs at 37 °C, shaking at 1,400 RPM. Next, 70 µL of pure N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA, Sigma-Aldrich) was added and incubated under the same conditions for 30 minutes. After derivatization, 100 µL of each sample was placed in a glass vial and used for GC-MS analysis.

GC-MS analysis was performed using an Intuvo 9000 GC System (Agilent) with an autosampler. The instrument was automatically tuned according to the manufacturer's instructions prior to each data collection sequence. For all samples, 0.5 µL was injected. The first 7 minutes after injection were not monitored to discard solvent peaks. Data collection was done over a retention time of 40 minutes.

Data analysis was done using the MassHunter Qualitative Analysis Navigator software (Agilent). For scan mode data, peak extraction and integration were done using the Agile 2 algorithm, filtering for peaks that had an area size at least 1% of the largest peak. The peak at 21.90-21.95 minutes was used as the ribitol

peak. The peak at 24.75-24.85 minutes was used as the homocitrate peak, based on chromatograms of a pure homocitrate standard. The homocitrate peak area was divided by the ribitol peak area to normalise for differences in instrument sensitivity between samples. For selected-ion monitoring mode, the Agile 2 algorithm was again used with default settings for peak extraction and integration. We saw no difference in the retention time of homocitrate in scan or SIM mode GC-MS. SIM-mode analysis therefore also used the peak at 24.75-24.85 minutes as that for homocitrate. Standard curve analysis and peak to concentration calculations for both scan and SIM-mode GC-MS were performed in Excel (Microsoft), using linear regression.

2.6.8 Acetylene reduction assays on Rlv3841 cultures

This protocol is based on recent work by Min et al. [336]. Strains of Rlv3841 were first grown on TY slopes with appropriate antibiotics for three days. Slopes were then washed three times with UMS without carbon or nitrogen sources using centrifugation at 4,000 RCF for 20 minutes. Washed cells were used to inoculate 10 mL liquid cultures in UMS with 10 mM glucose to OD_{600} 0.4. Taurine and IPTG inducers were added as appropriate, both to 10 mM. These liquid cultures were then incubated for 2 hours in an O_2 cabinet, with shaking at 200 RPM in a 1% O_2 atmosphere without tops, to equilibrate O_2 concentration. Following O_2 equilibration, liquid cultures were sealed with airtight neoprene lids and a 10% acetylene atmosphere was created by syringe extraction and injection as described in section . Cultures were then incubated for 20 hours with shaking at 200 RPM at 28 °C. Ethylene and acetylene readings were taken as described in section by gas chromatography, and an OD_{600} reading was also taken for each culture.

2.7 Computational methods

2.7.1 Statistical analyses

All statistical analyses were performed using Prism 8 (GraphPad Software). Significant differences between two groups were determined by Student's t-test. When analysing more than two groups, one-way or two-way ANOVA was used as specified for each figure. Dunnett's multiple comparisons post-hoc test correction was used. A p-value less than 0.05 was considered statistically significant.

2.7.2 Bioinformatic analyses

Protein homology calculations were performed using MUSCLE with default settings [367]. Alignments were viewed using Jalview [368]. Protein domain identification was carried out with the NCBI Conserved Domain Search tool using default settings and full results mode [369, 370]. Sequencing trace alignments were performed in Benchling, as were anaerobox and K-box motif searches.

2.7.3 Transcription start site (TSS) analysis

The TSS data set referenced in this work was produced by Vinoy Ramachandran and colleagues. Partial method details were recently published [2]. A publication discussing the work in detail is forthcoming. Data can be found in full on the NCBI SRA database, BioProject number PRJNA667846. Transcription start site locations were viewed and analysed using IGViewer [371, 372].

3

Activity of hFixL-FxkR-FixK and FnrN *ex planta*

Contents

3.1 Introduction

Regulation by oxygen (O_2) in rhizobia is essential for their symbioses with plants and involves multiple O_2 sensing systems [1]. Three systems are widespread in rhizobia: the FixLJ-FixK pathway, the similar hFixL-FxkR-FixK pathway, and

FnrN. Studies to date indicate that some rhizobia segregate these sensors into separate, redundant pathways [373, 374]. In others, the sensors appear to be merged into a single hierarchical pathway [188, 204]. To examine the relationship between these sensors, we studied the model organism *Rhizobium leguminosarum* biovar *viciae* 3841 (Rlv3841) [338]. Rlv3841 was chosen as sequence data suggested that it employs both a hFixL-FxkR-FixK pathway and FnrN, making it a useful model for how the two interact [152]. Past work in our group found that disrupting *fnrN* reduced nitrogen fixation to only 10% of WT, but disrupting the hFixL-FxkR-FixK pathway had little to no effect on fixation [362]. However, other work in our group had also shown that *fnrN* was induced under relatively mild microaerobic conditions (1% O_2), suggesting the hFixL-FxkR-FixK system was active and inducing the gene [366]. This was in line with a study in Rlv VF39 which also reported microaerobic *fnrN* induction [154]. The FixLJ-FixK pathway is known to be active at headspace concentrations as high as 5% O_2, but the O_2 sensitivity of the hFixL-FxkR-FixK pathway had not been reported [98, 104, 119].

The hFixL-FxkR-FixK and FnrN O_2 regulation systems in Rlv3841 could be in two arrangements (Figure 3.1). A parallel arrangement of hFixL-FxkR-FixK and FnrN would produce redundancy, whereas an arrangement in series would create hierarchy between the two systems. Our goal was therefore to understand how the two interact in Rlv3841 and to provide insight into why they coexist. In this chapter, we studied O_2 regulation in free-living Rlv3841 in microaerobic conditions. We began by surveying the transcriptional regulation elements upstream of the *hfixL*, *fxkR*, *fixK* and *fnrN* genes, and the arrangement of these genes in the genome. Many O_2 regulation genes and their targets were found to contain one or more anaerobox operators in their upstream region, notably including *fnrN* and two *fixNOQP* operons. The *fixNOQP* operon is a commonly studied target of O_2 regulation in rhizobia which encodes a high-affinity cbb_3-type terminal oxidase required for respiration during symbiosis [102, 103, 105]. As expected from the presence of anaeroboxes, *fnrN* and the two *fixNOQP* operons were induced under microaerobic conditions. hFixL-FxkR-FixK and FnrN can both induce anaerobox-regulated genes

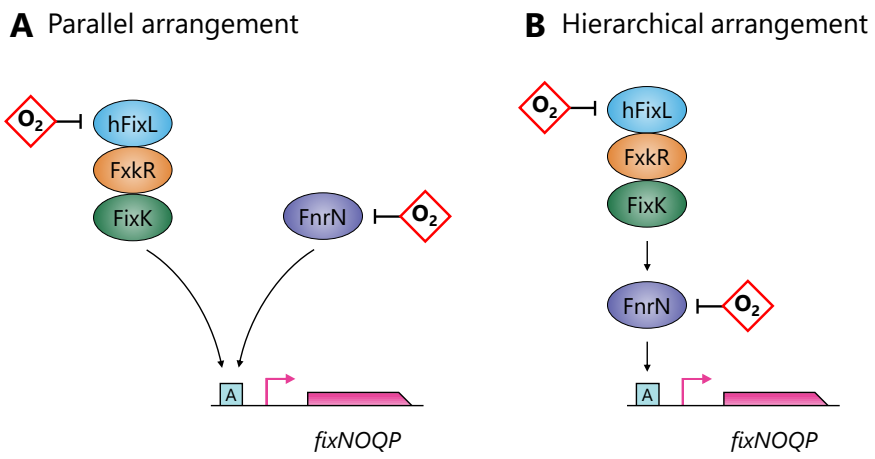


Figure 3.1: Possible arrangement of the hFixL-FxkR-FixK and FnrN oxygen regulation systems in Rlv3841. Rlv3841 encodes both a hFixL-FxkR-FixK TCS pathway and the FnrN transcription factor. Based on past findings, both should be able to regulate anaerobox-controlled genes, such as *fixNOQP* shown in the examples above. These two oxygen regulation systems could be in a (**A**) parallel arrangement or (**B**) hierarchical arrangement. In a parallel arrangement, one system could compensate for loss of activity from the other, creating redundancy. In a hierarchical system, loss of either system could critically impair induction of anaerobox-regulated genes. Note details of the hFixL-FxkR-FixK pathway have been omitted for simplicity, see Figure 3.4 for details.

[177, 230, 375]. Therefore, we next studied how disrupting the hFixL-FxkR-FixK pathway or FnrN affected microaerobic induction of *fnrN*, to investigate how the two sensors might cross-regulate. Microaerobic induction of *fnrN* was critically reliant on the hFixL-FxkR-FixK system, confirming it is functional in Rlv3841 and active at 1% O_2. FnrN itself had only a minimal impact on *fnrN* expression, indicating that under relatively mild microaerobic conditions, the hFixL-FxkR-FixK pathway and not FnrN is a key regulator of anaerobox genes. Disrupting *fnrN* had no significant impact on *fnrN* expression, suggesting FnrN was minimally active. Finally, expression of *fixNOQP* was studied. This largely reproduced our *fnrN* expression findings; microaerobic induction in Rlv3841 is driven primarily by the hFixL-FxkR-FixK pathway, with FnrN playing only a small role. Our study of both sensors is continued *in planta* in the next chapter (Chapter 4).

3.2 Results

3.2.1 Arrangement of O_2 regulation genes and operators on the genome and megaplasmids of Rlv3841

We began by studying the genome of Rlv3841 to identify potential homologs of known O_2 regulation genes and search for the presence of anaerobox and K-box operators regulating these genes (Figure 3.2). Rlv3841 has a single chromosome whose gene names start with RL, and six megaplasmids pRL7-12 whose gene names start with e.g. pRL9 [152]. O_2 regulation genes are found on the genome, on the main symbiotic megaplasmid pRL10 and on megaplasmid pRL9. Rlv3841 encodes two copies of *hfixL*, which we named $hfixL_9$ (pRL90020 on pRL9) and $hfixL_c$ (RL1879 on the chromosome), with 54.9% identity at the protein level. Rlv3841 also contains two homologs (58% identity) of *fxkR*, $fxkR_9$ (pRL90026) and $fxkR_c$ (RL1881). It has three putative *fixK* genes, which we designated $fixK_{9a}$ (pRL90019), $fixK_{9b}$ (pRL90025) and $fixK_c$ (RL1880). The $fixK_{9a}$ and $fixK_{9b}$ sequences have 53% amino acid identity, whilst $fixK_c$ shares 38% and 47% identity with these proteins, respectively. Both $fixK_{9a}$-$hfixL_9$ and $fixK_c$-$hfixL_c$ appear to form operons (Figures 3.2B and 3.2D). Rlv3841 has one copy of *fnrN*, on its chromosome (RL2818).

We also searched for two commonly O_2 regulated targets in rhizobia, the *fixNOQP* and *fixGHIS* operons, both required for nitrogen fixation. We found three putative *fixNOQP* operons in Rlv3841, which we labelled $fixNOQP_9$ (pRL90018-16), $fixNOQP_{10}$ (pRL100205-207) and $fixNOQP_c$ (RL0536-33). The closely related strain *R. leguminosarum* bv. *viciae* VF39 encodes two *fixNOQP* operons, and either was able to sustain nitrogen fixation activity [104]. The plasmid-encoded *fixNOQP* operons in Rlv3841 are near-identical (>90% protein identity) but diverge from the $fixNOQP_c$ operon, with which they share approximately 50% identity. The assembly machinery for *fixNOQP* is typically encoded by the *fixGHIS* operon [108, 109]. This operon is also usually anaerobox-regulated, and two *fixGHIS* operons (>90% protein identity) were found in Rlv3841 [171, 177, 376].

As well as apparently forming operons, most of these O_2 regulation genes and their targets are clustered together in Rlv3841. We identified four main clusters,

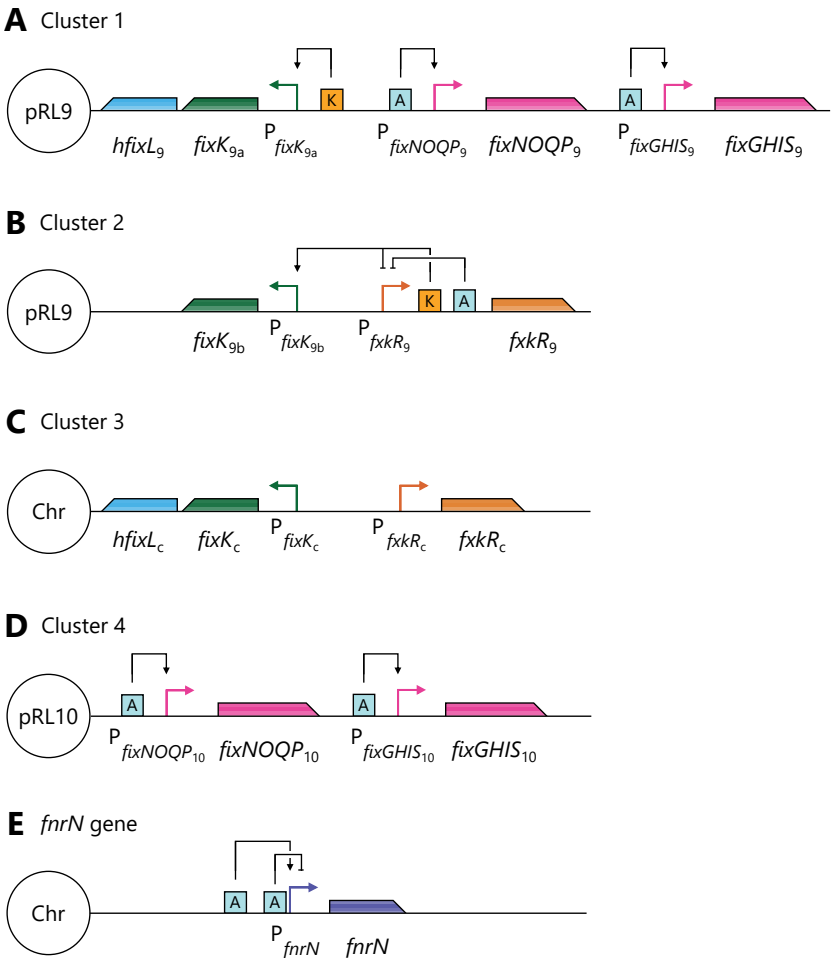


Figure 3.2: Oxygen regulation genes in Rlv3841 form clusters with anaerobox and K-box operators which suggest regulatory cross-talk and auto-regulation. Rlv3841 has genes encoding a hFixL-FxkR-FixK pathway and FnrN, both O_2 sensing systems. Multiple homologs of some oxygen regulation genes are present, and many are arranged in clusters. Anaeroboxes (bound by FixK/FnrN) are indicated by blue 'A' squares, and K-boxes (bound by FxkR) are indicated by orange 'K' squares. Right-angle arrows indicated transcription start sites (TSS). **(A)** Cluster 1 on megaplasmid pRL9; *fixK*$_{9a}$ forms an operon with *hfixL*$_9$, regulated by a K-box. Adjacent to this operon is *fixNOQP*$_9$, followed immediately by *fixGHIS*$_9$, both regulated by anaeroboxes. **(B)** Cluster 2, also on pRL9, contains *fixK*$_{9b}$ and *fxkR*$_9$. A K-box in their intergenic region likely regulates *fixK*$_{9a}$ but may also repress *fxkR*$_9$ when bound. The anaerobox downstream of the *fxkR*$_9$ TSS may also repress *fxkR*$_9$ expression. **(C)** Cluster 3 on the chromosome contains *fxkR*$_c$, *fixK*$_c$ and *hfixL*$_c$. Unlike the similar clusters on pRL9, the intergenic region of this cluster contains no anaerobox or K-box operators. **(D)** Cluster 4 on pRL10 has *fixNOQP*$_{10}$ and *fixGHIS*$_{10}$ immediately adjacent and anaerobox-regulated, like their homologs in cluster 1. **(E)** The *fnrN* gene is not part of a cluster and is positively and negatively regulated by distal and proximal anaeroboxes, respectively. Graphs are not to scale. Location details for TSS, anaeroboxes and K-boxes are given in Table 3.1

which we numbered 1-4. Cluster 1 (Figure 3.2A), on pRL9, contains the putative *fixK*$_{9a}$-*hfixL*$_9$ operon immediately adjacent to *fixNOQP*$_9$ and *fixGHIS*$_9$. Cluster 2 (Figure 3.2B), also on pRL9, contains the *fxkR*$_9$ and *fixK*$_{9b}$ genes. Cluster 3 (Figure 3.2C) is on the chromosome and contains the putative *fixK*$_c$-*hfixL*$_c$ operon adjacent to *fxkR*$_c$. Cluster 4 (Figure 3.2D) on pRL10 contains no O_2 regulation genes but has *fixNOQP*$_{10}$ immediately adjacent to *fixGHIS*$_{10}$, as in cluster 1. In summary, sequence data suggested that Rlv3841 employs both a hFixL-FxkR-FixK pathway and FnrN to regulate nitrogen fixation genes in response to O_2 concentration.

To understand how these genes might be regulated, we next searched their upstream regions for operators: anaeroboxes indicating FixK/FnrN control and K-boxes indicating FxkR control. Results are summarised in Figure 3.2. Previous work in our group has mapped the transcription start site (TSS) of Rlv3841 genes, allowing us to determine the location of these operators relative to the TSS of each gene (Table 3.1) [377]. Note that putative anaeroboxes were also found upstream of many other Rlv3841 genes, not included in this table.

Cluster	Gene name	TSS coordinate	Anaerobox 1 location	Anaerobox 2 location	K-box location
Cluster 1	*fixK*$_{9a}$ (pRL90019)	19878	×	×	-62
	fixNOQP$_{9}$ (pRL90018-16)	19721	-33	×	×
	fixGHIS$_{9}$ (pRL90015-12A)	15908	-32	×	×
Cluster 2	*fixK*$_{9b}$ (pRL90025)	27090	×	×	-62
	fxkR$_{9}$ (pRL90026)	27146	×	+38	+6
Cluster 3	*fixK*$_{c}$ (RL1880)	1977111	×	×	×
	fxkR$_{c}$ (RL1881)	1977167	×	×	×
Cluster 4	*fixNOQP*$_{10}$ (pRL100205-207)	206214	-34	×	×
	fixNOQP$_{10}$ (pRL100208-10A)	210004	-32	×	×
Not clustered	*fnrN* (RL2818)	2978390	-34	-2	×
	fixNOQP$_{c}$ (RL0536-33)	579729	×	×	×

Table 3.1: Location of transcription start sites, anaeroboxes and K-boxes for select oxygen regulation genes in Rlv3841. Locations of anaeroboxes and K-boxes are given relative to the TSS of each gene respectively. Anaeroboxes more than 90 bp upstream of the TSS are not included. The location of activating anaeroboxes is well conserved across the genes shown above (between -32 and -34 relative to the TSS). Likewise, both megaplasmid-encoded *fixK* copies have their respective K-boxes in the same position relative to the TSS (-62). The anaerobox downstream (+38) of the *fxkR*$_9$ transcription start site likely represses it when bound. A single K-box is shared between *fxkR*$_9$ and *fixK*$_{9b}$ (see Figure 3.2). It is correctly located (-62) to induce *fixK*$_{9b}$ when bound, and its location suggests it simultaneously represses *fxkR*$_9$ (+6). The operators downstream of the *fxkR*$_9$ TSS may create a negative feedback loop in the Rlv3841 hFixL-FxkR-FixK pathway. The downstream anaerobox also suggests FnrN repression of hFixL-FxkR-FixK via repression of *fxkR*$_9$ transcription, but this requires further study.

Cluster 1 contains a K-box controlling *fixK*$_{9a}$-*hfixL*$_9$, located -62 relative to the TSS, and an anaerobox controlling *fixNOQP*$_9$ located -33 relative to the TSS (Figure 3.2A). Immediately downstream of *fixNOQP*$_9$ is *fixGHIS*$_9$, also with a single anaerobox (-32). Past microarray work in our group has compared transcription from Rlv3841 genes in free-living cells vs bacteroids [377]. All three genes in cluster 1 were significantly upregulated; *fixK*$_{9a}$ 7.1-fold up (p=0.012), *fixN*$_9$ 38.1-fold up (p=0.010) and *fixG*$_9$ 9.3-fold up (p=0.022). Each operator appears to regulate only one gene or operon in cluster 1, which is what past studies have typically reported [1].

Cluster 2 appears to contain a more unusual arrangement that has not previously been reported, with one operator potentially serving both to activate and repress gene expression (Figure 3.2B). Like cluster 1, both an anaerobox and a K-box are present between the *fixK*$_{9b}$ and *fxkR*$_9$ open reading frames. The location of the K-box relative to the TSS of *fixK*$_{9b}$ (-62) is identical to that of the K-box upstream of *fixK*$_{9a}$, suggesting both these *fixK* homologs are induced by FxkR. However, past microarray work found no induction of *fixK*$_{9b}$ in 21-day old bacteroids (1.0-fold up, p=0.992) [377] and non-significant induction in 28-day old bacteroids (23.8-fold up, p=0.062). The gene is likely induced by FxkR, but this effect may not be as strong as *fixK*$_{9a}$. Instead, the K-box in cluster 2 may act mainly to repress gene expression; it is downstream of the *fxkR*$_9$ TSS (+6). Although not previously reported, this suggests FxkR binding to this K-box simultaneously induces *fixK*$_{9b}$ and represses *fxkR*$_9$, making FxkR negatively auto-regulated. Alternatively, the K-box may have a negligible impact on *fixK*$_{9b}$ expression and act purely to auto-regulate *fxkR*$_9$ expression. Also downstream of the *fxkR*$_9$ TSS is an anaerobox. This is more than 90 bp upstream of the *fixK*$_{9b}$ TSS. It may participate in *fixK*$_{9b}$ induction, but this is unlikely as anaeroboxes in Rlv3841 are consistently located closer (-32-34 bp) to their nearest TSS. It is more likely that FixK and/or FnrN binding at this anaerobox further represses *fxkR*$_9$. Anaerobox-mediated repression has been reported in the past, with *fnrN* genes, in which one anaerobox is located close enough to the *fnrN* TSS to sterically block transcription.

Cluster 3, on the chromosome, surprisingly has no anaerobox or K-box sites that conform to established motifs (Figure 3.2C). One or both may still be present, but neither *fxkR*$_c$ (1.6-fold up, p=0.140) nor *fixK*$_c$-*hfixL*$_c$ (1.3-fold up, p=0.145) was upregulated in 21-day old bacteroids relative to free-living Rlv3841. The balance of evidence is therefore that genes in this cluster are not O_2 regulated. It is unknown to what extent the genes are expressed at all without such induction. Although the genes do not appear O_2 regulated, they may still act as O_2 regulators. The Rlv3841 *fnrN* gene is located on the chromosome and does not form a cluster with other O_2 regulation genes (Figure 3.2E). Two anaeroboxes were found upstream of *fnrN*. A similar dual-anaerobox arrangement exists in Rlv UPM791, where FnrN positively and negatively auto-regulates its own expression [205]. Binding of FnrN to the distal anaerobox induces *fnrN* transcription and binding to the proximal anaerobox represses it. Auto-activation of FnrN has also been reported in *Rhizobium etli* CNPAF512 [131]. It therefore appeared likely that FnrN is auto-regulated in Rlv3841. Of note, FixK also binds anaeroboxes so could also regulate *fnrN*. However, this had not been investigated prior to this study. A study in *R. leguminosarum* VF39 found that microaerobic expression of *fnrN* also requires RpoN [229]. This finding has not been replicated elsewhere, and its significance remains unclear. Work in *R. etli* CNPAF512 showed that *fnrN* is not controlled by RpoN in that organism [131]. Rlv3841 encodes one putative *rpoN* gene (RL0422), but we found no RpoN binding sites upstream of the Rlv3841 *fnrN* transcription start site. RpoN therefore does not appear to be required for *fnrN* expression in Rlv3841.

We also found a cluster of O_2-regulated genes on pRL10 (Figure 3.2D), containing *fixNOQP*$_{10}$ and *fixGHIS*$_{10}$. Both have a single anaerobox upstream of their TSS, at relative positions near-identical to the anaeroboxes regulating *fixNOQP*$_9$ and *fixGHIS*$_9$. Microarray data supported O_2 upregulation of *fixN*$_{10}$ (119.7-fold up, p=0.003), although *fixG*$_{10}$ upregulation was not significant (3.5-fold up, p=0.068) [377]. Other microarray studies found both genes were strongly upregulated [378]. In line with findings in other rhizobia, the *fixNOQP* and *fixGHIS* operons on pRL9 and pRL10 are likely all anaerobox-regulated [156, 181, 229]. The *fixNOQP*$_c$

operon is an exception, as we found no evidence of an anaerobox upstream and microarray work showed it was not induced in nodules (1.7-fold upregulated, p=0.101). Given this lack of O_2 regulation and poor homology to the other *fixNOQP* operons, $fixNOQP_c$ may not be functional or may have diverged to serve a different, currently unknown, function.

Taken together, these findings from sequence information suggested that Rlv3841 has a functional hFixL-FxkR-FixK pathway and *fnrN* gene. There was also evidence of auto-regulatory mechanisms that had not yet been reported. On cluster 2, $fxkR_9$ may be directly auto-repressed via a K-box, and indirectly auto-repressed by an anaerobox. The hFixL-FxkR-FixK pathway in Rlv3841 may also be partially auto-activating, as the $fixK_{9a}$-$hfixL_9$ operon in cluster 1 is under K-box control. FxkR activity would therefore upregulate *hfixL* transcription. Finally, there were multiple routes by which hFixL-FxkR-FixK and FnrN could regulate each-other via FnrN/FixK binding of anaeroboxes controlling *fnrN*, $fxkR_9$ and possibly $fixK_{9b}$. A schema of O_2 regulation in Rlv3841 is given in Figure 3.3.

3.2.2 Microaerobic conditions induce *fnrN* and both *fixNOQP* operons in free-living Rlv3841

To confirm that the anaerobox operators we identified in the genome of Rlv3841 were functional, we studied the response of *fnrN*, $fixNOQP_9$ and $fixNOQP_{10}$ expression to microaerobic conditions (1% O_2) in free-living cells. All three have at least one upstream anaerobox so were expected to be induced. To study expression, promoter fusions were made to the fluorescent marker *syfp2* in a pOPS0786 backbone, and these reporter plasmids were conjugated into wild-type (WT) Rlv3841. A reporter plasmid was also assembled with the promoter of Rlv3841 *nifH*, a component of the nitrogenase complex [43, 295, 309]. This has no upstream anaerobox but is regulated by the O_2-sensing NifA transcription factor [280, 281, 283]. NifA becomes active under low O_2 conditions but in most rhizobia activity has only been reported in the near-anoxic core of nodules [98, 379–381]. However, work in *E. meliloti* has suggested the protein may already be active in the early stages

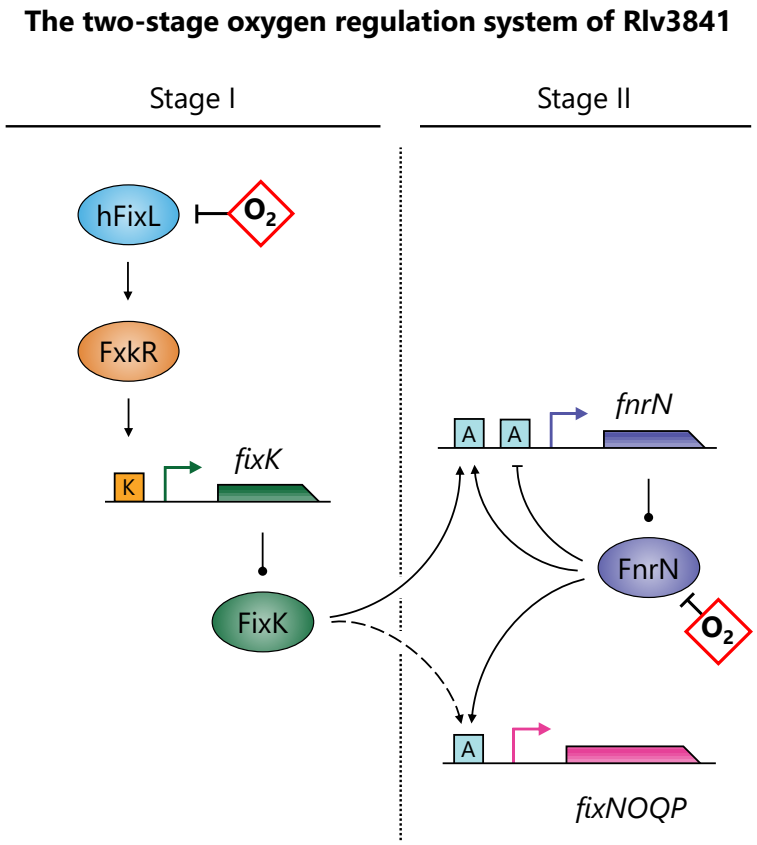


Figure 3.3: Rlv3841 senses O_2 via both a hFixL-FxkR-FixK pathway and the FnrN protein. Oxygen is shown in red diamonds. Proteins are shown as ovals, operator sites as squares and genes as pointed rectangles. TSS are shown as right-angled arrows. Line endings indicate activation (arrows), inhibition (blunt end) and translation (circle). The single pathway formed by the two sensors acts in two stages. Stage I starts under microaerobic conditions and can function outside the nodule. In this stage, hFixL is active but FnrN is not. hFixL activates FxkR, which binds to the K-box operator (orange "K" squares) to induce expression of *fixK*. FixK binds to anaerobox operators (blue "A" squares) to induce expression, including upstream of *fixNOQP* (dashed line) and *fnrN*. Once oxygen in the bacteria reaches near-anaerobic levels, FnrN becomes active and stage II begins. Like FixK, FnrN binds anaeroboxes. It auto-regulates *fnrN* both positively and negatively and induces *fixNOQP* expression.

of nodule development [382]. There was therefore a possibility that *nifH* would respond to microaerobic conditions in free-living cells. Later work in this thesis (see Chapter 5) suggests Rlv3841 NifA is active under microaerobic conditions if native transcriptional regulation is bypassed.

We observed microaerobic induction of *fnrN*, $fixNOQP_9$ and $fixNOQP_{10}$, indicating their upstream anaeroboxes are functional (Figure 3.3). Expression of *nifH* did not increase under microaerobic conditions, indicating Rlv3841 *nifA* is

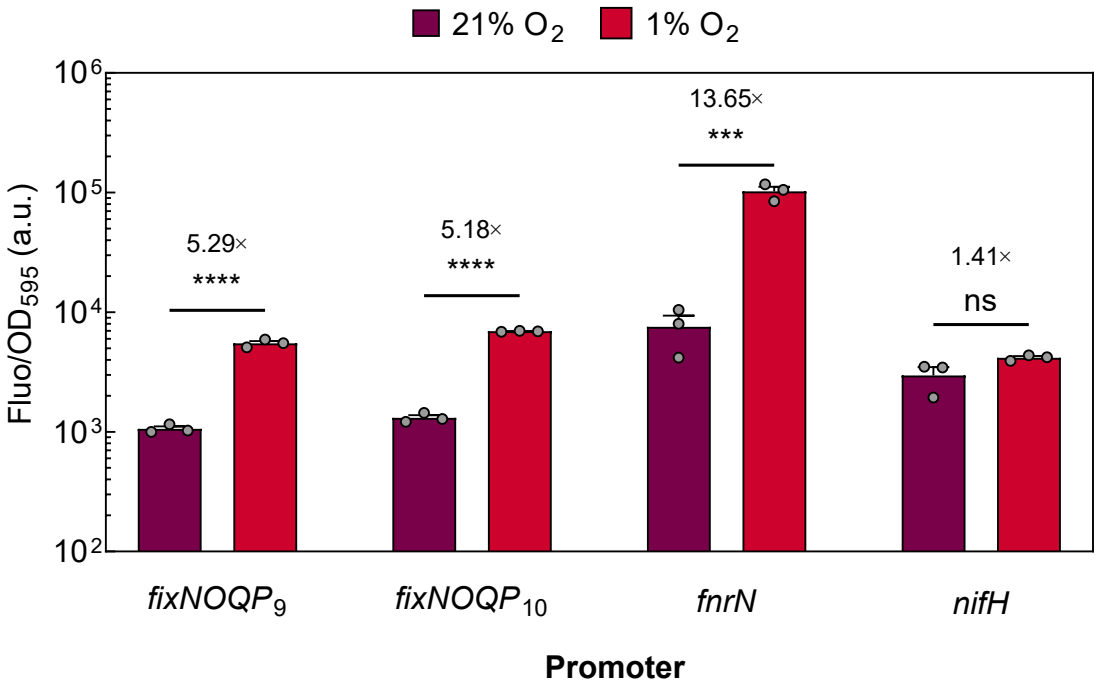


Figure 3.4: Microaerobiosis induces *fixNOQP* and *fnrN* genes in free-living Rlv3841. Promoter fusions to *syfp2* of $fixNOQP_9$ (OPS1267), $fixNOQP_{10}$ (OPS1287) and *fnrN* (OPS1296) were used to measure the activity of these promoters in free-living Rlv3841 WT at 1% O_2 (red bars) relative to 21% O_2 (mauve bars). Activity of all three promoters, measured as fluorescence normalised by OD_{595}, increased under microaerobic conditions. A positive control (OPS1294) showed no impact on OD-normalised fluorescence due to the microaerobic environment (data not shown). An induction fold of 5× was recorded for both *fixNOQP* operons, but *fnrN* showed more than double this fold change indicating stronger induction. No effect of O_2 concentration on *nifH* expression (OPS1268) was observed, indicating no NifA activity. Values are plotted on a logarithmic scale. Data are averages (±SEM) from three biological replicates, ns (not significant) $P \geq 0.05$; ***$P < 0.001$; ****$P < 0.0001$; by Student's t test.

not expressed or inactive under these conditions. We observed an induction fold in *fnrN* almost double that of either *fixNOQP* operon, indicating more induction under these conditions.

These findings agreed with the sequence data analysed in section 3.2.1 above and show that the anaeroboxes located upstream of *fnrN*, $fixNOQP_9$ and $fixNOQP_{10}$ in Rlv3841 are functional.

3.2.3 The hFixL-FxkR-FixK pathway is critical for microaerobic *fnrN* induction in free-living Rlv3841

Rlv3841 *fnrN* contains two anaeroboxes in its promoter region and was induced under microaerobic conditions. Both FixK and FnrN bind anaeroboxes and could be responsible for this effect [167, 201]. Microaerobic induction of *fnrN* (Figure 3.4) could therefore be due to FnrN auto-activation and/or induction by the hFixL-FxkR-FixK pathway (Figure 3.3). The hFixL-FxkR-FixK pathway is known to be active at relatively high O_2 concentrations, including in free-living rhizobia under microaerobic conditions [132, 139]. The O_2 concentration at which FnrN is active has not been investigated. To determine the respective importance of the hFixL-FxkR-FixK pathway and FnrN for microaerobic *fnrN* expression, we studied this reporter in Rlv3841 mutants of the hFixL-FxkR-FixK pathway or FnrN (Figure 3.5).

In a double *hfixL* mutant (LMB496; $hfixL_9$::ΩSpec $hfixL_c$:pK19 single recombination), free-living microaerobic expression of *fnrN* was reduced to 25% of its WT level. The single mutant of $hfixL_9$ (LMB495; $hfixL_9$::ΩSpec) individually reproduced most of this decrease whilst the single mutant of $hfixL_c$ (LMB403; $hfixL_c$:pK19 single recombination) did not decrease *fnrN* expression. This suggests $hfixL_9$ is the critical hFixL protein under free-living microaerobic conditions, with $hfixL_c$ playing little to no role. In other rhizobia, hFixL acts through the FxkR intermediate to induce *fixK* expression, which in turn induces anaerobox-regulated genes [129, 132]. Rlv3841 has two *fxkR* homologs, $fxkR_c$ and $fxkR_9$ [152]. Only one of these, $fxkR_9$, was found to be upregulated in past microarray work [377]. This homolog also has both an anaerobox and K-box upstream of its open reading frame (ORF), unlike $fxkR_c$ which has neither. We therefore speculated that $fxkR_9$ is the main FxkR homolog in Rlv3841 and deleted the gene to produce strain OPS1808 ($\Delta fxkR_9$). Like the double *hfixL* mutant, the $\Delta fxkR_9$ strain critically reduced induction of *fnrN* under free-living microaerobic conditions. This finding supports the role of $FxkR_9$ as the mediator of hFixL O_2 regulation in Rlv3841, in agreement with studies in other rhizobia. Some activity remained from P_{fnrN} in the $fxkR_9$ mutant and may be due in part to $FxkR_c$. However, this was minimal, supporting our theory

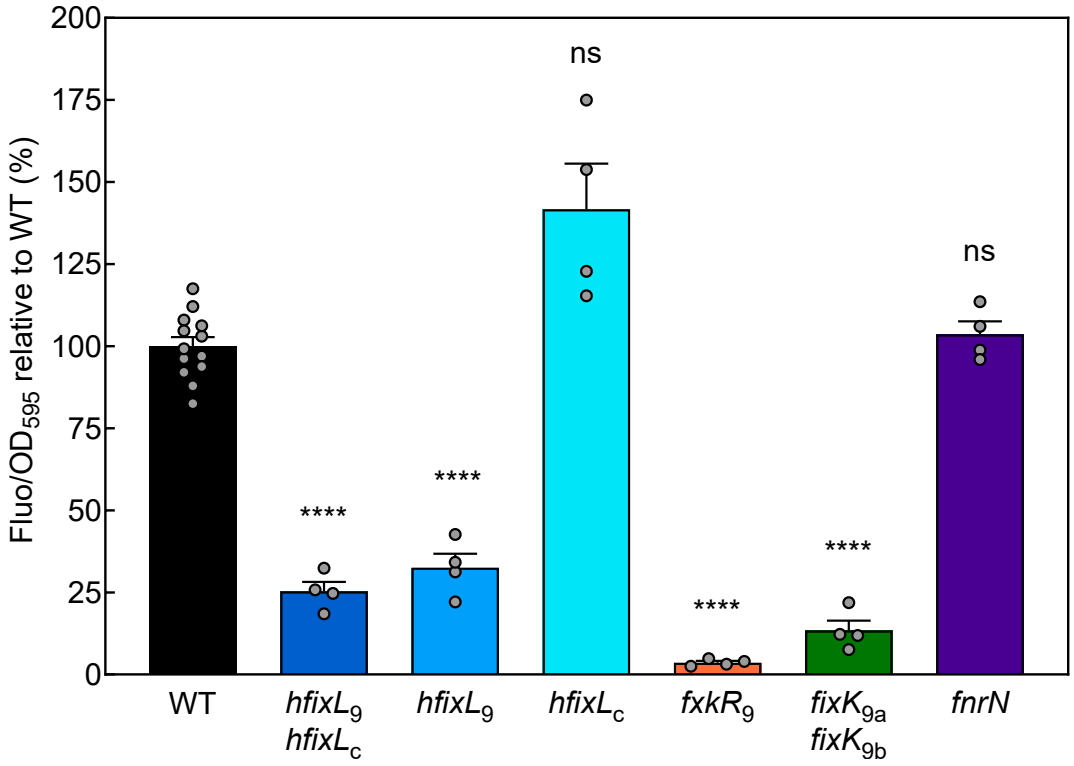


Figure 3.5: Under microaerobic (1% O_2) conditions in free living cells, the hFixL-FxkR-FixK pathway and not FnrN is a key activator of *fnrN*. The Rlv3841 *fnrN* promoter was fused to *syfp2* (pOPS0980) and conjugated into Rlv3841 WT and O_2 regulation mutants (*hfixL*$_9$ *hfixL*$_c$, LMB496; *hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403; *fxkR*$_9$, OPS1808; *fnrN*, LMB648; *fixK*$_{9a}$ *fixK*$_{9b}$, OPS2500). Fluorescence and OD_{595} measurements were taken after cells were grown under microaerobic conditions (1% O_2). Individual values (Fluo/OD_{595}) are normalised such that the WT average is 100%. Activity from the *fnrN* promoter was critically reduced or nearly abolished in the double *hfixL* mutant and double *fixK* mutant, as well as the *hfixL*$_9$ and *fxkR*$_9$ single mutant backgrounds. The single *hfixL*$_c$ mutant did not decrease activity, nor did the *fnrN* mutant. Data are averages (±SEM) from at least four biological replicates. Data for figures 3.5, 3.6 and 3.7 were analysed as a single set. Statistical tests are differences relative to wild-type (WT) expression; ns (no significant decrease) $P \geq 0.05$; ****$P < 0.0001$; by two-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

that *fxkR*$_9$ is the main *fxkR* homolog in Rlv3841. We also investigated the role of the final member of the hFixL-FxkR-FixK pathway, FixK. Rlv3841 has three *fixK* homologs, two on pRL9 and one on the chromosome (Figure 3.2). Like the *fixNOQP*$_c$ operon, *fixK*$_c$ does not appear to have an upstream anaerobox, shows lower homology to the other two *fixK* homologs and was not induced *in planta* when studied in past microarray work [377]. We therefore only disrupted *fixK*$_{9a}$ and *fixK*$_{9b}$ to eliminate FixK activity in Rlv3841, producing strain OPS2500 (Δ*fixK*$_{9a}$ *fixK*$_{9b}$:pRU877 single recombination). In line with its essential role in the hFixL-FxkR-FixK pathway, disrupting FixK activity in Rlv3841 critically impaired *fnrN* induction under microaerobic free-living conditions. As with the *fxkR*$_c$ homolog, *fixK*$_c$ may produce some activity but this is likely minimal when compared to the two pRL9 *fixK* homologs. The other two genes in cluster 3 (Figure 3.2C) with *fixK*$_c$, *hfixL*$_c$ and *fxkR*$_c$, also had little if any effect on the activity of P$_{fnrN}$. Lastly, we studied the role of FnrN auto-regulation. Past work has shown that FnrN is able to both auto-activate *fnrN* by binding to the distal anaerobox, and auto-repress *fnrN* by binding to the proximal anaerobox [205]. However, a mutant of *fnrN* (LMB648, *fnrN*::ΩTet) had no effect on expression of *fnrN* at 1% O_2, indicating FnrN auto-regulation does not occur under free-living microaerobic conditions.

In summary, induction of *fnrN* under microaerobic conditions in Rlv3841 is critically reliant on the hFixL-FxkR-FixK pathway, with no discernible FnrN auto-regulation taking place. Although Rlv3841 has multiple homologs of *hfixL*, *fxkR* and *fixK*, eliminating the pRL9-borne homologs was sufficient in each case to critically impair the activity of the hFixL-FxkR-FixK pathway. The chromosome-based homologs of this pathway thus appear to have little if any role in microaerobic *fnrN* induction.

3.2.4 Microaerobic induction of *fixNOQP* in free-living Rlv3841 relies primarily on the hFixL-FxkR-FixK pathway, with minor FnrN involvement

Past work in our group had shown that loss of *fnrN* but not the *hfixL* genes critically impaired nitrogen fixation in Rlv3841 [362]. This suggested that the hFixL-FxkR-

FixK pathway played a relatively minor role compared to FnrN in inducing key genes required for nitrogen fixation. To investigate the relative importance of the two different O_2 sensing systems, we chose to study the expression of the *fixNOQP*$_9$ and *fixNOQP*$_{10}$ operons. The *fixNOQP* operon encodes a high-affinity *cbb*$_3$-type terminal oxidase required for respiration during symbiosis [102, 103, 105]. It is typically regulated by an anaerobox [38]. Some rhizobia encode multiple redundant terminal oxidases controlled by different regulators, but no alternatives appear to be encoded by Rlv3841 [110, 111]. The strain therefore likely relies entirely on *fixNOQP* for respiration during symbiosis, and genes important for its expression are also important for nitrogen fixation activity.

Activity from the *fixNOQP*$_9$ (Figure 3.6) and *fixNOQP*$_{10}$ (Figure 3.7) reporter plasmids described in section 3.2.2 was studied in Rlv3841 strains with disrupted O_2 regulation genes, under microaerobic conditions. In Rlv3841 WT, both are induced under these conditions, albeit less strongly than *fnrN* (Figure 3.4). The double *hfixL* mutant severely reduced microaerobic induction, resulting in 17% of WT *fixNOQP*$_9$ expression and minimal expression of *fixNOQP*$_{10}$. The single *hfixL*$_9$ mutant significantly reduced expression of both operons whilst the *hfixL*$_c$ mutant only reduced expression of *fixNOQP*$_9$, to 71% of WT. This indicates *hfixL*$_9$ is the dominant homolog and *hfixL*$_c$ plays only a minor role in *fixNOQP* induction, as was the case for microaerobic *fnrN* induction (Figure 3.5).

In the Rlv3841 *fxkR*$_9$ mutant, expression of both *fixNOQP* operons was reduced to less than 25% of WT, as expected given that FxkR is required to transmit low O_2 sensing by hFixL into *fixK* induction [129]. As with the *fnrN* reporter, remaining expression from the *fixNOQP*$_9$ and *fixNOQP*$_{10}$ reporters in the *fxkR*$_9$ mutant may be due to redundancy via the *fxkR*$_c$ homolog, or the result of background FixK and/or FnrN activity. The Rlv3841 double *fixK*$_{9a}$ *fixK*$_{9b}$ mutant significantly reduced the expression of both *fixNOQP* operons, but >50% of WT activity remained. This contrasts with the *fnrN* reporter, where the double *fixK* mutant reduced activity to <15% of WT. One possibility is that the *fixK*$_c$ homolog sustains a high level of *fixNOQP* expression, but it is unclear why this was not the case for *fnrN* expression.

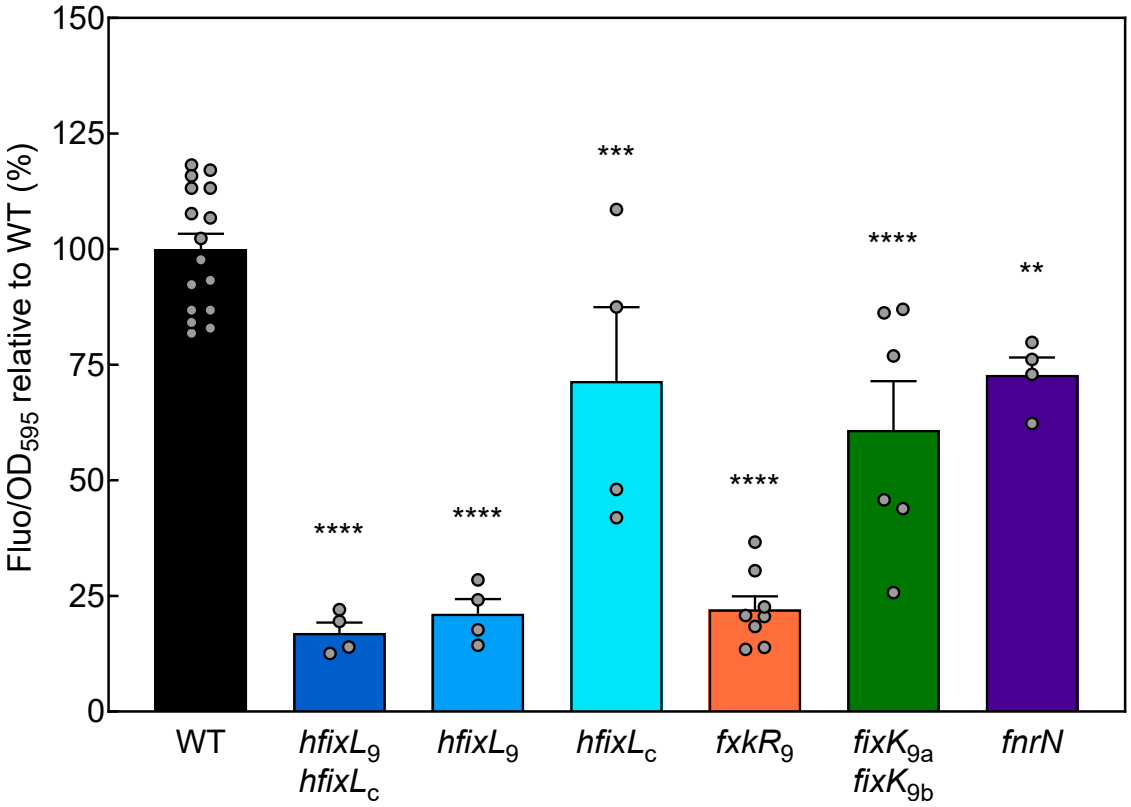


Figure 3.6: Under microaerobic (1% O_2) conditions in free living cells, the hFixL-FxkR-FixK pathway and not FnrN is a key activator of *fixNOQP*$_9$. The Rlv3841 *fixNOQP*$_9$ promoter was fused to *syfp2* (pOPS0978) and conjugated into Rlv3841 WT and O_2 regulation mutants (*hfixL*$_9$ *hfixL*$_c$, LMB496; *hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403; *fxkR*$_9$, OPS1808; *fnrN*, LMB648; *fixK*$_{9a}$ *fixK*$_{9b}$, OPS2500). Fluorescence and OD_{595} measurements were taken after cells were grown under microaerobic conditions (1% O_2). Individual values (Fluo/OD_{595}) are normalised such that the WT average is 100%. Like the *fnrN* promoter, activity from the *fixNOQP*$_9$ promoter was critically reduced in the double *hfixL* mutant and single *hfixL*$_9$ mutant, as well as the *fxkR*$_9$ mutant. Activity was slightly reduced in the *hfixL*$_c$ and *fnrN* single mutants. Unlike the *fnrN* promoter, the *fixNOQP*$_9$ promoter retained substantial (>60%) activity in the double *fixK*$_{9a}$ *fixK*$_{9b}$ mutant. Activity was nevertheless significantly decreased in this mutant. Data are averages (±SEM) from at least four biological replicates. Data for figures 3.5, 3.6 and 3.7 were analysed as a single set. Statistical tests are differences relative to wild-type (WT) expression; ns (no significant decrease) $P \geq 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; by two-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

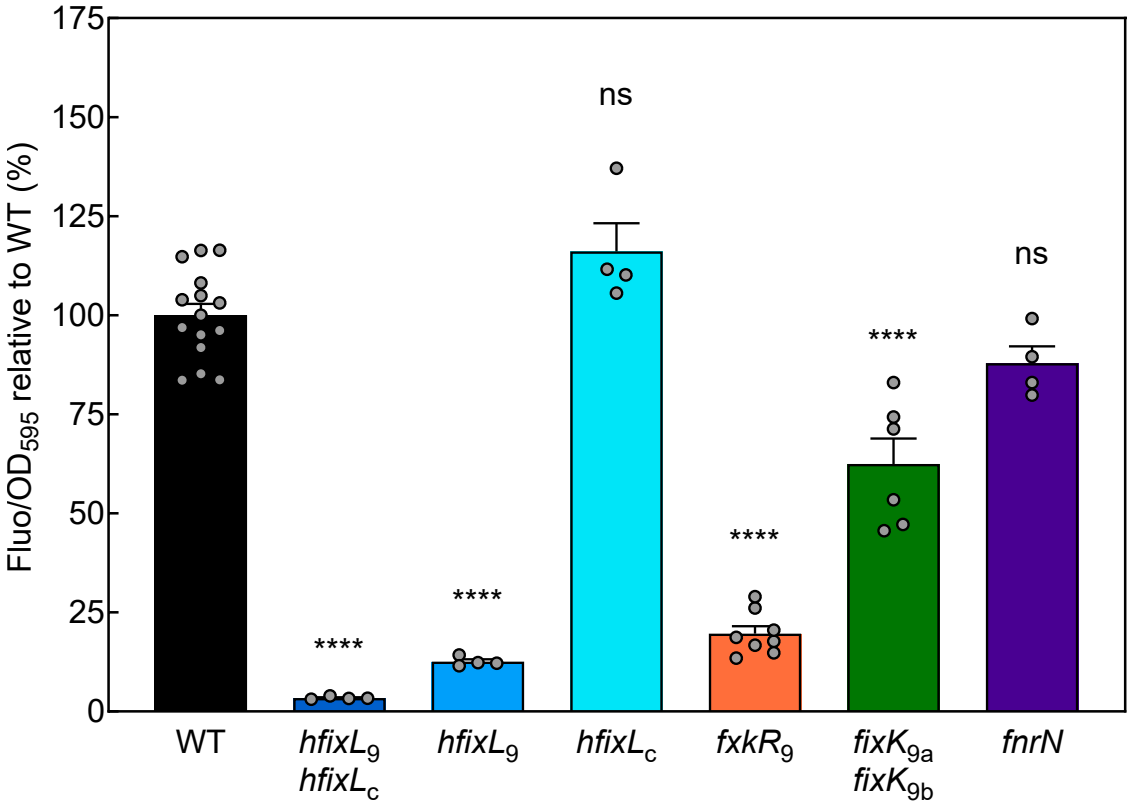


Figure 3.7: Under microaerobic (1% O_2) conditions in free living cells, the hFixL-FxkR-FixK pathway and not FnrN is a key activator of *fixNOQP*$_{10}$. The Rlv3841 *fixNOQP*$_{10}$ promoter was fused to *syfp2* (pOPS0977) and conjugated into Rlv3841 WT and O_2 regulation mutants (*hfixL*$_9$ *hfixL*$_c$, LMB496; *hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403; *fxkR*$_9$, OPS1808; *fnrN*, LMB648; *fixK*$_{9a}$ *fixK*$_{9b}$, OPS2500). Fluorescence and OD_{595} measurements were taken after cells were grown under microaerobic conditions (1% O_2). Individual values (Fluo/OD_{595}) are normalised such that the WT average is 100%. Like the *fnrN* promoter, activity from the *fixNOQP*$_{10}$ promoter was critically reduced or nearly abolished in the double *hfixL* mutant and single *hfixL*$_9$ mutant, as well as the *fxkR*$_9$ mutant. Activity was not significantly reduced in the *hfixL*$_c$ and *fnrN* single mutants. Like the *fixNOQP*$_9$ promoter, the *fixNOQP*$_{10}$ promoter was significantly decreased in the double *fixK*$_{9a}$ *fixK*$_{9b}$ mutant but retained substantial (>60%) activity. Data are averages (±SEM) from at least four biological replicates. Data for figures 3.5, 3.6 and 3.7 were analysed as a single set. Statistical tests are differences relative to wild-type (WT) expression; ns (no significant decrease) $P \geq 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; by two-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

Taken together, our results show that the hFixL-FxkR-FixK pathway is a key regulator of *fixNOQP* expression under free-living microaerobic conditions. However, we also found that disrupting *fnrN* significantly decreased expression of $fixNOQP_9$. Expression of $fixNOQP_{10}$ was also decreased in the *fnrN* mutant, but this change was not significant (p=0.441). This contrasts with the *fnrN* reporter, where disrupting *fnrN* had no effect on expression. In summary, under free-living microaerobic conditions the hFixL-FxkR-FixK pathway is a key regulator of *fnrN* and both $fixNOQP_9$ and $fixNOQP_{10}$. However, there is also evidence that FnrN plays a role in regulating the *fixNOQP* operons.

3.2.5 Plasmid-borne expression of $hfixL_9$ from the P_{lac} promoter only partially complemented the Rlv3841 double $hfixL_9$ $hfixL_c$ mutant

Our $hfixL_9$ (LMB495) mutant showed the largest phenotypic effect out of the two single *hfixL* mutants, and we therefore attempted to complement the strain in which this gene had been disrupted. This was complicated by the fact that $hfixL_9$ appears to form an operon with $fixK_{9a}$. We first attempted complementation via a Tn7-integrated $hfixL_9$ expression cassette in which $hfixL_9$ was under P_{lac} control with a non-native ribosome binding site (RBS), using the pOGG280 backbone. This construct assembled in *E. coli* but could not be successfully conjugated into Rlv3841. We theorized that $hfixL_9$ was being produced to a toxic level in Rlv3841 due to poor LacI repression. We attempted to rectify this problem by driving $hfixL_9$ from the native $fixK_{9a}$ promoter and RBS instead, again with the pOGG280 backbone. However, this construct could not be transformed into *E. coli*, suggesting the native $fixK_{9a}$ promoter was causing toxic levels of $hfixL_9$ production. Finally, we sought to strike a balance between these two approaches by using the P_{lac} promoter but the native $hfixL_9$ RBS, in a non-integrating pOGG250 backbone. This construct was successfully assembled in *E. coli* and conjugated into the Rlv3841 double *hfixL* mutant, producing strain OPS2467. We tested for complementation using the *fnrN* and $fixNOQP_9$ reporters under microaerobic conditions (Figure 3.8).

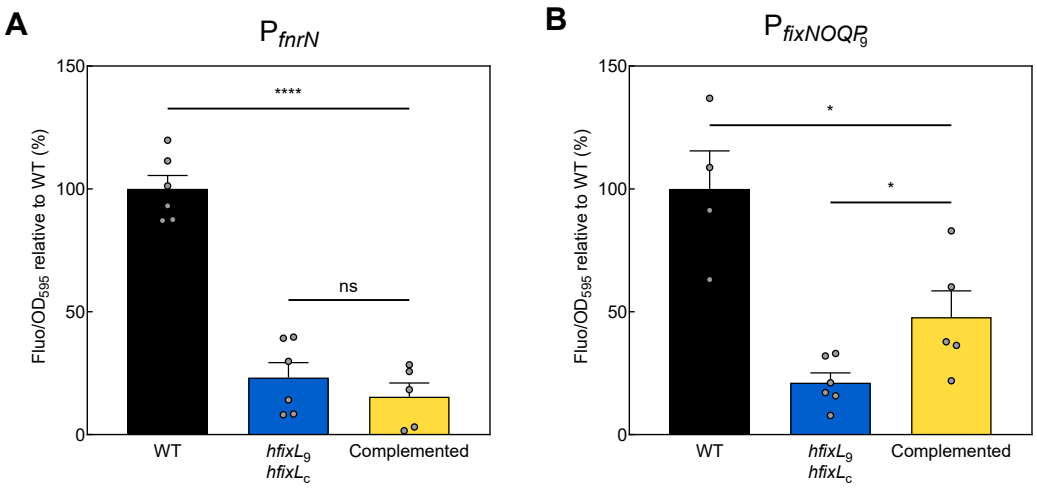


Figure 3.8: Complementation of the Rlv3841 *hfixL*$_9$ *hfixL*$_c$ double mutant (LMB496). Rlv3841 WT is shown in black, the double *hfixL*$_9$ *hfixL*$_c$ mutant (LMB496) in blue and the complemented strain in yellow (OPS2467). Graphs show activity from the **(A)** *fnrN* promoter and **(B)** *fixNOQP*$_9$ promoter in each strain. Promoters are fused to *syfp2* to give plasmids pOPS0980 and pOPS0978 respectively. Fluorescence and OD_{595} measurements were taken after cells were grown under microaerobic conditions (1% O_2). Individual values (Fluo/OD_{595}) are normalised such that the WT average is 100% for both promoters. Complementation had no significant effect on the expression of *fnrN*. Expression of *fixNOQP*$_9$ was significantly increased relative to the mutant but on average restored only 50% of WT activity. Data are averages (±SEM) from at least four biological replicates, ns (not significant) $P \geq 0.05$; *$P < 0.05$; by Student's t test.

We found no significant increase in *fnrN* expression in the complemented strain relative to the double *hfixL* mutant (Figure 3.8A). A significant increase in expression of *fixNOQP*$_9$ was detected in the complemented strain relative to the double *hfixL* mutant (Figure 3.8B). However, this restored only 50% of WT activity. Only partial complementation of the double *hfixL* mutant was therefore attained. It is likely that the combination of P_{lac} with the native *hfixL*$_9$ RBS avoided the toxicity seen in earlier constructs in part by reducing *hfixL*$_9$ production, at the cost of producing insufficient protein to achieve full complementation.

3.3 Discussion

Much of the transcriptional changes undergone by rhizobia during symbiosis are regulated in response to microaerobic conditions. Past studies have shown that this often occurs via an operator motif called an anaerobox, which enables promoters to be induced by the hFixL-FxkR-FixK (or FixLJ-FixK) pathway and/or the FnrN protein, both active under low O_2 conditions. In Rlv3841, anaerobox sequences are present upstream of *fnrN* and two *fixNOQP* operons, amongst many other genes (Figure 3.2 and Table 3.1). Expression of *fnrN*, $fixNOQP_9$ and $fixNOQP_{10}$ was induced under microaerobic conditions in free-living Rlv3841, indicating that these anaerobox sequences are functional (Figure 3.4). Rlv3841 contains the genes for both the hFixL-FxkR-FixK pathway and FnrN (Figure 3.3). Our results show that the hFixL-FxkR-FixK pathway is active at 1% O_2 in free-living Rlv3841 (Figure 3.5-3.7). This agrees with past findings that the FixLJ-FixK pathway is active under relatively aerobic conditions (>1% O_2) [98, 104, 119]. As generally assumed, our results suggest hFixL becomes active at a similar range of O_2 concentrations as FixL.

Some FnrN activity was also detected, but this was far lower than from the hFixL-FxkR-FixK pathway (Figure 3.5-3.7). The O_2 sensitivity of FnrN has not yet been determined, but it is likely that the FnrN protein remains mostly inactive at the 1% O_2 used in our free-living experiments. This suggested that FnrN could play a more important role *in planta*, where O_2 concentrations are far lower and the protein would be more active. Work in the next chapter (Chapter 4) confirmed this theory.

Rlv3841 has multiple homologs of *hfixL*, *fxkR* and *fixK* (Figure 3.2). We found that many of these contribute little if any activity under free-living microaerobic conditions (Figure 3.5-3.7). Further work will be needed to determine whether this is due to changes in protein sequence or gene expression, or both. All three O_2 regulation genes in cluster 3 ($fxkR_c$, $fixK_c$ and $hfixL_c$) appear to be largely dispensable. Unlike the similar cluster 2, cluster 3 has neither an anaerobox nor a K-box. Genes in this cluster are therefore not expected to be induced under microaerobic conditions. Thus, poor expression currently seems the most likely explanation for the lack of activity from genes in cluster 3. However, we have no

reason to believe these genes do not encode functional protein products, and these may induce targets other than those studied in this work.

The hFixL-FxkR-FixK and FnrN O_2 sensing systems of Rlv3841 are studied in more detail under *in planta* conditions in Chapter 4. Several other avenues for free-living work are not explored in this thesis and could form the basis for future work. Cluster 2 has an unusual anaerobox and K-box arrangement that suggests auto-regulation of $fxkR_9$, the main *fxkR* homolog in Rlv3841 (Figure 3.2 and Table 3.1). To determine whether auto-regulation takes place, a reporter fusion of the region upstream of $fxkR_9$ could be built and studied under microaerobic conditions in Rlv3841 WT and a strain with a disrupted hFixL-FxkR-FixK pathway. Multiple versions of the $fxkR_9$ reporter could also be built in which the anaerobox or K-box are removed, to determine their respective impact. Should auto-regulation be confirmed by these studies, it would be of interest to add this mechanism to the model we recently published, to investigate its impact on O_2 regulation in Rlv3841 [2]. The K-box controlled operon likely formed by $fixK_{9a}$-$hfixL_9$ creates another, indirect, auto-regulatory loop in the hFixL-FxkR-FixK pathway that would also be interesting to model.

Cluster 2 also contains $fixK_{9b}$, likely under K-box control. It is not clear what the respective importance of $fixK_{9a}$ and $fixK_{9b}$ is; the individual $fixK_{9b}$ mutant (LMB374, $fixK_{9b}$:pRU877 single recombination) could be used with the reporters produced in this study to investigate this. A reporter fusion for $fixK_{9b}$ could also be built to investigate whether this is indeed regulated by the K-box, despite the intervening $fxkR_c$ TSS. Past microarray work in our group suggests this arrangement may have been preventing microaerobic induction [377].

Broadening the scope of the work, another avenue of interest would be to investigate 'non-traditional' roles for K-box and anaerobox operators. Most studies consider the anaerobox as a unidirectional operator that influences the expression of only one gene or operon, but work by Batut et al. indicated that it could influence transcription both upstream and downstream of its location [182]. The arrangement of cluster 2 in Rlv3841 also suggests that it could act to repress genes

besides *fnrN* (Figure 3.2B). It is possible that the K-box which we have assigned to the $fixK_{9a}$ promoter also partially upregulates $fixNOQP_9$ when bound, and vice versa the anaerobox upstream of $fixNOQP_9$ may upregulate $fixK_{9a}$. This could be studied by building $fixK_{9a}$ and $fixNOQP_9$ reporters in which these operators were deleted. Bi-directional regulation by these operators may be partially responsible for the extensive clustering of O_2 regulated genes in Rlv3841, notably including the O_2 regulation genes themselves. Also of interest would be a whole-genome search for genes in Rlv3841 which contain an anaerobox or K-box downstream of their TSS, indicating a repressing effect. Past studies of hFixL-FxkR-FixK/FixLJ-FixK in rhizobia have typically focused on which genes are induced by these pathways [1]. Both pathways act in the early stages of symbiosis, and it would be of interest to understand not just what genes are induced when symbiosis begins but also which are repressed.

4

Activity of hFixL-FxkR-FixK and FnrN during symbiosis

Contents

4.1 Introduction

During symbiosis, legume plants host rhizobia in dedicated root nodules which provide the optimal environment for symbiotic nitrogen fixation [95, 383]. To prevent deactivation of nitrogenase, nodules tightly regulate the concentration of oxygen in their internal environment [30, 34, 57]. O_2 is captured and shuttled to bacteroids by plant leghaemoglobins [37, 64, 70, 73]. The concentration of remaining

free O_2 in the core nitrogen fixation zone of nodules is as low as 20-50 nM [56, 59]. This is several orders of magnitude lower than the headspace O_2 concentration we used for microaerobic free-living work.

Rhizobia undergo a radical lifestyle change after nodule entry to survive and fix nitrogen in these conditions [25, 35]. In indeterminate nodules, such as those produced by *Pisum sativum* (pea), rhizobia are initially free-living upon entry [96, 384]. They then undergo irreversible lifestyle changes as they move from the nodule tip to its core [85, 86]. Beginning in zone II and accelerating in the II-III interzone, rhizobia terminally differentiate into quasi-organelle bacteroids specialized for nitrogen fixation [385, 386]. Zone III of indeterminate nodules contains differentiated bacteroids which are actively fixing nitrogen [380]. Some studies have investigated where inside the nodule O_2 regulation systems become active, but this remained poorly understood. Soupène et al. reported that in *E. meliloti*, FixLJ-regulated *fixK* was induced at the II-III interzone [98]. Schlüter et al. found that *fixN* in Rlv VF39 was induced in zone III [104]. No work had reported on the spatial localization of FnrN activity.

In Chapter 3, we studied the hFixL-FxkR-FixK and FnrN O_2 regulation systems in free-living Rlv3841 under microaerobic (1% O_2) conditions. We found that both systems were active, but the hFixL-FxkR-FixK pathway was substantially more important than FnrN under those conditions. This was contrary to past work in our group, which showed that FnrN was critical for Rlv3841 nitrogen fixation, whilst strains with a disrupted hFixL-FxkR-FixK pathway retained over 50% of WT activity [362]. We therefore theorized that FnrN was still deactivated by O_2 under these microaerobic conditions but was active during symbiosis.

In this chapter, we studied the role of the hFixL-FxkR-FixK pathway and FnrN inside nodules infected with Rlv3841. Expression of *fnrN* and the $fixNOQP_9$ and $fixNOQP_{10}$ operons was again used as an indirect measure of the activity of these O_2 regulation systems. We found that *fnrN* is induced by both hFixL-FxkR-FixK and FnrN. Both were required to attain full *fnrN* expression. Expression of *fixNOQP* is also reduced when the hFixL-FxkR-FixK pathway is disrupted but relies primarily

on FnrN for induction inside nodules. We measured nitrogenase activity from a range of Rlv3841 O_2 mutants and found that this very closely aligned with their *fixNOQP* expression level, as expected given the importance of this operon for respiration during symbiosis. Finally, we used confocal microscopy to map the activity of these pathways inside nodules. We found that the different O_2 sensitivity of the hFixL and FnrN O_2 sensors produce spatially distinct expression patterns for both *fnrN* and the *fixNOQP* operons. Our results confirm past reports of the importance of both O_2 sensor systems for successful symbiosis, and show that hFixL-FxkR-FixK and FnrN are not separate and redundant systems but form a single, hierarchical regulation cascade [38, 123, 154, 188].

4.2 Results

4.2.1 FnrN and hFixL-FxkR-FixK are both important for expression of anaerobox-controlled genes in Rlv3841 during symbiosis

In free-living Rlv3841, we found that the hFixL-FxkR-FixK system played an important role in the microaerobic (1% O_2) induction of *fnrN* (Figure 3.5) and both *fixNOQP* operons (Figures 3.6 and 3.7). FnrN had little to no impact. This contrasted with past work in our group which showed that FnrN was the more important O_2 regulator in Rlv3841 during symbiosis [362]. FnrN is closely related to the *E. coli* FNR protein which is active only under anaerobic conditions [175, 176, 202]. We therefore speculated that FnrN was more O_2 sensitive than the hFixL-FxkR-FixK pathway, and only active at the nanomolar O_2 concentrations in nodules [54, 56, 59]. Under these conditions, the protein might become a more active inducer of anaerobox-controlled genes than the hFixL-FxkR-FixK pathway, making it the more important regulator. To investigate this possibility, we studied regulation by FnrN and the hFixL-FxkR-FixK pathway *in planta* during symbiosis.

We began by studying the regulation of *fnrN in planta* (Figure 4.1A). The double *hfixL* mutant reduced *fnrN* expression to 28% of WT levels. The hFixL-FxkR-FixK pathway therefore continues to play an important regulatory role *in planta*. However, the *fnrN* mutant similarly reduced *fnrN* expression, to 22% of WT levels. This contrasts sharply with our results in free-living microaerobic conditions, where we found no evidence of auto-activation (Figure 3.5). FnrN auto-activation is thus an important regulatory effect during symbiosis but not under microaerobic conditions.

This supports a model in which FnrN is substantially more O_2 sensitive than hFixL and does not become fully active until Rlv3841 is inside the nodule.

To determine whether any residual FixK production was taking place even in the double *hfixL* mutant, we also studied the activity of a $fixK_{9a}$ reporter produced in the past in our group (Figure 4.2). Expression of *fixK* is driven by the hFixL-FxkR TCS, and no other regulators of the gene have been reported to date in rhizobia [1]. Activity from this promoter was abolished in the double *hfixL* mutant, indicating

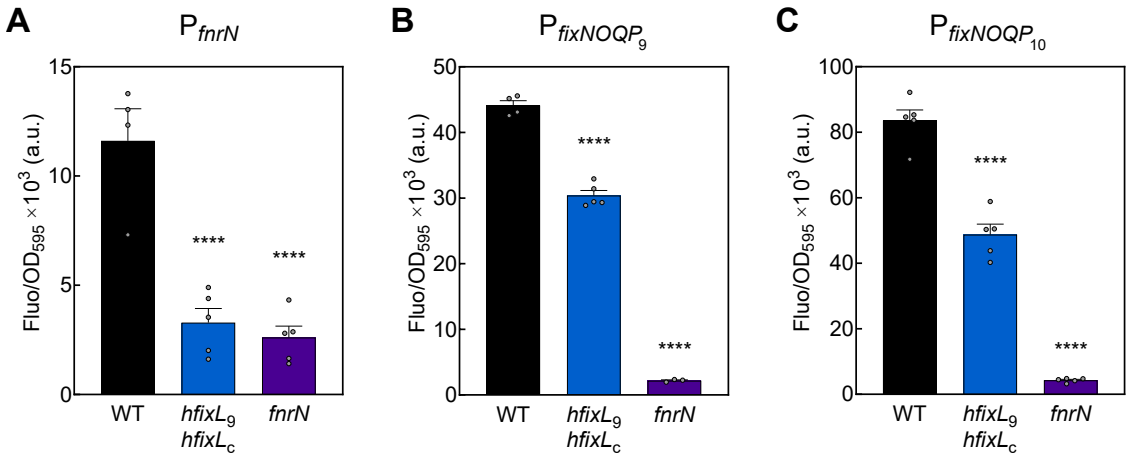


Figure 4.1: *In planta*, *fnrN* is both auto-regulated and controlled by the hFixL-FxkR-FixK pathway, whilst the *fixNOQP* operons are primarily controlled by FnrN. Rlv3841 WT and mutant strains (*fnrN*, LMB648; *hfixL*$_9$ *hfixL*$_c$, LMB496) containing promoter fusions to *syfp2* for *fnrN* (pOPS0980), *fixNOQP*$_9$ (pOPS0978) and *fixNOQP*$_{10}$ (pOPS0977) were inoculated on plants and bacteroids isolated for measurements. Expression in bacteroids of **(A)** *fnrN* is impaired in the *fnrN* background where auto-activation cannot take place. Expression of *fnrN* is similarly impaired in the double *hfixL* mutant, indicating the hFixL-FxkR-FixK pathway also plays an important role in symbiotic *fnrN* induction. Expression of **(B)** *fixNOQP*$_9$ and **(C)** *fixNOQP*$_{10}$ is significantly reduced in the double *hfixL* mutant and almost abolished in the *fnrN* mutant. Thus, both FnrN and the hFixL-FxkR-FixK pathway play an important role in the expression of all three genes. Note scale differences, there was substantially more activity from the *fixNOQP* promoters than from the *fnrN* promoter. Data are averages (±SEM) from at least three plants, ****P $<$ 0.0001; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

complete disruption of the hFixL-FxkR TCS in this background (Figure 4.2). We found no evidence that some activity was retained from the promoter due to other inducing mechanisms. Taken together, our results show that expression of *fnrN* is driven both by auto-activation and the hFixL-FxkR-FixK pathway, and both are required to attain full WT-level expression of *fnrN*.

Next, we investigated the role of FnrN and hFixL in *fixNOQP* expression during symbiosis. We found that nodules formed by the *fnrN* mutant expressed both *fixNOQP* operons at only 5% of WT (Figures 4.1B and 4.1C). The FnrN sensor is thus critical for *fixNOQP* expression inside the nodule. In nodules infected by the double *hfixL* mutant, expression of *fixNOQP*$_9$ and *fixNOQP*$_{10}$ was reduced to 68% and 58% of WT, respectively. *In planta*, *fnrN* is the key regulator of *fixNOQP* expression, with a minor role played by the hFixL-FxkR-FixK pathway.

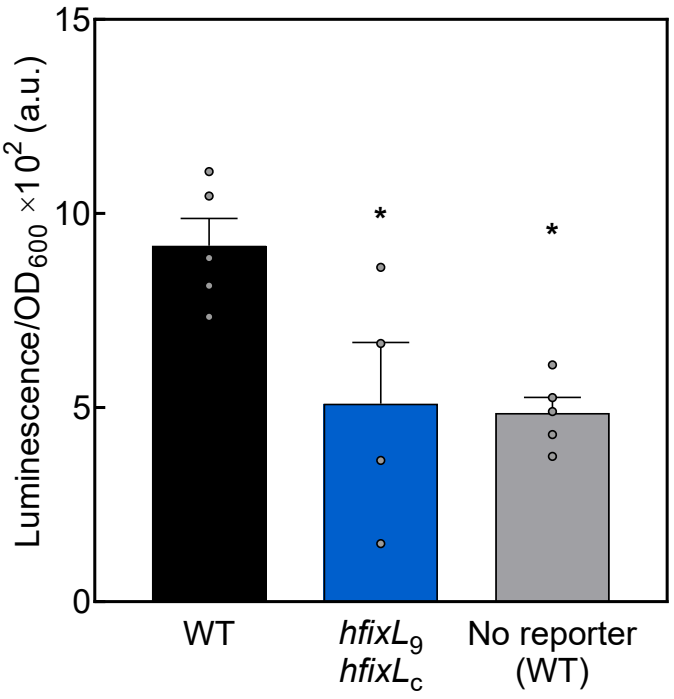


Figure 4.2: The hFixL-FxkR TCS is required for *in planta* $fixK_{9a}$ expression in Rlv3841. A reporter (pOPS0136) was used with the $fixK_{9a}$ promoter fused to the *luxCDABE* operon. The promoter was active in isolated Rlv3841 WT bacteroids (black, OPS0376), but in double *hfixL* mutant bacteroids (blue, OPS0528) no luminescence above background (WT without a reporter, grey) was recorded. Data are averages (±SEM) from at least four plants, *P $<$ 0.05; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

In summary, we found that FnrN is critical during symbiosis for both *fnrN* and *fixNOQP* expression, but the hFixL-FxkR-FixK pathway also plays a significant role.

4.2.2 FnrN is critical for nitrogen fixation in Rlv3841, with some contribution from the hFixL-FxkR-FixK pathway

O_2 regulation is required for successful nitrogen fixation during symbiosis in rhizobia [36–38]. Past work in our group had shown that FnrN was critical for nitrogenase activity in Rlv3841 in symbiosis with *P. sativum* [362]. This aligns well with our finding that FnrN is critical for *fixNOQP* expression, which Rlv3841 appears to rely on for respiration *in planta* [152]. We assessed the impact of disrupting *fnrN* as well as genes of the hFixL-FxkR-FixK pathway on symbiotic nitrogen

fixation. Plants were inoculated with Rlv3841 O_2 regulation mutants, and after 21 days nitrogenase activity was measured using acetylene reduction assays. Nodules were also photographed as defects in symbiosis often alter nodule morphology. Correctly developed nodules containing Rlv3841 WT are elongated (more egg-shaped than spherical) by 21 days after inoculation and have a reddish hue, due to high leghaemoglobin content (Figure 4.3B). As expected, the *fnrN* mutant was critically impaired in nitrogen fixation, reducing acetylene at only 15% of the WT level (Figure 4.3A). Plants inoculated with this mutant produced only small and unelongated nodules that were pale or brown, indicative of poor development and low leghaemoglobin production (Figure 4.3D).

Plants inoculated with either individual *hfixL* mutants or the double mutant were also impaired in nitrogen fixation but retained approximately 75% of WT acetylene reduction activity (Figure 4.3A). No morphological changes were observed in nodules containing the double *hfixL* mutant (Figure 4.3C). The $fxkR_9$ mutant impaired acetylene reduction rates but the decrease was not quite significant (p=0.0584). The $FxkR_c$ homolog is likely at least partially active and sufficiently produced to rescue hFixL regulation in the absence of $FxkR_9$. In the triple *fnrN* $hfixL_9$ $hfixL_c$ mutant (LMB673; *fnrN*::ΩTet $hfixL_9$::ΩSpec $hfixL_c$:pK19) only negligible levels of fixation were recorded. This reinforces the importance of O_2 regulation by both the hFixL-FxkR-FixK pathway and FnrN, suggesting no additional regulators exist which induce anaerobox controlled genes in Rlv3841 during symbiosis. In summary, FnrN is critical for symbiotic nitrogen fixation activity in Rlv3841, but the hFixL-FxkR-FixK pathway is also an important contributor and is required to attain a WT level of fixation.

4.2.3 FnrN complementation

Disruption of *fnrN* in Rlv3841 critically impaired nitrogen fixation activity in Rlv3841 and produced an altered nodule morphology indicative of poor development. We chose to complement this mutant with a cassette containing the *fnrN* ORF under its native promoter. This cassette was cloned into a pOGG280 backbone

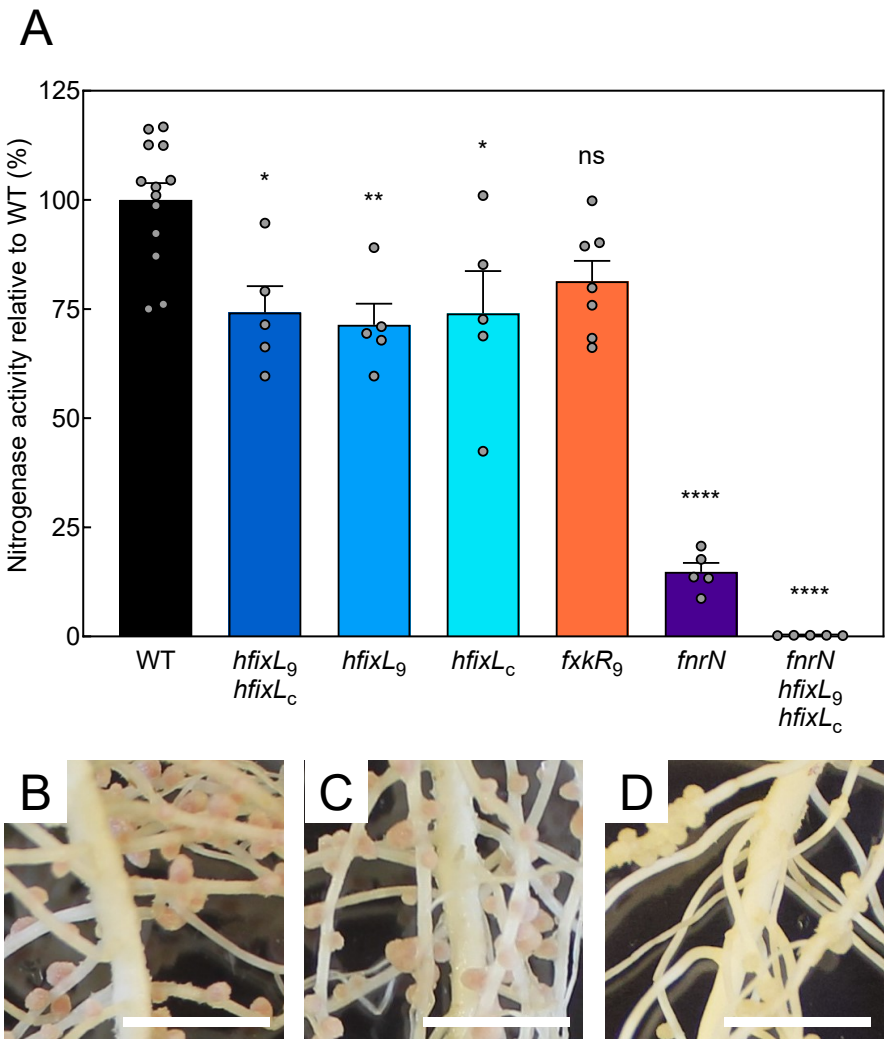


Figure 4.3: Effect of Rlv3841 oxygen regulation mutants on nodule morphology and acetylene reduction rates. (A) Acetylene reduction rates of Rlv3841 mutant strains, normalised by WT activity (5.8 µmoles ethylene plant^{-1} hr^{-1}, 16.8×10^{-3} µmoles ethylene mg^{-1} of nodules hr^{-1}). Knocking out the *hfixL* genes (blue) in combination (*hfixL*$_9$ *hfixL*$_c$, LMB496) and individually (*hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403) only slightly reduced fixation. The single *fxkR*$_9$ mutant (orange, OPS1808) did not significantly reduce fixation (p=0.0584), possibly because of redundancy through the *fxkR*$_c$ homolog. The *fnrN* mutant (purple, LMB648) critically reduced fixation. The mutant lacking both FnrN and hFixL-FxkR-FixK function (*fnrN hfixL*$_9$ *hfixL*$_c$, LMB673) fixed at only a negligible rate. Data for the LMB673 mutant was collected by Graham Hood and is used with permission [362]. Rates are normalised per plant to total mass of nodules. Data are averages (±SEM) from at least five plants, ns (not significant) $P \geq 0.05$; *$P < 0.05$; **$P < 0.01$, ****$P < 0.0001$ by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons. Photos of nodules colonized by **(B)** WT, **(C)** the double *hfixL* knockout and **(D)** the *fnrN* knockout. Scale bar, 1 cm.

and genomically integrated by Tn7 into Rlv3841 with a disrupted *fnrN* gene (LMB648) to produce strain OPS2260. This complementation strategy restored 88% of WT acetylene reduction activity, a level not significantly different from WT (Figure 4.4A). The complemented strain also produced nodules morphologically indistinguishable from WT. By both metrics, complementation was therefore successful with this strategy.

4.2.4 The hFixL-FxkR-FixK pathway is the main O_2 regulation system early in symbiosis, followed by a critical role for FnrN

Legume nodules create a large internal O_2 gradient, with semi-aerobic conditions at their tip and near-anoxic conditions as low as 20 nM O_2 in the central nitrogen fixing zone [96, 387]. This gradient is typically split into four zones (Figure 4.5A) containing different O_2 concentrations and rhizobia in different stages of differentiation [93, 96, 386]. Given the apparent differences in O_2 sensitivity of the hFixL-FxkR-FixK pathway and FnrN, we theorized that they could act at different locations inside the nodule. To study where these systems were active, we used confocal microscopy to map the spatial expression of *fnrN* and *fixNOQP* in nodules.

Expression of *fnrN* in nodules infected with Rlv3841 WT (Figure 4.5B) was visible throughout all nodule zones. This included expression in infection threads in zone I, indicating that low O_2 induction of *fnrN* begins when Rlv3841 first enters the nodule and before the bacteria have differentiated into bacteroids. By contrast, *fnrN* expression in zone I was greatly reduced in nodules infected with the double *hfixL* mutant (Figure 4.5B). This suggests the O_2 concentration in the relatively aerobic environment of zone I is sufficiently low to activate the hFixL-FxkR-FixK pathway. In the absence of this pathway, some *fnrN* expression was retained in zone II and interzone II-III, but this was weaker than WT. Minimal *fnrN* expression was observed in zone III in the *hfixL* double mutant.

In the *fnrN* mutant, expression of *fnrN* appeared to be localized primarily in infection threads, around the entire periphery of the nodule (Figure 4.5B). Nodules

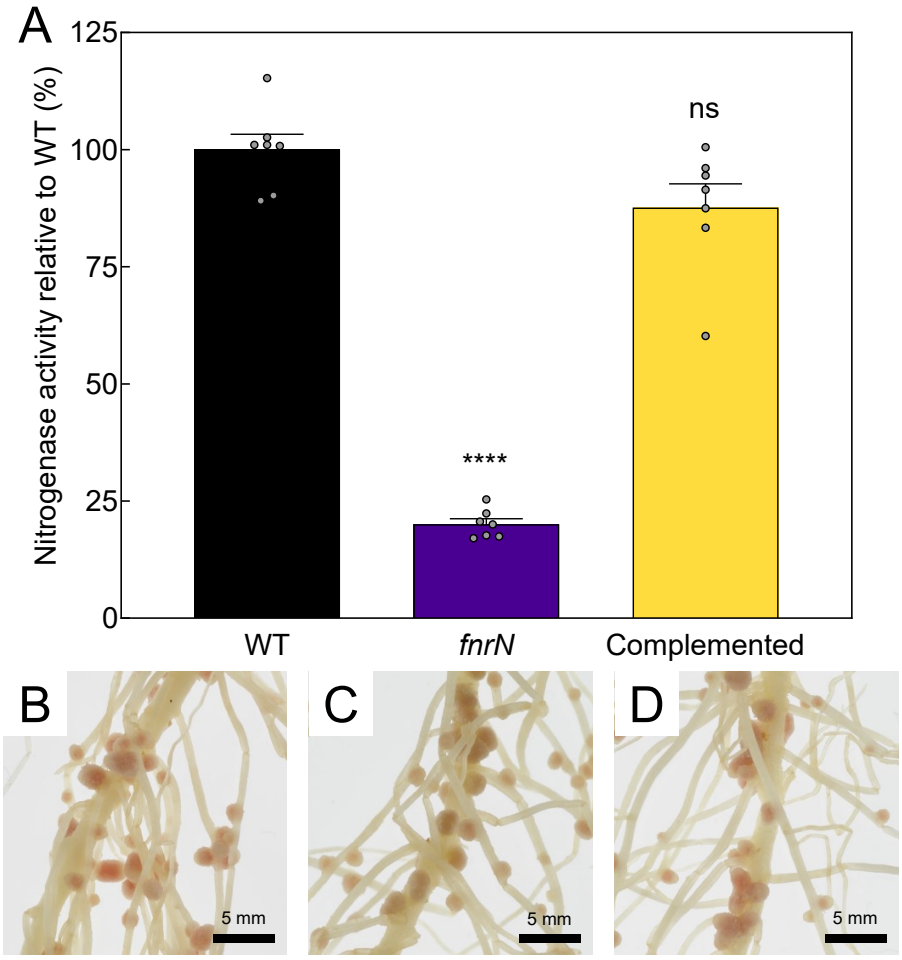


Figure 4.4: Complementation of the Rlv3841 *fnrN* mutant (LMB648). (A) Acetylene reduction rates, normalised by WT activity (6.1 µmoles ethylene plant^{-1} hr^{-1}, 16.0×10^{-3} µmoles ethylene mg^{-1} of nodules hr^{-1}). The activity of the *fnrN* mutant (purple, LMB648) was 20% of Rlv3841 WT. The complemented strain (yellow, OPS2260) fixed at a rate not significantly different from WT. Nodules colonized by Rlv3841 **(B)** WT, **(C)** the *fnrN* mutant and **(D)** the complemented *fnrN* mutant. WT nodules are elongated and have a red hue, indicating significant leghaemoglobin content. Nodules formed by the *fnrN* mutant are spherical and brown, indicating poor development and little to no leghaemoglobin production. Complementation restored WT nodule morphology. Data are averages (±SEM) from seven plants, ns (not significant) $P \geq 0.05$; ****$P < 0.0001$; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

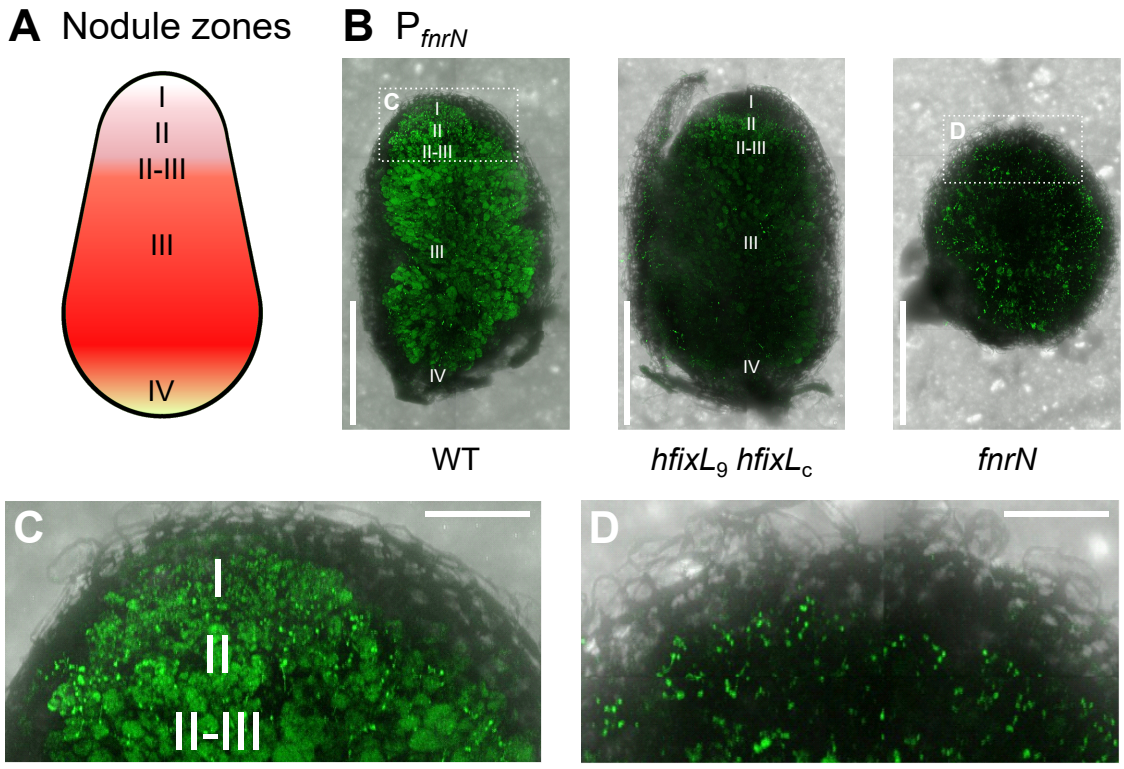


Figure 4.5: Spatial expression pattern of *fnrN* in nodules infected with Rlv3841 WT and mutants. (A) Schematic representation of an indeterminate nodule formed by *P. sativum*. Zone I contains undifferentiated rhizobia in infection threads. Rhizobia enter plant cells in zone II and undergo substantial differentiation towards becoming bacteroids in the II-III interzone. Zone III is the main nitrogen fixing zone. Zone IV contains bacteroids which are beginning to senesce. **(B)** Nodule cross sections showing expression of *fnrN* when inoculated with strains of Rlv3841 (Tn7 integrated *syfp2* promoter fusion: WT, OPS2429; $hfixL_9$ $hfixL_c$ double mutant, OPS2435; *fnrN* mutant, OPS2432). Expression begins immediately in zone I in nodules inoculated with WT; see C for a close-up of the region highlighted in white. A similar level of expression is present across all zones. When inoculated with the double *hfixL* mutant, expression began in zone II and was highest in this zone. In nodules inoculated with the *fnrN* mutant, expression was observed in infection threads around the periphery of the nodule; see D for a close up of the region highlighted in white. This mutant does not form mature nodules, and the normal zones are therefore unlikely to be fully developed. **(C)** Magnified view of *fnrN* expression in WT bacteria in zone I of the nodule. **(D)** Magnified view of *fnrN* expression in the infection threads of a nodule inoculated with the *fnrN* mutant. Scale bar; 1 mm **(B)**, 0.25 mm (C and D). All images were captured and processed using identical parameters; see Chapter 2 for details.

infected by this mutant were severely impaired in their development, failed to elongate and contained little to no leghaemoglobin (Figure 4.3D).

Free O_2 concentration is unlikely to drop as much in these nodules as it does in fully developed nodules. It is therefore noteworthy that the hFixL-FxkR-FixK

pathway is nevertheless active, suggesting even poorly developed nodules produce a sufficiently low O_2 concentration to activate the pathway.

Expression patterns of *fixNOQP*$_9$ (Figure 4.6A) and *fixNOQP*$_{10}$ (Figure 4.6B) were similar. In Rlv3841 WT, expression of both started abruptly in the II-III interzone of nodules, in agreement with past studies [98, 104, 388, 389]. This abrupt start was absent in nodules infected with the double *hfixL* mutant, indicating it requires the hFixL-FxkR-FixK pathway (Figure 4.6C, D). Without hFixL-FxkR-FixK, expression of *fixNOQP*$_9$ and *fixNOQP*$_{10}$ started gradually after the II-III interzone, presumably driven by FnrN. Expression was also weaker than in the WT. In the *fnrN* mutant, we observed minimal expression of *fixNOQP*$_9$. This may be due in part to the poor development of these nodules, but the expression of *fnrN* in the *fnrN* mutant (Figure 4.5B) indicates that the hFixL-FxkR-FixK pathway is still relatively active in these underdeveloped nodules. The lack of *fixNOQP* expression in the *fnrN* mutant therefore indicates that hFixL-FxkR-FixK cannot directly induce much *fixNOQP* expression in zone III of mature nodules. Although FnrN is the main driver of *fixNOQP* expression, the hFixL-FxkR-FixK pathway is required for full *fnrN* expression and also plays an important role in *fixNOQP* expression, albeit indirectly.

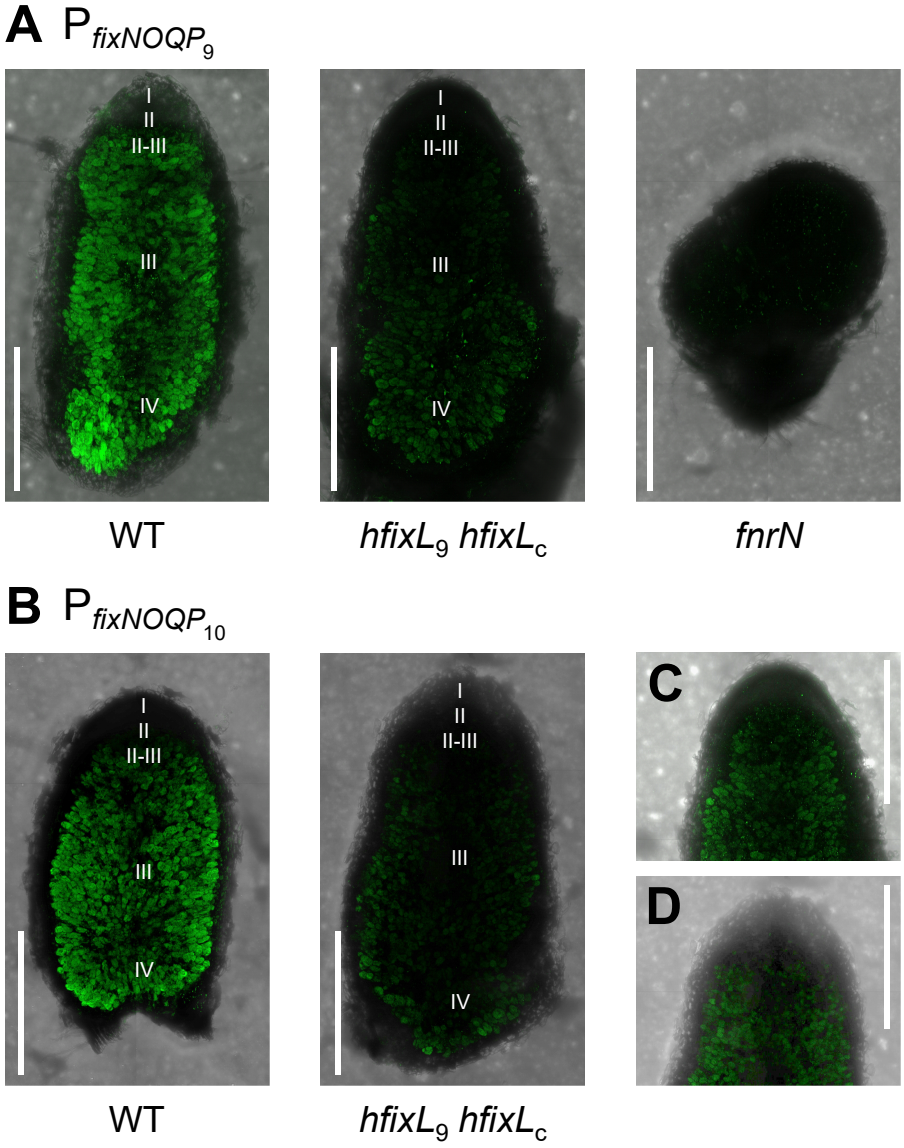


Figure 4.6: Spatial expression pattern of the *fixNOQP* operons in nodules infected with Rlv3841 WT and mutants. (A) Expression of *fixNOQP*$_9$ in strains of Rlv3841 (Tn7 integrated *syfp2* promoter fusion: WT, OPS2428; *hfixL*$_9$ *hfixL*$_c$ double mutant, OPS2434; *fnrN* mutant, OPS2431). In nodules inoculated with WT, expression starts abruptly at the II-III interzone. In the double *hfixL* mutant, expression is reduced and begins gradually and at a point more proximal to the root. Almost no expression is found in *fnrN* mutant nodules. **(B)** Expression of *fixNOQP*$_{10}$ in Rlv3841 WT and the double *hfixL* mutant followed a similar pattern as *fixNOQP*$_9$ (pJP2 reporter plasmid *syfp2* promoter fusion: WT, OPS2468; *hfixL*$_9$ *hfixL*$_c$, OPS2469). Expression begins at the II-III interzone. In the double *hfixL* mutant, expression again begins at a point more proximal to the root, in zone III of the nodule, and is reduced. **(C)** and **(D)** are areas of the *hfixL* double mutant reporter images (for *fixNOQP*$_9$ and *fixNOQP*$_{10}$ respectively) with their brightness and contrast altered to better display the distribution of fluorescence. Images within sets A and B were captured and processed using identical parameters; see Chapter 2 for details. Fluorescence intensity across the A and B image sets, and across C and D, should not be compared as intensity was normalised.

4.3 Discussion

Legumes create a near-anoxic environment in their nodules, with O_2 concentrations as low as 20-50 nM [56, 59]. As well as enabling nitrogenase activity, these conditions demand that rhizobia, which are aerobes, undergo a drastic lifestyle transition to survive symbiosis. In Rlv3841, we found two O_2 sensing systems with previously reported roles in regulating this lifestyle transition: the hFixL-FxkR-FixK pathway and the FnrN protein [1]. Our initial microaerobic work *in vitro* showed that the hFixL-FxkR-FixK pathway is active and regulates both *fnrN* and two *fixNOQP* operons (see Chapter 3). In these non-symbiotic conditions, *fnrN* had little impact. FnrN is closely related to the *E. coli* FNR protein which is active only under anoxic conditions [177, 230, 375]. We therefore reasoned that FnrN activity was more intolerant of O_2 than the hFixL sensor and thus was still largely inactivated at the 1% O_2 headspace concentration we used for our work *ex planta*. Given the near-anoxic conditions reported in nodules, we theorized FnrN would become more active in this context.

FnrN and the hFixL-FxkR-FixK pathway both regulate anaerobox-controlled genes. The two systems could therefore act redundantly, with disruption in one partially or fully rescued by activity from the remaining regulator. This has been reported in other rhizobia and was the prevailing paradigm through which these systems were viewed [373, 374]. However, there were also reports that O_2 regulation systems could act hierarchically, with the less O_2 sensitive system inducing expression of the more sensitive system [188, 204]. Our results show that hFixL-FxkR-FixK and FnrN operate in such a hierarchical arrangement in Rlv3841. Expression of *fnrN in planta* was significantly decreased when *fnrN* was disrupted (Figure 4.1A). This agrees with the dual-anaerobox arrangement in the *fnrN* promoter region of Rlv3841, reported to enable auto-regulation in other rhizobia [131, 152, 205]. However, disrupting the hFixL-FxkR-FixK pathway reduced *fnrN* expression to a near-identical degree. Thus, although *fnrN* is auto-regulated, full expression *in planta* also requires hFixL-FxkR-FixK, putting this pathway upstream of FnrN in a hierarchical arrangement.

Whilst both O_2 regulation systems had a similar impact on *fnrN* expression, FnrN was substantially more critical for *fixNOQP* expression than the hFixL-FxkR-FixK pathway. Our work in free-living Rlv3841 showed that the pathway can induce $fixNOQP_9$ and $fixNOQP_{10}$ (see Chapter 3), but during symbiosis FnrN is the major inducer of these operons. This was further supported by our study of nitrogenase activity in O_2 regulation mutants of Rlv3841 (Figure 4.3). Disruption of *fnrN* critically reduced nitrogenase activity, as expected given the poor expression of *fixNOQP* in this mutant (Figure 4.1). The hFixL-FxkR-FixK pathway also reduced nitrogen fixation, likely due to poor *fnrN* expression which indirectly reduces *fixNOQP* expression.

Our confocal work illustrated how the different O_2 sensitivity and hierarchical arrangement of hFixL-FxkR-FixK and FnrN is spatially translated inside the nodule by its internal O_2 gradient. The less O_2 sensitive hFixL-FxkR-FixK pathway was active as early as the infection threads of zone I, where it induced *fnrN*. Expression of the *fixNOQP* operons began only in the II-III interzone, suggesting this is where FnrN becomes fully active.

Taken together with our work on free-living Rlv3841 in microaerobic conditions (see Chapter 3), our results *in planta* show that the hFixL-FxkR-FixK pathway and FnrN form a combined, hierarchical O_2 regulation cascade. In the earliest stages of symbiosis, the hFixL-FxkR-FixK pathway becomes active and induces transcription of *fnrN*. FnrN also auto-activates. This is followed by FnrN-controlled induction of the *fixNOQP* operons, required for respiration in the near-anoxic conditions of the nodule core. Both parts of the combined hFixL-FxkR-FixK cascade are required for full expression of *fnrN* and *fixNOQP*, and hence to attain WT nitrogen fixation activity during symbiosis. A full map of the confirmed and possible regulatory connections present between the pathways is given in Figure 4.7.

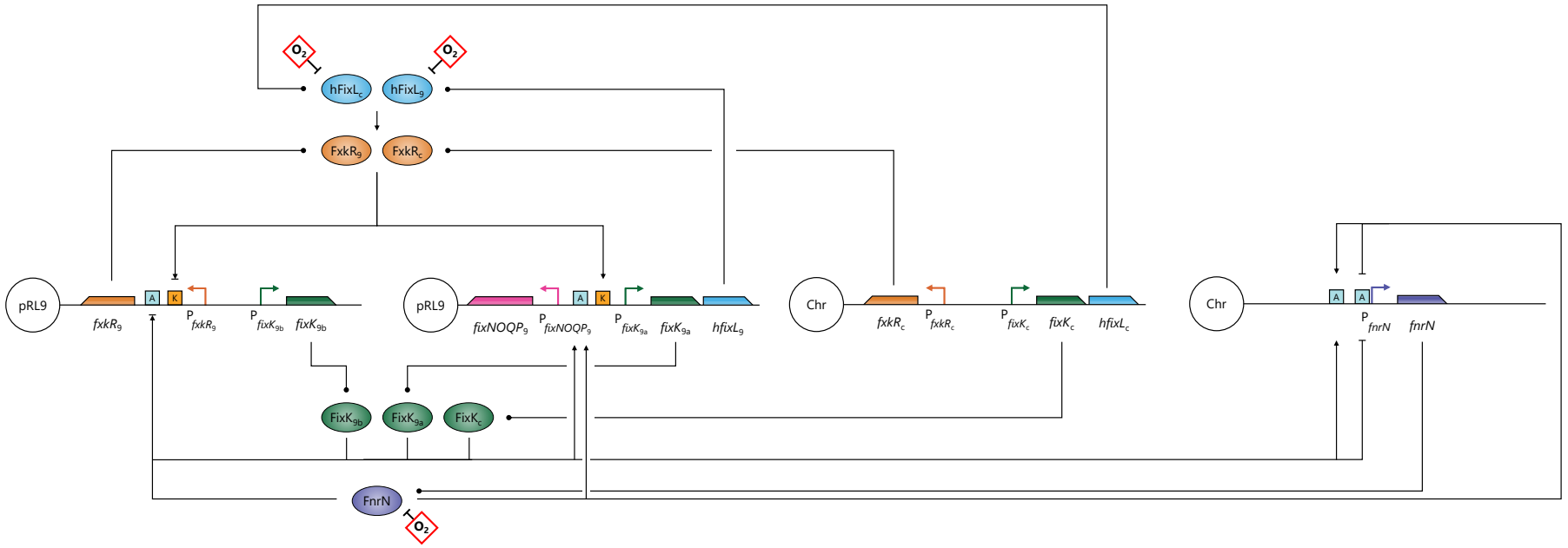


Figure 4.7: Full map of known and potential connections between the hFixL-FxkR-FixK and FnrN O_2 regulation systems in Rlv3841. Oxygen is shown in red diamonds. Proteins are shown as ovals, operator sites as squares and genes as pointed rectangles. TSS are shown as right-angled arrows. Line endings indicate activation (arrows), inhibition (blunt end) and translation (circle). Where a connection may cause both activation and inhibition, an arrow pointing to a blunt end is used.

5

Repurposing NifA as a tool to engineer nitrogen fixation

Contents

5.1 Introduction

All symbiotic rhizobia regulate the activation of nitrogen fixation using the O_2-sensing NifA transcription factor. Its central role in inducing the expression of nitrogen fixation genes makes it an attractive engineering target [18]. Artificial

control of this single protein can be sufficient to activate all the machinery required for nitrogen fixation in some species [328, 336]. However, although there has been substantial work on the regulation and activity of NifA proteins in free-living diazotrophs, relatively little work has investigated rhizobial NifA proteins [1]. Studies have consistently shown that NifA activity is highly cross-compatible; NifA from one organism can function and even enable nitrogen fixation in other diazotrophic species [243–247]. Because of this cross-compatibility, one strategy put forward to engineer nitrogen fixation is to bypass native regulation in a diazotroph by expressing *nifA* from another species.

A NifA variant of particular interest is *A. caulinodans* ORS571 NifA ($NifA_{ORS571}$). *A. caulinodans* can perform free-living nitrogen fixation, which suggests $NifA_{ORS571}$ is active at intracellular O_2 concentrations attainable under free-living conditions [239, 278]. This would make the protein an attractive regulator, potentially able to activate nitrogen fixation in rhizobia such as Rlv3841 which normally fix only under symbiotic conditions. This strategy remains largely untested. It could also yield important biological insights about native mechanisms controlling rhizobial NifA proteins. Much about these proteins remains unknown, particularly with regards to their protein-level regulation. Rhizobial NifA proteins are known to sense O_2 via an inter-domain linker (IDL), but the O_2 concentration at which this regulation operates has not been determined [244, 280]. When we began this work, it was generally assumed that rhizobial NifA proteins, including the NifA protein of Rlv3841 ($NifA_{Rlv3841}$), would only be active at the nanomolar O_2 concentrations found in nodules. Also largely unexplored is the function of the N-terminal GAF domain found in most rhizobial NifA proteins [1]. Studies of the NifA GAF domain in free-living diazotrophs such as *K. pneumoniae* have consistently shown that it serves a regulatory role [252, 263–265]. However, GAF domains are poorly conserved and the signal they respond to can vary drastically between species [261, 262]. GAF domains in rhizobial NifA proteins are believed to regulate protein activity, but have received little attention. It was therefore unknown what the function of the $NifA_{Rlv3841}$ GAF domain is. Transcription of

rhizobial *nifA* is better understood; it is typically auto-regulated, and in some cases partially induced under microaerobic conditions [41, 234]. Work in Rlv UPM791 indicated that *nifA* was only expressed under symbiotic conditions, and this is generally assumed to hold for all symbiotic rhizobia [293]. One notable exception is *B. japonicum*, which expresses *nifA* even under free-living conditions [295, 296]. Because of this strict transcriptional regulation, coupled with protein-level O_2 sensing, our initial assumption was that $NifA_{Rlv3841}$ would likely not be active under free-living conditions, even if transcriptional regulation was bypassed.

Only one other limitation is currently known to prevent Rlv3841 from fixing nitrogen under free-living conditions; the lack of a *nifV* gene [336]. This encodes a homocitrate synthase, a cofactor required for nitrogenase activity [390, 391]. Engineering both NifV and NifA activity into free-living Rlv3841 could therefore potentially enable nitrogen fixation outside of symbiosis.

This chapter begins by studying the activity of two $NifA_{Rlv3841}$ variants: the full-length protein and $NifA_{Rlv3841}$ with its GAF domain truncated ($NifA_{Rlv3841}$ ΔGAF). Native transcriptional regulation is bypassed using inducible expression vectors, thereby isolating regulation to protein-level mechanisms. Our aim was to determine whether any activity could be detected outside of symbiosis, and what if any effect the deletion of the GAF domain had. We also studied $NifA_{ORS571}$, to determine if this is functional in Rlv3841 and able to induce transcription from native NifA-regulated promoters in Rlv3841. Some activity was detectable from all three proteins. $NifA_{ORS571}$ consistently produced the highest level of activity, and removal of the GAF domain consistently produced less activity than $NifA_{Rlv3841}$. We observed no $NifA_{ORS571}$ activity in its native host, potentially due to native mechanisms which were not bypassed. We compared the activity of all three NifA variants under a variety of fixed nitrogen and O_2 concentrations to determine their sensitivity to these signals. Both $NifA_{Rlv3841}$ and $NifA_{ORS571}$ showed signs of deactivation in high fixed nitrogen concentrations, but more work is needed to confirm this effect was due to protein-level regulation. Minimal response to O_2 concentration was observed, likely because the lowest O_2 concentration tested, 1%, still inactivated all three

NifA variants. After establishing free-living NifA activity in Rlv3841, we engineered NifV activity into the organism. By combining NifV and NifA activity in Rlv3841, we obtained detectable levels of free-living nitrogen fixation.

5.2 Results

5.2.1 A dual-plasmid system to study NifA activity

To facilitate the combinatorial testing of multiple NifA variants with multiple reporter promoters, we created a two-plasmid system (Figure 5.1). The first plasmid, the reporter, hosts *gfp* under the control of a NifA-activated promoter; GFP fluorescence thus signals NifA activity. We selected the native Rlv3841 P_{fixA} and P_{nifH} promoters for use in Rlv3841, as these are commonly used and known to be highly upregulated by NifA [377, 381]. The pME6041 backbone was chosen for these reporter plasmids as it has a pVS1 origin of replication that is compatible with a wide range of other plasmids [355]. The second plasmid, the *nifA* vector, hosts a *nifA* variant under control of an inducible promoter. Inducible control was used as the substantial changes in gene expression caused by the protein could significantly impair cell growth. By placing *nifA* under the control of an inducible promoter, activity could be restrained outside of experimental work. Three *nifA* variants were selected for study: full-length $nifA_{\text{Rlv3841}}$, $nifA_{\text{Rlv3841}}$ from which the GAF domain had been removed ($nifA_{\text{Rlv3841}}$ ΔGAF, Δ1-141 amino acids inclusive) and full-length $nifA_{\text{ORS571}}$.

5.2.2 Characterizing the behaviour of the pLMB51 backbone

Plasmid pLMB51 was initially selected for use as the backbone for the *nifA* expression vector. It is a taurine-inducible low-copy number plasmid which had been reported to work well in Rlv3841 [356]. We began by confirming the functionality of taurine induction from the pLMB51 backbone in Rlv3841 and *A. caulinodans* by using the pLMB51-derived pLMB134 vector. This contains *gfp-mut3* under control of the TauR-regulated *tauA* (P_{tauA}) promoter, inducible by taurine. TauR is encoded in the pLMB51 backbone. We sought to determine the dynamic range and appropriate taurine concentrations to use in both organisms, and study whether the presence of taurine caused any growth defect.

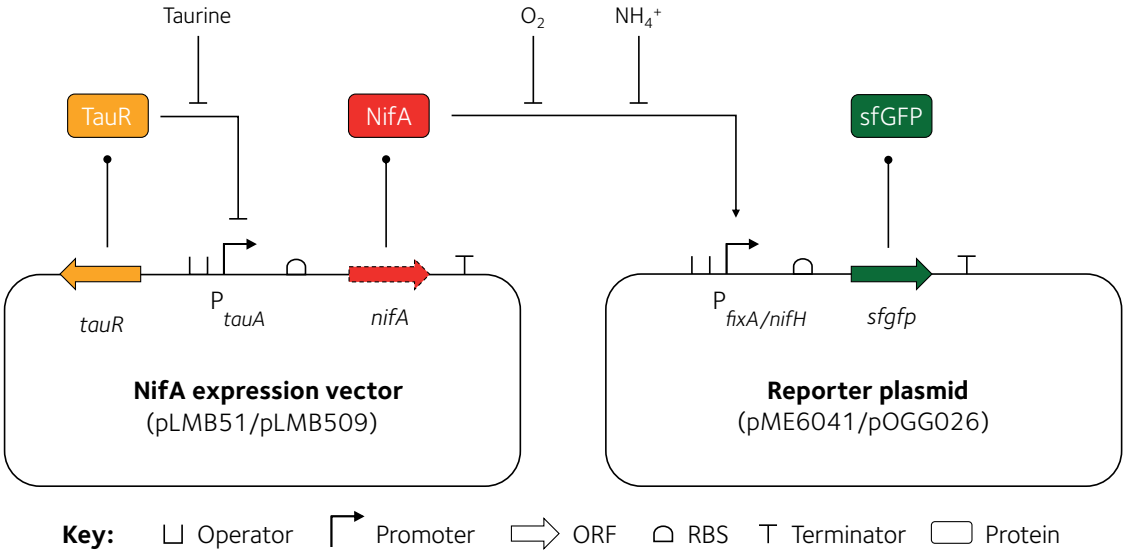


Figure 5.1: Dual plasmid system used to characterize NifA variants. Two plasmids were used to facilitate the combinatorial testing of multiple NifA variants with multiple reporter promoters. The first plasmid was used to express *nifA* variants (red, dashed outline) under the taurine-inducible P_{tauA} promoter, controlled by TauR. The second plasmid contained the NifA-controlled reporter promoter, either P_{fixA} or P_{nifH}, driving expression of the *sfgfp* fluorescent protein. Lines indicate interactions: blunt ends show deactivation/repression, arrows show induction and circular ends show translation. Symbols based on the SBOL visual v2 standard [392, 393], see key for details. ORF, open reading frame; RBS, ribosome binding site.

Behaviour of the pLMB51 backbone in Rlv3841

In Rlv3841, the pLMB51 backbone behaved broadly as described by its original creators (Figure 5.2) [356]. No significant increase in fluorescence was detected in strains containing the uninduced plasmid, indicating minimal leaky expression (Figure 5.2A). The taurine dose-response curve largely matched that previously reported, with 0.5 mM taurine sufficient for induction and a response that began to plateau past 1 mM taurine (Figure 5.2B). Increasing taurine concentration to 10 mM improved fluorescence, but only an incremental further increase was observed when 20 mM taurine was used. We therefore opted to used 10 mM taurine for subsequent induction experiments, which produced an 8.5-fold increase in fluorescence relative to the uninduced plasmid. Taurine toxicity has been reported, so the growth of Rlv3841 with the plasmid at a variety of taurine concentrations was studied (Figure 5.2C). We saw no impact on growth at any taurine concentration tested.

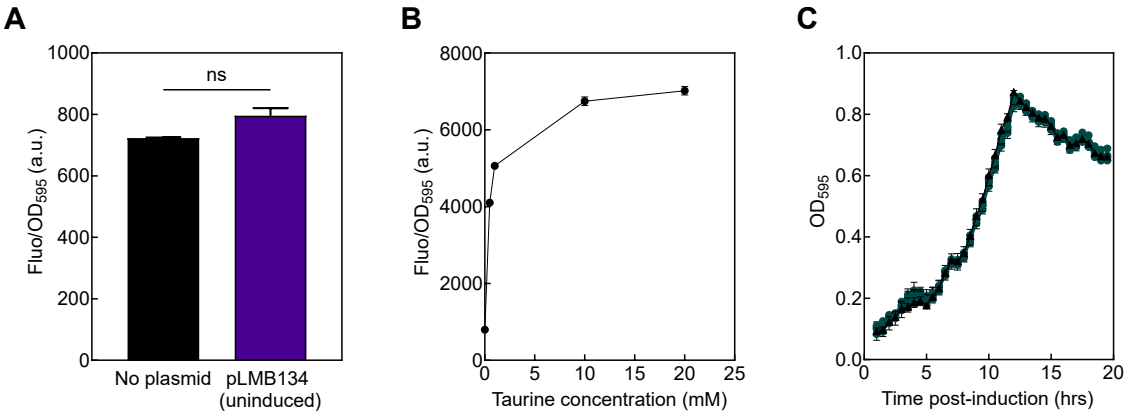


Figure 5.2: Characterization of taurine induction from the pLMB51 backbone in Rlv3841. Testing was carried out using plasmid pLMB134, a pLMB51 backbone into which the fluorescent reporter *gfp-mut3* was inserted under control of the P_{tauA} promoter. **(A)** Comparison of Fluo/OD_{595} in Rlv3841 without any plasmid (black) and with the uninduced pLMB134 plasmid (purple). The plasmid produced no significant increase in fluorescence in the absence of taurine. **(B)** Fluo/OD_{595} produced by pLMB134 in Rlv3481 at a range of taurine concentrations. The lowest taurine concentration tested, 0.5 mM, was enough for induction. Induction began to plateau after 1 mM taurine. **(C)** Growth of Rlv3841 without pLMB134 under the range of taurine concentrations tested. No difference in growth was observed between cultures growing without taurine (black triangles) and those growing with 0.5-20 mM taurine (teal circles). Note the close overlap of these data sets occludes in large part the data points from taurine-induced cultures. Experiments grown in a 3% O_2 headspace. Data are averages from n = 2 biological replicates; error bars represent SEM. Statistical analysis by Student's t test; ns (not significant) $P \geq 0.05$.

Behaviour of the pLMB51 backbone in *A. caulinodans*

We also studied the behaviour of pLMB134 in *A. caulinodans*. No past results were available for the behaviour of the pLMB51 backbone in this host. We found that the plasmid was functional but differed notably in several key induction characteristics when compared to its behaviour in Rlv3841 (Figure 5.3). A significant increase in fluorescence was observed when the uninduced plasmid was inserted, indicating leaky expression (Figure 5.3A). However, more taurine was required for induction in *A. caulinodans*, with concentrations below 10 mM producing a minimal increase in fluorescence (Figure 5.3B). This may indicate worse taurine uptake in *A. caulinodans* than Rlv3841. Clear induction was however recorded at 10 mM and increasing concentration to 20 mM taurine further improved fluorescence. Despite higher leakiness, the induction fold of pLMB134 in *A. caulinodans* was better than in

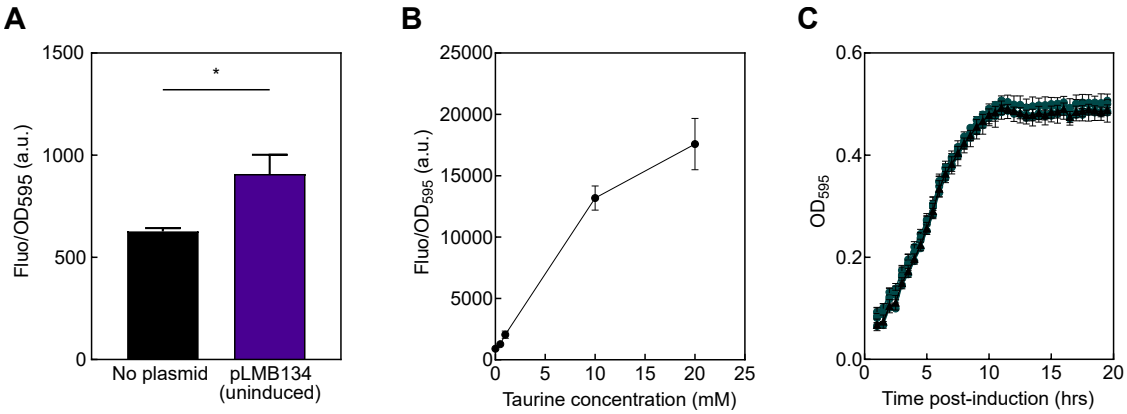


Figure 5.3: Characterization of taurine induction from the pLMB51 backbone in *A. caulinodans*. Testing was carried out using the pLMB51-derived pLMB134 plasmid, see text for details. **(A)** Comparison of Fluo/OD_{595} in *A. caulinodans* without any plasmid (black) and with the uninduced pLMB134 plasmid (purple). The plasmid produced a slight but significant increase in fluorescence even in the absence of taurine, indicating leaky expression. **(B)** Fluo/OD_{595} produced by pLMB134 in *A. caulinodans* at a range of taurine concentrations. Minimal changes in fluorescence were observed at 1 mM taurine, but induction was observed at 10 and 20 mM. **(C)** Growth of *A. caulinodans*, without pLMB134, under the range of taurine concentrations tested. No difference in growth was observed between cultures growing without taurine (black triangles) and those growing with 0.5-20 mM taurine (teal circles). Note the close overlap of these data sets occludes in large part the data points from taurine-induced cultures. Experiments grown in a 3% O_2 headspace atmosphere. Data are averages from at least n = 2 biological replicates; error bars represent SEM. Statistical analysis by Student's t test; *P < 0.05.

Rlv3841, reaching 14.5-fold at 10 mM. As in Rlv3841, the presence of taurine did not inhibit the growth of *A. caulinodans* (Figure 5.3C).

5.2.3 Activity of $NifA_{Rlv3841}$ variants expressed from pLMB51 in Rlv3841

Having established the functionality of taurine induction via the pLMB51 backbone in Rlv3841, the backbone was used to build plasmids with taurine-inducible full-length and ΔGAF truncated $nifA_{Rlv3841}$. These constructs were inserted into an Rlv3841 *nifA* mutant (Δ*nifA*, OPS1737) in combination with a reporter vector (P_{nifH}, pOPS1178) to study $NifA_{Rlv3841}$ activity (Figure 5.4). Cells were grown in UMS media (see Chapter 2) with 10 mM ammonium chloride and a headspace O_2 concentration of 3%. This O_2 concentration was used as it had been reported to be optimal for $NifA_{ORS571}$ [185]. We found that the plasmid containing full-length

nifA$_{\text{Rlv3841}}$ (pOPS1009) increased fluorescence relative to the reporter-only strain even in the absence of taurine (Figure 5.4A). This suggests some leaky expression of *nifA*$_{\text{Rlv3841}}$ was occurring. Adding taurine did not further increase induction. Reporter activation therefore appears unaffected by *nifA*$_{\text{Rlv3841}}$ induction in this system, which may be partly due to the low induction fold of the pLMB51 backbone in Rlv3841. No increase in fluorescence over background levels was observed from the plasmid containing truncated *nifA*$_{\text{Rlv3841}}$ ΔGAF (pOPS1104); we found no evidence of activity from this variant (Figure 5.4B).

The reporter vector on its own appeared to slightly impair the growth of Rlv3841 Δ*nifA* (Figures 5.5A and 5.5B), but there was no significant change in mean generation time or maximum OD_{595} reached (Figures 5.5C and 5.5D). Adding either *nifA* vector in combination with the reporter plasmid further impaired growth (Figures 5.5A and 5.5B), resulting in a significant drop in the maximum OD_{595} (Figure 5.5D). Mean generation time remained unaffected in strains containing both the reporter and *nifA* vector (Figure 5.5C). With both *nifA* vectors, taurine induction did not further impair growth (data not shown). Expression of both *nifA* variants may be slightly impairing growth, but the lack of any taurine response also suggests the vectors themselves could be responsible for part of the growth defect.

In many species, the activity of NifA is regulated in response to fixed nitrogen concentration. We therefore studied the behaviour of the dual-plasmid system at lower ammonium chloride concentrations to see whether this influenced NifA activity. Full-length NifA$_{\text{Rlv3841}}$ showed no taurine-response at 10 mM ammonium chloride, but significant induction was observed at 0.5 mM ammonium chloride and in its absence (Figures 5.6A and 5.6C). Activity thus increased as the ammonium chloride concentration dropped, suggesting the activity of NifA$_{\text{Rlv3841}}$ was directly or indirectly regulated by the presence of fixed nitrogen. In line with our earlier results, we measured no response to taurine induction from the truncated NifA$_{\text{Rlv3841}}$ ΔGAF variant at any of the three ammonium chloride concentrations tested (Figures 5.6B and 5.6C), further suggesting the GAF domain truncation inactivates the protein.

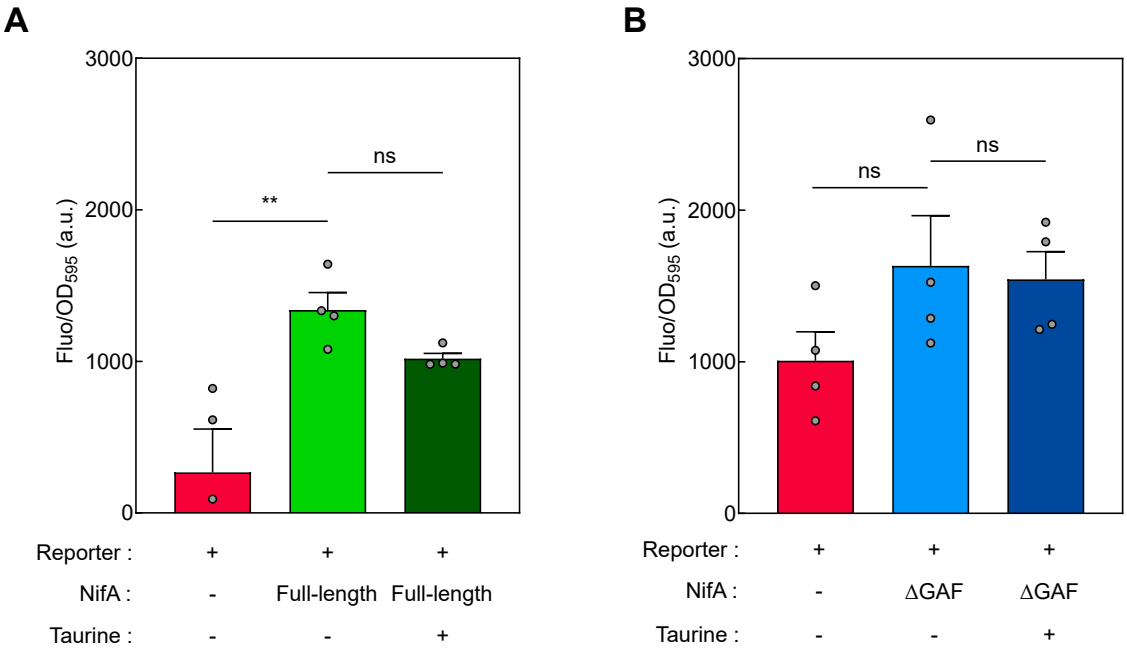


Figure 5.4: Activity of full-length $NifA_{Rlv3841}$ and $NifA_{Rlv3841}$ ΔGAF in Rlv3841. Fluorescence from the Rlv3841 P_{nifH} reporter plasmid (pOPS1178) in strains derived from Rlv3841 Δ*nifA* (OPS1737). Background fluorescence from Rlv3841 Δ*nifA* containing no plasmids has been subtracted from all values. **(A)** In red, fluorescence produced by the strain (OPS2080) carrying only the P_{nifH} reporter. In light green, significantly more fluorescence was produced by the strain (OPS2081) carrying the reporter and $nifA_{Rlv3841}$ in the pLMB51 backbone (pOPS1009). In dark green, induction with taurine did not significantly increase fluorescence. **(B)** In red, fluorescence produced by the strain carrying only the reporter. In light blue, no significant increase in fluorescence was observed in the strain (OPS2082) carrying the reporter and $nifA_{Rlv3841}$ ΔGAF in the pLMB51 backbone (pOPS1104). In dark blue, induction with taurine did not significantly increase fluorescence. Note data for (A) and (B) were carried out using different plate reader gains and $Fluo/OD_{595}$ values are not comparable. Experiments grown in a 3% O_2 headspace. Data are averages from n = 4 biological replicates; error bars represent SEM. Note one data point for the reporter-only strain (red) is below background for (A). Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; $^{**}P < 0.01$.

Studies have shown that rhizobial NifA proteins are repressed by high O_2 concentration through a redox-sensing cysteine-rich cluster in their IDL region [244, 276, 280]. However, the O_2 concentrations at which rhizobial NifA proteins are active remains mostly unknown. To determine whether a lower O_2 concentration might improve activity, we reduced the headspace O_2 concentration from 3% to 1% O_2. Some activity was still observed (Figure 5.7) but combining the burden of the dual-plasmid system with that imposed by lower O_2 concentration led to

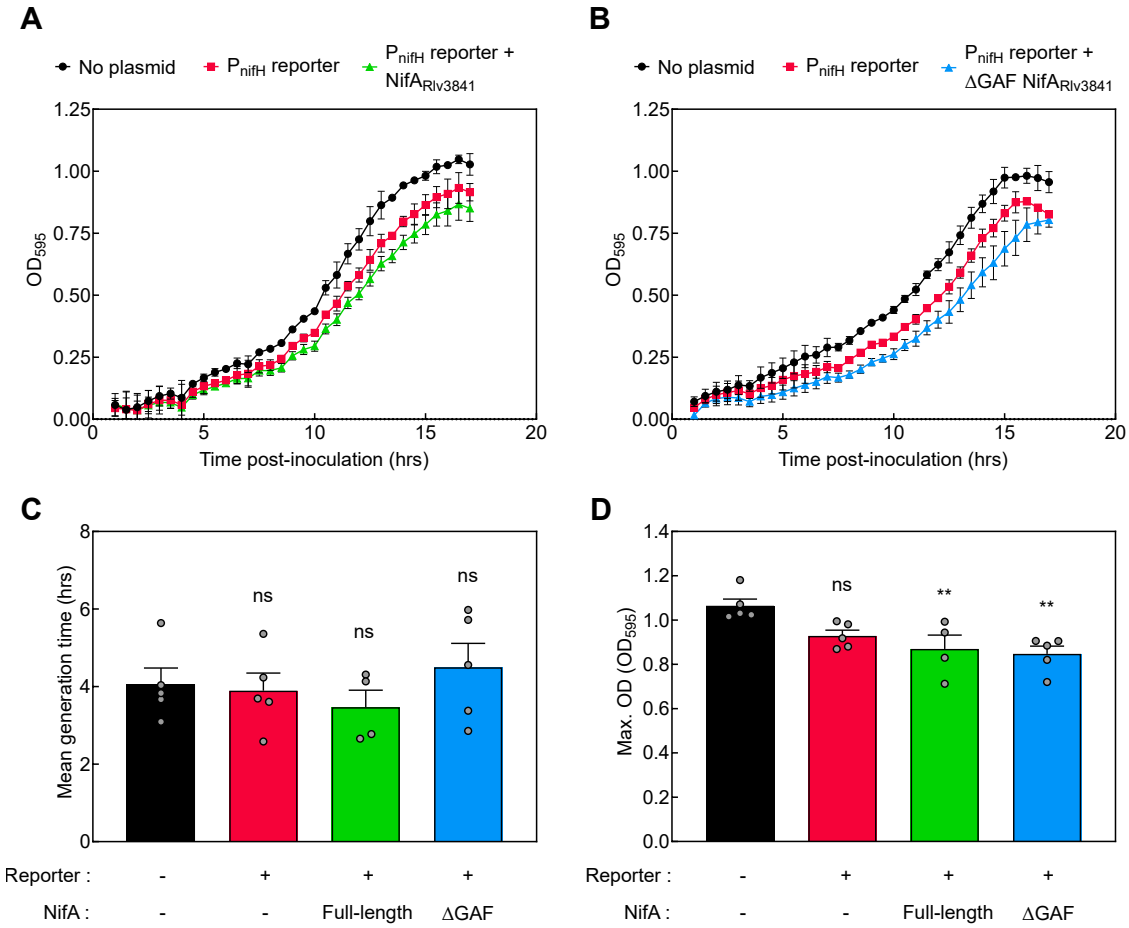
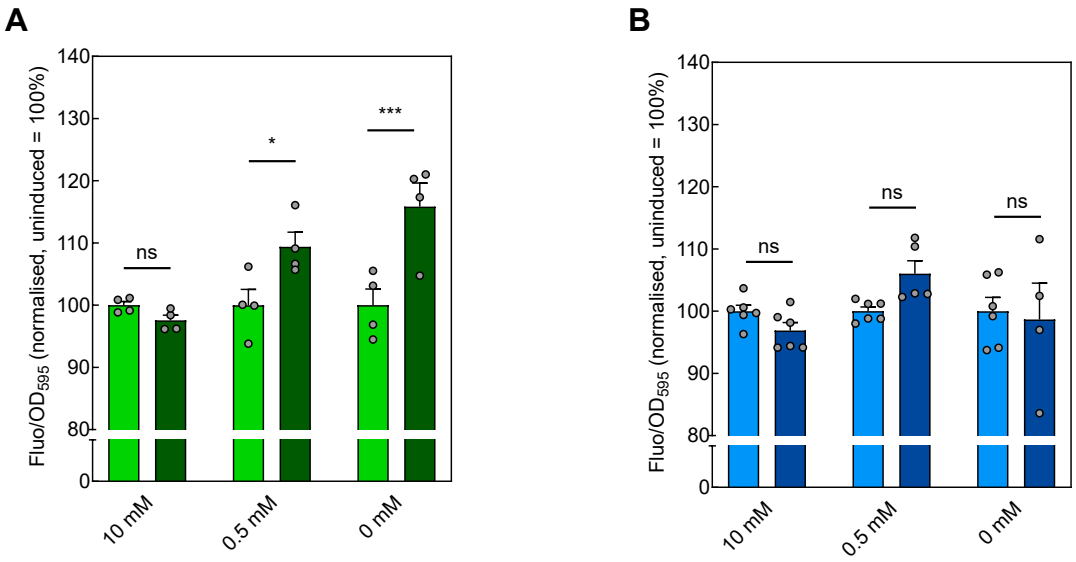


Figure 5.5: Growth of Rlv3841 strains containing the P_{nifH} reporter and *nifA* expression vectors derived from pLMB51. (A, B) Growth of Rlv3841 Δ*nifA* carrying no plasmids (OPS1737, black circles), carrying the P_{nifH} reporter plasmid (OPS2080, red squares) and carrying both the reporter and **(A)** the pLMB51-derived $nifA_{\mathrm{Rlv3841}}$ expression vector (OPS2081, green squares) or **(B)** the pLMB51-derived $nifA_{\mathrm{Rlv3841}}$ ΔGAF expression vector (OPS2082, blue squares). A slight retardation of growth was observed when the reporter plasmid was present, which was exacerbated by the presence of either *nifA* expression vector. Taurine induction did not alter growth (data not shown). (C, D) Characteristics of the growth curves given in (A, B). Black, no plasmids; red, P_{nifH} reporter only; green, $nifA_{\mathrm{Rlv3841}}$ vector; blue, $nifA_{\mathrm{Rlv3841}}$ ΔGAF vector. No significant difference was found in **(C)** the mean generation time of the strains. However, **(D)** the maximum OD_{595} reached by strains was significantly lower in those containing either *nifA* expression vector. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 4 biological replicates; error bars represent SEM. Statistical tests are differences relative to the strain without plasmids, performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; **$P < 0.01$.



C

	Induction fold, taurine uninduced vs induced	
Ammonium chloride concentration	**$NifA_{Rlv3841}$**	**$NifA_{Rlv3841}$ ΔGAF**
10 mM	0.976× (±0.017)	0.969× (±0.035)
0.5 mM	1.094× (±0.063)	1.060× (±0.045)
0 mM	1.158× (±0.084)	0.987× (±0.109)

Figure 5.6: Activity of endogenous Rlv3841 NifA variants at different nitrogen concentrations. Activity of the Rlv3841 P_{nifH} reporter (pOPS1178) in media with three different ammonium chloride concentrations, in the presence of the **(A)** pLMB51-derived $nifA_{Rlv3841}$ vector (pOPS1009) in green or **(B)** $nifA_{Rlv3841}$ ΔGAF vector (pOPS1591) in blue. Lighter shading, no taurine induction; darker shading, 10 mM taurine. Values normalised at each ammonium chloride concentration such that the uninduced average is 100%. Full-length $nifA_{Rlv3841}$ showed no activity at 10 mM ammonium chloride, but a significant increase in fluorescence was observed at 0.5 mM and 0 mM ammonium chloride. The truncated $nifA_{Rlv3841}$ ΔGAF variant showed no activity at any of the concentrations tested. **(C)** Induction fold changes based on averages from the data in **(A)** and **(B)**. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 4 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; *$P < 0.05$; ***$P < 0.001$.

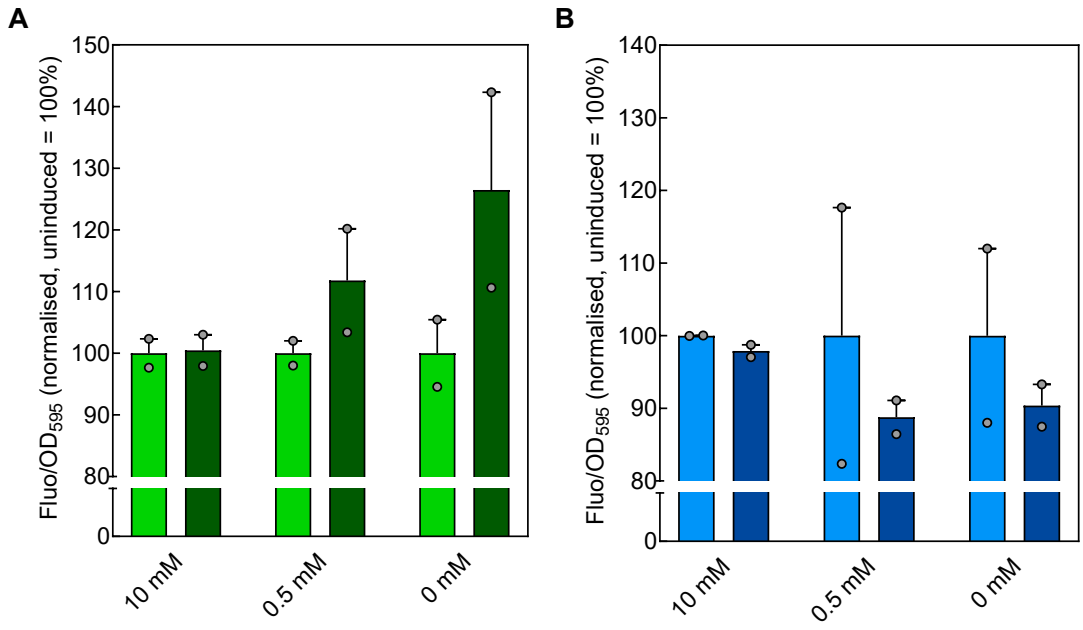


Figure 5.7: Activity of $NifA_{Rlv3841}$ variants at different nitrogen concentrations in a 1% O_2 headspace. Activity of the P_{nifH} reporter (pOPS1178) in cultures grown in a 1% O_2 headspace, in media with three different ammonium chloride concentrations, activated by **(A)** $NifA_{Rlv3841}$ and **(B)** $NifA_{Rlv3841}$ ΔGAF produced from pLMB51-derived vectors (pOPS1009 and pOPS1591 respectively). Lighter shading, no taurine induction; darker shading, 10 mM taurine. Values normalised at each ammonium chloride concentration such that the uninduced average is 100%. Cultures grew poorly under these conditions and experiments were discontinued. There was no evidence of increased NifA activity in this lower headspace O_2 concentration. Data are averages from n = 2 biological replicate; error bars represent SEM.

a critical growth defect. Many biological replicates failed to grow entirely after inoculation in microtiter plates. These experiments were therefore not pursued. Results collected did not suggest a drastic increase in NifA activity, but it was not possible to determine whether this was due to insufficiently low O_2 concentration or due to the burden effect.

5.2.4 Characterizing the behaviour of the pLMB509 backbone

Given the limited NifA-mediated reporter induction measured when $nifA_{Rlv3841}$ variants were expressed from pLMB51, we sought to improve NifA activity in the dual-plasmid system by increasing *nifA* expression. Although the pLMB51-based

NifA expression vectors did impair growth, this was not aggravated under taurine induction, suggesting increased levels of NifA production were possible without critically impairing growth. To increase *nifA* expression, the pLMB509 backbone was selected. This backbone encodes a taurine-inducible system like that in pLMB51. However, pLMB509 has a pBBR-MCS origin of replication which results in a higher copy number, potentially leading to higher *nifA* expression. Unlike pLMB51, pLMB509 also does not contain the *parABCD* system, improving its compatibility with other plasmids. The behaviour of the pLMB509 backbone was characterized using an *mcherry* fluorescent reporter inserted under P_{tauA} control (pLMB719).

Behaviour of the pLMB509 backbone in Rlv3841

The behaviour of the pLMB509 backbone was similar to pLMB51 in Rlv3841 (Figure 5.8). As with the pLMB134 plasmid, the presence of the uninduced plasmid did not significantly increase fluorescence (Figure 5.8A). Fluorescence also increased at 0.5 mM taurine, the lowest concentration tested, and largely flattened out beyond 1 mM as was the case for pLMB134 (Figure 5.8B). However, two notable differences were observed. First, in Rlv3841 the pLMB509 backbone exhibited a much larger dynamic range than did the pLMB51 backbone. At 10 mM taurine, pLMB719 produced an induction of over 200-fold relative to the uninduced plasmid, in contrast to the 8.5-fold induction produced by pLMB134.

The second notable difference was in the growth effect of the plasmid. The presence of pLMB719, regardless of the taurine concentration, produced a longer lag phase in Rlv3841, suggesting a burden effect from the backbone (Figure 5.8C). This contrasts with the negligible impact on growth caused by pLMB134 (data not shown).

Behaviour of the pLMB509 backbone in *A. caulinodans*

As in Rlv3841, in *A. caulinodans* the behaviour of the pLMB509 backbone was similar to that of the pLMB51 backbone. Like pLMB134, pLMB719 produced a significant increase in fluorescence in *A. caulinodans* even when not induced (Figure 5.9A). The taurine-response curve was also similar, with relatively little increase in

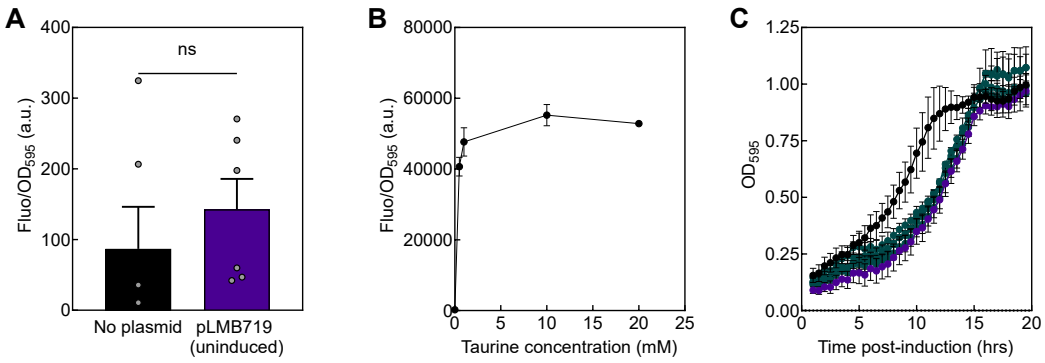


Figure 5.8: Characterization of taurine induction from the pLMB509 backbone in Rlv3841. The behaviour of the pLMB509 backbone was studied in Rlv3841 using derivative plasmid pLMB719, in which *mcherry* has been cloned under control of the taurine-inducible promoter P_{tauA}. **(A)** Comparison of Fluo/OD_{595} in Rlv3841 without a plasmid (black) and with the uninduced pLMB719 plasmid (purple). The plasmid produced no significant increase in fluorescence when not induced, indicating low leakiness in Rlv3841. **(B)** Fluo/OD_{595} produced by pLMB719 in Rlv3481 at a range of taurine concentrations. The lowest taurine concentration tested, 0.5 mM, was enough for induction. Increasing taurine concentration beyond 1 mM produced little additional fluorescence, matching our findings with pLMB134. **(C)** Growth of Rlv3841 without a plasmid (black), with uninduced pLMB719 (purple), and with pLMB719 induced at 0.5-20 mM taurine (teal). The uninduced plasmid caused an initial lag before the start of exponential growth. This impact was not exacerbated at any of the taurine concentrations tested, suggesting a burden effect caused by the pLMB509 backbone rather than by *mcherry* production. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 3 biological replicates; error bars represent SEM. Statistical analysis by Student's t test; ns (not significant) $P \geq 0.05$.

fluorescence from pLMB719 below 10 mM taurine and substantial induction at 10 mM and 20 mM (Figure 5.9B). Unlike the improvement in dynamic range observed in Rlv3841, in *A. caulinodans* pLMB719 produced a fold-change in fluorescence near-identical to that produced by pLMB134. Also unlike Rlv3841, we observed no change in the growth of *A. caulinodans* when pLMB719 was present, regardless of the taurine concentration (Figure 5.9C).

Comparing $NifA_{Rlv3841}$ activity from the pLMB51 and pLMB509 backbones

Having established the induction characteristics of the pLMB509 backbone, we sought to compare it directly to pLMB51 as a *nifA* expression vector in the context of our dual-plasmid system. The activity of full-length $nifA_{Rlv3841}$ on the Rlv3841

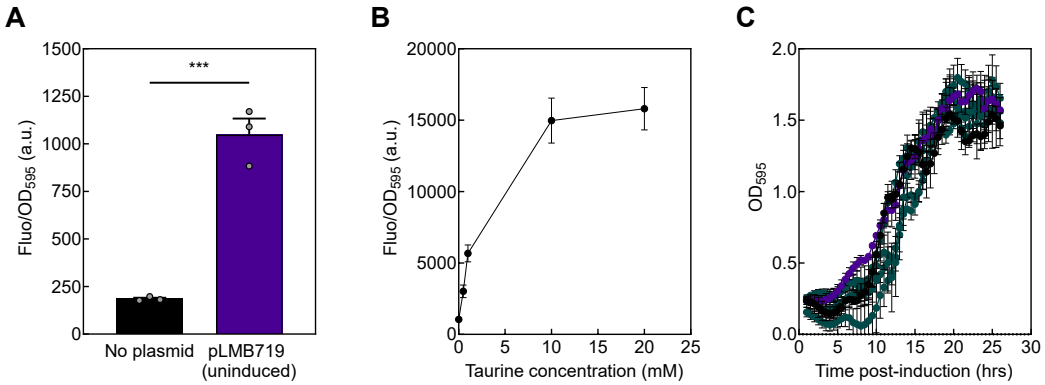


Figure 5.9: Characterization of taurine induction from the pLMB509 backbone in *A. caulinodans*. The behaviour of the pLMB509 backbone was studied in *A. caulinodans* using pLMB719 (see text for details). **(A)** Comparison of Fluo/OD_{595} in *A. caulinodans* without (black) and with the uninduced pLMB719 plasmid (purple). Even without taurine induction the plasmid produced a significant increase in fluorescence, indicating leaky expression. **(B)** Fluo/OD_{595} produced by pLMB719 in *A. caulinodans* at a range of taurine concentrations. Induction was relatively weak until 10 mM taurine was used, then plateaued after that point. **(C)** Growth of *A. caulinodans* without a plasmid (black), with uninduced pLMB719 (purple), and with pLMB719 induced at 0.5-20 mM taurine (teal). Neither the presence of the plasmid nor its induction by taurine significantly retarded growth. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 3 biological replicates; error bars represent SEM. Statistical analysis by Student's t test; ***P < 0.001.

P_{nifH} and P_{fixA} reporters was compared when expressed from either a pLMB51 or pLMB509 backbone. As our earlier results found maximum activity in the absence of ammonium chloride, the comparison was carried in media without any. Both backbones were induced with 10 mM taurine. Induction was stronger with P_{fixA}, but results were generally identical across the two reporters. As expected, the pLMB509 background led to higher uninduced and induced NifA activity, likely due to its higher copy number (Figures 5.10A and 5.10B). Whilst the pLMB51 backbone results in a better fold-change in reporter activity when comparing induced vs. uninduced conditions, the pLMB509 backbone produces higher total activity under induced conditions (Figure 5.10C). We also compared the growth of strains carrying the two different plasmids (Figure 5.11). To accommodate the required number of samples, multiple plates were required and grown inside an oxygen cabinet and growth was therefore not monitored at regular intervals (see section 2.6.2 for details).

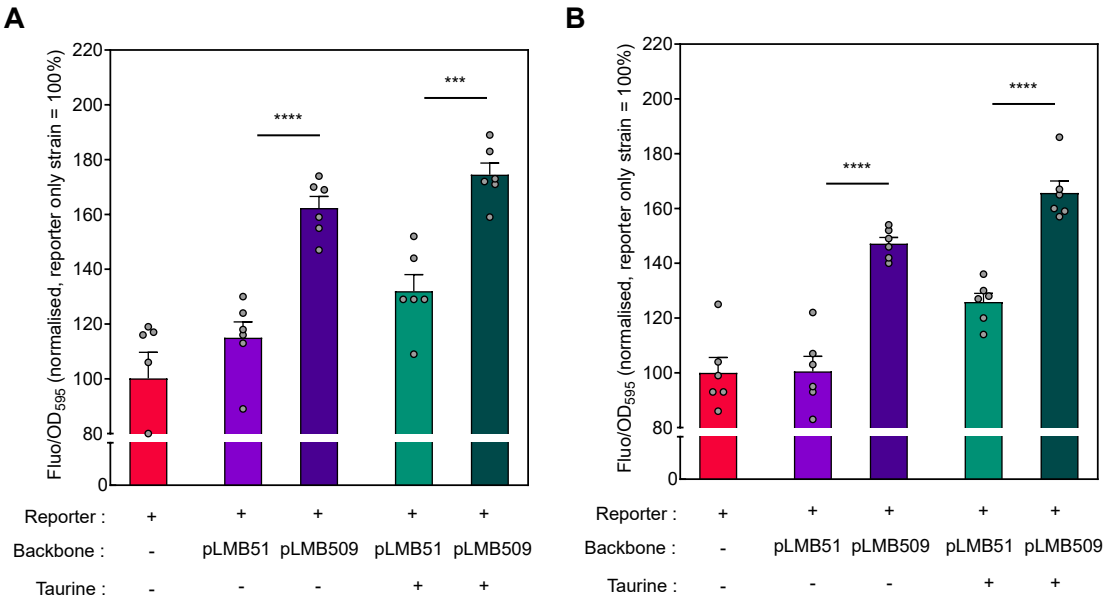
Based on endpoint readings, neither the pLMB51 nor pLMB509 based *nifA*$_{\text{Rlv3841}}$ expression vectors created a large growth defect. A significant decrease in maximum OD_{595} relative to the reporter-only strain was observed only in cultures carrying *nifA*$_{\text{Rlv3841}}$ in a pLMB51 backbone and the P_{fixA} reporter (Figure 5.11A). Thus, we saw no evidence that the increased NifA activity produced by the pLMB509 backbone caused toxicity. Based on these findings, we chose pLMB509 as the backbone for use in subsequent experiments.

5.2.5 Activity of *A. caulinodans* NifA expressed from pLMB509

Activity of NifA$_{\text{ORS571}}$ in Rlv3841

Before proceeding with a direct comparison of the three NifA variants, we studied NifA$_{\text{ORS571}}$ activity in Rlv3841 Δ*nifA*, expressed from the pLMB509 backbone (pOPS0983) (Figure 5.12). We opted to use the native Rlv3841 P_{fixA} reporter as this had shown more activity (Figure 5.10). We observed a more substantial growth defect in Rlv3841 cultures carrying this plasmid than we had observed when native *nifA*$_{\text{Rlv3841}}$ variants were expressed from either the pLMB51 or pLMB509 backbone (Figure 5.12B). The presence of the *A. caulinodans* NifA expression vector significantly decreased the maximum OD_{595} reached by cultures (Figure 5.12D). Mean generation time increased, but there was substantial variation in the growth of the dual-plasmid strain, and the difference was not statistically significant (Figure 5.12C). The defect did not worsen when the plasmid was taurine-induced (data not shown). Given the high level of uninduced NifA activity previously recorded by pLMB509, this growth effect is likely caused at least in part by *A. caulinodans nifA* expression.

As a result of the growth defect, we measured a decrease in the Fluo/OD_{595} of cultures containing the uninduced NifA vector when compared to cultures containing only the P_{fixA} reporter (Figure 5.12A). However, despite this growth defect, taurine induction significantly increased Fluo/OD_{595}, indicating that NifA$_{\text{ORS571}}$ was active in Rlv3841. As this activity was recorded using the Rlv3841 P_{fixA} promoter, it



C

Reporter / *nifA* vector backbone (strain)	Fold change	
	Reporter and $nifA_{Rlv3841}$ vector, 0 mM vs 10 mM taurine	Reporter without $nifA_{Rlv3841}$ vector vs reporter with induced $nifA_{Rlv3841}$ vector
P_{nifH} / pLMB51 (OPS1824)	1.253× (±0.062)	1.258× (±0.070)
P_{nifH} / pLMB509 (OPS2279)	1.128× (±0.030)	1.658× (±0.102)
P_{fixA} / pLMB51 (OPS1779)	1.150× (±0.072)	1.322× (±0.128)
P_{fixA} / pLMB509 (OPS2398)	1.076× (±0.035)	1.746× (±0.153)

Figure 5.10: Comparison of $nifA_{Rlv3841}$ activity expressed from a pLMB51 vs pLMB509 backbone. Rlv3841 Δ*nifA* carrying the **(A)** P_{fixA} (pOPS1177) and **(B)** P_{nifH} (pOPS1178) reporters in combination with $nifA_{Rlv3841}$ vectors. All values normalised such that activity from strains containing only the reporter is 100% (red). Strains containing both a reporter plasmid and a $nifA_{Rlv3841}$ vector are shown in purple (not taurine induced) and teal (taurine induced). Lighter shading represents strains containing $nifA_{Rlv3841}$ in the pLMB51 backbone (pOPS1009), darker shading represents strains containing $nifA_{Rlv3841}$ in the pLMB509 backbone (pOPS1591). **(C)** Induction fold changes from **(A)** and **(B)** in table format. Uninduced pLMB51 produced little to no increase in fluorescence relative to Rlv3841 containing only a reporter, indicating minimal leaky expression. Uninduced pLMB509 produced a substantial increase in fluorescence relative to Rlv3841 containing only a reporter, indicating leaky expression. The highest fluorescence from both reporters was observed in combination with the induced pLMB509-derived $nifA_{Rlv3841}$ vector. Thus, the pLMB51-derived $nifA_{Rlv3841}$ vector is less leaky in the absence of taurine, but the pLMB509-derived vector produces a stronger signal in the presence of taurine. Experiments grown in a 3% O_2 headspace. Data are averages from n = 6 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ***P < 0.001; ****P < 0.0001.

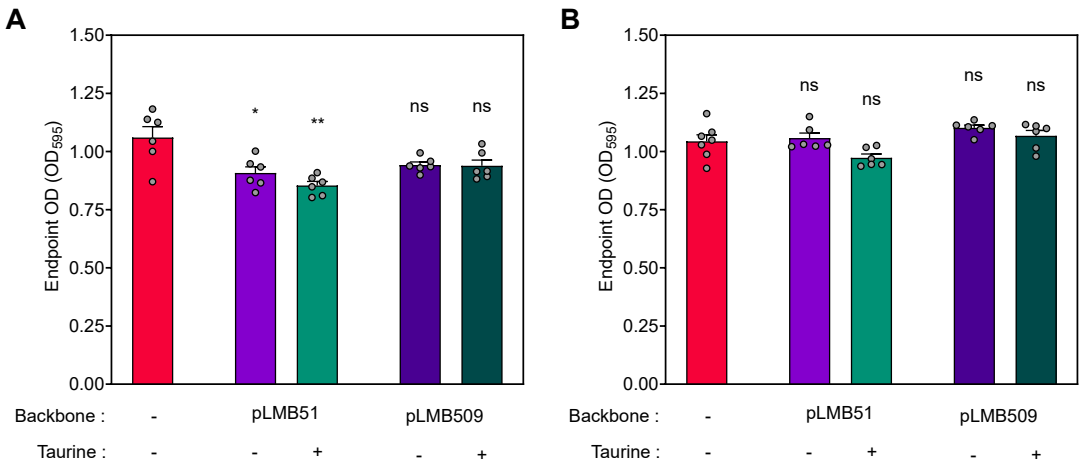


Figure 5.11: Comparison of endpoint OD_{595} reached by Rlv3841 cultures containing a combination of reporter plasmids and *nifA*$_{Rlv3841}$ expression vectors. Endpoint OD_{595} reached by Rlv3841 Δ*nifA* cultures with the **(A)** P_{fixA} (pOPS1177) and **(B)** P_{nifH} (pOPS1178) reporters on their own (red), with and without a pLMB51- (lighter shading) or pLMB509-derived (darker shading) *nifA*$_{Rlv3841}$ expression vector (pOPS1009 and pOPS1591 respectively). Uninduced cultures shown in purple, taurine induced cultures in teal. Only the pLMB51-derived *nifA*$_{Rlv3841}$ vector in combination with the P_{fixA} reporter significantly reduced endpoint OD_{595}. None of the strains exhibited a major growth defect compared to strains carrying only a reporter plasmid. Experiments grown in a 3% O_2 headspace. Data are averages from n = 6 biological replicates; error bars represent SEM. Statistical tests are differences relative to the reporter-only strain, performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; *$P < 0.05$; **$P < 0.01$.

further demonstrates that the protein is at least partially able to activate the expression of heterologous nitrogen fixation genes.

We next investigated the nitrogen sensitivity of NifA$_{ORS571}$. Significant activity was observed at every ammonium chloride concentration tested (Figure 5.13A). Like Rlv3841 NifA, activity increased as the ammonium concentration decreased (Figure 5.13B). This agrees with past studies showing that NifA$_{ORS571}$ is repressed in the presence of fixed nitrogen [239, 278]. However, with errors included all three induction folds overlap, indicating no significant change as nitrogen concentration decreased.

Activity of NifA$_{ORS571}$ endogenously expressed in *A. caulinodans*

The activity of NifA$_{ORS571}$ in Rlv3841 at 10 mM ammonium chloride could partially be due to the absence of its native repressors. We therefore studied the activity of the

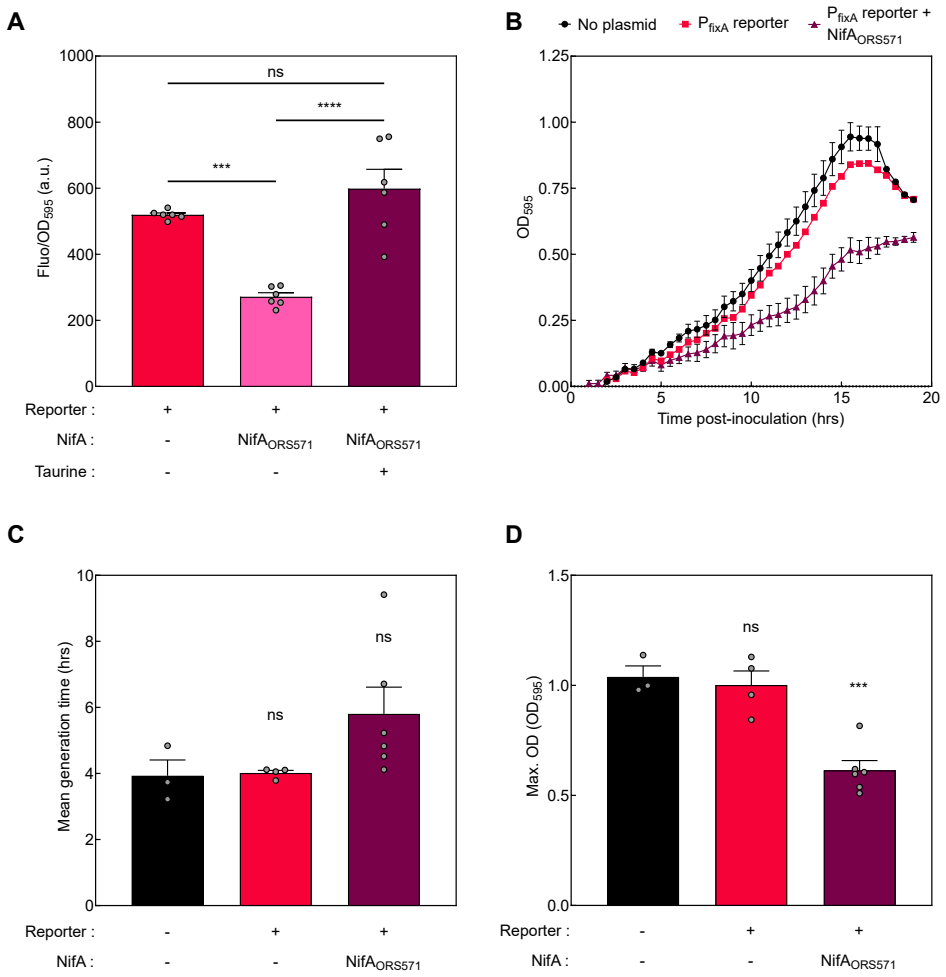
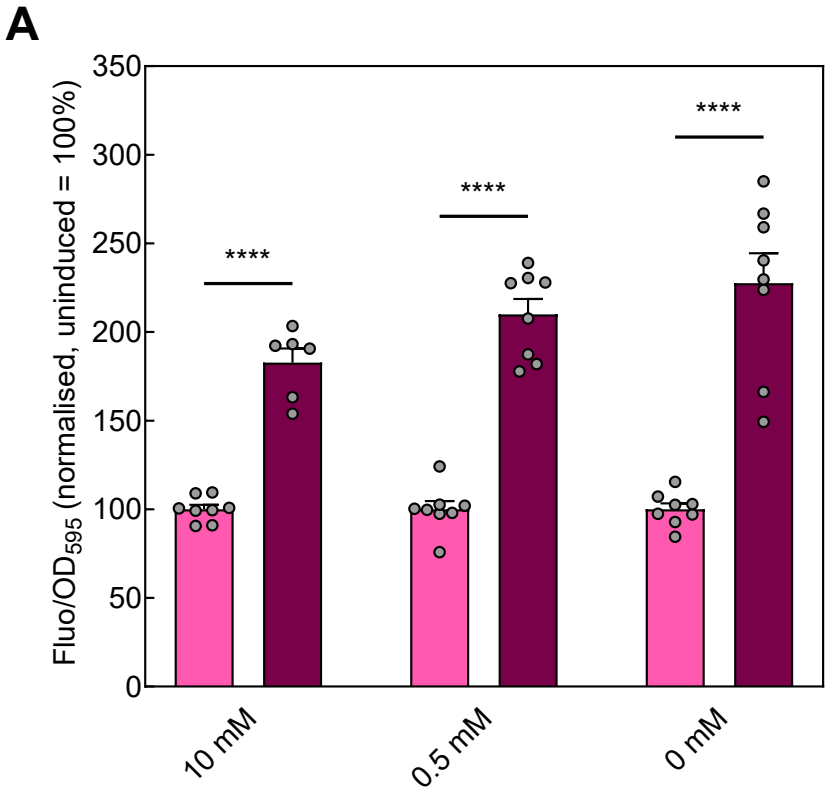


Figure 5.12: Activity of NifA$_{ORS571}$ produced from a pLMB509 backbone in Rlv3841. (A) Fluorescence from the Rlv3841 P$_{fixA}$ reporter plasmid (pOPS1177) in Rlv3841 Δ*nifA* (OPS1737) derived strains. In red, the strain containing only the reporter (OPS2080). The strain (OPS2612) carrying both the reporter and the *nifA*$_{ORS571}$ vector (pOPS0983) is given in pink when not induced and in purple when induced. Adding the *nifA*$_{ORS571}$ vector without taurine significantly decreased the Fluo/OD$_{595}$ values, largely due to a growth defect (see panel B). Taurine induction increased Fluo/OD$_{595}$, indicating NifA$_{ORS571}$ activity. **(B)** Growth of Rlv3841 Δ*nifA* strains without plasmids (black circles), with only the reporter plasmid (red squares) and both the reporter plasmid and *nifA*$_{ORS571}$ vector (purple diamonds). A slight retardation of growth was observed when the reporter plasmid was present. The pLMB509-derived *nifA*$_{ORS571}$ vector critically impaired growth. Induction with taurine did not aggravate this impairment (data not shown). **(C)** Mean generation time and **(D)** maximum OD$_{595}$ of strains studied in (A). Neither growth metric was significantly impacted by the presence of the reporter plasmid. Mean generation time was highly variable when the *nifA*$_{ORS571}$ vector was present, indicating inconsistent growth, although the data are not significantly different from Rlv3841 Δ*nifA* growth. The maximum OD$_{595}$ reached by cultures was significantly impaired by the presence of the *nifA*$_{ORS571}$ vector. Experiments grown under a 3% O_2 headspace atmosphere. Data are averages from at least n = 3 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant); ***P < 0.001; ****P < 0.0001.



B

Ammonium chloride concentration	Induction fold, 0 mM vs 10 mM taurine
10 mM	1.828× (±0.214)
0.5 mM	2.100× (±0.344)
0 mM	2.275× (±0.482)

Figure 5.13: Activity of $NifA_{ORS571}$ in media with different ammonium chloride concentrations. (A) Activity of the P_{fixA} reporter (pOPS1177) activated by $NifA_{ORS571}$ (pOPS0983) in media with three different ammonium chloride concentrations, in Rlv3841 Δ*nifA* (OPS1737). Values normalised at each concentration such that the uninduced average is 100% (pink). Cultures induced with 10 mM taurine are shown in purple. Significant induction was detectable at all three ammonium chloride concentrations. **(B)** Induction fold changes based on averages from the data in (A). Induction fold increased as the ammonium chloride concentration dropped, indicating increased $NifA_{ORS571}$ activity. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 6 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ****$P < 0.0001$.

protein in its native host with a reporter based on the *A. caulinodans* P_{fixA} promoter fused to *sfgfp* in the pOGG026 backbone (pOPS0999). This backbone was used as it had already been validated for use in *A. caulinodans* by past unpublished work in our group. As seen in Rlv3841, the reporter plasmid and dual-plasmid system impaired growth, although neither significantly changed the mean generation time or maximum OD_{595} without taurine induction (Figures 5.14C and 5.14D). When the dual-plasmid system was induced, this resulted in a significant impairment in both the maximum OD_{595} reached and the mean generation time of these cultures. Unlike our results in Rlv3841, this indicates that the increase in $nifA_{ORS571}$ expression was detrimental to *A. caulinodans* growth. We saw no increase in Fluo/OD_{595} in *A. caulinodans* when the dual plasmid system was present, regardless of whether it was induced (Figure 5.14A).

We also tested the system under the same range of ammonium chloride concentrations used in Rlv3841, but likewise saw no induction under any of these conditions (Figure 5.15). There was therefore no evidence that $NifA_{ORS571}$ was active when expressed in its native host. Some activity may be present but masked by the growth impairment caused by the dual-plasmid system. However, given its activity in Rlv3841, these results suggest that native regulators in *A. caulinodans* are inactivating the $NifA_{ORS571}$ protein under the conditions studied.

5.2.6 High-throughput NifA activity studies

Having established activity from Rlv3841 NifA and *A. caulinodans* NifA in Rlv3841 when expressed from the pLMB509 backbone, we moved to high-throughput experiments to directly compare these proteins and better understand how they responded to O_2 and fixed nitrogen levels. More details about the experimental procedures used can be found in Chapter 2. In brief, incubation in an oxygen cabinet instead of a plate reader allowed multiple plates to be grown simultaneously. This expanded the number of strains and conditions that could be studied in parallel, at the cost of removing the kinetic measurements made for earlier experiments incubated in a plate reader. A single endpoint reading of all samples was instead taken.

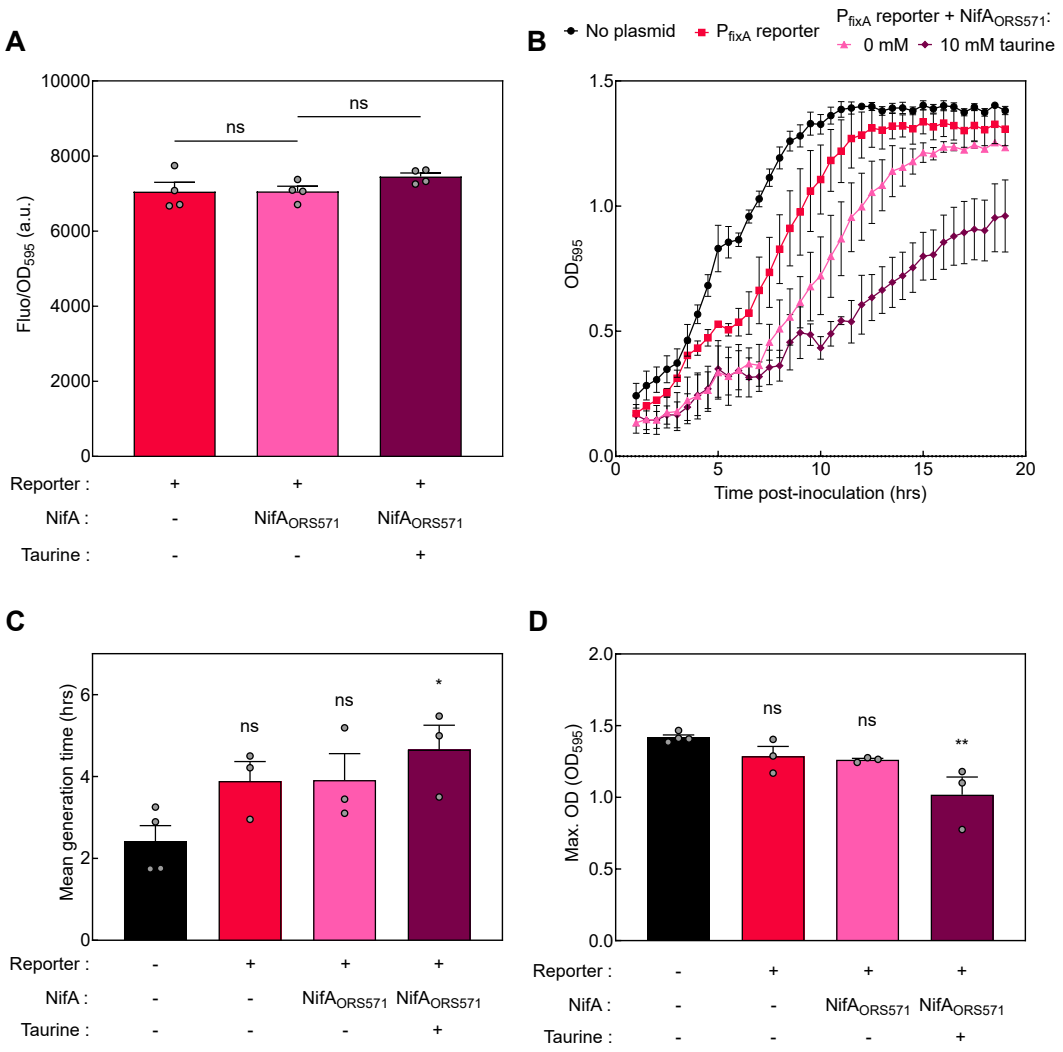


Figure 5.14: Activity of $\text{NifA}_{\text{ORS571}}$ in *A. caulinodans*. **(A)** Fluorescence from *A. caulinodans* Δ*nifA* derived strains containing combinations of an *A. caulinodans* reporter and $nifA_{\text{ORS571}}$ vector. In order, strains contain only the *A. caulinodans* P_{fixA} reporter plasmid (red, pOPS0999), both the P_{fixA} reporter and the uninduced (pink) or induced (purple) pLMB509-derived $nifA_{\text{ORS571}}$ vector (pOPS0983). Adding the reporter plasmid on its own increased fluorescence; however, neither the induced nor uninduced $nifA_{\text{ORS571}}$ expression vector further increased it, suggesting the $\text{NifA}_{\text{ORS571}}$ protein was inactive. **(B)** Growth of *A. caulinodans* Δ*nifA* strains without any plasmid (black circles), with only the reporter plasmid (red squares) and both the reporter plasmid and $nifA_{\text{ORS571}}$ vector, uninduced (pink triangles) or induced with 10 mM taurine (purple diamonds). Growth appeared increasingly impaired as more plasmids were present and a particularly strong impact was observed when the $nifA_{\text{ORS571}}$ vector was induced. **(C)** Mean generation time and **(D)** maximum OD_{595} reached by cultures in **(B)**. Both the reporter plasmid and the uninduced $nifA_{\text{ORS571}}$ vector appeared to slightly impair the mean generation time and maximum OD_{595} of cultures relative to *A. caulinodans* Δ*nifA*, but there was no statistically significant difference in our data. However, inducing the $nifA_{\text{ORS571}}$ vector did significantly impair the mean generation time and maximum OD_{595} of cultures carrying both plasmids. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 3 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; *$P < 0.05$; **$P < 0.01$.

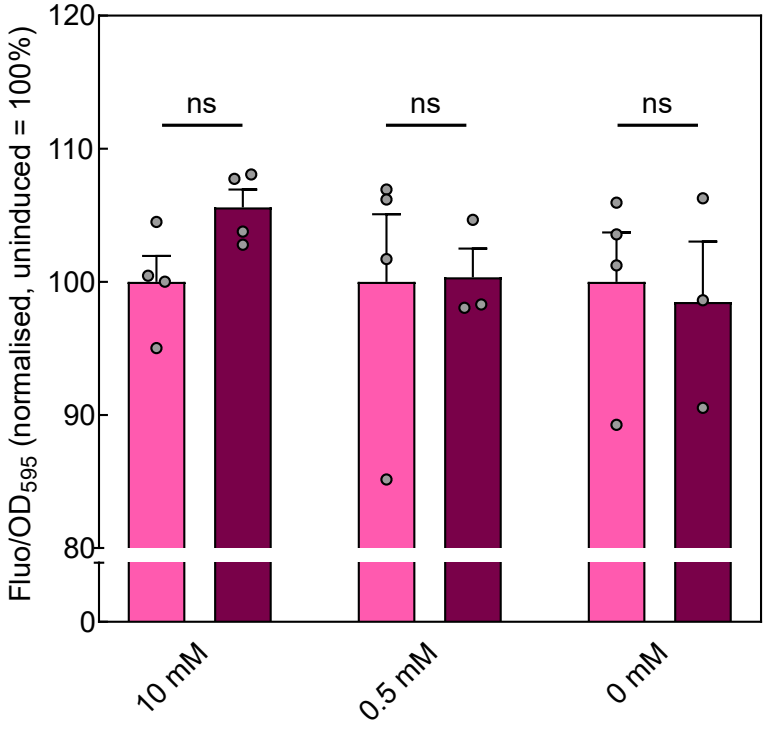


Figure 5.15: Activity of endogenous NifA$_{ORS571}$ at different nitrogen concentrations. Activity in *A. caulinodans* Δ*nifA* of the *A. caulinodans* P$_{fixA}$ reporter (pOPS0999) in combination with the pLMB509-derived *nifA*$_{ORS571}$ vector (pOPS0983). Activity tested in media with three different ammonium chloride concentrations. In pink, no taurine induction; in purple, 10 mM taurine. Values normalised at each ammonium chloride concentration such that the uninduced average is 100%. No evidence of *nifA*$_{ORS571}$ activity was found at any of the ammonium chloride concentrations tested. Experiments grown in a 3% O_2 headspace. Data are averages from at least n = 3 biological replicates; error bars represent SEM. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant).

Maximum OD reached by dual-plasmid strains under high-throughput growth conditions

Plates inside the plate reader were shaken in double-orbital mode at the highest available speed, 700 RPM. This shaking regimen was not possible inside the oxygen cabinet with the equipment available. Only orbital-mode shaking was available, and the highest practical speed was 300 RPM. This did not necessarily represent a downgrade in aeration, given the larger tray used to shake plates and that shaking was not being frequently paused for kinetic measurements. The maximum OD_{595} reached by cultures had consistently been more sensitive to burden effects and we therefore investigated this in various strains grown in high-throughput

conditions. Cultures carrying the P_{nifH} reporter in combination with the Rlv3841 NifA expression plasmid attained a significantly lower maximum OD_{595} than did those carrying only the P_{nifH} reporter (Figure 5.16B). No significant difference in maximum OD_{595} reached was observed in any strains carrying the P_{fixA} reporter plasmid alone or in combination with a NifA expression vector (Figure 5.16A). However, these cultures consistently attained a lower maximum OD_{595} reading, potentially because of stronger *gfp* expression from this reporter causing a higher burden. Maximum OD_{595} readings were spread over a substantially larger range than had been observed inside the plate reader, particularly in strains carrying the P_{fixA} reporter. This is likely due in part to the orbital shaking mode, which aggravates any cell clumping that may be happening due to burden effects. Shaking does appear to have been better overall, as the growth defect caused by the $nifA_{\text{ORS571}}$ vector was not seen in these experiments. Instead, the results suggest that the combination of the P_{fixA} reporter and $nifA_{\text{Rlv3841}}$ expression plasmid causes the most burden, as these cultures reached the lowest maximum OD_{595}.

In summary, the P_{fixA} reporter caused a higher burden than did the P_{nifH} reporter, and $nifA_{\text{Rlv3841}}$ caused the highest burden out of any of the *nifA* expression vectors. Despite these effects, minimal differences were observed in the maximum OD_{595} reached by the various strains tested, allowing direct comparisons to be made.

Activity of NifA variants in response to changing levels of fixed nitrogen

The activity of the three NifA variants was compared at a range of nitrogen concentrations using both the P_{fixA} and P_{nifH} reporter. Multiple protocol and analysis changes were made to incorporate learnings from the preliminary low-throughput experiments described above. As seen earlier, directly comparing absolute Fluo/OD_{595} readings across cultures is unreliable when there are large differences in growth (Figure 5.12). Such differences in growth were inevitable across the range of ammonium chloride concentrations to be tested. Simultaneously, high uninduced expression from the pLMB509 backbone made taurine uninduced vs induced reporter activity a poor indicator of NifA activity. To resolve both these

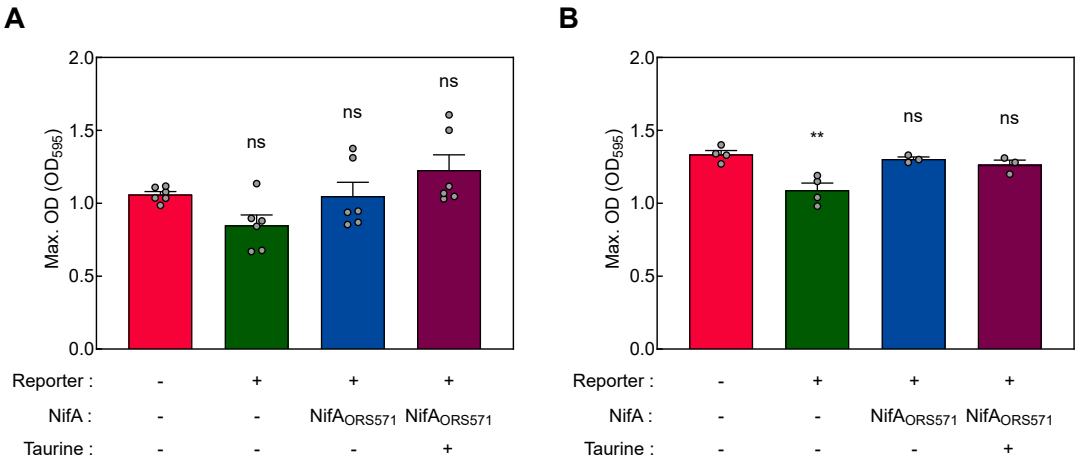


Figure 5.16: Maximum OD reached by Rlv3841 Δ*nifA* strains carrying reporter and *nifA* vectors. Maximum OD reached by Rlv3841 Δ*nifA* (OPS1737) derived strains in high-throughput experiments containing the **(A)** P_{fixA} or **(B)** P_{nifH} reporter (pOPS1177 and pOPS1178 respectively). Strains shown carry only the reporter (red), the reporter in combination with a taurine-induced $nifA_{\text{Rlv3841}}$ vector (pOPS1591, green), $nifA_{\text{Rlv3841}}$ ΔGAF vector (pOPS0982, blue) or $nifA_{\text{ORS571}}$ vector (pOPS0983, purple). No significant difference in maximum OD_{595} reached was observed in any of the strains carrying the P_{fixA} reporter plasmid. Cultures carrying the P_{nifH} reporter in combination with the $nifA_{\text{Rlv3841}}$ vector attained a significantly lower maximum OD_{595} than did those carrying only the P_{nifH} reporter, but the difference was small. Experiments grown in a 3% O_2 headspace. For each strain, at least n = 3 biological replicates were measured; error bars represent SEM. Statistical tests are differences relative to the reporter-only strain, performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant) $P \geq 0.05$; **$P < 0.01$.

issues, we standardized the $FLUO/OD_{595}$ reading of each dual-plasmid culture to that from a reference constitutive promoter, at each different ammonium chloride concentration tested (see Chapter 2 for details) [394]. This converts absolute reporter activity into a relative value that is a percentage of the activity of the reference promoter under those conditions. This standardization compensates for differences in growth due to ammonium chloride concentration. By comparing the $Fluo/OD_{595}$ of induced dual-plasmid strains from which reporter-only $Fluo/OD_{595}$ had been subtracted, leaky activity of NifA from the pLMB509 backbone was also taken into account.

Based on the activity of the P_{fixA} reporter, we observed that both $NifA_{\text{Rlv3841}}$ and $NifA_{\text{ORS571}}$ showed the highest activity in the absence of ammonium chloride

(Figure 5.17A).

In both proteins, this activity decreased as ammonium chloride concentration increased. Activity was highest from $NifA_{ORS571}$. Of note, some activity above background was recorded from truncated $NifA_{Rlv3841}$ ΔGAF, whereas none had been detected in earlier experiments using the pLMB51 backbone. Higher expression from the pLMB509 backbone therefore appears to have increased the sensitivity of the system. The level of activity detected remained lower than that from the full-length protein. In line with earlier experiments, we detected no activity from either $NifA_{Rlv3841}$ protein in media with high (5-10 mM) ammonium chloride concentrations.

The P_{nifH} reporter plasmid produced substantially less fluorescence from all NifA variants tested (Figure 5.17B). Past microarray work in our group has found that both genes are induced to a similar extent under symbiotic conditions [377]. The P_{fixA} promoter may be stronger than P_{nifH}, possibly to ensure sufficient read-through activity to express *nifA* itself downstream of the *fixABCX* cluster. However, recent work in our group has shown that the strength of NifA-regulated promoters in Rlv3841 is significantly influenced by elements that can be hundreds of base pairs upstream of the transcription start site [395]. Indeed, NifA binding sites are known to be able to influence transcription from promoters that are 1,000 bp distant [257, 258]. It is therefore most likely that the difference in strength between the P_{fixA} and P_{nifH} reporters is due to the absence of upstream elements which we did not include in our cloning. Although the lower activity of the P_{nifH} reporter makes it difficult to discern a meaningful ammonium chloride dose-response, both $NifA_{Rlv3841}$ variants again show no activity under high fixed nitrogen concentration. Like the P_{fixA} reporter, the highest P_{nifH} reporter activity was observed in the presence of the $NifA_{ORS571}$ protein.

Activity of NifA variants in response to changing concentrations of O_2

Rhizobial NifA proteins can indirectly sense O_2 concentration via an iron-binding cluster in their IDL [244, 276, 280]. However, to date the exact O_2 conditions

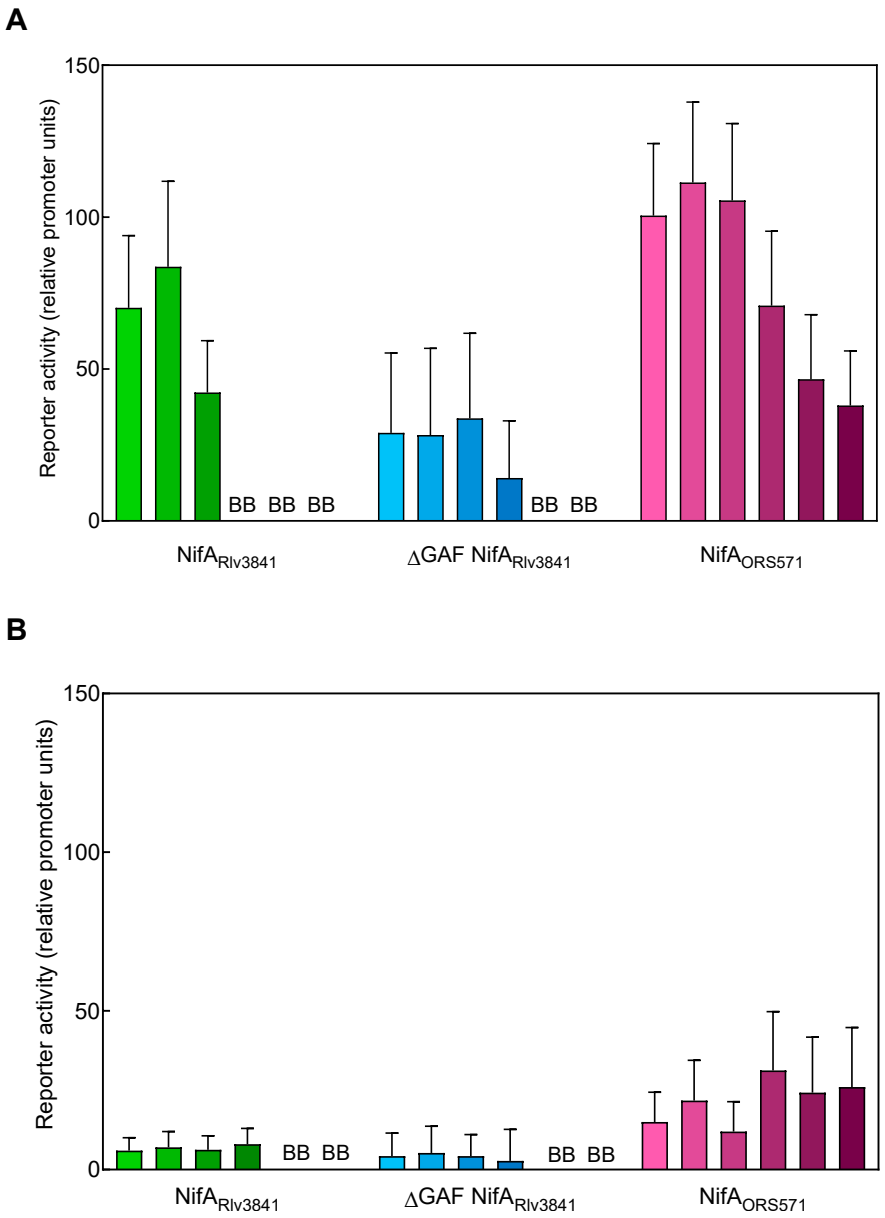


Figure 5.17: Response of NifA variant activity to ammonium chloride concentration. All *nifA* variants expressed in Rlv3841 Δ*nifA* (OPS1737) from pLMB509-derived plasmids induced with 10 mM taurine; $nifA_{\mathrm{Rlv3841}}$ (pOPS1591, green), $nifA_{\mathrm{Rlv3841}}$ ΔGAF (pOPS0982, blue) and $nifA_{\mathrm{ORS571}}$ (pOPS0983, pink). Darker shading represents higher ammonium chloride concentration. Activity at each ammonium chloride concentration is normalised to a reference to compensate for differences due to growth (see text for details). **(A)** Activity from the Rlv3841 P_{fixA} reporter (pOPS1177). Both the full-length and $\mathrm{NifA}_{\mathrm{Rlv3841}}$ ΔGAF proteins were inactive at higher (5, 10 mM) ammonium chloride concentrations. $\mathrm{NifA}_{\mathrm{ORS571}}$ also became less active at higher ammonium chloride concentrations but retained some activity even at 10 mM. Truncated $\mathrm{NifA}_{\mathrm{Rlv3841}}$ ΔGAF was less active than full-length $\mathrm{NifA}_{\mathrm{Rlv3841}}$. Highest reporter activity was observed in combination with $\mathrm{NifA}_{\mathrm{ORS571}}$. **(B)** Activity from the Rlv3841 P_{nifH} reporter (pOPS1178). In line with earlier experiments, activity from this reporter was substantially lower than from the P_{fixA} reporter. No clear effect of nitrogen concentration on the activity of the three NifA variants could be determined from this reporter. Experiments grown in a 3% O_2 headspace. Background reporter activity has been subtracted from all values; BB, below background. Values are an average of n = 8 (P_{fixA}) and n = 4 (P_{nifH}) biological replicates, error bars represent SEM. Individual data points are not shown for clarity.

under which these proteins are optimally active is unclear. Early work with *A. caulinodans* reported that optimal nitrogen fixation occurred at 3% O_2, and we therefore selected this concentration for our experiments [185, 186]. However, more recent work reported optimal activity at 1% O_2 [336]. We investigated the sensitivity of our NifA variants to O_2 concentrations of 1% and 21% O_2 (Figure 5.18). Replicating our nitrogen sensitivity work, the P_{nifH} reporter gave insufficient signal to make a meaningful comparison between the two O_2 concentrations (Figure 5.18B). By contrast, fluorescence data gathered with the P_{fixA} reporter showed significantly higher activity from both full-length Rlv3841 and *A. caulinodans* NifA at 1% O_2 compared to 21% O_2 (Figure 5.18A). As expected, we observed minimal activity from $NifA_{Rlv3841}$ ΔGAF, and no significant response to O_2 concentration. Whilst these results confirm the O_2 sensitivity of both proteins, the relatively slight increase in activity recorded suggests that both proteins remain largely inactive under the range of O_2 concentrations tested.

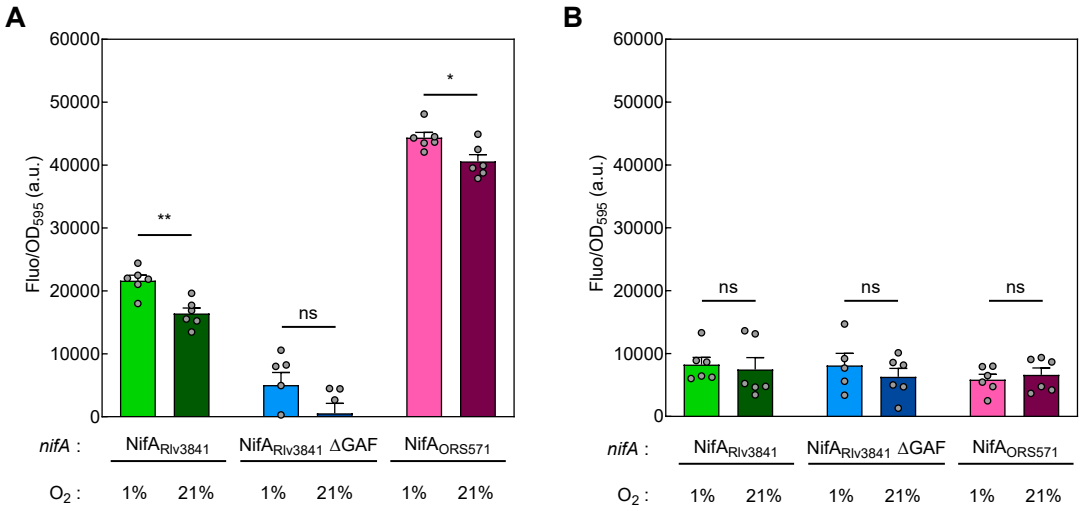


Figure 5.18: Activity of NifA variants at 1% and 21% O_2. Activity in Rlv3841 Δ*nifA* (OPS1737) of reporters in combination with pLMB509-derived $nifA_{\text{Rlv3841}}$ (pOPS1591, green), $nifA_{\text{Rlv3841}}$ ΔGAF (pOPS0982, blue) and $nifA_{\text{ORS571}}$ (pOPS0983, pink) vectors under a 1% (lighter shading) and 21% (darker shading) headspace O_2 concentration. **(A)** Activity from the Rlv3841 P_{fixA} reporter (pOPS1177). There was a small but significant increase in the activity of $NifA_{\text{Rlv3841}}$ and $NifA_{\text{ORS571}}$ at 1% O_2 relative to 21% O_2. There was no significant difference in the activity of $NifA_{\text{Rlv3841}}$ ΔGAF. In line with other experiments, activity was highest from $NifA_{\text{ORS571}}$, and lowest from the $NifA_{\text{Rlv3841}}$ ΔGAF protein. **(B)** Activity from the P_{nifH} reporter (pOPS1178). In line with earlier experiments, activity from this reporter was substantially lower than that from the P_{fixA} reporter. No significant effect of O_2 concentration was found on any of the NifA variants using this reporter. Data are averages from at least n = 5 biological replicates, error bars represent SEM. Note some data points are below the X-axis and not visible on the graph. Statistical tests performed by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons; ns (not significant); *P < 0.05; **P < 0.01.

5.2.7 Colony morphology effects

Morphological effect of $nifA_{\text{Rlv3841}}$ expression from the pLMB509 backbone in Rlv3841

We observed a morphological change in Rlv3841 colonies containing full-length Rlv3841 *nifA* in the pLMB509 backbone (Figure 5.19). Rlv3841 colonies are typically hemispherical in shape and mucoid in appearance, due in part to their production of exopolysaccharides [396]. Rlv3841 Δ*nifA* retained this morphology, and no morphological change was observed when the mutant carried the P_{fixA} reporter (OPS1778). However, the strain (OPS2398) carrying both the P_{fixA} reporter and $nifA_{\text{Rlv3841}}$ in pLMB509 produced mesa-shaped, dry colonies even in the absence of

taurine. This change was not observed with the pLMB51-based *nifA*$_{\text{Rlv3841}}$ expression vector, nor in strains carrying either *nifA*$_{\text{ORS571}}$ or *nifA*$_{\text{Rlv3841}}$ ΔGAF on pLMB509. Thus, the effect was specific to the full-length *nifA*$_{\text{Rlv3841}}$ gene and occurred only when this was expressed at a relatively high level from the pLMB509 backbone. This suggests that overexpression of *nifA*$_{\text{Rlv3841}}$ causes a change in transcription that *nifA*$_{\text{ORS571}}$ does not, indicating a difference in their regulatory activity.

Morphological effect of *nifA*$_{\text{ORS571}}$ expression from the pLMB509 backbone in *A. caulinodans*

We also investigated whether any morphological changes were taking place in *A. caulinodans* when its native *nifA*$_{\text{ORS571}}$ was overexpressed from the pLMB509 backbone (Figure 5.20). However, we saw no changes in the strain carrying the P$_{fixA}$ reporter plasmid, nor in the strain carrying both this reporter and *nifA*$_{\text{ORS571}}$ in pLMB509. This is despite the leakiness of pLMB509 in *A. caulinodans* (Figure 5.8). One possibility is that native protein-level mechanisms continue to repress NifA$_{\text{ORS571}}$ activity in *A. caulinodans*. This agrees with our earlier finding that the protein was not active in its native host despite native transcriptional control being overridden. These mechanisms may not exist or be active in Rlv3841 under *in vitro* conditions, resulting in aberrant regulation and hence altered colony morphology when native transcriptional control is bypassed, producing NifA outside of symbiosis.

5.2.8 Activity of NifA$_{\text{Rlv3841}}$ in *A. caulinodans*

Several past studies have shown that NifA proteins can be cross-compatible, with NifA from one organism able to partially or fully complement the absence of *nifA* in a non-native host [243–247]. Our finding that NifA$_{\text{ORS571}}$ was active in Rlv3841, and indeed surpassed the activity of the native protein, agreed with these findings. We next attempted to complement an *A. caulinodans* Δ*nifA* mutant using *nifA*$_{\text{Rlv3841}}$. Work in our group is currently ongoing to engineer rhizopine-controlled *nifA* expression in *A. caulinodans* to create artificial cross-kingdom signalling [335]. As part of this work, there was an interest in determining whether heterologous NifA proteins produced in *A. caulinodans* could further bypass native

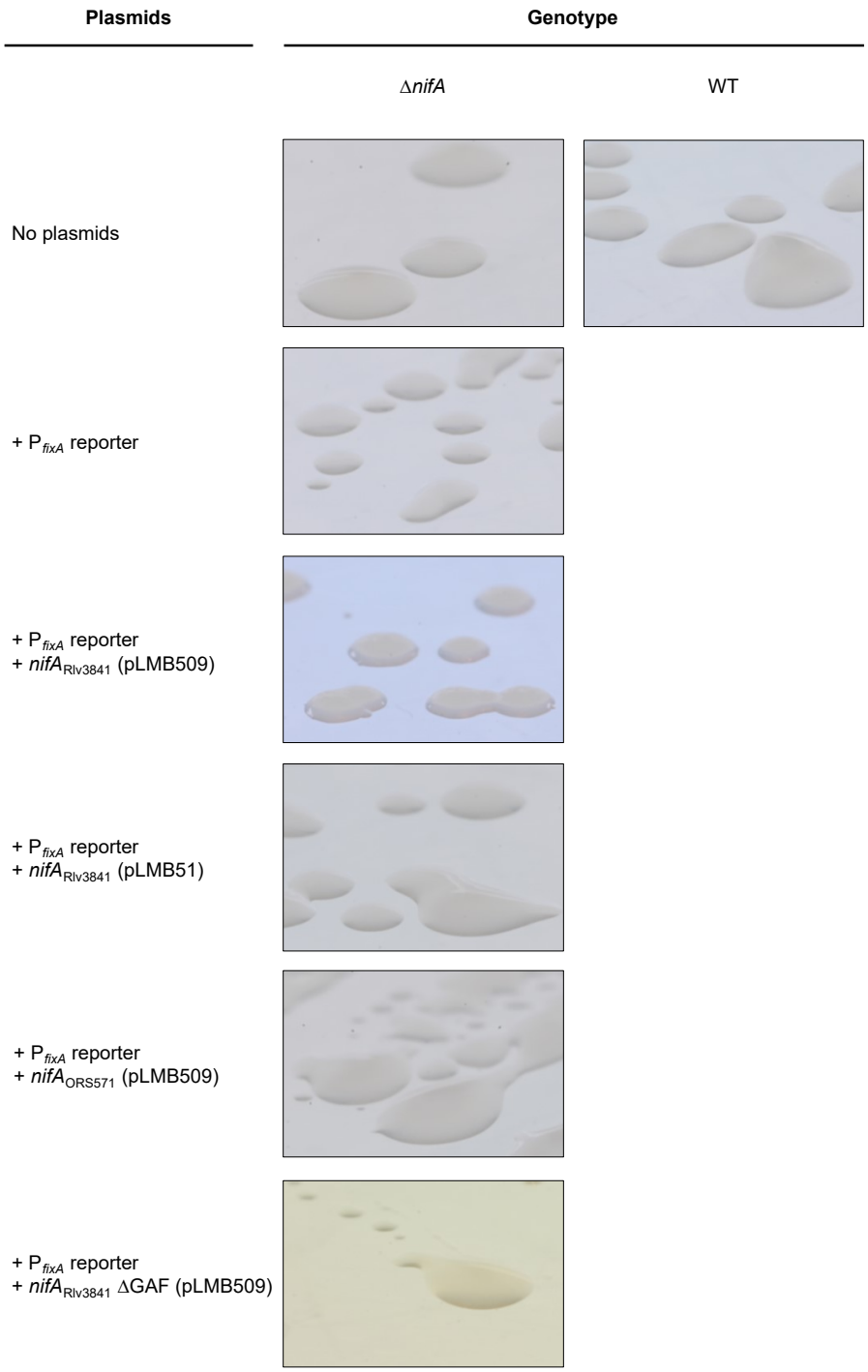


Figure 5.19: Morphology of Rlv3841 colonies with reporter plasmids and *nifA* vectors. All colonies grown on UMS agar. WT colonies have a "gloopy" morphology, forming droplet-like hemispheres; no morphology change was observed in Rlv3841 Δ*nifA* (OPS1737). The P_{fixA} reporter plasmid (pOPS1177) also had no effect. However, uninduced $nifA_{\mathrm{Rlv3841}}$ in a pLMB509-derived plasmid (pOPS1591) altered colony morphology, forming dry, flattened, mesa-like structures. This morphology suggests a substantial reduction in the production of exopolysaccharides from the bacteria, normally responsible for their "gloopy" morphology. This effect was not observed in colonies containing uninduced $nifA_{\mathrm{Rlv3841}}$ in a pLMB51-derived plasmid (pOPS1009), suggesting insufficient protein was produced from this vector. Neither uninduced $nifA_{\mathrm{Rlv3841}}$ ΔGAF nor $nifA_{\mathrm{ORS571}}$ in pLMB509-derived plasmids altered colony morphology (pOPS0982 and pOPS0983 respectively).

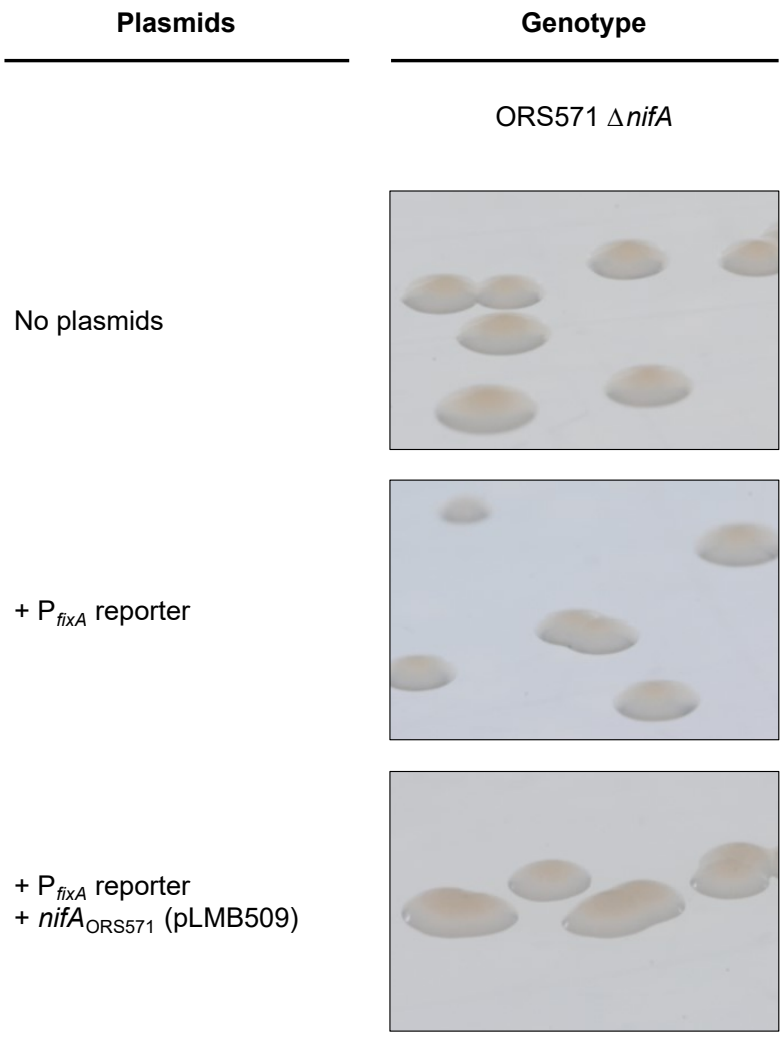


Figure 5.20: Morphology of *A. caulinodans* colonies containing reporter and *nifA* vectors. All colonies grown on UMS agar. *A. caulinodans* colonies typically form hemispherical colonies that are smaller and less viscous than Rlv3841 colonies. The presence of the *A. caulinodans* P_{fixA} reporter plasmid (pOPS0999) alone did not alter colony morphology. No difference was also observed when both the reporter plasmid and the uninduced $nifA_{ORS571}$ pLMB509-derived plasmid (pOPS0983) were present.

NifA regulatory mechanisms. We therefore opted to use a rhizopine-inducible vector to make the results of our work directly comparable to results already gathered as part of these efforts.

The vector selected, pOPS0889, had already been shown in unpublished work in our group to produce sufficient $nifA_{ORS571}$ when induced to enable nitrogen fixation in *A. caulinodans* Δ*nifA*. We inserted both full-length and ΔGAF truncated $nifA_{Rlv3841}$ into pOPS0889 under rhizopine control to produce plasmids pOPS1724

and pOPS1725, respectively. We then studied the ability of these plasmid to complement free-living nitrogen fixation in *A. caulinodans* Δ*nifA* (Figure 5.21).

In line with unpublished work in our group, we found that *A. caulinodans* Δ*nifA* retained some acetylene reduction activity. When complemented with a plasmid containing the native *nifA*, acetylene reduction activity was restored to WT levels, also in line with past results in our group. However, we recorded no significant increase in acetylene reduction when the Δ*nifA* mutant was complemented with $nifA_{\text{Rlv3841}}$. Thus, the protein appears insufficiently active in *A. caulinodans*, or fails to activate certain key genes required for fixation. Of note, $nifA_{\text{Rlv3841}}$ ΔGAF abolished all detectable nitrogen fixation activity in *A. caulinodans* Δ*nifA*. This may indicate that the protein retains the ability to bind the upstream activator sequence (UAS) of gene targets, but sterically inhibits expression as it is inactive. Loss of fixation activity may also be due to a burden effect caused by misfolding of the protein.

5.2.9 Engineering NifV activity in Rlv3841

Having demonstrated taurine inducible NifA activity in Rlv3841, we set out to combine this with NifV activity. NifV catalyses the production of homocitrate, an essential cofactor for the nitrogenase complex [391]. In symbiotic nitrogen fixing rhizobia including Rlv3841, *nifV* is typically no longer encoded by the bacteria, which instead relies on its plant host for homocitrate production [4]. Engineering free-living nitrogen fixation in Rlv3841 therefore requires that NifV activity be reintroduced.

Establishing a GC-MS based homocitrate assay

We began by establishing an assay for the detection of homocitrate in biological samples. We chose a gas chromatography-mass spectrometry (GC-MS) based assay, as this enabled cell lysates to be sampled directly without the need to purify NifV. We initially performed detection using scan-mode GC-MS (see Chapter 2 for details) [397]. A pure standard was used to determine the retention time of homocitrate

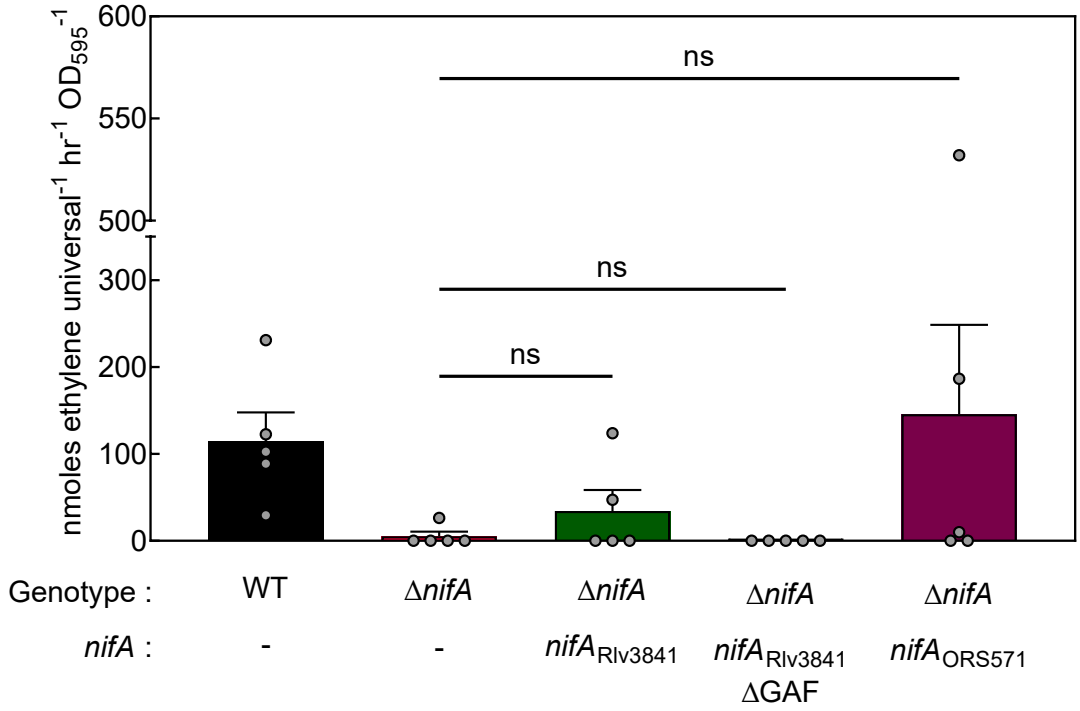


Figure 5.21: Acetylene reduction activity from *A. caulinodans* strains with plasmid-based *nifA* variant expression. Activity from *A. caulinodans* WT is shown in black. Deletion of the *nifA* gene (red) eliminated activity in all but one biological replicate. Expression of $nifA_{\text{Rlv3841}}$, $nifA_{\text{Rlv3841}}$ ΔGAF and $nifA_{\text{ORS571}}$ was re-introduced by placing these variants under rhizopine-inducible control in the pOPS0889 backbone (vectors pOPS1724, pOPS1725 and pOPS1581 respectively). Expression of $nifA_{\text{Rlv3841}}$ (green) produced activity in some biological replicates, whilst no activity was observed in strains expressing $nifA_{\text{Rlv3841}}$ ΔGAF (blue). Neither vector significantly increase acetylene reduction activity above the level of *A. caulinodans* Δ*nifA*. Expression of $nifA_{\text{ORS571}}$ produced activity in some replicates, including levels comparable to or in excess of those from *A. caulinodans* WT, but two replicates also failed to fix. For each strain, n = 5 biological replicates were measured; error bars represent SEM.

(Figure 5.22). Homocitrate produced a distinct, single peak at approximately 24.8 minutes after sample injection, and no background was detected in a blank sample (Figure 5.22A). The spectrum for this peak contained several major ions, with the strongest signal produced by an ion with a mass-to-charge (m/z) ratio of 287 (Figure 5.23). A standard curve was then produced (Figure 5.24). A linear response to homocitrate concentration was observed across the range of concentrations used. However, the lowest concentration tested, 9 ng/µL, produced no detectable homocitrate peak in scan-mode GC-MS. The detection limit of scan-mode GC-MS

for homocitrate was therefore between 18 ng/µL and 9 ng/µL.

In preliminary biological assays (data not shown), we could detect no homocitrate peak by scan-mode GC-MS. To improve sensitivity, we therefore switched to selected ion monitoring (SIM) mode GC-MS. In this mode the instrument only detects the level of a small, pre-set list of ions, improving sensitivity. We began by determining what ion to use for SIM-mode homocitrate detection. Based on our scan-mode spectrum results (Figure 5.23), we initially selected the 147 m/z, 287 m/z and 377 m/z ions.

All three ions produced a peak at the correct retention time with a pure homocitrate sample (Figure 5.25). The best signal was produced by the 287 m/z ion (Figure 5.25C), followed closely by the 377 m/z ion (Figure 5.25E). The 147 m/z ion produced a weaker signal (Figure 5.25A). Minimal background was observed in the blank sample at either 147 m/z or 287 m/z (Figures 5.25B and 5.25D). Monitoring at 377 m/z gave more background across the chromatogram (Figure 5.25F). This ion also produced a peak in the blank sample at a retention time immediately preceding that of homocitrate. This peak was not fully distinct from the homocitrate peak in the pure sample. We therefore selected the 287 m/z ion for SIM-based homocitrate detection as this produced a distinct peak with the best signal-to-noise ratio.

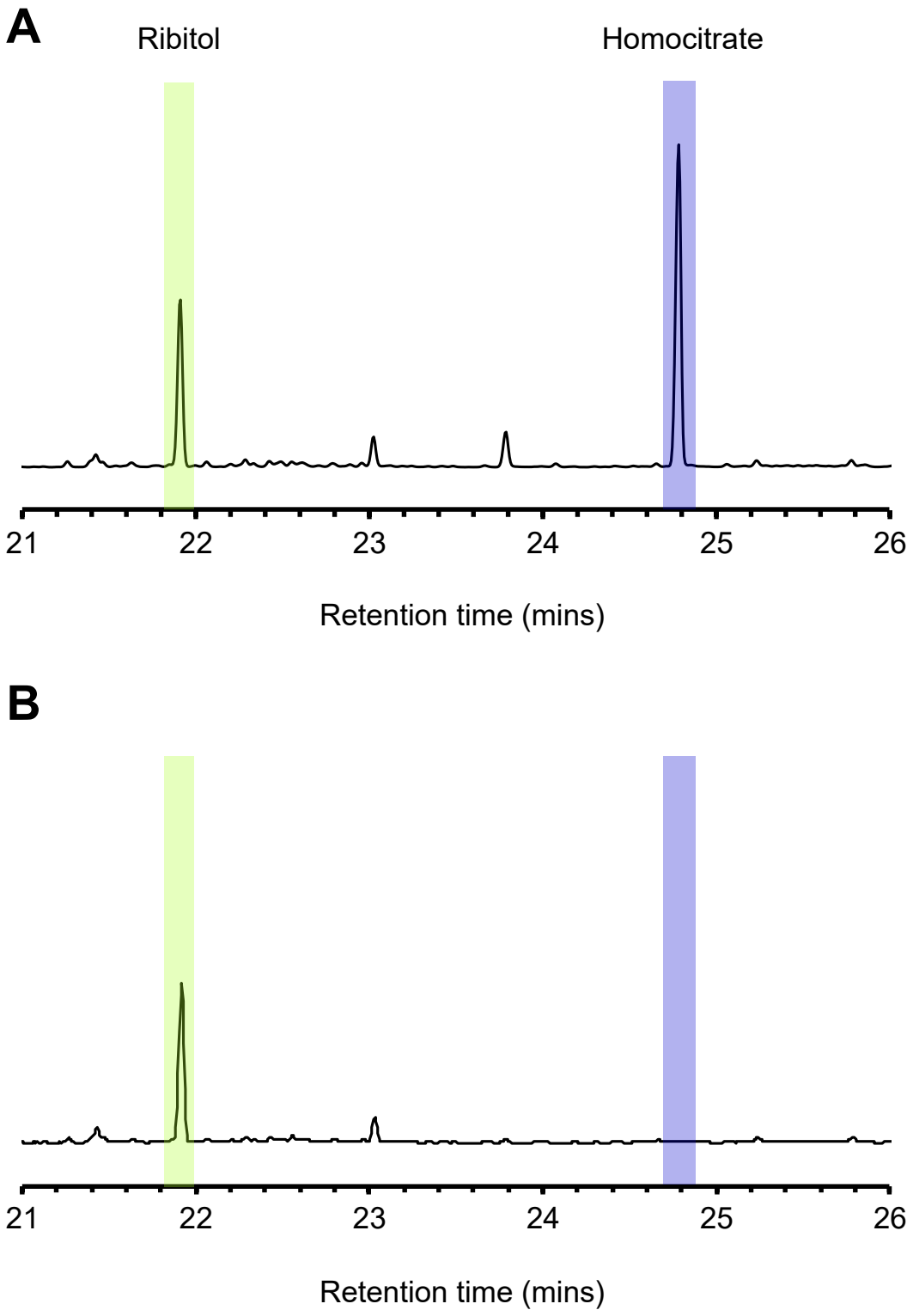


Figure 5.22: Homocitrate detection by scan-mode GC-MS. (A) Representative scan-mode chromatogram of a pure homocitrate standard. The ribitol peak (green) typically had a retention time of 21.90-21.95 minutes. Homocitrate (blue) typically had a retention time of 24.75-24.85 minutes. **(B)** Representative scan-mode chromatogram of a blank containing no homocitrate.

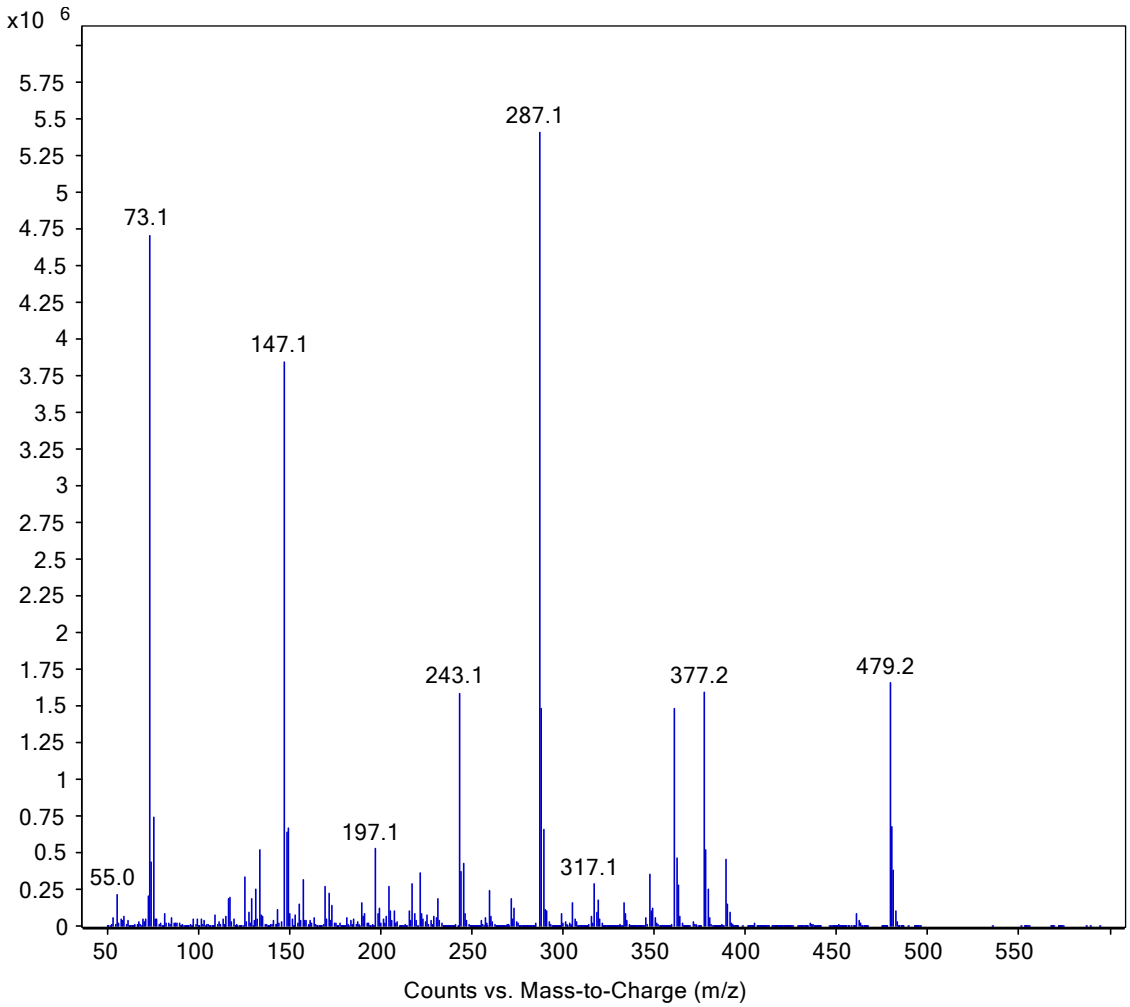


Figure 5.23: Ions detected in a homocitrate peak by scan-mode GC-MS. Representative ions detected in the homocitrate peak of a pure homocitrate sample by scan-mode GC-MS. The strongest signal was from an ion with a m/z ratio of 287. Strong signals were also emitted by ions with m/z ratios of 73, 147, 243, 377 and 479.

A SIM-mode homocitrate standard curve using the 287 m/z ion was produced with stocks ranging from 4.5×10^{2} ng/µL to 9.0×10^{-6} ng/µL (Figure 5.26). All produced a signal, indicating that the detection limit using SIM-mode was far below that of scan-mode and below the concentrations tested. The two highest concentrations used overloaded the instrument, creating a non-linear concentration response (Figure 5.26A). Excluding these two points produced a linear concentration response with a good fit (R^2=0.9989) (Figure 5.26B). However, the slope of this response was not representative of the lowest four concentrations used (including the blank, which produced no peak) (Figures 5.26C and 5.26D). As biological

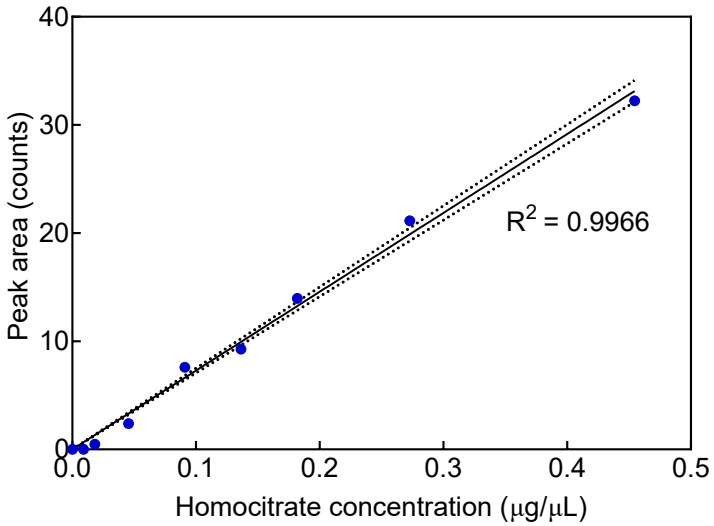


Figure 5.24: Homocitrate standard curve produced by scan-mode GC-MS. A linear relationship between peak area and concentration was maintained across most of the range of homocitrate concentrations tested. However, no peak was detected at the lowest homocitrate concentration used, 9 ng/µL. The lower detection limit for homocitrate in our GC-MS instrument in scan-mode therefore appeared to be above 9 ng/µL but below 18 ng/µL, the next concentration used. Black line: linear trendline, dotted lines represent 95% confidence intervals. Data from two separately diluted standard curves, with one measurement each per sample. SEM error bars are too small to be visible on this scale.

readings were within the range covered by these four concentrations, we based our standard curve on these four data points when converting SIM-mode measurements to homocitrate concentrations.

Determination of homocitrate concentration in nitrogen fixing *A. caulinodans*

Having established a more sensitive SIM-mode GC-MS homocitrate assay, we confirmed biological homocitrate detection using free-living nitrogen fixing *A. caulinodans*, which is known to produce its own homocitrate under these conditions. Homocitrate could be detected in both the WT strain (Figure 5.27A) and, to a lesser extent, the Δ*nifA* strain (Figure 5.27B). However, we observed that this signal degraded rapidly, making repeated technical measurements unreliable (Figure 5.28). No such degradation was observed with pure standards. The effect may be due in part to lactonization of homocitrate, which could be prevented through the use of alkaline conditions [398]. We found that less homocitrate was present

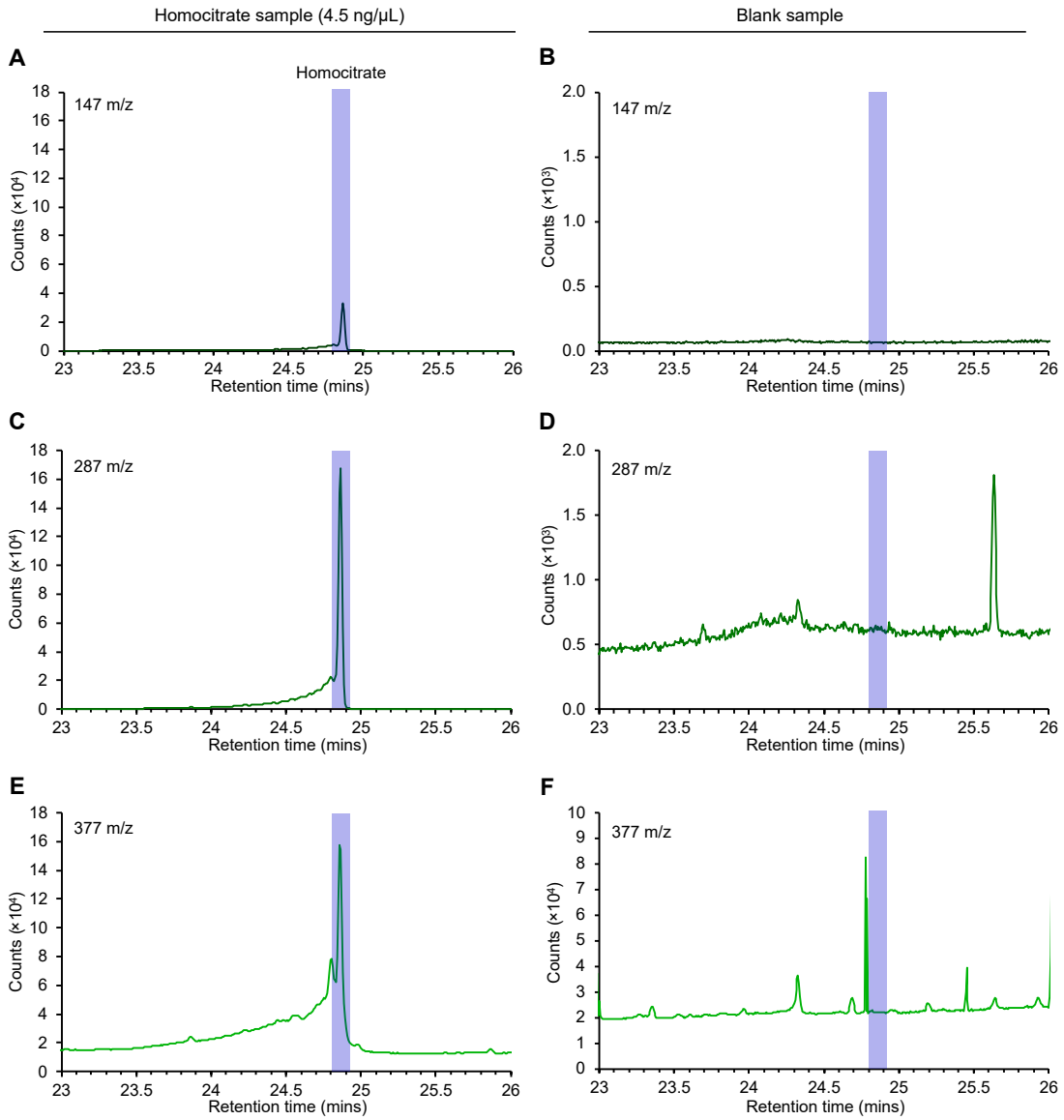


Figure 5.25: Homocitrate detection by selected ion monitoring (SIM) gas-chromatography mass spectrometry (GC-MS). Based on initial scan-mode GC-MS measurements of homocitrate, three ions were chosen for selected ion monitoring mode, with m/z ratios of **(A, B)** 147 (dark green), **(C, D)** 287 (green) and **(E, F)** 377 (light green). (A, C, E) Representative SIM-mode chromatograms of a pure homocitrate standard. The strongest signal was produced by the 287 m/z ion, followed closely by the 377 m/z ion. The 147 m/z ion also gave a signal, but this was far weaker. (B, D, F) Representative SIM-mode chromatograms of a blank sample. Both the 287 and the 147 m/z ions had minimal background noise, and no peaks from these ions were observed in the vicinity of the homocitrate retention time. The 377 m/z ion produced a far higher background signal, and a peak was observed very near the retention time of homocitrate. Based on these findings, the 287 m/z ion was used for selected ion monitoring detection of homocitrate.

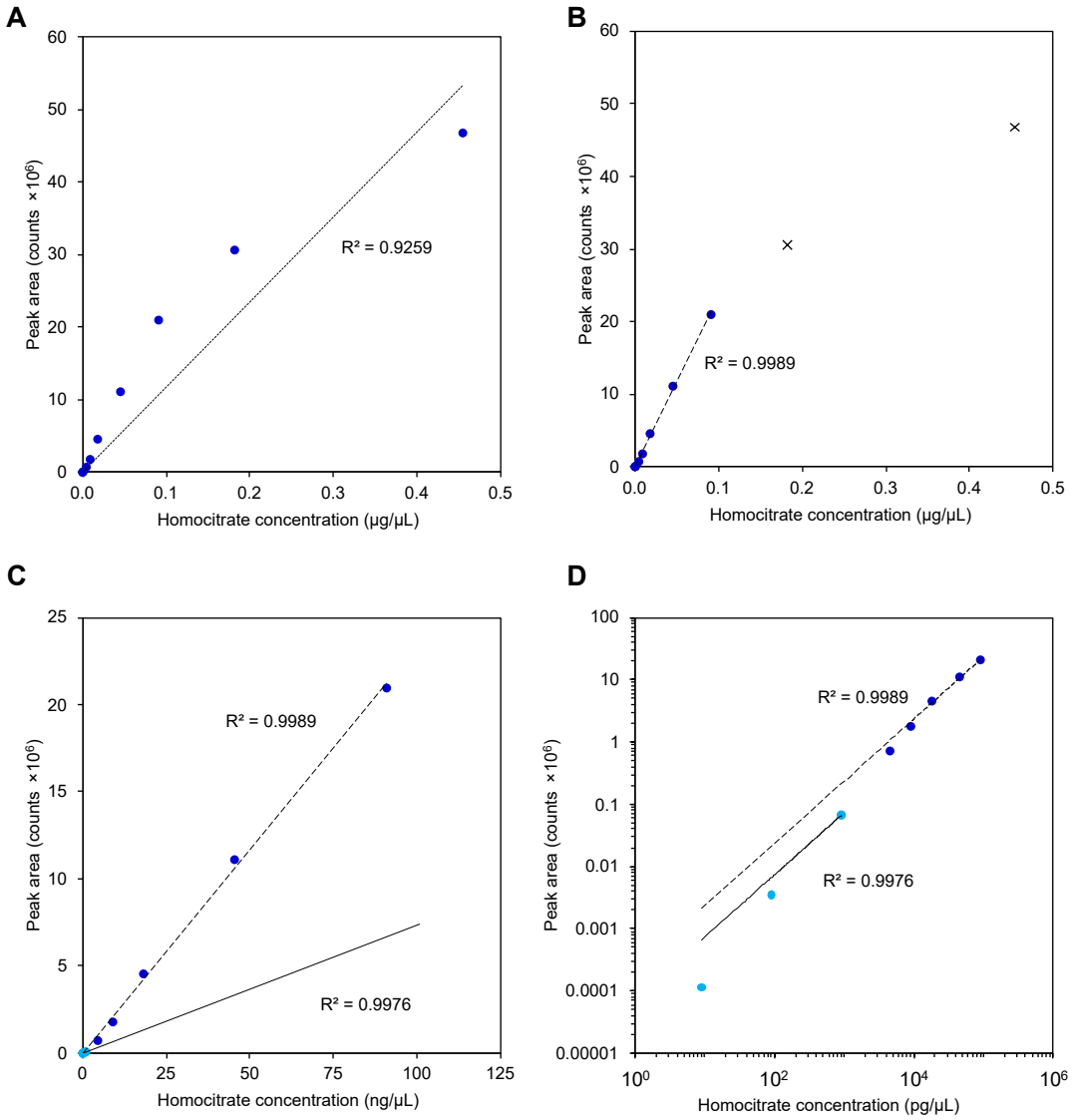


Figure 5.26: Homocitrate standard curve produced by GC-MS selected ion monitoring at 287 m/z. (A) Saturation occurred in the two highest concentrations tested (0.45 µg/µL and 0.18 µg/µL), leading to a non-linear relationship between homocitrate concentration and peak area (dotted line). **(B)** With the top two concentrations excluded (crosses), a good linear relationship emerged (dashed line). **(C)** However, this relationship was not representative of the first four concentrations tested, 0-0.9 ng/µL (light blue). The solid line represents the linear relationship between these points. Biological readings were all within the range of the first four concentrations, and only these were therefore used as the standard curve for SIM-mode GC-MS concentration calculations. **(D)** Values from **(C)** plotted on a log-log scale. Note the blank value cannot be shown on this graph. Data are averages from two separately diluted standard curves, with one measurement each per sample.

in the Δ*nifA* mutant, although insufficient replicates were collected to confirm this discrepancy (Figure 5.29).

Engineering NifV activity in Rlv3841

Three plasmids were used for *nifV* expression in Rlv3841. Two were constructed for this work using the pOGG281 backbone. The pOGG281 was chosen as it could be genomically integrated using the Tn7 system to minimize cell burden, and its resistance marker was compatible with pLMB509-based vectors. To further limit cell burden, *nifV* was placed under control of an inducible promoter. Two versions of the plasmid were built, with *nifV* under control of the P_{lac} or $P_{A1lacO1}$ inducible promoter. Given its relatedness to Rlv3841, the *A. caulinodans nifV* gene ($nifV_{ORS571}$) was chosen and used without codon optimization. A third plasmid, predating this work, was also used. This contained *Azotobacter vinelandii nifV* ($nifV_{A.\ vinelandii}$) under the constitutive control of a P_{neo} promoter, in a pJP2neo backbone. This was also compatible with the pLMB509 backbone but could not be integrated.

All three plasmids produced sufficient homocitrate in Rlv3841 cultures for GC-MS detection (Figure 5.30). One or more unidentified compounds in Rlv3841 produced peaks with a retention time only slightly lower than that of homocitrate, but which were still resolvable during peak extraction. The highest concentration of homocitrate was detected in Rlv3841 carrying the pJP2neo-based $nifV_{A.\ vinelandii}$ vector (Figure 5.31). We also observed poor growth from this strain (data not shown), likely indicating that this level of NifV activity was toxic. All three systems produced OD-standardized homocitrate levels far above those found in *A. caulinodans*, suggesting NifV activity was sufficient for nitrogenase activity.

5.2.10 Acetylene reduction activity in free-living Rlv3841

Having established homocitrate production in Rlv3841, we next combined this with our established pLMB509-based *nifA* expression vectors to elicit nitrogen fixation

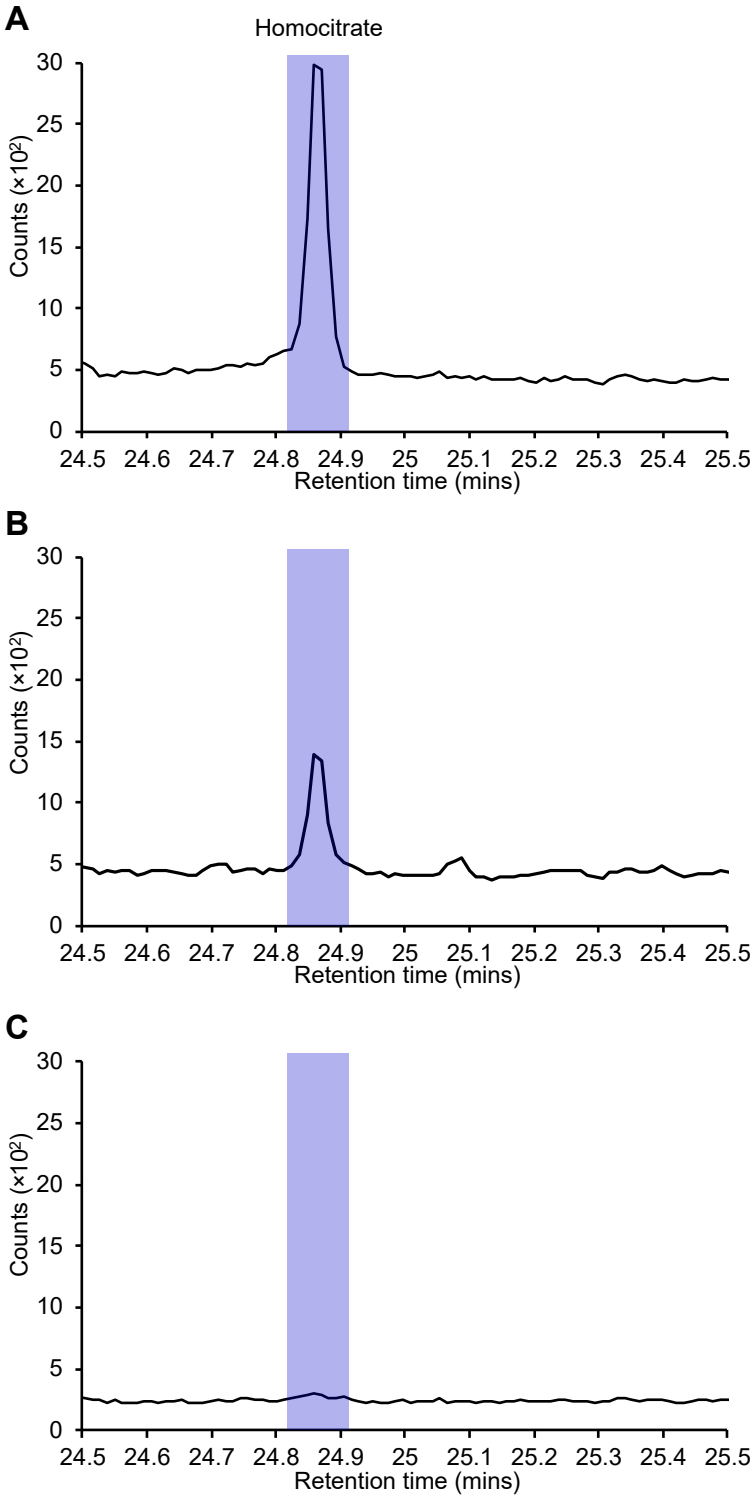


Figure 5.27: Representative homocitrate SIM-mode GC-MS chromatograms of *A. caulinodans* strains. Representative chromatograms of **(A)** *A. caulinodans* WT, **(B)** *A. caulinodans* Δ*nifA* and **(C)** a blank sample. Homocitrate was detectable in both *A. caulinodans* strains, and no substantial background was present in the blank sample.

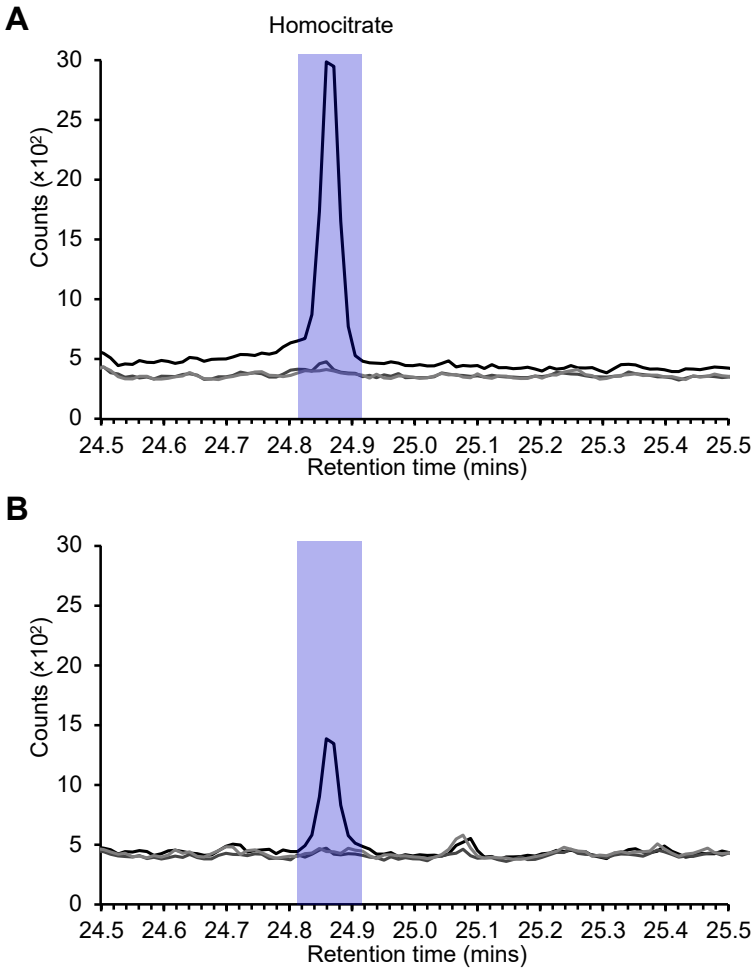


Figure 5.28: Degradation of GC-MS homocitrate peak in *A. caulinodans* samples. Chromatograms of repeated readings of *A. caulinodans* **(A)** WT and **(B)** Δ*nifA* by SIM-mode GC-MS. Earlier readings are indicated by a darker colour, later readings by a lighter colour. Readings are separated by approximately 80 minutes. WT homocitrate readings were consistently higher than those from the Δ*nifA* strain, but in both cases a sharp degradation in signal over time was observed. In the third and final reading, no homocitrate could be detected from either sample.

from free-living Rlv3841. Strains were grown up according to a recently published protocol and acetylene reduction activity measured (see Chapter 2 for details) [336].

As expected, Rlv3841 WT produced no acetylene reduction activity, nor was any detected in Rlv3841 WT carrying any of the three *nifV* expression vectors (Figure 5.32). Acetylene reduction was detected from Rlv3841 Δ*nifA* with the combination of induced pLMB509-based $nifA_{\text{Rlv3841}}$ and a Tn7-integrated $P_{A1lacO1}$-$nifV_{\text{ORS571}}$ cassette, but not with the P_{lac}-$nifV_{\text{ORS571}}$ cassette nor the pJP2neo-based $nifV_{A.\ vinelandii}$ expression vector. In Rlv3841 Δ*nifA* with induced pLMB509-based

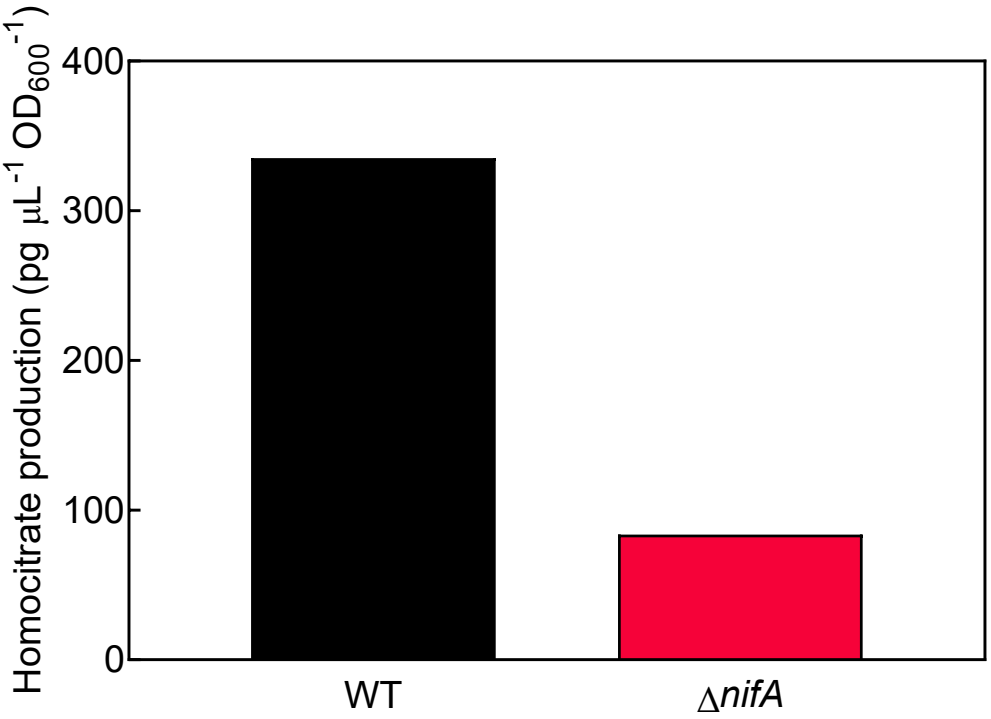


Figure 5.29: Homocitrate concentration in *A. caulinodans* strains. Homocitrate levels determined by SIM-mode GC-MS. More homocitrate concentration was recorded in *A. caulinodans* WT (black) than in the Δ*nifA* mutant (red). Because of sample degradation (see Figure 5.28), only the first reading for each strain was used.

$nifA_{\text{ORS571}}$, activity was detected only in combination with the Tn7-integrated $P_{A1lacO1}$-$nifV_{\text{ORS571}}$ cassette and the pJP2neo-based $nifV_{A.\ vinelandii}$ expression vector. However, acetylene reduction activity was minimal in all three combinations where it could be detected. Further, several biological replicates for all three combinations showed no activity. Consequently, although some free-living nitrogen fixation could be detected, it is not possible to determine whether there is a significant difference in the activity enabled by the various *nifA* and *nifV* expression systems used.

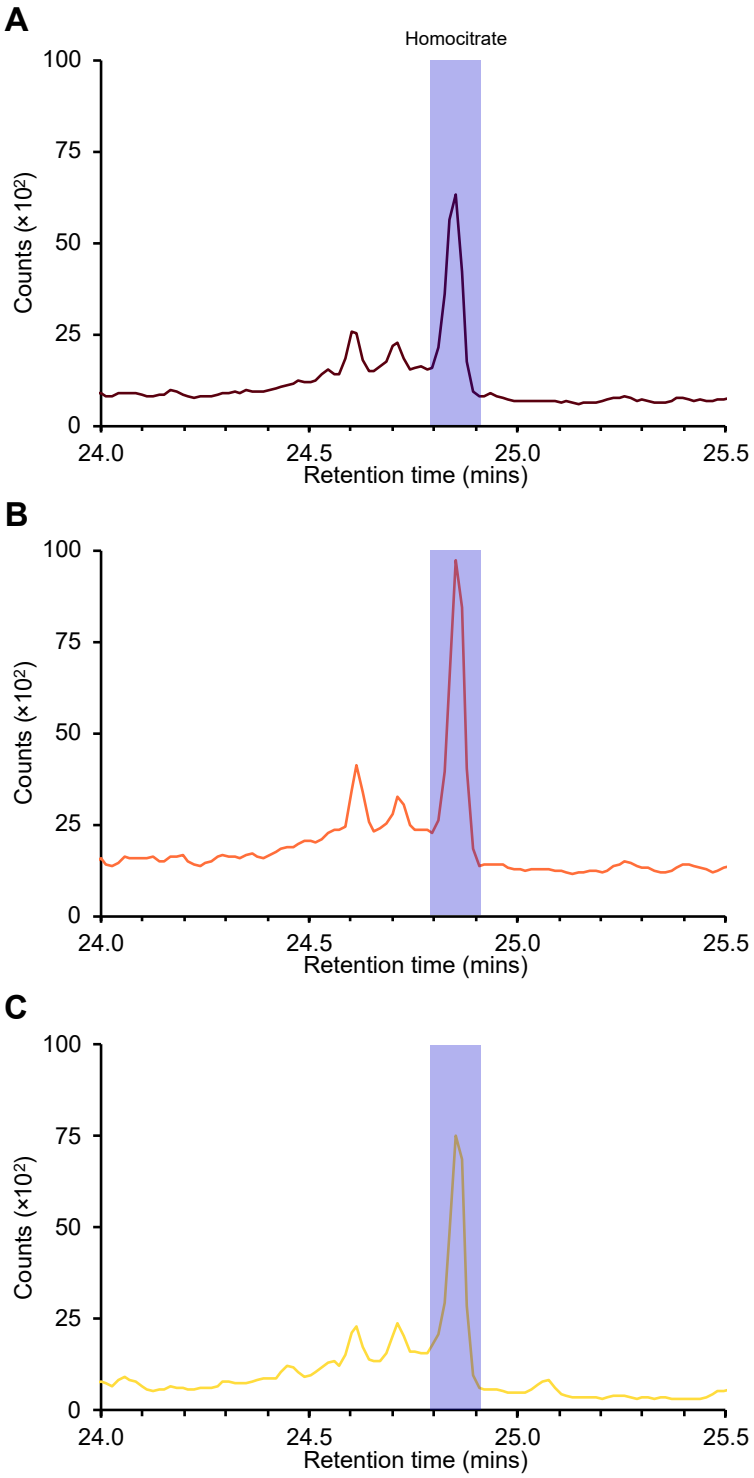


Figure 5.30: Representative homocitrate SIM-mode GC-MS chromatograms of Rlv3841 with *nifV* expression vectors. Representative SIM-mode GC-MS chromatograms of Rlv3841 Δ*nifA* (OPS1737) with **(A)** Tn7-integrated P_{lac}-$nifV_{ORS571}$ (OPS2607, dark orange), **(B)** Tn7-integrated $P_{A1lacO1}$-$nifV_{ORS571}$ (OPS2608, orange) and **(C)** P_{neo}-$nifV_{A.\ vinelandii}$ on a pJP2neo derived vector (OPS2604, yellow). Homocitrate production was detectable from all three strains (blue highlight). Minimal differences were observed in homocitrate peak area.

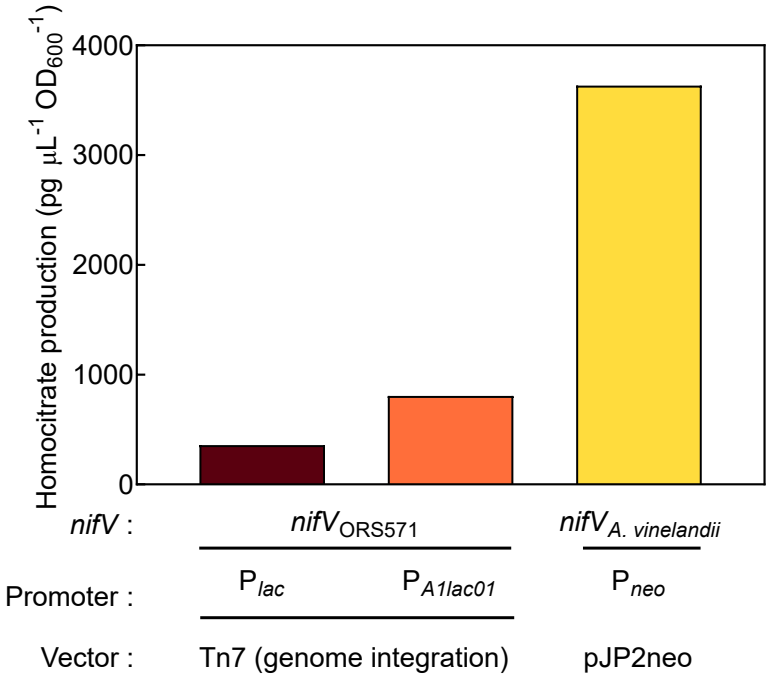


Figure 5.31: Homocitrate concentration in Rlv3841 with *nifV* expression vectors. Homocitrate levels determined by SIM-mode GC-MS. A similar level of homocitrate was produced by strains with P_{lac}-$nifV_{\mathrm{ORS571}}$ (dark orange), $P_{A1lacO1}$-$nifV_{\mathrm{ORS571}}$ (orange) and P_{neo}-$nifV_{A.\ vinelandii}$ (yellow), designated OPS2607, OPS2608 and OPS2604, respectively. One reading was taken for each strain.

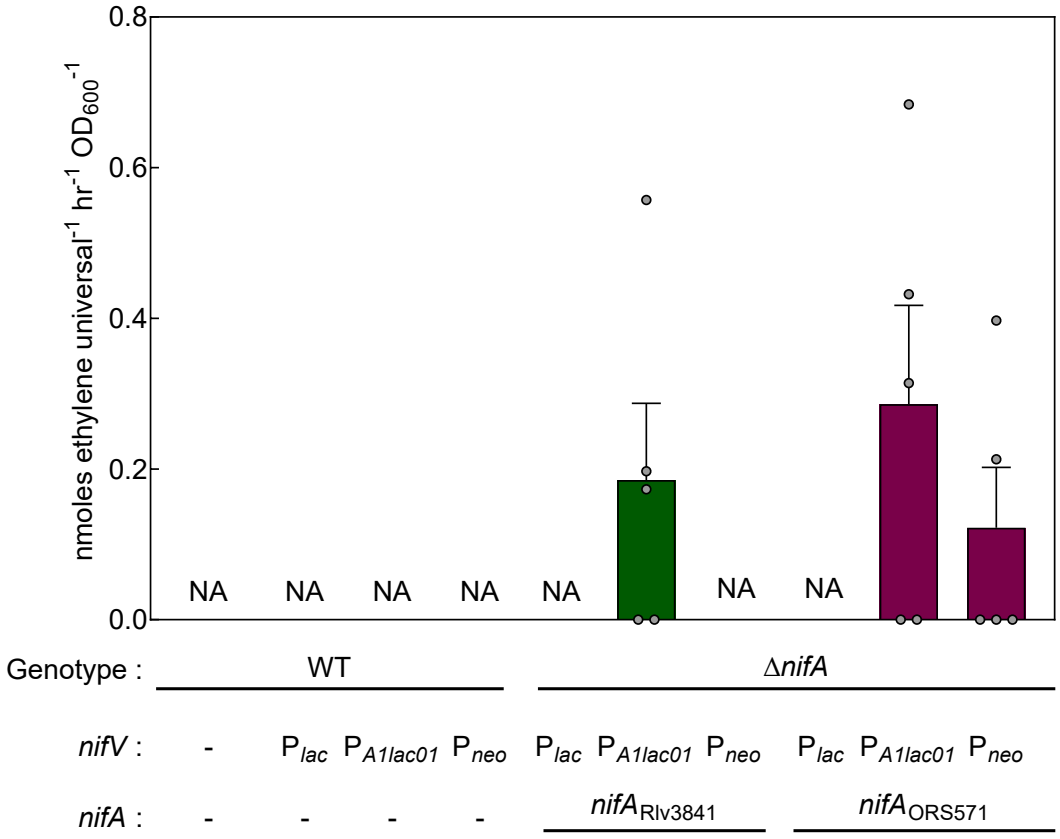


Figure 5.32: Acetylene reduction activity in Rlv3841 strains containing *nifA* and *nifV* vectors. No activity was detected in Rlv3841 WT, with or without introduced *nifV* expression. In an Rlv3841 Δ*nifA* host, activity was detected in the presence of the Tn7-integrated $P_{A1lacO1}$-$nifV_{\mathrm{ORS571}}$ cassette in combination with either the $nifA_{\mathrm{Rlv3841}}$ or $nifA_{\mathrm{ORS571}}$ pLMB509-based expression vectors (pOPS1591 and pOPS0983 respectively). Some activity was also detectable when the $nifA_{\mathrm{ORS571}}$ expression vector was paired with the P_{neo}-$nifV_{A.\ vinelandii}$ cassette in a pJP2neo backbone (pLMB504). In all three combinations where activity was recorded, several biological replicates showed no fixation. NA, no activity detected. n = 5 biological replicates were tested.

5.3 Discussion

Its central role as the activator of nitrogen fixation makes NifA a key target for engineering efforts which seek to modify how this activation is regulated. NifA is regulated at both the transcriptional and protein level. Its transcriptional regulation has been studied in the past and found to be relatively well conserved across symbiotic diazotrophs [1, 41]. However, protein-level regulation is less well studied, and in particular the protein-level regulation of NifA from symbiotic rhizobia remains poorly understood. We bypassed native transcriptional regulation to study the activity of rhizobial NifA proteins and how this is regulated.

Some activity was detectable when $NifA_{Rlv3841}$ was expressed from a taurine-induced pLMB51 backbone, but this was limited (Figure 5.4A). Switching to the pLMB509 backbone improved activity, at the cost of increased uninduced expression (Figure 5.10). This was particularly evident in the colony morphology changes caused in Rlv3841 by overproduction of native NifA from the uninduced plasmid (Figure 5.18). Given the high *nifA* expression expected from induced pLMB509 but the relatively low activity recorded in all conditions, expression may have been excessive in this backbone, and one or more of the NifA variants may have been misfolding and/or forming inclusion bodies. We did not explore this possibility, but this would be important if pLMB509 was to be used in further NifA work. Both pLMB51 and pLMB509 required notably more taurine and produced more uninduced expression in *A. caulinodans*, suggesting their behaviour can vary significantly between hosts (Figures 5.3 and 5.9). Despite producing sufficient NifA for activity to be detectable, the high uninduced expression and inter-host variability of the pLMB509 backbone are important drawbacks. The plasmid is unlikely to be suitable for large-scale studies involving multiple host organisms, and its high basal expression can cause substantial growth defects even in the absence of taurine, as we observed in both Rlv3841 and *A. caulinodans* (Figures 5.9 and 5.10).

Both full-length $NifA_{Rlv3841}$ and $NifA_{ORS571}$ were active under several *in vitro* conditions tested (Figure 5.17). Past work has shown that some related *Rhizobium* species have no NifA GAF domain, and indeed its deletion from *E. meliloti* NifA

increased the protein's activity [244, 277, 339]. It was therefore theorized that the domain was dispensable in rhizobial NifA proteins [1, 262, 276]. However, the truncated $NifA_{Rlv3841}$ ΔGAF protein showed little to no activity, indicating that the GAF domain is important for NifA activity in this species (Figures 5.4 and 5.17). In some non-symbiotic diazotrophs, the GAF domain is required for NifA activity as it relieves inhibition by NifL [240, 252, 263, 264]. Many non-symbiotic diazotrophs have no NifL protein, and the role of the NifA GAF domain is not known in these organisms. Activity from the truncated Rlv3841 NifA protein remained low outside its native host and we therefore have no evidence that the domain acts to de-repress NifA from a native regulator. The consistent loss of activity across hosts may also be due to protein misfolding in the absence of the GAF domain, which has been previously reported [262]. Our data suggests the protein retains the ability to fold, as limited activity was recorded in Rlv3841 (Figure 5.17). Further, the total loss of nitrogen fixation in *A. caulinodans* Δ*nifA* expressing $nifA_{Rlv3841}$ ΔGAF suggests the inactive protein was folding sufficiently to bind and sterically repress the expression of key genes (Figure 5.19). Further work on this avenue of the project could investigate the activity of *R. leguminosarum* bv. *trifolii* NifA ($NifA_{Rlv\ trifolii}$) in Rlv3841; this protein has no GAF domain but is functional. Symbiotic nitrogen fixation in Rlv3841 Δ*nifA* expressing $nifA_{Rlv\ trifolii}$ would suggest that whatever function the native domain performs is not critical for fixation, or the opposite. More in-depth studies could be performed to identify more nuanced changes that may occur when $nifA_{Rlv\ trifolii}$ regulates nitrogen fixation in Rlv3841. For example, confocal microscopy of P_{nifH} or P_{fixA} expression in nodules with Rlv3841 Δ*nifA* expressing $nifA_{Rlv\ trifolii}$ may show an altered pattern of expression that could be informative of the native GAF domain function.

Despite detecting NifA activity in several conditions, it is likely that both the Rlv3841 and *A. caulinodans* proteins remained largely inactive in most conditions tested. The P_{fixA} and P_{nifH} promoters used as reporters are typically very highly upregulated during symbiotic nitrogen fixation, with induction folds over 100 reported from microarray work in Rlv3841 [377]. In our hands, we observed activity

from these reporters that was on par with the constitutive J23115 promoter, a relatively weak constitutive promoter in Rlv3841 (Figure 5.17). Fluorescent reporter work with pLMB51 and pLMB509 showed that the latter produced significantly more protein when induced, as expected given its higher copy number (Figures 5.2 and 5.8). However, we observed only a relatively small increase in NifA activity when expressing it from the pLMB509 backbone instead of the pLMB51 backbone, suggesting most NifA produced is inactivated at the protein-level (Figure 5.10). Another possibility is poor activity from the Rlv3841 P_{fixA} and P_{nifH} reporters. Sequences upstream of the promoter regions cloned may be required for full activity, as work in our group has demonstrated [395]. Alternatively, there may be currently unknown regulators in Rlv3841 which repress the promoters under free-living conditions, resulting in low activity in the conditions we tested. This could be investigated by using non-native promoters, such as the *A. caulinodans* P_{fixA} or P_{nifH} reporters in Rlv3841, which may escape native regulation and therefore be more active.

The two main known regulators of NifA proteins are fixed nitrogen concentration and O_2 concentration. We observed increased activity from both NifA proteins studied as fixed nitrogen concentration dropped (Figure 5.17). This was expected from the *A. caulinodans* protein, which is known to sense and be repressed by high levels of fixed nitrogen at the protein level [239, 278]. Unlike $NifA_{Rlv3841}$, activity from $NifA_{ORS571}$ was detectable even at 10 mM ammonium chloride, indicating limited nitrogen sensitivity (Figures 5.10 and 5.11). This agrees with the current consensus that a substantial portion of fixed nitrogen NifA regulation in this organism occurs at the transcriptional level [166, 239, 284, 336]. The nitrogen sensitivity of $NifA_{Rlv3841}$, and other NifA proteins from rhizobia which fix only under symbiotic conditions, had not previously been studied. It was generally theorized that these proteins would not be subjected to fixed nitrogen repression, to enable fixation to continue in a symbiotic context where high levels of fixed nitrogen would be produced. However, we observed a marked increase in Rlv3841 NifA activity as the level of ammonium chloride decreased, suggesting the protein was

regulated at the protein level by fixed nitrogen levels (Figure 5.17). It is unknown how this regulation operates. Past work with NifA proteins has shown that the GAF domain plays a role in regulating the protein in response to fixed nitrogen levels. We found no evidence that the absence of the GAF domain altered the fixed nitrogen sensitivity of truncated Rlv3841 NifA (Figure 5.17).

The poor activity of the three NifA variants in Rlv3841, and their sensitivity to fixed nitrogen levels, may be partially due to the activity of other proteins (or lack thereof). One candidate for this indirect modulation is the RpoN σ^{54} factor which NifA requires to function [259, 309, 399]. Min et al. recently reported that expressing *rpoN* as well as *nifA* could improve the latter's activity when both were expressed from a plasmid [336]. This suggests that RpoN production is a limiting factor in $\text{NifA}_{\text{ORS571}}$ activity in their conditions and may be upregulated during nitrogen fixation. Microarray work in Rlv3841 to date has suggested *rpoN* is upregulated in 21-day old bacteroids compared to free-living bacteria, albeit only slightly (1.95-fold upregulated, p=0.049) [377]. Expressing $rpoN_{\text{Rlv3841}}$ from a vector as well as *nifA* could thus improve on the activity recorded in this study. This is unlikely to influence the nitrogen response we observed, as there is currently no evidence to suggest *rpoN* expression or RpoN activity would be upregulated when Rlv3841 is grown under low fixed nitrogen conditions. RpoN is more typically upregulated in low-O_2 conditions, and indeed an anaerobox is located upstream of $rpoN_{\text{Rlv3841}}$ [152, 191]. Thus, upregulation of $rpoN_{\text{Rlv3841}}$ may account for some of the increased NifA activity observed at 1% O_2 relative to 21% O_2.

Another candidate for indirect modulation of NifA activity is the NtrC protein (*glnG*), particularly under nitrogen-limiting conditions. NtrC has not been extensively studied in rhizobia but has received attention in the free-living diazotroph *K. pneumoniae* [400]. NtrC is activated under nitrogen-limiting conditions by NtrB [401, 402]. Like NifA, NtrC activates transcription from σ^{54} promoters [403, 404]. Both appear to induce transcription in part by bending DNA in the region of the promoter, but bind separate UAS [269, 405, 406]. Despite binding different sites, studies have shown that NtrC is able to supplement NifA activity, inducing

NifA-regulated promoters [407]. Notably, Sundaresan et al. found that *E. coli* NtrC could activate an *E. meliloti* P_{nifH} reporter [243]. Rlv3841 contains both *glnG* and *ntrB* homologs [152]. Thus, some of the reporter activity we observed may have been due to NtrC regulation, particularly under low-nitrogen conditions. Expression from the P_{fixA} and P_{nifH} promoters may have been limited by NtrC availability, explaining why large increases in *nifA* expression did not produce correspondingly large increases in reporter expression. Further work is necessary to determine whether $NifA_{Rlv3841}$ is directly sensing fixed nitrogen levels or whether its activity is being indirectly regulated. Studying reporters in an Rlv3841 Δ*nifA* Δ*glnG* double mutant background would help clarify what, if any, effect NtrC has on P_{fixA} or P_{nifH} activity.

Alongside regulation by fixed nitrogen levels, both NifA proteins are regulated by O_2, and this likely accounts for most of the inhibition taking place under the *in vitro* conditions used. Both proteins contain a cysteine-rich cluster in their IDL that enables the proteins to respond directly to O_2 concentration [280–282]. Due to equipment and biological constraints, it was not possible to reproduce the nanomolar O_2 concentrations in which $NifA_{Rlv3841}$ operates during symbiosis. By contrast, *A. caulinodans* is able to fix nitrogen under free-living conditions, with optimal activity occurring in the 1-3% O_2 range [185, 186, 336]. It was therefore possible that $NifA_{ORS571}$ was substantially less sensitive to O_2 than $NifA_{Rlv3841}$. We observed an increase in the activity of both proteins at 1% vs 21% O_2 (Figure 5.18). However, this change was minimal, and may have been partly due to increased *rpoN* expression under the action of the hFixL-FxkR-FixK system at 1% O_2. Based on the limited range of O_2 concentrations tested in this work, we therefore found no evidence that $NifA_{ORS571}$ is significantly less O_2 sensitive than $NifA_{Rlv3841}$. To enable free-living fixation at relatively high O_2 levels, *A. caulinodans* is known to create a very low internal O_2 environment [168]. No measurements of internal O_2 levels have been published, but it is possible that during free-living nitrogen fixation $NifA_{ORS571}$ operates in a microaerobic environment comparable to that experienced by $NifA_{Rlv3841}$ inside bacteroids. Thus, there may be relatively little

difference in the O_2 sensitivity of the two proteins, and both were likely mostly inactivated by O_2 even at 1% O_2.

Past work has demonstrated in many different contexts that NifA proteins are cross-compatible across different species, with the ability to activate nitrogen fixation genes and even nitrogen fixation in non-native hosts [243–247]. In Rlv3841, $NifA_{ORS571}$ was able to activate both the Rlv3841 P_{fixA} and P_{nifH} reporters (Figures 5.12 and 5.17). We also consistently observed more activity from $NifA_{ORS571}$ in Rlv3841 than from the native $NifA_{Rlv3841}$ protein (Figure 5.17). This may be due to inherent differences in activity but is more likely a reflection of regulatory differences. As well as potentially being less inhibited at the fixed nitrogen and O_2 levels tested here, $NifA_{ORS571}$ may not be subject to native regulatory elements inhibiting the activity of $NifA_{Rlv3841}$. In contrast, we observed no $NifA_{ORS571}$ activity when the protein was expressed from a pLMB509 backbone in its native host (Figures 5.14 and 5.15). A significant growth burden was imposed by this NifA expression plasmid, particularly when induced (Figure 5.14B). It is possible $NifA_{ORS571}$ activity was obscured by the effects of this burden. However, the lack of activity also suggests that native regulatory elements, absent in Rlv3841, repressed $NifA_{ORS571}$ in its native host.

Despite the activity of $NifA_{ORS571}$ in Rlv3841, several findings highlight potential limitations to cross-host NifA compatibility. Overexpression of $nifA_{ORS571}$ in its native host enabled fixation from a Δ*nifA* mutant (Figure 5.21). However, we found no significant increase in nitrogen fixation when the *A. caulinodans* Δ*nifA* mutant was complemented with $NifA_{Rlv3841}$, suggesting the protein is not or insufficiently active in this host to enable fixation. Further, the morphological changes observed in Rlv3841 Δ*nifA* colonies expressing native $NifA_{Rlv3841}$ but not $NifA_{ORS571}$ suggests some genes in the NifA regulon were being differentially activated by the native protein compared to the non-native NifA (Figure 5.19). The nature of these differences was not pursued in this project but could be readily investigated through techniques such as RNA-seq. Imperfect cross-compatibility would have important implications for engineering efforts which employ non-native NifA proteins to bypass

native regulatory mechanisms. The symbiotic context may be better suited for studying the cross-compatibility of NifA variants in Rlv3841. One avenue to further explore cross-compatibility and the potential of altered *nifA* regulation would be to integrate $nifA_{\text{Rlv3841}}$ and $nifA_{\text{ORS571}}$ in the genome of Rlv3841 Δ*nifA* using the Tn7 system and measure its symbiotic nitrogen fixation. Genomic integration would likely alleviate much of the leaky production seen in pLMB509 and reduce cell burden. Expression could be regulated by the native *nifA* promoter or placed under the control of a constitutive or anaerobox-controlled promoter. Such work would help determine whether $\text{NifA}_{\text{ORS571}}$ could activate symbiotic nitrogen fixation in Rlv3841.

SIM-mode GC-MS was sufficiently sensitive to assay homocitrate synthase activity without the need for NifV purification (Figure 5.26). Although degradation of samples was observed over time, the technique enabled readings of homocitrate concentration from cell extract samples, allowing us to determine the level present in nitrogen-fixing *A. caulinodans* (Figures 5.28 and 5.29). All three *nifV* expression vectors tested produced homocitrate in Rlv3841 (Figure 5.30). Homocitrate concentrations in Rlv3841 were substantially in excess of the level found in nitrogen fixing *A. caulinodans*, as expected given the inducible overexpression of *nifV* (Figure 5.31).

Recent work has shown that symbiotic diazotrophs can fix nitrogen in free-living conditions with refactored nitrogen fixation gene clusters under artificial inducible control [336]. We combined inducible *nifA* and *nifV* expression in Rlv3841 and found that acetylene reduction activity was detectable in some of the combinations tested (Figure 5.32). This is in contrast to a recent study by Min et al. which was unable to engineer free-living nitrogen fixation in Rlv3841 [336]. We relied on the native nitrogen-fixation machinery of Rlv3841, whereas the approach of Min et al. was based on inserting nif clusters from free-living diazotrophs into rhizobia. Some of the genes in these clusters, or their protein products, may not have been functional in Rlv3841. Many other factors, including growth conditions, could also have contributed to the lack of nitrogen fixation, and more work would be needed to identify the failure mode. Min et al. observed activity in most of the organisms into which *nif* clusters were transferred, suggesting Rlv3841 may be unusually

resistant to this approach. We found that free-living Rlv3841 nitrogen fixation activity was lower than that in *A. caulinodans*, and inconsistent across biological repeats. Further investigation will be needed to robustly determine whether any difference in fixation activity exists between the various *nifV* expression systems, and between $NifA_{Rlv3841}$ and $NifA_{ORS571}$. The acetylene reduction activity observed from strains with the pJP2neo-based $nifV_{A.\ vinelandii}$ vector was relatively low given the high concentrations of homocitrate detected from this vector. However, this may be due to the growth defect observed from this vector; excessive NifV activity may be causing toxicity. Given that all three *nifV* systems produced a similar, likely excessive, homocitrate concentration in Rlv3841, nitrogen fixation activity is likely limited by NifA activity. Other factors, such as the internal O_2 concentration of Rlv3841, may also play a role in determining its ability to support free-living nitrogen fixation. Free-living acetylene reduction assays in this study were performed in UMS media with glucose as the carbon source (see Chapter 2 for details). In the future, using succinate as the carbon source instead would more closely mimic conditions in nodules and force cells to consume more O_2, potentially creating better conditions for free-living fixation [408]. In summary, our findings demonstrate that altered *nifA* regulation can enable free-living nitrogen fixation in Rlv3841. Activity was however limited, and there are several potential avenues for future engineering work to improve on the acetylene reduction levels observed in this study.

6 Discussion

Contents

O_2 regulation is essential for rhizobia to establish a successful symbiosis with their legume partners [36–38]. Several O_2 regulation systems have been studied in rhizobia, and all appear to employ multiple systems [41, 115]. NifA appears to be ubiquitous, typically paired with some combination of the FixLJ-FixK, hFixL-FxkR-FixK and FnrN O_2 sensing systems [121–124, 232]. The model *Rhizobium* Rlv3841 employs three O_2 sensing systems: the hFixL-FxkR-FixK pathway, FnrN and NifA. We studied all three systems in Rlv3841 through a combination of studies under free-living and *in planta* conditions. Aspects of all three systems were characterized to better understand their behaviour, interactions and potential as tools for engineering nitrogen fixation.

6.1 O_2 regulation by the hFixL-FxkR-FixK pathway and FnrN in Rlv3841

The first area we explored was the interaction of O_2 regulation systems. Sequence data and past work has consistently shown that rhizobia not only employ multiple systems but also that these interact [1]. However, to date the relationship between coexisting systems had not been well studied, and the functional importance of these relationships was unknown. When we began this work, it was unclear whether multiple O_2 regulation systems in the same organism were redundant or hierarchical, and how their inter-connectedness manifested during symbiosis. To investigate this, we studied the hFixL-FxkR-FixK and FnrN O_2 regulation systems of Rlv3841 under free-living conditions and *in planta*. hFixL appears to be more O_2 tolerant than either FnrN or NifA; it is active in free-living bacteria under microaerobic conditions and *in planta* beginning in zone I of nodules. The FnrN protein is largely inactive under free-living microaerobic conditions and only becomes fully active from the II-III interzone onwards in nodules. Thus, we found that the hFixL-FxkR-FixK pathway is active in the earliest stages of symbiosis, followed by FnrN as the bacteria move to areas of the nodule proximal to the root. Both regulate genes required for effective symbiosis. FnrN is critical for expression of *fixNOQP* and nitrogen fixation activity. Indirectly, the hFixL-FxkR-FixK pathway also plays an important role by inducing *fnrN* expression under microaerobic conditions, priming it for auto-activation in the central nitrogen fixing zone. We also worked with collaborators to model regulation by the hFixL-FxkR-FixK pathway and FnrN in Rlv3841 [2]. This work suggests that the induction of *fnrN* by the hFixL-FxkR-FixK pathway prevents bistability in the response of Rlv3841 to low O_2, thereby ensuring all cells commit to *fixNOQP* expression in the central nitrogen fixing zone. Taken together, our findings show that hFixL-FxkR-FixK and FnrN act as a single regulatory pathway which integrates both O_2 sensors. Both systems are required for full nitrogen fixation activity, and their different sensitivity to O_2 give rise to spatiotemporally distinct roles during symbiosis.

Like Rlv3841, it is common for other rhizobia to employ multiple O_2 regulation systems. There are several possible advantages that may have contributed to the evolution of multiple systems and their inter-connection. The presence of several O_2 regulation systems creates redundancy, a feature often found in key regulatory pathways to improve their robustness [183, 188, 373, 374]. Redundancy is not limited to the FixLJ-FixK, hFixL-FxkR-FixK and FnrN systems; a recently discovered strain of *B. japonicum* was shown to encode redundant homologs of *nifA* [313]. Despite closely integrating the hFixL-FxkR-FixK pathway and FnrN, partial redundancy remains in Rlv3841. The strain encodes multiple homologs of *hfixL*, *fxkR* and *fixK* [152]. To abolish nitrogen fixation activity, we found that both *hfixL* and *fnrN* must be disrupted.

Besides redundancy, the use of multiple inter-connected systems can also provide other advantages. Our results in Rlv3841 demonstrate that each O_2 sensor plays an important, distinct role. Rhizobia experience a drop in O_2 concentration of at least three orders of magnitude as they transition from a free-living lifestyle in soil to terminally differentiated bacteroids in nodules [56, 59]. Integrating two sensors into a single cascade in Rlv3841 enables a nuanced, gradual response across the entire range of O_2 concentrations experienced during symbiosis. This is not limited to Rlv3841; both *A. caulinodans* and *E. meliloti* do not encode *fnrN* but have the FixLJ TCS, and use it to induce transcription of auto-activating *nifA* [124, 239, 284]. This arrangement has striking similarities to that of hFixL-FxkR-FixK and *fnrN* in Rlv3841. There is no evidence that *nifA* is anaerobox or K-box regulated in Rlv3841 [152, 395]. Integration of multiple O_2 sensors is not limited to rhizobia. A similar dual-sensor arrangement has previously been described in *Rhodopseudomonas palustris*, which combines a FixLJ-FixK pathway with the FnrN homolog AadR [409–411]. *R. palustris* is not symbiotic but is noted for its ability to grow under both aerobic and anaerobic conditions [412]. The combined pathway in *R. palustris* was shown to provide fine-tuned regulation, allowing the organism to adapt to the large range of O_2 concentrations it experiences. Using multiple sensors

to produce a more nuanced and responsive O_2 regulation system thus appears common in organisms that experience a wide range of O_2 concentrations.

Finally, the prevalence of multi-sensor O_2 regulation in rhizobia may also have arisen in response to competitive fitness pressures. Legume plants can sanction rhizobia based on their nitrogen fixation activity [413–416]. We speculate the bacteria may also be selected based on the speed with which they are able to adapt to life inside nodules and begin productively fixing nitrogen. This would create pressure for strains to rapidly demonstrate their effectiveness to their legume host. Past work has suggested that one of the benefits of FnrN compared to FixLJ-FixK is that it is more responsive to O_2 concentration, providing more flexible regulation [205, 206, 417]. Shifting *nifA* regulation from the level of transcription to protein-level control, as seen in some rhizobia, may also serve to produce more responsive regulation. It has been suggested legumes decrease the pH of symbiosomes as a method to force rhizobia to metabolize succinate and produce ammonia [418]. Both these activities counteract acidification but suggest regulation must act rapidly to allow the rhizobia to survive. Taken together, these findings indicate that integrating multiple O_2 regulation systems to produce more nuanced control may speed up the symbiotic transition, providing a competitive advantage.

6.2 Activity of Rlv3841 NifA outside symbiosis

We also studied the central activator of nitrogen fixation, NifA [41, 52, 234–236]. The protein has been well-studied in free-living diazotrophs, notably *K. pneumoniae*, but not in symbiotic diazotrophs such as Rlv3841 [1, 237]. Unlike the NifA protein of free-living diazotrophs, NifA in symbiotic diazotrophs can directly sense O_2 via an inter-domain linker region [244, 280]. NifA is generally considered the final activator of nitrogen fixation in bacteroids, active only in the core of nodules [98, 233, 379–381]. The O_2 concentration in this part of nodules is at nanomolar levels [56, 59]. Thus, it is generally thought that diazotrophs which only fix inside plants have a NifA protein that is inactivated at all but these extremely low O_2 concentrations. However, some work in *E. meliloti*, which fixes only during symbiosis, indicated its NifA protein was already active during early nodule development [382]. Other work has suggested that the NifA of symbiotic diazotrophs can be active outside the nodule environment [313, 336]. When we began this work, it was therefore unclear what the O_2 sensitivity of rhizobial NifA proteins is. Another important area of uncertainty in symbiotic NifA regulation is the function of the N-terminal GAF domain that most rhizobial NifA proteins have [260–262, 277]. Studies in free-living diazotrophs have consistently found a regulatory role for the NifA GAF domain in these organisms [240, 252, 263, 264]. These studies also showed significant variation in what signal the domain responds to [271, 275]. The GAF domain is very poorly conserved across species, so no reliable inferences can be drawn from sequence information [41, 166]. In summary, the O_2 sensitivity of rhizobial NifA proteins was unknown, and the GAF domain likely represented a separate, largely unknown NifA regulation mechanism. As the central activator of nitrogen fixation, NifA is a very attractive control target. Understanding how it is regulated, by O_2 and other factors, could therefore have significant implications for ongoing efforts to engineering the *Rhizobium*-legume symbiosis.

We used a dual-plasmid system to bypass native *nifA* regulation and produce $NifA_{Rlv3841}$ and $NifA_{ORS571}$ inside free-living cells. We were able to detect activity from both proteins, albeit weakly. Our results thus suggest $NifA_{Rlv3841}$ can be

very weakly active outside of the nodule, at 1-21% O_2. Attempts to increase NifA activity by using a higher-copy vector for *nifA* expression produced only a minor improvement. NifA activity therefore did not appear to be limited by *nifA* expression. Activity increased when O_2 concentration was decreased, and when ammonium chloride concentration was decreased. However, both produced relatively small increases in NifA activity, and more work will be needed to confirm that these effects were not due to activity from a third party, such as a protein other than NifA. In particular, the response of $NifA_{Rlv3841}$ to ammonium chloride concentration does not accord with the current paradigm of nitrogen fixation regulation in symbiotic diazotrophs. It is generally thought that bacteroids should be blind to fixed nitrogen levels to enable them to sustain high levels of nitrogen fixation during symbiosis. One possibility is that the weak response of $NifA_{Rlv3841}$ to fixed nitrogen may have been produced by NtrC activation rather than NifA activation [407]. Another possibility is that NifA activity in Rlv3841 was constrained by insufficient RpoN concentration. Work by others has shown that overexpressing *rpoN* can improve NifA activity [336]. NifA has generally been found to be cross-compatible across diazotrophs [243–247]. In line with this, $NifA_{ORS571}$ was active in Rlv3841. However, we found no significant increase in nitrogen fixation in an *A. caulinodans* Δ*nifA* strain expressing $nifA_{Rlv3841}$. This may indicate incomplete cross-compatibility but may also be remediated by simultaneous *rpoN* overexpression.

We also investigated the role of the $NifA_{Rlv3841}$ GAF domain. Studies have found that deleting the GAF domain increases activity from some rhizobial NifA proteins, and abolishes it in others [244, 262, 276]. We found that deleting the GAF domain nearly abolished activity from $NifA_{Rlv3841}$. Past studies in free-living diazotrophs has found roles for the GAF domain in regulating NifA activity based on fixed nitrogen and O_2 concentration [244, 266, 268–271]. We found that deleting the domain did not influence the response of $NifA_{Rlv3841}$ ΔGAF to ammonium chloride or O_2 concentration. Our results suggest $NifA_{Rlv3841}$ is misfolded or unable to escape deactivation when the GAF domain is removed.

Some nitrogen fixation was detectable in free-living Rlv3841 with engineered *nifA* and *nifV* expression. As shown by recent work, this confirms that diazotrophs which ordinarily fix only under symbiotic conditions can be made to fix under free-living conditions [336]. However, the minimal levels detected show that much more engineering of Rlv3841 is needed to achieve robust free-living fixation. It is likely that NifA activity was a key limiting factor in our work, as homocitrate levels were substantially higher than those measured inside *A. caulinodans* during free-living diazotrophic growth.

6.3 Future work

Our work with the hFixL-FxkR-FixK and FnrN O_2 regulation systems of Rlv3841 has focused on the regulation of *fixK*, *fnrN* and the *fixNOQP* operons, which have been studied in the past [38]. This has elucidated how the two systems inter-connect in Rlv3841, but much remains to be discovered about the rest of their regulon. We found that *fixNOQP* expression in Rlv3841 is heavily reliant on both systems, and this may be sufficient to explain most of the loss of nitrogen fixation activity when they are disrupted. However, studies have shown that far more anaerobox-controlled genes exist besides those which are commonly used to study O_2 regulation in rhizobia [124–126]. Investigating this regulon via RNA-seq of Rlv3841 WT and O_2 regulation mutants could shed light on other O_2 regulated targets during symbiosis. Unpublished work by Patricia R. Rosas found that disrupting *fnrN* in *R. etli* CFN42 not only downregulated common targets such as *fixNOQP*, but also upregulated some genes, many of which are of unknown function [419]. This upregulation may be indirect, or may indicate that anaerobox-controlled repression, as seen in Rlv3841 with *fnrN* and likely $fxkR_c$, also controls other genes.

Past insertion sequencing (INseq) work in our group has found that a large number of genes in Rlv3841 are essential for growth under 1% O_2 but have no apparent impact on symbiosis [408]. With the exception of *fnrN*, the chromosomal O_2 regulation genes studied in this work generally had little to no importance for symbiotic nitrogen fixation, but may play a more important role in regulating other genes at 1% O_2. INseq results also found over 700 genes that were important for growth at 1% O_2 and symbiotic nitrogen fixation, indicating substantial overlap in the genes required for survival in these conditions.

RNA-seq could be used to determine whether FixK and FnrN have different anaerobox motif preferences. Our work in Rlv3841 found that FnrN is critical for symbiotic *fixNOQP* expression, whilst the hFixL-FxkR-FixK pathway was not. This may be due to differences in the *in planta* activity of these systems, but could also reflect divergent anaerobox binding preferences. If FixK and FnrN do have divergent preferences, this could create semi-separate regulons, with some

anaerobox-regulated genes preferentially induced by hFixL-FxkR-FixK during early symbiosis, whilst others are induced from the II-III interzone onwards by FnrN. Understanding which genes are part of which regulon in such a scenario would provide important information about the changes rhizobia undergo as they adapt to life inside nodules, and the order in which they happen.

Other pathways involved in symbiotic oxygen regulation remain to be discovered and understood. The NtrR protein of *S. meliloti* regulates microaerobic *nif* and *fix* gene induction as well as several unrelated metabolic functions, but it is unknown what signal the protein responds to [420]. An intermediate metabolite of purine biosynthesis represses *fixK* in *E. meliloti* through an as yet unidentified pathway [421]. A purine metabolism mutant of *R. etli* was found to express the *fixNOQP* operon under free-living conditions [422, 423]. This and other connections likely integrate as-yet unidentified signals important in the regulation of nitrogen fixation, which may include signals from the legume host. These additional inputs no doubt play important roles in supplementing, modifying or balancing oxygen regulation, and understanding them will be essential to arrive at a holistic picture of nitrogen fixation control in rhizobia.

Broader questions also remain about the variety of O_2 regulation mechanisms found in rhizobia. Rlv3841 employs the hFixL-FxkR-FixK pathway as well as FnrN and integrates the two O_2 regulation systems into a single hierarchical cascade. Past work suggests rhizobia near-universally encode multiple O_2 regulation systems, and inter-connect them [1]. Our work in Rlv3841 suggests that systems with different O_2 sensitivities are connected hierarchically, but more work in other organisms will be needed to confirm whether this is the typical situation. Another largely unexplored question is whether the diversity of O_2 regulation arrangements in rhizobia produces corresponding functional diversity. Different species use different O_2 sensors and connect them in different ways. This may be the result of convergent evolution in a context with frequent horizontal gene transfer events: O_2 regulation in rhizobia may appear superficially varied, but these arrangements may produce broadly similar hierarchical functionality [424–426]. Alternatively, these variations

may reflect genuine functional differences driven by diverse regulatory needs. These may be arising from different soil or environmental conditions, or different nodule O_2 gradients created by legume species.

Several avenues need to be explored to better understand the behaviour we observed in our NifA work in Rlv3841. Much is still unclear about how $NifA_{Rlv3841}$ is regulated. NtrC has been reported to activate NifA targets, and may be responsible for some of the effect of ammonium chloride concentration on reporter activity we observed [243, 407]. Our work could be repeated in an Rlv3841 *ntrC* mutant to investigate this possibility. The role of the $NifA_{Rlv3841}$ GAF domain remains largely unknown. As deletion of the domain appears to disable $NifA_{Rlv3841}$, future work should take a more nuanced approach. Nested deletions of the domain could be created and integrated into the chromosome of Rlv3841 Δ*nifA* via Tn7, then tested *in planta* for activity. This could be used to determine whether any deletions of the domain retain activity, after which these variants could be tested outside of the plant. As part of this work, Rlv *trifolii* NifA could also be integrated. This has no GAF domain, so represents a naturally occurring GAF-deleted NifA variant that could shed light on the role of the domain [339]. However, it is not known whether the protein is able to activate Rlv3841 promoters. Another avenue to explore would be determining what $NifA_{Rlv3841}$ binds. Past work on NifA GAF domains has found that they can interact with both small molecules and other proteins, so both would need to be investigated in $NifA_{Rlv3841}$ [240, 252, 263, 264, 271]. It is possible that the target bound by $NifA_{Rlv3841}$ GAF is only present in nodules, so this work may need to be performed on isolated bacteroids. Once the GAF domain binding target is identified, the function of this binding could be studied by point mutations in the GAF domain, disrupting binding without entirely disabling $NifA_{Rlv3841}$ activity.

Although minimal, the nitrogen fixation activity we detected in free-living Rlv3841 is useful as a proof-of-concept. To improve on this activity, increasing NifA activity is a priority, from $NifA_{Rlv3841}$ or another NifA variant. One option would be to co-express native Rlv3841 *rpoN*, which past work has shown can improve NifA activity [336]. Another change that may be beneficial is the use

of genomic integration instead of plasmid-based expression. We noticed burden effects with several vector-host combinations in this work, which genomic integration may alleviate. We saw no evidence that inducing *nifA* expression from pLMB509 impaired growth in Rlv3841, but this could occur if NifA activity was improved. Using genomic integration will also improve stability, facilitating *in planta* work.

Currently, it is unclear whether NifA or FnrN is the more O_2 sensitive, but it is likely that $NifA_{Rlv3841}$ is largely inactivated under free-living O_2 concentrations. Determining the O_2 sensitivity of $NifA_{Rlv3841}$ will be important to understand whether it is useful in the context of free-living fixation, or whether another NifA variant should be sought. Since Rlv3841 is aerobic, there is a limit to the O_2 concentrations that can be tested in this host without impairing respiration. Determining the O_2 sensitivity of $NifA_{Rlv3841}$ may therefore best be done in a facultative anaerobe like *E. coli* [399]. This would likely remove any native rhizobial NifA regulation mechanisms but allow a full range of O_2 concentrations to be tested.

Using NifA from one organism to regulate nitrogen fixation in another represents a potentially useful strategy to engineer nitrogen fixation but has not yet been systematically explored. If $NifA_{Rlv3841}$ is too O_2 sensitive to be used for free-living nitrogen fixation, high-throughput screening could be used to search for an alternative NifA protein. Studying a variety of NifA variants in Rlv3841, in free-living and symbiotic conditions, may not only find more active variants but also yield important insights about how protein-level regulation of NifA varies across species. Another approach would be to create a semi-rational library of NifA variants, produced through combinatorial assembly of GAF, AAA+ and HTH domains from different NifA proteins. This library could then be screened in Rlv3841 for NifA reporter output or free-living nitrogen fixation activity.

Finally, improving free-living nitrogen fixation activity in a symbiotic diazotroph like Rlv3841 may be possible through a directed evolution approach. In nitrogen-free media, nitrogen fixation should be a growth-advantageous phenotype. Thus, an engineered strain that weakly fixes under free-living conditions could be put through serial passaging with intermittent mutagenesis to improve fixation activity, useful

as both an engineering strategy but also with the potential to provide new insights into what changes to *nifA* and the strain's genome more broadly lead to improved fixation. However, it is likely that the level of nitrogen fixation attained in this work does not yet provide a sufficient growth advantage for such a strategy to be applied.

Biological nitrogen fixation has been extensively studied because of its importance to the biosphere and its potential for improving agriculture in the future [5–7]. The growth of many crop plants is limited by nitrogen availability [9–11]. This can be remedied with man-made nitrogen fertilizer, but the current production and use of these fertilizers is not sustainable [12–15]. In symbiosis with rhizobia, legumes can produce their own nitrogen fertilizer [3, 4]. Engineering this symbiosis into widely grown crops could make agriculture much more sustainable, and efforts have been underway for over half a century to achieve this [17–19, 21]. At the heart of biological nitrogen fixation is the nitrogenase complex, which catalyses the conversion of atmospheric dinitrogen into ammonia [39, 40]. The nitrogenase complex evolved over a billion years ago, when O_2 concentration in the atmosphere was negligible, and is inactivated in the presence of O_2 [49, 54]. Modern day biological nitrogen fixation revolves around accommodating this limitation [50, 79]. In legumes, it has led to the evolution of nodules, organic bioreactors which create the optimal near-anaerobic conditions for rhizobia to fix nitrogen [29, 30]. In turn, rhizobia have evolved a host of regulators and mechanisms which allow them to adapt and survive inside these nodules to fix nitrogen [1, 81–83, 115]. Until a new nitrogenase complex can be engineered which is adapted to the current composition of the atmosphere, understanding O_2 regulation in rhizobia is an essential part of engineering biological nitrogen fixation [20, 327]. Legume nodules create not merely a single low-O_2 zone but a gradient of O_2 concentrations, which rhizobia have evolved to exploit through inter-connected, multi-sensor regulation systems [92, 96, 387]. Only by understanding these systems will it be possible to transplant the *Rhizobium*-legume symbiosis into other crops.

Appendices

A
Publications

Contents

A.1 Appendix 1: Oxygen regulatory mechanisms of nitrogen fixation in rhizobia

CHAPTER NINE

Oxygen regulatory mechanisms of nitrogen fixation in rhizobia

Paul J. Rutten* **and Philip S. Poole**
Department of Plant Sciences, University of Oxford, Oxford, United Kingdom
*Corresponding author: E-mail: paul.rutten@plants.ox.ac.uk

Contents

Advances in Microbial Physiology, Volume 75
ISSN 0065-2911
https://doi.org/10.1016/bs.ampbs.2019.08.001

Abstract

Rhizobia are α- and β-proteobacteria that form a symbiotic partnership with legumes, fixing atmospheric dinitrogen to ammonia and providing it to the plant. Oxygen regulation is key in this symbiosis. Fixation is performed by an oxygen-intolerant nitrogenase enzyme but requires respiration to meet its high energy demands. To satisfy these opposing constraints the symbiotic partners cooperate intimately, employing a variety of mechanisms to regulate and respond to oxygen concentration. During symbiosis rhizobia undergo significant changes in gene expression to differentiate into nitrogen-fixing bacteroids. Legumes host these bacteroids in specialized root organs called nodules. These generate a near-anoxic environment using an oxygen diffusion barrier, oxygen-binding leghemoglobin and control of mitochondria localization. Rhizobia sense oxygen using multiple interconnected systems which enable a finely-tuned response to the wide range of oxygen concentrations they experience when transitioning from soil to nodules. The oxygen-sensing FixL-FixJ and hybrid FixL-FxkR two-component systems activate at relatively high oxygen concentration and regulate *fixK* transcription. FixK activates the *fixNOQP* and *fixGHIS* operons producing a high-affinity terminal oxidase required for bacterial respiration in the microaerobic nodule. Additionally or alternatively, some rhizobia regulate expression of these operons by FnrN, an FNR-like oxygen-sensing protein. The final stage of symbiotic establishment is activated by the NifA protein, regulated by oxygen at both the transcriptional and protein level. A cross-species comparison of these systems highlights differences in their roles and interconnections but reveals common regulatory patterns and themes. Future work is needed to establish the complete regulon of these systems and identify other regulatory signals.

1. Introduction

1.1 Role of oxygen in the *Rhizobium*-legume symbiosis

Rhizobia are soil dwelling α- and β-proteobacteria that form symbiotic partnerships with legume plants in which they fix inert N_2 to biologically accessible NH_3 in return for a carbon source (Oldroyd, Murray, Poole, & Downie, 2011; Poole, Ramachandran, & Terpolilli, 2018). The symbiosis has been intensely studied as a potential alternative to nitrogen fertilizers (Burén, López-Torrejón, & Rubio, 2018; Conway, 2000; Mus et al., 2016; Postgate, 1974). Nitrogen is a limiting factor in the growth of many crop plants and the use of these fertilizers has enabled dramatic improvements in yield (Canfield, Glazer, & Falkowski, 2010; Dobermann & Cassman, 2005; Vicente & Dean, 2017). However, the Haber-Bosch process producing nitrogen fertilizers consumes 1%–2% of the world's total energy supply. Their over-use is also causing global environmental damage due to excess fixed nitrogen and its reactive by-products entering ground water, oceans and the atmosphere (Erisman, Sutton, Galloway, Klimont, &

Winiwarter, 2008; Gruber & Galloway, 2008; Smil, 2001; Vitousek et al., 1997). Paradoxically the cost of these fertilizers makes them unaffordable in many areas, creating vast inequalities in crop yields across the world (Vitousek, Menge, Reed, & Cleveland, 2013). Their application might be reduced by increased use of legumes and supplementation with artificial rhizobial-crop symbioses or plants engineered to perform nitrogen fixation (Allen et al., 2017; Burén & Rubio, 2018; Cocking, Stone, & Davey, 2005; López-Torrejón et al., 2016; Rogers & Oldroyd, 2014). This would help meet the agricultural demands of a surging world population more sustainably and equitably but will require a comprehensive and detailed understanding of the mechanisms by which oxygen regulates nitrogen fixation (Alexandratos & Bruinsma, 2012; Geddes et al., 2015; Oldroyd & Dixon, 2014).

Rhizobia taking part in the symbiosis undergo a dramatic change in lifestyle during their move from soil into the symbiotic environment (reviewed in Downie, 2014; Murray, 2011). To initiate symbiosis the bacteria migrate toward plant roots, attach and subsequently enter them (reviewed in Andrews & Andrews, 2017; Oldroyd et al., 2011; Wheatley & Poole, 2018). Where rhizobial entry occurs, plants form specialized organs called nodules which create an environment in which nitrogen fixation can occur (see Gage, 2002; Oldroyd & Downie, 2008 for reviews). Rhizobia within nodules subsequently infect plant cells (Vance, 1983). Infected plant cells enclose rhizobia into vesicles called symbiosomes (Bassett, Goodman, & Novacky, 1977; Roth & Stacey, 1989). In these intracellular compartments the bacteria differentiate into bacteroids, a highly specialized quasi-organelle form optimized for nitrogen fixation (reviewed by Martin, Uroz, & Barker, 2017; Prell et al., 2009). This shift from their free-living lifestyle in soil to nitrogen fixing bacteroids in nodules requires dramatic changes in bacterial gene expression. These must be carefully coordinated to coincide with the rhizobium's progress along this transition. Oxygen regulation is at the heart of these changes (Appleby, 1984; Fischer, 1994; Jacques Batut & Boistard, 1994). Free-living rhizobia in soil experience oxygen concentrations up to atmospheric levels, whilst concentrations in nodules are at nanomolar levels (Layzell et al., 1993, pp. 393–398). Oxygen concentration drops throughout this transition from a free-living soil bacterium to a nitrogen fixing bacteroid. This drop is driven by plant processes that reduce oxygen inside the nodule and produce near-anoxic conditions in symbiosomes.

1.2 Importance of oxygen in nitrogen fixation

All rhizobia rely on an enzyme complex known as the nitrogenase to perform nitrogen fixation, and all use a molybdenum-containing form of the complex (Rubio & Ludden, 2005; Seefeldt, Hoffman, & Dean, 2009). Nitrogenase is commonly encoded by a single *nifHDK* operon and contains two components (Burgess & Lowe, 1996; Dixon & Kahn, 2004). One of these is the Fe protein, a homodimeric dinitrogenase reductase encoded by *nifH* (Hu & Ribbe, 2013). The second is the MoFe heterotetramic dinitrogenase protein encoded by *nifDK* (Georgiadis et al., 1992; Shah & Brill, 1977). Both proteins contain iron-sulfur clusters required for electron transfer during the enzyme's activity but which are highly intolerant of oxygen (De Maagd et al., 1994; Imlay, 2006; Shaw, 1984). The nitrogenase complex is therefore only stable and functional under near anoxic conditions. Simultaneously, nitrogen fixation by nitrogenase is a highly energy intensive process (Marchal & Vanderleyden, 2000). In many species the enzyme relies on an electron bifurcating complex encoded by *fixABCX* to supply it with low-potential electrons (Edgren & Nordlund, 2004; Ledbetter et al., 2017). Bacteroids must respire at very high levels to meet this energy demand. Symbiotic nitrogen fixation therefore paradoxically requires a very low oxygen environment permitting nitrogenase function but simultaneously demands a high oxygen supply to meet its energy requirements (Udvardi & Poole, 2013). Rhizobia and legumes have evolved several strategies to overcome this paradox, many relying on their close cooperation.

1.3 Regulation of oxygen in plant nodules

The oxygen concentration experienced by bacteroids is determined and regulated by the legume host. Three main legume mechanisms have been identified which adjust nodule oxygen levels and facilitate its supply to bacteroids.

The first of these is a barrier in the cortical layer of nodules that is impermeable to oxygen diffusion (for reviews see Minchin, 1997; Witty & Minchin, 1990; Witty, Minchin, Skot, & Sheehy, 1986). Only nanomolar concentrations of oxygen are present beneath the diffusion barrier (King et al., 1988; Kuzma, Hunt, & Layzell, 1993). Further, the permeability of the layer is variable and controlled by the plant (for mechanistic details of this control see Dakora & Atkins, 1989; Draper, 1986; Witty, Skøt, & Revsbech, 1987). This enables legumes to regulate the influx of oxygen into nodules and maintain a stable internal oxygen concentration despite

significant environmental stresses (Bergersen, 1996; Frank R Minchin, Sheehy, & Witty, 1985; Thumfort, Atkins, & Layzell, 1994; Wei & Layzell, 2006). Studies have shown that the barrier is highly responsive, with permeability able to adjust within minutes (Denison & Kinraide, 1995).

A second mechanism regulating nodule oxygen is the production of leghemoglobin, giving the organs their characteristic red hue (Downie, 2005; Minchin, James, & Becana, 2008). Leghemoglobins are typically monomeric haemoproteins of 16 kDa similar to human myoglobin but employing a different oxygen binding mechanism (Kundu, Trent, & Hargrove, 2003). Compared to myoglobin, leghemoglobins have far higher affinity for oxygen and typically show very fast binding but relatively slow release. The importance of leghemoglobin has been highlighted by studies linking its expression to that of the nitrogenase complex. In a leghemoglobin knockout plant, bacteroid nitrogenase expression was eliminated (Ott et al., 2005). This connection appears to run in both directions, with a very different pattern of leghemoglobin expression found in nodules infected by non-fixing rhizobia (Kawashima, Suganuma, Tamaoki, & Kouchi, 2001). In line with these results, a strong correlation between the leghemoglobin content and nitrogen fixation activity of nodules has been demonstrated (Dakora, 1995).

Leghemoglobin protects nitrogenase by binding free oxygen, produces a high concentration of leghemoglobin-bound oxygen to meet respiratory demands and enables a high oxygen flux to bacteroids. At the low oxygen concentrations found in nodules there is insufficient free oxygen to support adequate bacteroid respiration. However leghemoglobin is itself present at a concentration several orders of magnitude higher than free oxygen (Ott et al., 2005). Most oxygen within nodules is therefore bound by leghemoglobin and present at a relatively high concentration, with the remainder buffered at nanomolar levels. This bound oxygen cannot damage the nitrogenase but is accessible to bacteroids and enables sufficient respiration for nitrogen fixation. Leghemoglobin thus simultaneously limits free oxygen to prevent nitrogenase damage and generates a high concentration of 'safe' oxygen for bacteroids.

Leghemoglobin binding of oxygen also facilitates its diffusion within the nodule (Wittenberg, Bergersen, Appleby, & Turner, 1974; Wittenberg & Wittenberg, 2011, pp. 177–199). At the nanomolar concentration present in nodules, unaided diffusion of oxygen would be unable to meet the oxygen flux requirements of respiring bacteroids. Adequate diffusion of oxygen is also essential for its even distribution. Modeling has suggested that in the

absence of leghemoglobin a significant oxygen gradient would exist between the interior and periphery of plant cells infected with bacteroids (Sheehy, Minchin, & Witty, 1985). This would likely result in bacteroids at plant cell edges receiving excess oxygen whilst those deeper inside would receive too little. Leghemoglobin ensures both a high oxygen flux and an even distribution by facilitating oxygen diffusion within nodules.

The effects of leghemoglobin may be fine-tuned by the existence of isoenzymes of the protein. Recent work found two isoenzymes in peas which had different oxygen affinities and spatial expression patterns within nodules (Kawashima et al., 2001). This indicates leghemoglobins could play a role in creating and shaping the oxygen gradient of indeterminate nodules, but this has received little attention to date.

A third mechanism in nodules appears to specifically control the oxygen levels of infected plant cells (Bergersen, 1994). Mitochondria in these cells employ high affinity terminal oxidases and localize to areas at the cell periphery adjacent to intercellular air pockets, where oxygen influx is likely high (Schulze, 2004). This is thought to act as an additional oxygen barrier, with mitochondrial oxygen consumption protecting symbiosomes from the influx of oxygen in these areas. It has also been proposed that this serves to adjust fixation rates in response to oxygen availability. A drop in oxygen influx could reduce mitochondrial ATP production, in turn reducing the rate of nitrogen fixation by bacteroids (Millar, Day, & Bergersen, 1995). This is in line with a proposed model in which oxygen supply is a limiting factor for nitrogen fixation and used by plants to up- or down-regulate the process (Layzell & Hunt, 1990; Neo & Layzell, 1997; Oresnik, Atkins, & Layzell, 1995).

The combination of these plant mechanisms produces a nodule environment supporting rhizobial nitrogen fixation. In determinate nodules such as those of soybean, the plant creates a uniformly low oxygen concentration throughout the nodule (Wang et al., 2010). All rhizobia within determinate nodules reversibly differentiate to fix nitrogen (Mergaert et al., 2006; Montiel et al., 2017; Van de Velde et al., 2010). Indeterminate nodules, such as those of pea and alfalfa, instead contain both free-living and terminally differentiated nitrogen fixing rhizobia (Arrese-Igor, Royuela, Lorenzo, Felipe, & Aparicio-Tejo, 1993; Oono, Denison, & Kiers, 2009; van de Wiel et al., 1990; Weisbach, Walther, Hartwig, & Nosberger, 1999). In plants with indeterminate nodules the oxygen diffusion barrier is not present at the nodule apex and oxygen is able to freely enter here, creating a longitudinal oxygen gradient (Dixon & Kahn, 2004; Wycoff et al., 1998).

Undifferentiated bacteria at the tip experience relatively high oxygen concentrations whilst terminally differentiated bacteroids in the central nitrogen fixing zone are in a near anoxic environment. Rhizobial differentiation from the tip to the core is an ongoing process during the growth of indeterminate nodules. The nodule oxygen gradient is used by the bacteria to regulate their differentiation. Four zones, shown in Fig. 1, have been delineated within indeterminate nodules (Gourret & Fernandez-Arias, 1974; Mylona, Pawlowski, & Bisseling, 1995). Zone I, at the tip of the nodule, has an oxygen concentration similar to the soil and contains undifferentiated bacteria. Zone II holds rhizobia preparing to infect plant cells, contained in so-called infection threads that direct their movement in the nodule (reviewed by Gage & Margolin, 2000). Above Zone III is a key area known as the II-III interzone where bacteria are released from infection threads and infect plant cells (Vasse, De Billy, Camut, & Truchet, 1990). This has been found to coincide with a sharp drop in oxygen concentration and a concurrent induction of

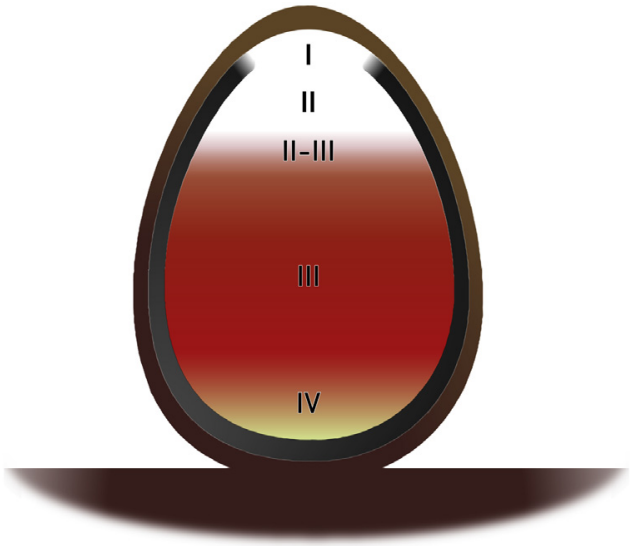


Fig. 1 ***Structure of an indeterminate nodule.*** Legumes such as pea and alfalfa form indeterminate nodules containing free-living rhizobia and terminally differentiated bacteroids. The nodule cortex and root are shown in brown. The oxygen diffusion barrier (black) is absent at the apex of the nodule, creating an oxygen gradient inside. This can be divided into multiple zones. Zone I has free-living rhizobia and oxygen levels similar to surrounding soil. Zone II contains infection threads with bacteria preparing to enter plant cells. At the II-III interzone rhizobia infect plant cells to begin differentiating into bacteroids and oxygen drops sharply. Free oxygen concentration in zone III is at nanomolar levels. This zone houses plant cells infected with nitrogen fixing bacteroids and contains high levels of leghaemoglobin, indicated in red. Zone IV contains senescing cells which are no longer fixing nitrogen. Nodule layers and zones are not to scale.

several genes important for nitrogen fixation (Hernando, Palacios, Imperial, Ruiz-Argueso, & Ruiz-Argüeso, 1995; Romanov et al., 1995; Soupène, Foussard, Boistard, Truchet, & Batut, 1995). Zone III represents the primary nitrogen fixation zone, whilst Zone IV contains plant cells which are beginning to senesce (Timmers et al., 2000).

1.4 Adaptation to nodule oxygen conditions by rhizobia

Free oxygen in much of the nodule is at nanomolar levels and rhizobia have evolved to survive and respire under these conditions during nitrogen fixation. Key to this adaptation is the ability of bacteroids to access leghemoglobin-bound oxygen (Preisig, Anthamatten, & Hennecke, 1993). The concentration of this is several orders of magnitude higher than that of free oxygen. To access this oxygen, symbiotic bacteria employ a specialized cbb_3-type terminal oxidase encoded by the *fixNOQP* operon in their respiratory chain which has very high affinity for oxygen (Delgado, Bedmar, & Downie, 1998; Preisig, Zufferey, Thöny-Meyer, Appleby, & Hennecke, 1996). The complex was first identified in *Bradyrhizobium japonicum* and is broadly required for nitrogen fixation in symbiotic rhizobia (Andreas Schlüter et al., 1997; Kopat et al., 2017; Lee et al., 2008; Renalier et al., 1987). Assembly and maturation of this symbiotic terminal oxidase requires the *fixGHIS* operon (Kahn et al., 1989; Koch, Winterstein, Saribas, Alben, & Daldal, 2000). The *fixNOQP* and *fixGHIS* operons are often similarly regulated and located in proximity to each other. A notable exception is *Azorhizobium caulinodans* in which individual mutants of both operons retained significant nitrogen fixation activity (Mandon, Kaminski, & Elmerich, 1994; Karine Mandon et al., 1993). This suggests the species employs multiple terminal oxidases performing a similar function.

1.5 Mechanisms of oxygen regulation in rhizobia

This review will cover oxygen regulation of nitrogen fixation in several species to highlight variations as well as common themes across rhizobia. The focus will be on strains of five well studied species; *Sinorhizobium meliloti*, *Azorhizobium caulinodans*, *Bradyrhizobium japonicum*, *Rhizobium etli* and *Rhizobium leguminosarum* bv. *viciae*. The divergence of rhizobial species predates the evolution of their legume hosts and the nitrogen fixing partnership (Turner & Young, 2000). Symbiotic capability does not therefore descend from a single ancestor but has instead been horizontally transferred extensively between rhizobial strains and species (Barcellos, Menna, Batista, & Hungria, 2007; Rogel, Hernández-Lucas, Kuykendall, Balkwill, &

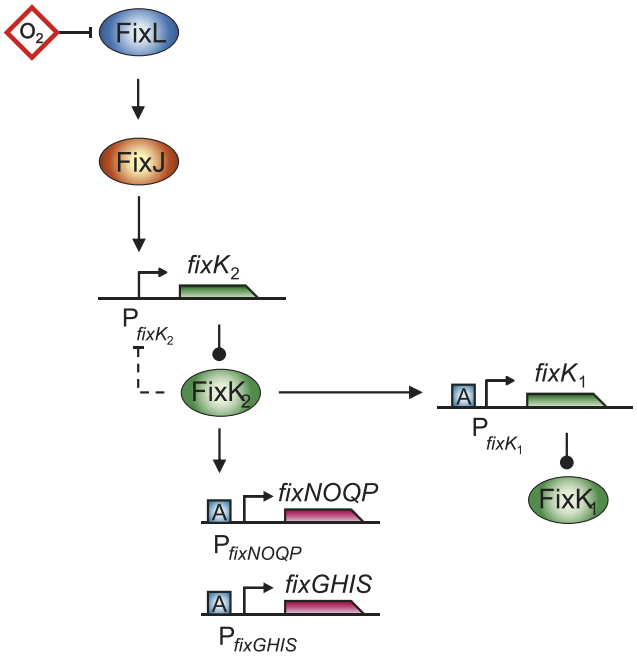
Martinez-Romero, 2001). Successful transfers have relied on the rapid adaptation of pre-existing bacterial regulatory mechanisms under plant selection to arrive at nodulation and fixation competent rhizobia (Masson-Boivin, Giraud, Perret, & Batut, 2009). It appears pre-existing host regulators were commonly repurposed to control symbiotic nitrogen fixation. As a result, whilst a core set of symbiotic genes are conserved across all symbiotic diazotrophs, the regulation of these genes varies significantly at short evolutionary ranges, at both the species and strain level (Martínez-Romero, 2009). It is difficult to determine the cause and significance of these variations; some may result from differing external pressures, whilst others reflect the diversity of host regulation which existed prior to horizontal transfer of symbiotic functions.

Regulation mechanisms have been ordered according to the oxygen concentration at which they activate. The FixL-FixJ and hybrid FixL-FxkR systems are active at relatively high oxygen concentration and are thought to be one of the first oxygen-sensing mechanisms to act during the symbiotic transition. Both in turn lead to production of the FixK protein which is not oxygen sensitive itself. The oxygen concentration at which FnrN homologs operate is less well understood but appears to fall between that of FixK and NifA, although some overlap is likely. The NifA factor is active only at very low oxygen concentrations and is one of the final regulators of bacteroid differentiation. The extensive interconnection of these regulators will also be discussed. Finally, remaining gaps in the literature will be evaluated with a view to identifying key areas for future work.

2. The FixL-FixJ and hybrid FixL-FxkR cascades

The FixL-FixJ (FixLJ) regulatory cascade is an oxygen-responsive two-component system (TCS) (Fischer, 1994). TCSs are composed of a sensor-regulator protein pair and are ubiquitously used by bacteria to respond to a range of environmental conditions (reviewed by Stock, Park, Surette, & Levit, 1995; West & Stock, 2001). Variants of the cascade across several rhizobial species are shown in Fig. 2. The FixL oxygen concentration sensor activates the FixJ regulator under microaerobic conditions, which in turn induces expression of genes controlled by the cascade. The FixLJ system was first discovered in *S. meliloti* as the regulatory mechanism responsible for microaerobic *fixNOQP* expression (David et al., 1988). FixLJ variants are also present in *A. caulinodans*, *B. japonicum*, *R. etli* and *R. leguminosarum*

Bradyrhizobium japonicum



Azorhizobium caulinodans

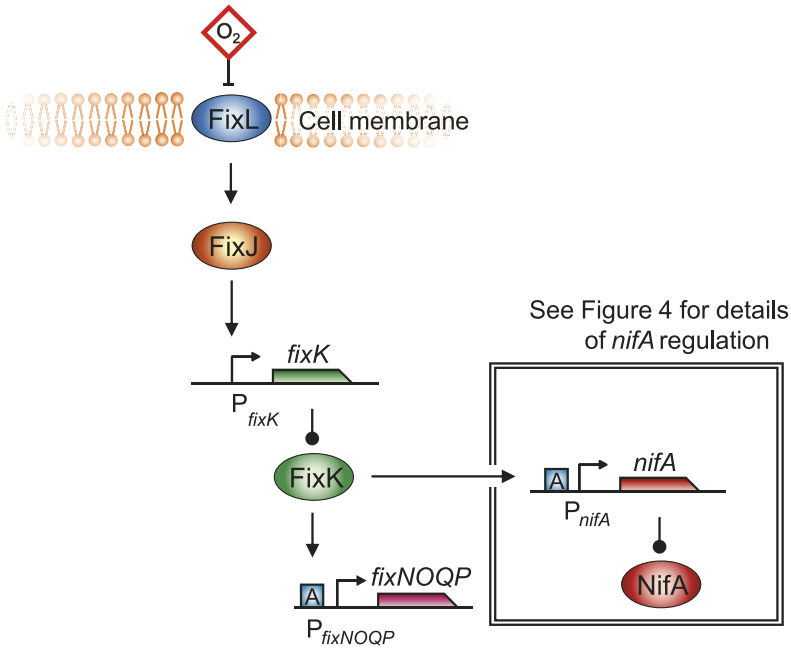
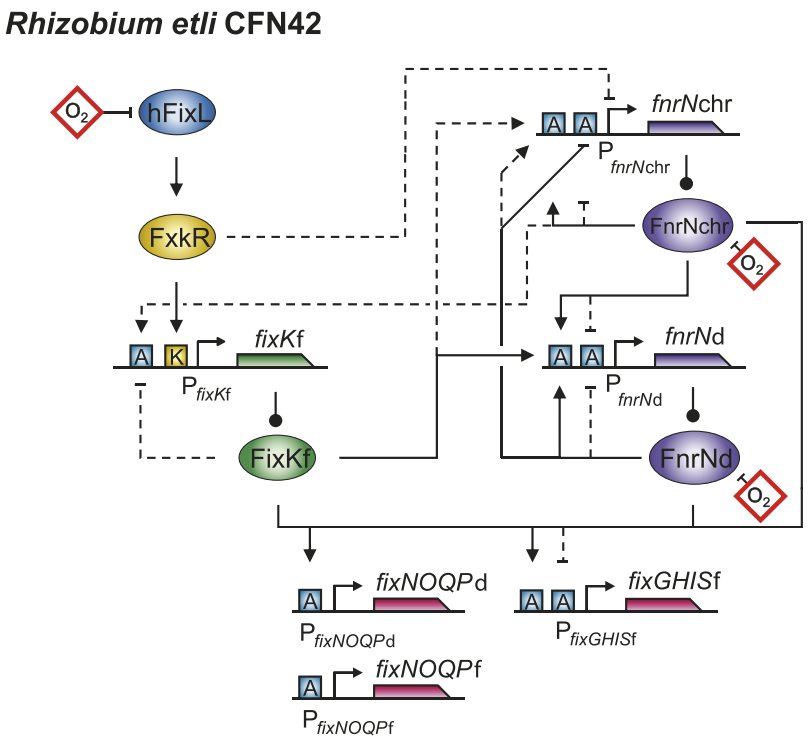


Fig. 2 (*continued*).



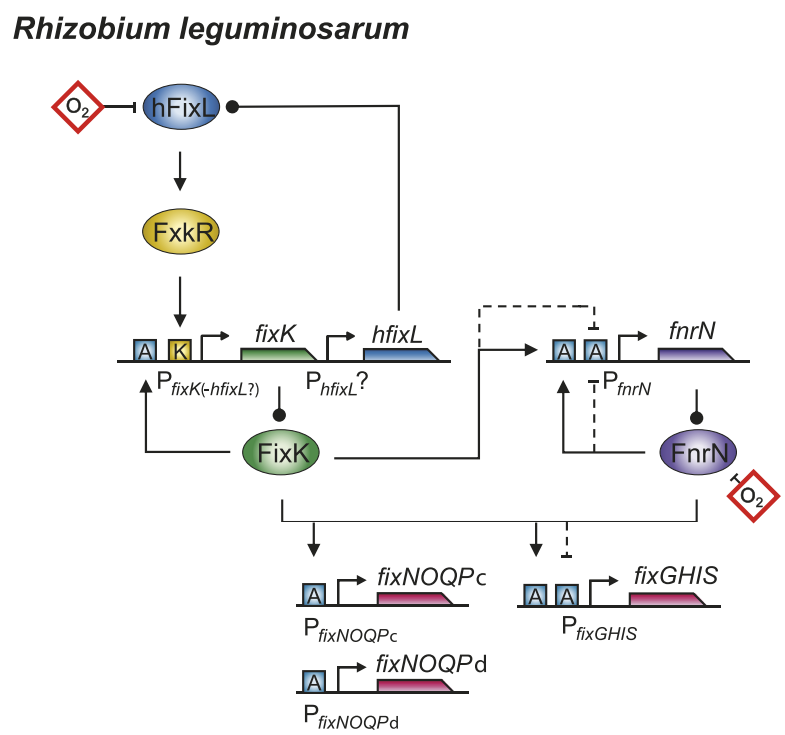


Fig. 2 (*continued*).

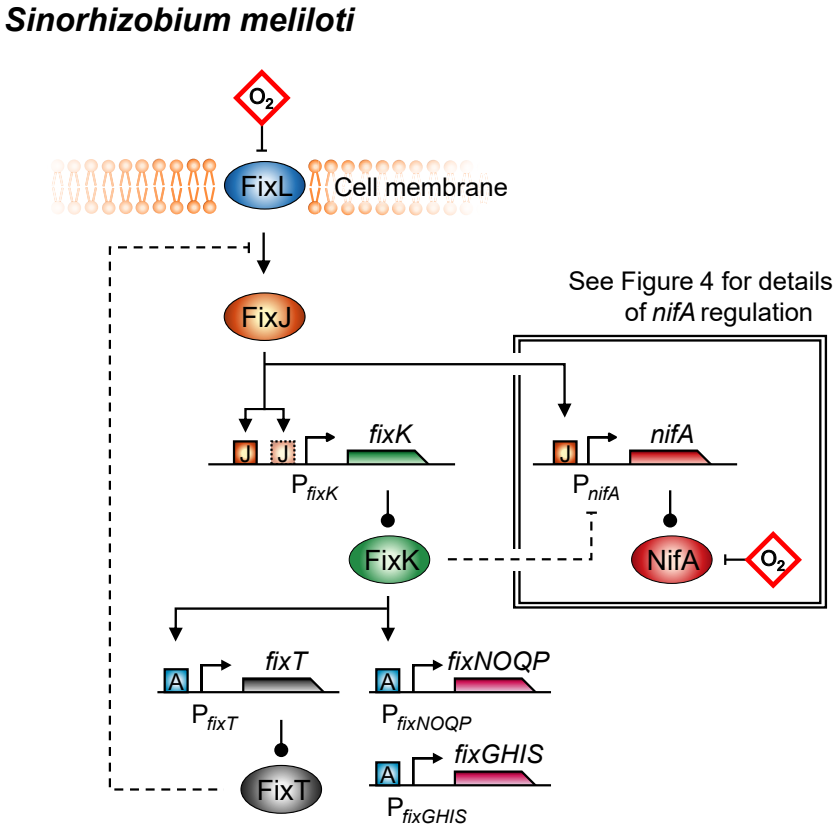


Fig. 2 ***Oxygen regulatory networks controlling rhizobial nitrogen fixation genes.*** Oxygen is indicated in red diamonds. Lines ending in arrows or bars indicate activating and repressing regulation, respectively. Lines end at operators where these are known or predicted from sequence information. Dashed lines represent theorized connections or where the mechanism of action is unclear. Operator sites are shown as boxes. Anaeroboxes are shown in blue and marked with 'A'. FixJ and FxkR (K-box) operators are shown in orange and marked with 'J' and 'K' respectively. *S. meliloti* FixJ binds different operator sequences upstream of *nifA* and *fixK*, and in the latter binds a second low-affinity non-consensus site (faded orange, dotted outline). Pointed rectangles are genes or operons. Lines ending in circles indicate translation. Proteins are shown as ellipses. FixL proteins thought to be membrane associated are indicated. The map shown for *Rhizobium leguminosarum* corresponds to Rlv VF39: Rlv 3841 instead has two copies of hFixL whilst Rlv UPM791 has no functional FixK homolog and two copies of FnrN. See text for details of regulatory connections and Fig. 4 for details of *nifA* regulation.

bv. *viciae* strains (Denise Anthamatten & Hennecke, 1991; Kaminski & Elmerich, 1991; Patschkowski, Schlüter, & Priefer, 1996). Of note, no homolog of the cascade has been found in non-symbiotic diazotrophs, suggesting its role is intimately tied to nitrogen fixation in a symbiotic context.

The FixLJ cascade controls only a small number of genes directly, commonly including *fixK* and in certain species *nifA* (Fischer, 1994). However, if indirect targets of the system are included its regulon is one of the largest of any TCS studied to date (Bobik, Meilhoc, & Batut, 2006; Hertig

et al., 1989; Socorro Mesa et al., 2008). These indirect targets vary at the species and strain level but typically include the *fixNOQP* and *fixGHIS* operons, controlled via FixK.

Recent findings suggest two forms of the FixL protein exist, each corresponding to a different cascade. The first form of FixL is found in *A. caulinodans*, *S. meliloti* and *B. japonicum* (Foussard et al., 1998). FixL proteins from these species are approximately 55% homologous and around 55 kDa in size. A second form of FixL appears to be employed by species including *R. leguminosarum* bv. *viciae* VF39 (Rlv VF39) and *R. etli* CFN42 (Andreas Schlüter et al., 1997; Girard et al., 2000). This FixL variant is larger, at a size of 70 kDa, and shows over 85% homology within these species. It has less than 40% identity to FixL variants from the first group. These two FixL forms appear to act on two different cascades, as evidenced by the presence or absence of a *fixJ* homolog encoding the traditional partner of FixL. Homologs of *fixJ* are present in all species with the first type of FixL but none have been found in Rlv VF39 or *R. etli* CFN42 (Zamorano-Sánchez et al., 2012). A significantly altered system appears to operate in these organisms. FixL retains its oxygen-sensing role but regulates targets such as *fixK* through a partner called FxkR, discussed in more detail in Section 2.2.2.

The importance of FixL has been found to correlate with its form. In species with the canonical FixLJ cascade it plays a crucial role, whereas in species using the second form FixL is largely dispensable. FixL is required for nitrogen fixation by *S. meliloti* and both free-living and symbiotic fixation by *A. caulinodans* (David et al., 1988; Kaminski & Elmerich, 1991). Disruption of *fixL* or *fixJ* in *B. japonicum* led to a 90% drop in nitrogen fixation activity (Denise Anthamatten & Hennecke, 1991). By contrast, null mutations of their *fixL* homolog reduced nitrogen fixation activity by only 50% in Rlv VF39 and had a minimal effect on fixation by *R. etli* CFN42 (D'hooghe et al., 1995; Girard et al., 2000; Patschkowski et al., 1996).

FixL type varies at the strain level so is not necessarily uniform for a given species. A distribution of both FixL forms exists across *R. etli* strains, with the second variant found to be more common (Girard et al., 2000). The *R. etli* CNPAF512 strain for instance employs FixL and FixJ (D'hooghe et al., 1995). However, contrary to other species with FixLJ, disabling the cascade in this strain reduced but did not abolish nitrogen fixation (Moris, Dombrecht, Xi, Vanderleyden, & Michiels, 2004). Whilst the pathway still senses oxygen this no longer appears to be done by FixL and the regulatory targets of FixLJ are largely unknown in this strain. Of note, different FixL forms and their respective cascades appear mutually compatible. Sequence information suggests both systems are present in *S. meliloti* SM11, and may also be found

in the *S. meliloti* 1021 strain (Reyes-González et al., 2016; Zamorano-Sánchez et al., 2012). In the latter strain the hybrid FixL cascade may function outside of symbiosis but this has yet to be further investigated (Trzebiatowski, Ragatz, & De Bruijn, 2001).

2.1 Structures

2.1.1 FixL structure

FixL is an oxygen sensor which activates the FixJ transcription factor by phosphorylation under microaerobic conditions. FixL proteins from the first group have two main components, shown in Fig. 3 (Gong et al., 1998). At the N-terminus is a sensory Per-Arnt-Sim (PAS) domain which contains an oxygen-binding heme group (Green, Crack, Thomson, & LeBrun, 2009). The C-terminus contains a histidine kinase (HK) module composed of a dimerization and histidine phosphotransfer domain and a catalytic ATP binding domain. This module is responsible for signal transmission to FixJ by phosphorylation (Yamada et al., 2009). The FixL proteins of *A. caulinodans* and *S. meliloti* have multiple predicted transmembrane helices and are likely membrane bound or associated, with the PAS domain sensing extracellular oxygen concentration (David et al., 1988; Stigter, 1994). The first full FixL structure was recently determined using the *B. japonicum* protein and this is instead cytoplasmic (Wright et al., 2018).

Under microaerobic conditions release of oxygen from the PAS domain causes activation of the FixL kinase function. Activation is mediated by intramolecular signaling through conformational changes in both the sensor PAS domain and the coiled-coil region connecting it to the HK module (Hao, Isaza, Arndt, Soltis, & Chan, 2002; Wright et al., 2018). Once active the HK autophosphorylates FixL on a conserved histidine residue. The phosphate is then transferred to an aspartate residue on FixJ, switching it to its active conformation (for details of the FixL-FixJ interaction see Rodgers & Lukat-Rodgers, 2005; Sousa, Tuckerman, Gonzalez, & Gilles-Gonzalez, 2007). Oxygen-bound FixL also has phosphatase activity on FixJ, thereby decreasing the cascade's background activity under aerobic conditions (Lois, Weinstein, Ditta, & Helinski, 1993). In summary FixL strictly regulates FixJ-dependent transcription in response to a microaerobic oxygen concentration.

2.1.2 FixJ structure

FixJ is a transcription factor that induces expression of FixLJ cascade targets when phosphorylated by FixL under microaerobic conditions (Galinier

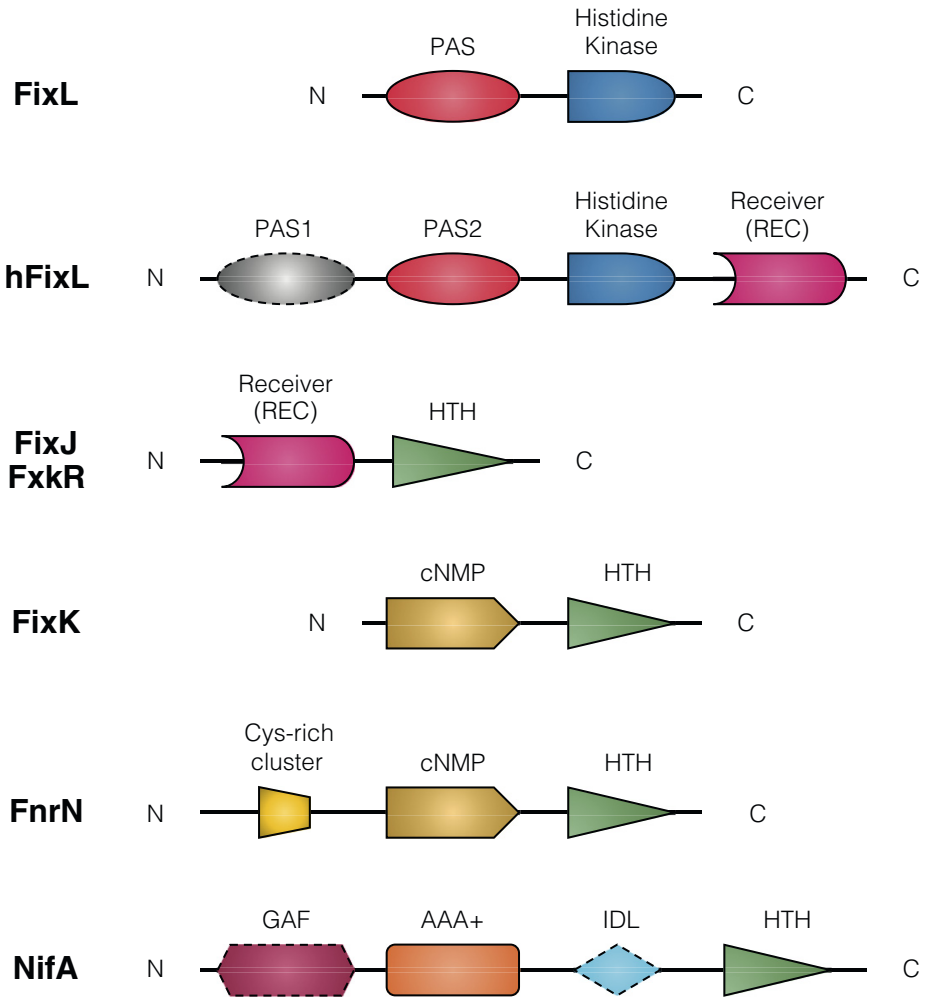


Fig. 3 ***Components of rhizobial oxygen regulation proteins.*** Ellipses are PAS domains: PAS and PAS2 (red) in FixL and hFixL respectively are oxygen sensing, but PAS1 (grey) in hFixL is not. Histidine kinase modules are dark blue, receiver (REC) domains are pink and helix-turn-helix domains (HTH) are shown as green triangles. FixJ and FxkR have the same domains but belong to the NarL/FixJ and OmpR/PhoB families, respectively. cNMP domains are brown and the cysteine-rich cluster at the FnrN N-terminus is shown in yellow. NifA contains a GAF (purple), AAA+ (orange) and HTH domain and an inter-domain linker (IDL, light blue diamond). NifA domains with a dotted outline are not found in some species. Lengths are not to scale.

et al., 1994). The protein is approximately 22 kDa and shows sequence conservation of around 50% across *A. caulinodans*, *S. meliloti* and *B. japonicum* (Fischer, 1994). FixJ has features typical of a TCS regulator protein, with an N-terminal receiver domain (REC) and a C-terminal helix-turn-helix (HTH) transcription activating domain (Schumacher et al., 2000; Stock et al., 1995). Under aerobic conditions, the non-phosphorylated receiver domain inhibits the function of the transcription activating domain (Gouet et al., 1999). Under anaerobic conditions FixL phosphorylates the receiver, alleviating this repression. Derepression is due to phosphorylation-induced

conformational changes which propagate throughout the tertiary structure of the protein (for details, see Re, Bertagnoli, Fourment, Reyrat, & Kahn, 1994; Roche, Mouawad, Perahia, Samama, & Kahn, 2002). In *S. meliloti* phosphorylation causes FixJ to adopt an open configuration, thereby relieving steric inhibition of the transcription activating domain by the receiver domain (Birck et al., 1999). Simultaneously, phosphorylation leads to a conformational change in the receiver that exposes a dimerization interface in the domain (Da Re et al., 1999). The resulting oligomerization of FixJ has been shown in *S. meliloti* to significantly increase the protein's affinity for its target promoters.

2.1.3 Hybrid FixL structure

Certain rhizobial species employ a different form of FixL, called a hybrid FixL (hFixL). These include *R. etli* CFN42, Rlv VF39 and likely *R. leguminosarum* bv. *viciae* 3841 (Rlv 3841), with research to date having focused primarily on the first two species (Girard et al., 2000; Patschkowski et al., 1996; Young et al., 2006). The hFixL protein retains the ability to sense oxygen but regulates a non-canonical pathway, as these organisms have no FixJ homolog. hFixL combines structural elements of canonical FixL and FixJ; a domain homologous to the FixJ receiver domain is present at the protein's C-terminus (Guimarães et al., 2017). In Rlv VF39, this appended domain was shown to be required for microaerobic induction of genes regulated by hFixL (Boesten & Priefer, 2004). Alignments show the receiver domain contains an aspartate residue at position D573 that is analogous to the FixJ residue phosphorylated by canonical FixL. In the canonical FixJ protein, phosphorylation of its receiver domain exposes a dimerization interface important for its function. This does not appear to occur in hFixL and there is no evidence to suggest the protein dimerizes (Sousa, Tuckerman, Gondim, Gonzalez, & Gilles-Gonzalez, 2013).

In *R. etli* CFN42, hFixL contains not only this appended receiver domain but also a second PAS (PAS1) domain at its N-terminus alongside the canonical heme-PAS domain (PAS2) (Sousa et al., 2013). Many of the protein's functions and characteristics appear to be a result of inter-domain interactions. For instance, the protein's kinase activity required the new PAS1 domain. Furthermore, whilst PAS1 has no heme binding capability and no apparent mechanism to sense oxygen directly, the domain does modulate the oxygen affinity of the canonical PAS2. Wild-type (WT) *R. etli* CFN42 hFixL has one of the lowest measured oxygen affinities of any FixL variant, but deletion of PAS1 increased the protein's affinity for oxygen

8-fold (Sousa et al., 2013). At present it remains unknown whether this is mediated by protein binding to this domain or if PAS1 binds a small molecule to regulate hFixL. The protein's oxygen affinity was also modulated by the appended receiver domain. A D573 N mutation at the conserved phosphorylation site in this domain doubled the oxygen affinity of hFixL. These results suggest modulation of hFixL oxygen affinity is an important mechanism regulating the protein's function in *R. etli* CFN42 and involves its PAS1 and receiver domains.

No transmembrane domains are predicted in the hFixL proteins of *R. etli* CFN42 or Rlv VF39, unlike the canonical FixL proteins of *S. meliloti* and *A. caulinodans*. It is therefore believed that hFixL proteins are cytoplasmic and sense intracellular oxygen. This localization is supported by heterologous expression studies in *Escherichia coli* which showed the *R. etli* CFN42 protein was highly soluble even when overexpressed (Sousa et al., 2013). Of note, *B. japonicum* FixL appears to combine elements from both canonical and hybrid FixL protein types. Like the hFixL proteins of *R. etli* CFN42 and Rlv VF39, *B. japonicum* FixL is cytoplasmic and senses intracellular oxygen concentration (Wright et al., 2018). The *B. japonicum* protein also contains a second N-terminal PAS domain that likely regulates the protein's oxygen affinity as shown in *R. etli* CFN42 hFixL. This suggests classifying some FixL proteins as members of the canonical or hybrid group may be impossible, with proteins instead falling along a spectrum between these two forms.

It remains unknown what led to the evolution of hFixL proteins and what their benefits and regulatory implications are. In other organisms hybrid histidine kinase proteins in TCSs allow multiple signals to be integrated and it is suggested they enable more finely tuned regulation (Heermann & Jung, 2010). No studies have yet investigated this possibility in rhizobia. Studying organisms such as *S. meliloti* SM11 where both FixL forms act in parallel will be of interest as these may shed light on differences in the role of the two forms and lead to an improved understanding of their respective functions.

2.2 Role in oxygen regulation

2.2.1 The FixLJ cascade

The FixLJ cascade is active at a relatively high oxygen concentration and appears to be one of the earliest oxygen regulators of symbiotic establishment. In nodules, activation of the cascade likely occurs in the II-III interzone (Andreas Schlüter et al., 1997; Mylona et al., 1995; Soupène et al., 1995).

The most common target of FixLJ microaerobic induction is *fixK*, as found in *S. meliloti*, *B. japonicum* and *A. caulinodans* (David et al., 1988; Kaminski & Elmerich, 1991; Masson-Boivin et al., 2009; Nellen-Anthamatten & Rossi, 1998). In *S. meliloti*, FixLJ also directly activates *nifA* (Hertig et al., 1989). The FixJ DNA binding motif and its regulon have historically been extensively investigated in this species. Early attempts to identify a FixJ operator failed to find a binding motif common to both the *fixK* and *nifA* promoters. More recent work has demonstrated that each promoter has a different consensus sequence and their activation appears to proceed by a different mechanism (Ferrières & Kahn, 2002).

The first FixJ operator type, found in the *fixK* promoter, has a GTAGTTTCCC consensus sequence and is bound by a FixJ dimer. It shows high cooperativity; binding of a single FixJ protein promotes recruitment of a second monomer to form the active dimer. However, mutation studies demonstrated that this site is not critical for *fixK* induction in *S. meliloti*. Instead, downstream of this operator was found a second site that was critical for *fixK* induction (Waelkens et al., 1992). This second operator has a far lower affinity for FixJ than the first site, and its sequence shows no homology to either the *fixK* or *nifA* operators. A model has been proposed whereby the first, upstream consensus site recruits FixJ binding to this second, downstream operator. The first site does not appear to be essential and its elimination could be compensated by increased expression of FixJ (Galinier et al., 1994).

A second FixJ operator type is found in the *nifA* promoter of *S. meliloti*. *Ab initio* methods suggest the consensus sequence for this second FixJ operator is a semi-palindromic GTACGTAG motif. This appears to have a lower binding affinity than the consensus sequence of the first *fixK* operator. A sequence with poor homology to this motif is found upstream of *nifA* and was shown to be responsible for its regulation by FixJ (Ferrières & Kahn, 2002). Results indicate this second operator type binds multiple FixJ dimers, unlike the single dimer bound at the *fixK* site. This is supported by the finding in footprinting experiments that a very large region around the site was protected, and that multiple protein-DNA complexes were visible during gel shift titration. However, a FixJ mutant unable to form dimers was still able to induce *nifA* (Re et al., 1994). Induction of *nifA* by FixJ therefore does not appear to require its oligomerization at the promoter.

In line with their different FixJ operator sites and binding characteristics, the mechanism by which FixJ induces expression from the *fixK* and *nifA* promoters also differs. At the *fixK* promoter the FixJ receiver domain plays

a role in RNA polymerase recruitment (Ton-Hoang, Salhi, Schumacher, Da Re, & Kahn, 2001). A mutation of the domain interfering with this recruitment reduced activation of *fixK* 10-fold in *S. meliloti*. By contrast the mutation had no effect on FixJ induction of *nifA*. This suggests FixJ promotes transcription of *nifA* through a different, as yet undetermined, mechanism.

Recent genome-wide work in *S. meliloti* has begun to shed light on members of the FixLJ regulon besides the canonical *fixK* and *nifA* targets. This species contains a chromosome and two extrachromosomal replicons pSymA and pSymB (Weidner et al., 2013). It is believed the FixLJ cascade and its original regulon were initially introduced through horizontal transfer of pSymA (Ferrières, Francez-Charlot, Gouzy, Rouillé, & Kahn, 2004). Its regulon appears subsequently to have begun encompassing other targets in the *S. meliloti* genome. In accordance with this theory, putative FixJ operators were found to be very unevenly distributed, with a majority found on pSymA and a much smaller number on the chromosome. All pSymB sites were within coding regions so none are thought to be functional. However, it has been theorized that some FixJ operators in the coding regions of pSymA and the chromosome do have a functional role. A significant number of these have been found and they may serve as 'reservoirs', partially controlling the intracellular localization of FixJ to facilitate its diffusion to operators in spatially proximal promoters. Analysis of a selection of new putative FixJ targets confirmed that two genes involved in proline metabolism are also regulated by the FixLJ cascade. Several studies have demonstrated the importance of proline metabolism to symbiosis but its role remains poorly understood (Jimenez-Zurdo, 1995; Jiménez-Zurdo, García-Rodríguez, & Toro, 1997; King, Hojnacki, & O'Brian, 2000). It is likely that many processes controlled by FixLJ and important in symbiotic regulation remain to be discovered.

The cascade may also serve a function outside of nodules (Li, Xu, Ren, & Chen, 2010). In *A. caulinodans*, FixLJ activation occurs in free-living cultures even in an environment with atmospheric oxygen concentration (Kaminski & Elmerich, 1991; Kaminski, Mandon, Arigoni, Desnoues, & Elmerich, 1991; Loroch, Nguyen, & Ludwig, 1995). This may be due to the microaerobic internal environment created by this species to enable free-living nitrogen fixation, using multiple terminal oxidases (Kitts & Ludwig, 1994). In *B. japonicum*, which fixes only in symbiosis, free-living cells also showed FixLJ activation of gene targets (Sciotti, Chanfon, Hennecke, & Fischer,

2003). Induction was demonstrated when cells were exposed to a 5% O_2 atmosphere and activation gradually increased as oxygen dropped to 0.5%. Few studies have investigated the role of FixLJ outside symbiosis and the biological significance of this remains largely unknown.

2.2.2 The hFixL-FxkR cascade

Several rhizobia employ a hFixL protein which contains an appended domain homologous to the receiver domain of FixJ. These species typically have no FixJ homolog, and instead hFixL oxygen sensing is transmitted to gene induction by the FxkR (*fixK* regulating) protein. This situation has been studied primarily in *R. etli* CFN42 and Rlv VF39 (Andreas Schlüter et al., 1997; Girard et al., 2000). The *fixL* and *fixJ* genes commonly form an operon located near other nitrogen fixation genes, suggesting they were acquired together in a horizontal gene transfer event. By contrast, *fxkR* in *R. etli* CFN42 and Rlv VF39 is not in the vicinity of its *hfixL* signaling partner, suggesting it was not acquired in the same transfer event and may be derived from a pre-existing host regulator. Sequence information suggests Rlv 3841 also encodes two putative homologs of hFixL and one of FxkR (Young et al., 2006). *S. meliloti* SM11 encodes homologs of both the hybrid and canonical FixL proteins, demonstrating the proteins and their respective pathways are not mutually exclusive (Reyes-González et al., 2016). It is unknown at present whether any crosstalk occurs between such parallel systems. Unlike the FixLJ system that is generally critical for nitrogen fixation activity, mutations eliminating hFixL-FxkR regulation are found to have a limited effect on fixation. In *R. etli* CFN42, the pathway appears dispensable (Zamorano-Sánchez et al., 2012).

As in the traditional FixLJ cascade, hFixL regulates FxkR at the protein level by phosphorylating it under microaerobic conditions (Zamorano-Sánchez et al., 2012). Transcription of *fxkR* is thought to be constitutive. FxkR has a predicted weight of 27 kDa and a typical response regulator structure with an N-terminal receiver domain and a C-terminal HTH DNA-binding domain. These domains are analogous to those found in FixJ, but the two proteins are not related. FxkR belongs to the OmpR/PhoB family whereas FixJ belongs to the eponymous NarL/FixJ family. Based on studies of other OmpR/PhoB family proteins, it is thought that phosphorylation of the FxkR receiver domain induces a conformational change leading to protein dimerization. This dimer brings together two FxkR HTH domains which bind a DNA target sequence composed of a pair of direct-repeat half sites (Gao & Stock, 2009).

As expected given their different families it appears there is limited or no cross-talk between regulation by FxkR and FixJ. An *R. etli* CFN42 *fxkR* mutant could not be complemented by heterologous *S. meliloti* FixJ (Zamorano-Sánchez et al., 2012). It therefore appears hFixL cannot activate the canonical FixJ protein or FixJ is not functional in this strain. In contrast complementation of the *R. etli* CFN42 *fxkR* mutant was possible with Rlv VF39 *fxkR*, demonstrating the similarity of their hFixL-FxkR cascades.

In *R. etli* CFN42, microaerobic activation of *fixKf* is due to hFixL and requires the FxkR intermediary in a manner analogous to the role of FixJ in transmitting oxygen sensing by FixL. Analysis of a library of confirmed and putatively FxkR-regulated genes identified a GTTACA-N_4-GTTACA consensus binding motif, named the "K-box" (Zamorano-Sánchez et al., 2012). A K-box is found in front of the *fixKf* and *fixK* genes of *R. etli* CFN42 and Rlv VF39 respectively, and both are induced by hFixL-FxkR under microaerobic conditions. By contrast, no K-box element is found upstream of *fixK* genes in *S. meliloti* SM11 (Reyes-González et al., 2016). In this strain, the cascade instead functions to repress activation by the parallel FixLJ system. This repression appears to act by protein level inhibition of the TCS, but the mechanism and target of this repression have yet to be determined. It also remains unclear how hFixL senses oxygen in this strain as the protein has no oxygen-binding heme group (D'hooghe et al., 1995).

In summary, the hFixL-FxkR cascade performs the same function as the canonical FixLJ pathway. However, to date it has always been found with a system providing redundancy (see Section 4 for details) so is never essential for nitrogen fixation. It appears that when the pathway is present in parallel with the canonical system, hFixL-FxkR interacts with FixLJ and supresses its activation. Of note, sequence analysis putatively identified the presence of hFixL-FxkR variants in non-rhizobial species. This suggests that, unlike FixLJ, the hFixL-FxkR system is used as an oxygen sensing mechanism in a variety of contexts (Zamorano-Sánchez et al., 2012).

3. The FixK transcription factor

Most of the regulation exerted by both the hFixL-FxkR and canonical FixLJ cascades acts indirectly, through their induction of *fixK* (Bobik et al., 2006; Socorro Mesa, Bedmar, Chanfon, Hennecke, & Fischer, 2003). FixK proteins are transcription factors which show some 35–45% conservation across commonly studied strains. They act as intermediates, regulating

gene expression in response to oxygen sensing by the FixL and hFixL proteins. FixK is generally found to be crucial for nitrogen fixation, including in *S. meliloti*, *B. japonicum* and *A. caulinodans* (David et al., 1988; Kaminski & Elmerich, 1991; Nellen-Anthamatten & Rossi, 1998). Where FixK operates in conjunction with an FnrN-like regulator (see Section 4), it is generally non-essential (Granados-Baeza et al., 2007; Patschkowski et al., 1996). The FixK and FnrN regulons overlap significantly and the two proteins often appear to operate in a redundant fashion.

3.1 Structure of FixK

FixK is typically 27 kDa and is a member of the CRP/FNR superfamily of transcriptional regulators. This superfamily can be divided into three subgroups of CRP-like, NtcA-like and FNR-like proteins, the last of which FixK falls into (Crack, Green, Cheesman, Le Brun, & Thomson, 2007; Mesa, Hennecke, & Fischer, 2006; Stephen Spiro & Guest, 1990). FNR-like proteins can be further subdivided into three classes (Moris et al., 2004). Class IA proteins such as *E. coli* FNR directly sense oxygen and play an important role in the cellular response to microaerobic conditions (Crack & Le Brun, 2018; Guest, Green, Irvine, & Spiro, 1996). Class IB includes *B. japonicum* $FixK_1$ and the FnrN proteins which are also able to directly sense oxygen and regulate genes accordingly, discussed in more detail in Section 4 (Zamorano-Sánchez & Girard, 2015). Class IC is composed primarily of rhizobial FixK proteins including *B. japonicum* $FixK_2$, but these do not respond directly to oxygen at the protein level (Fischer, 1994). *B. japonicum* $FixK_2$ has been shown to respond to reactive oxygen species at the post-translational level, but it is unclear how widespread this regulation is in rhizobia and what role it plays in symbiosis (Socorro Mesa, Reutimann, Fischer, & Hennecke, 2009).

Two domains are conserved across all members of the CRP/FNR superfamily (Körner, Sofia, & Zumft, 2003; Stephen Spiro, 1994). At the C-terminus is a HTH domain which binds a DNA motif and interacts with RNA polymerase to regulate transcription (Socorro Mesa, Ucurum, Hennecke, & Fischer, 2005). This domain and its function are very well conserved across FNR-like proteins, notably reflected in the highly similar DNA motif sequence these proteins bind (Hernando et al., 1995; Socorro Mesa et al., 2008; Stephen Spiro, 1994). Many members of class IA, IB and IC FNR-like protein bind the so-called "anaerobox", a highly conserved palindromic TTGA-N_6-TCAA operator sequence. Protein binding to this motif can positively and negatively regulate expression, both upstream and

downstream of the anaerobox (Batut et al., 1989). At the N-terminus is a cyclic nucleotide-monophosphate-like binding domain (cNMP). However, FNR-like proteins diverge in the function of their N-terminal region. In members of class IA and IB this region contains a cysteine-rich motif which enables the proteins to directly sense oxygen and control gene expression accordingly (see Section 4). This oxygen sensing motif is absent in the FixK proteins of *S. meliloti*, *A. caulinodans*, *R. etli*, *R. leguminosarum* and $FixK_2$ of *B. japonicum* (Fischer, 1994; Nellen-Anthamatten & Rossi, 1998). Class IC FixK proteins therefore do not respond to oxygen at the protein level and instead transmit oxygen sensing by the FixLJ or hFixL-FxkR cascade to the level of downstream gene expression (Gutiérrez, Hernando, Palacios, Imperial, & Ruiz-Argüeso, 1997; Patschkowski et al., 1996). No evidence has been found to date that these proteins integrate other signals at the post-transcriptional level (Socorro Mesa et al., 2005).

3.2 Role of FixK in oxygen regulation

Across rhizobial species three members of the FixK regulon have been well studied (Fischer, 1994). The first are the *fixNOQP* and *fixGHIS* operons, respectively encoding a terminal oxidase with high oxygen affinity required for symbiosis and a complex required for its assembly (Koch et al., 2000; Preisig et al., 1996). Second, in *S. meliloti* FixK has been shown to repress *nifA*, which encodes the central activator of nitrogen fixation (discussed in more detail in Section 5). Lastly, autoregulation of *fixK* appears to be a very common mechanism in rhizobia.

In species encoding the canonical FixLJ cascade, FixK is generally critical for the microaerobic induction of both *fixNOQP* and *fixGHIS*. In *B. japonicum* and *S. meliloti* both operons have upstream anaeroboxes and are regulated by FixK (Batut, de Philip, Reyrat, Waelkens, & Boistard, 1993; Kahn et al., 1989; Preisig et al., 1993; Socorro Mesa et al., 2005). In *A. caulinodans* *fixNOQP* is likewise under FixK control but this species has no homolog of *fixS* and no anaerobox is present upstream of the *fixGHI* operon (Mandon et al., 1994; Karine Mandon et al., 1993). In contrast to its regulation in other symbiotic diazotrophs, expression of *fixGHI* in *A. caulinodans* is not controlled by FixK, does not respond to oxygen and occurs under free-living conditions. These differences probably reflect the organism's ability to perform free-living nitrogen fixation using multiple terminal oxidases (Dreyfus, Elmerich, & Dommergues, 1983; Dreyfus & Dommergues, 1981; Gebhardt, Turner, Gibson, Dreyfus, & Bergersen, 1984; Kitts & Ludwig, 1994).

In species employing the hFixL-FxkR cascade, FixK appears to be a non-essential regulator of *fixNOQP* (D'hooghe et al., 1995; Patschkowski et al., 1996). In Rlv VF39, a *fixK* mutant was largely unaffected in microaerobic induction of this operon (Andreas Schlüter et al., 1997; Boesten & Priefer, 2004). *R. etli* CFN42 encodes two copies of *fixK*, named *fixK*d and *fixK*f, respectively located on its pCFN42d and pCFN42f plasmids (Girard et al., 2000). Regulation of *fixK*f but not *fixK*d is under the control of the hFixL-FxkR cascade, and only FixKf appears to be important for microaerobic induction. *R. etli* CFN42 also encodes two copies of *fixNOQP*, *fixNOQP*d and *fixNOQP*f, both controlled by FixKf (Lopez et al., 2001). Expression of *fixNOQP*d is the more important of the two for nitrogen fixation, and requires FixKf. However, although the absence of hFixL suppressed *fixNOQP*f expression, *fixNOQP*d was still expressed at significant levels. Expression of *fixK*f may be induced by a second regulator, potentially the FnrNchr protein discussed in more detail in Section 4.

Less commonly, FixK also regulates expression of *nifA*, the central activator of nitrogen fixation (Dixon & Kahn, 2004; Fischer, 1994). In *S. meliloti*, FixK represses expression of *nifA* whilst FixJ activates it (Batut et al., 1989; Reyrat et al., 1993). The mechanism of FixK repression is unknown and may be indirect as no anaerobox is present upstream of *nifA*. In combination, these opposing regulatory mechanisms probably balance expression of *nifA*. In *A. caulinodans*, FixK instead activates expression of *nifA* (Hans Martin Fischer, 1996; Kaminski et al., 1991). In *B. japonicum*, *nifA* expression is under indirect $FixK_2$ control (Kullik et al., 1991; Socorro Mesa et al., 2008). In this species expression of $rpoN_1$, encoding the sigma factor σ^{54} required for NifA-mediated transcriptional activation, is controlled by $FixK_2$ (Michiels et al., 1998). The FixLJ-$FixK_2$ pathway therefore indirectly regulates the activity of NifA in this species. However, control by the pathway is not complete as a redundant $rpoN_2$ paralog exists which does not appear to be regulated by oxygen. A similar situation exists in *R. etli* CNPAF512 where only one of the *rpoN* paralogs is believed to be regulated by FixK. As expected from post-transcriptional regulation of NifA in this strain, the FixLJ cascade induced *nifH* through NifA upregulation but had no effect on *nifA* expression. Because the strain has three copies of *nifH*, abolishing FixLJ regulation reduced but did not eliminate its ability to fix nitrogen. Some rhizobial species thus appear to employ control of RpoN concentration by FixK as a tool to modulate NifA activity.

Autoregulation by FixK is also a common mechanism, regardless of the FixL cascade type. *E. coli* FNR auto-represses its expression by binding to a

sequence downstream of its own promoter, presumably sterically hindering transcription by RNA polymerase (Spiro & Guest, 1987b). In *B. japonicum* $FixK_2$ is negatively auto-regulated (Nellen-Anthamatten & Rossi, 1998; Reutimann, Mesa, & Hennecke, 2010; Socorro Mesa et al., 2005). The same effect has been shown in *S. meliloti* (Batut et al., 1989). The mechanism of FixK auto-repression in both these species remains poorly understood and two theories have been put forward. The first is a direct repression effect. Two putative FixK binding sites have been found upstream of *S. meliloti* *fixK* at -487 and -43 relative to the transcription start site (TSS) (Waelkens et al., 1992). It has been suggested that at high concentration FixK binds to both these sites to form a repression loop, resulting in downregulation of *fixK* expression. In a proposed alternative mechanism, repression is an indirect effect (Foussard et al., 1998). In *S. meliloti*, FixK induces expression of *fixT* (Bergès et al., 2001). FixT appears to act as an anti-kinase, repressing phosphorylation of FixL (Garnerone, Cabanes, Foussard, Boistard, & Batut, 1999). This inhibits the FixLJ cascade and in turn represses *fixK* induction, thus completing the FixK autoregulation loop (Garnerone, Foussard, Boistard, & Batut, 1999). However, this system remains poorly understood and it is unclear if the role of FixT is to close this loop or whether the protein acts to integrate another signal into FixLJ-mediated gene expression.

In *R. etli* CFN42, two regulatory inputs have been found at the level of *fixKf* expression. First, in line with the situation in other rhizobia, FixKf auto-represses its expression (Girard et al., 2000). Further, a CRP/FNR-type regulator encoded by *stoRd* was recently shown to repress *fixKf* expression (Granados-Baeza et al., 2007). A knockout of this gene appeared to enhance the nitrogen fixation of *R. etli* CFN42 but the function and importance of StoRd remain unclear. A homolog of *stoRd* was found in *S. meliloti* but eliminating this had no effect on nitrogen fixation, suggesting the protein's role varies significantly between species (Bobik et al., 2006). In Rlv VF39, autoregulation activates rather than represses *fixK* (Patschkowski et al., 1996). The *hfixL* and *fixK* genes may form an operon in this strain suggesting both are auto-activated directly by FixK and indirectly by hFixL. This operon arrangement has yet to be confirmed and the biological significance of such a feedback loop is unknown.

Recent work has begun to explore the wider FixK regulon and found that this extends far beyond the three common targets discussed above. In *B. japonicum*, the direct regulon of $FixK_2$ is reported to contain over 200 members, an order of magnitude more than that of FixJ (Socorro Mesa et al., 2003). A study using shotgun proteomics recently reported over

600 genes were uniquely expressed in *B. japonicum* under microoxic conditions (Fernández et al., 2019). In *S. meliloti*, FixK was found to have one of the largest regulons of any TCS studied to date (Bobik et al., 2006). The role of many of these targets and the importance of their regulation by FixK has yet to be established. Notable examples are the *B. japonicum* heme biosynthesis pathway genes *hemA, hemB* and *hemN*, all members of the FixK regulon (Page & Guerinot, 1995; Socorro Mesa et al., 2005). In rhizobia these genes are typically expressed under symbiotic conditions and *hemA* mutants abolish nitrogen fixation. However, they are not essential in *B. japonicum* which also expresses them under non-symbiotic conditions (McGinnis & O'Brian, 1995). In *R. etli* CFN42, FixKf also regulates the response to nitric oxide (NO) (Gómez-Hernández et al., 2011). NO signaling is implicated in symbiosis, but its role is unclear. As with most other members of its regulon, the importance of FixK heme biosynthesis control is therefore unclear. Further investigations of the FixK regulon are likely to reveal many additional functions regulated in response to oxygen concentration.

4. The FnrN transcription factor

The FnrN proteins are transcription factors which directly sense oxygen concentration and regulate rhizobia during symbiosis. *B. japonicum* $FixK_1$ belongs to the same group and will also be discussed in this section (Socorro Mesa et al., 2008). Like FixK, FnrN homologs are FNR-like proteins and typically induce expression of *fixNOQP* and *fixGHIS*. Both FnrN and FixK mediate transcriptional regulation by binding to anaerobox operators (Batut et al., 1989; Gamper, Zimmermann, & Haas, 1991; Kahn et al., 1989; Schlüter, Patschkowski, Unden, & Priefer, 1992). However, whilst FixK proteins are class IC and oxygen insensitive (see Section 3), FnrN homologs are of class IB and can directly sense oxygen (Guest et al., 1996; Stephen Spiro, 1994). The presence of FnrN correlates with that of the hFixL-FxkR-FixK pathway, but the protein appears to individually recapitulate the full oxygen sensing and regulating functions of that pathway. Of note, an *S. meliloti fixJ* mutant could be partially complemented by expression of *fnrN*, showing the protein also covers some of the function of the FixLJ cascade (Colonna-Romano et al., 1990). FnrN and FixK usually act in a redundant fashion, with organisms retaining at least some nitrogen fixation activity if either is individually eliminated (Patschkowski et al., 1996). This is in contrast to organisms employing only the canonical FixLJ-FixK cascade,

where it is usually essential for symbiotic nitrogen fixation (Fischer, 1994). Beyond introducing a degree of redundancy, the function of FnrN remains poorly understood. It may serve to produce more finely-tuned regulation, including by responding to a different oxygen concentration than the hFixL-FxkR TCS. FnrN may also provide more responsive regulation for rhizobia exposed to rapidly fluctuating oxygen concentrations (Colombo, Gutiérrez, Palacios, Imperial, & Ruiz-Argüeso, 2000; Sánchez-Cañizares et al., 2018).

FnrN was first found in Rlv VF39 during a search for endogenous regulators able to activate heterologous *S. meliloti fixNOQP* expression under microaerobic conditions (Colonna-Romano et al., 1990; Schlüter et al., 1992). Two copies of *fnrN* are also found in *R. leguminosarum* bv. *viciae* UPM791 (Rlv UPM791) and a homolog appears to be present in Rlv 3841 (Sánchez-Cañizares et al., 2018; Young et al., 2006). This suggests the protein is broadly conserved across strains of *R. leguminosarum*. Likewise, two FnrN homologs regulate nitrogen fixation in *R. etli* CFN42 and one in *R. etli* CNPAF512, suggesting FnrN proteins are also conserved across *R. etli* strains (Lopez et al., 2001; Moris et al., 2004). $FixK_1$ in *B. japonicum* is also an FNR-like class IB protein but no FnrN homologs have been found in *S. meliloti* or *A. caulinodans* to date (Socorro Mesa et al., 2008). Like *fxkR*, the lack of co-localization between *fnrN* and other symbiotic genes in Rlv VF39 and Rlv 3841 suggests it was originally a non-symbiotic regulator. This is supported by the wide distribution of FNR-like proteins, including in soil bacteria (Green, Scott, & Guest, 2001; Ray & Williams, 2006; Socorro Mesa et al., 2003; Vollack, Härtig, Körner, & Zumft, 1999). In the non-fixing marine bacterium *Dinoroseobacter shibae*, FNR-like proteins regulate its transition from aerobic to anaerobic growth and induce the expression of a high affinity terminal oxidase (Ebert et al., 2017).

Rlv VF39 employs both FixK, regulated by the hFixL-FxkR pathway, and FnrN. In this strain, mutation of *fnrN* was found to reduce nitrogen fixation activity by some 40% relative to WT (Colonna-Romano et al., 1990). A knockout of *hfixL* or *fixK* respectively resulted in 60% and 80% reduction of fixation activity in this strain (Patschkowski et al., 1996). A double *fixK fnrN* mutation eliminated all nitrogen fixation activity. Thus, the proteins are semi-redundant, parallel rather than hierarchical, and collaborate to activate expression of genes required for nitrogen fixation. In Rlv UPM791, there is no FixK homolog but a similar redundancy exists between its two *fnrN* homologs (Gutiérrez et al., 1997). Individual *fnrN* mutants retained nitrogen fixation activity whilst a double mutant abolished it (Gutiérrez,

Hernando, Palacios, Imperial, & Ruiz-Argüeso, 1998; Hernando et al., 1995). *R. etli* CFN42 also shows extensive redundancy, with complete elimination of nitrogen fixation only observed in a triple mutant of its two *fnrN* homologs and *hfixL* (Lopez et al., 2001). In *R. etli* CNPAF512, the *fnrN* mutant showed a severe reduction in nitrogen fixation, indicating a near-essential role for the gene in this strain (Moris et al., 2004). The role of the FixLJ system in this species is presently unknown, and it is possible this has been entirely replaced by FnrN as seen in Rlv UPM791.

A different system operates in *B. japonicum*, which contains both the class IC $FixK_2$ protein and the oxygen-sensing class IB $FixK_1$ protein. In contrast to the situation found in *R. leguminosarum* and *R. etli* strains, $FixK_1$ is not a redundant regulator and is under $FixK_2$ control in a hierarchical cascade (Nellen-Anthamatten & Rossi, 1998; Socorro Mesa et al., 2008). $FixK_2$ is required for nitrogen fixation in *B. japonicum* but $FixK_1$ has no effect on fixation, indicating the two perform very different functions in this species (Anthamatten, Scherb, & Hennecke, 1992; Terpolilli, Hood, & Poole, 2012). In summary, FnrN is a key regulator of nitrogen fixation in several species. It is commonly used as a parallel, semi-redundant system for the hFixL-FxkR cascade and in some strains replaces it entirely. However, important variations still exist in the protein's role across species, as demonstrated by the situation in *B. japonicum* where the class IB $FixK_1$ protein is not required for fixation and is under the control of $FixK_2$.

4.1 Structure of FnrN

FnrN proteins and $FixK_1$ in *B. japonicum* are FNR-like class IB transcription factors, capable of sensing and responding to oxygen concentration at the protein level. In a typical FNR-like protein oxygen is sensed by an N-terminal sensor domain (for reviews see Crack, Green, & Thomson, 2004; Unden & Trageser, 1991). Binding of oxygen triggers a conformational change that results in formation of DNA-binding FNR dimers via an interface at the C-terminal domain (Moore & Kiley, 2001; Moore, Mettert, & Kiley, 2006). The *E. coli* FNR protein is the best understood example of this model and studies have consistently shown that rhizobial class IB proteins are functionally very similar to this protein (Bates et al., 2000; Spiro & Guest, 1987b; Stephen Spiro & Guest, 1990).

Under anaerobic conditions *E. coli* FNR ligates a $[4Fe{-}4S]^{2+}$ iron-sulfur cluster (Sutton, Mettert, Beinert, & Kiley, 2004; Sutton, Stubna, et al., 2004). The cluster is coordinated by four cysteine residues (Spiro & Guest, 1987a; Trageser & Unden, 1989). Three of these are grouped in a so-called

cysteine-rich motif (Cys-X_2-Cys-X_5-Cys) at the N-terminus of FNR and the fourth is located at a conserved position in the central part of the protein. The cluster mediates FNR dimerization, the active form of the protein (Lazazzera, Bates, & Kiley, 1993). Exposure to oxygen deactivates FNR by converting the iron-sulfur cluster to $[2Fe{-}2S]^{2+}$, and sustained exposure causes the protein to unbind the cluster completely (Jervis & Green, 2007; Khoroshilova, Popescu, Munck, Beinert, & Kiley, 1997). FNR in *E. coli* therefore rapidly and stringently responds to intracellular oxygen concentration and forms active dimers only under low oxygen conditions. Rhizobial FnrN proteins and *B. japonicum* $FixK_1$ are a different class of FNR-like proteins but employ a very similar oxygen-sensing mechanism. Like *E. coli* FNR, three cysteine residues are grouped in a Cys-$X_{2/3}$-Cys-X_7-Cys motif at the N-terminus of these proteins, with the fourth found in a central position (Anthamatten et al., 1992; Colonna-Romano et al., 1990; Schlüter et al., 1992). It is therefore widely assumed rhizobial FnrN proteins respond to oxygen concentration through a mechanism very similar to *E. coli* FNR (Schlüter et al., 1992). This is supported by multiple complementation studies. The Rlv VF39 FnrN protein appears to be able to form heterodimers with *E. coli* FNR that could bind anaeroboxes and activate microaerobic induction of target genes. Rlv VF39 FnrN could complement an *E. coli fnr* mutant for growth on nitrate and promote anaerobic induction of several FNR targets including *narGHJI*, *nirB* and *fdnGHI* (Anthamatten et al., 1992). Complementation of the *E. coli fnr* mutant was also possible with *B. japonicum* $FixK_1$ but not its class IC $FixK_2$ protein. Conversely, *E. coli* FNR complemented an Rlv VF39 *fnrN* mutant for regulation of *fixNOQP*. In summary, *B. japonicum* $FixK_1$ and the rhizobial FnrN proteins show strong similarity to the *E. coli* FNR protein, and all appear to act as integrated oxygen-sensing gene regulation systems.

At their C-terminus, homologs of both class IC FixK and class IB FnrN proteins as well as *E. coli* FNR encode a very highly conserved HTH DNA-binding domain (Sawers, 1991; Stephen Spiro, 1994; Stephen Spiro & Guest, 1990). Under anaerobic conditions a dimerization interface is exposed in this domain and protein dimers are formed which can bind DNA and activate transcription. Key residues involved in DNA binding by this domain are conserved across the three proteins and all bind a near-identical palindromic TTGAT-N_4-ATCAA operator called an anaerobox in the promoters of the genes they regulate (Jayaraman, Gaston, Cole, & Busby, 1988; Jayaraman, Cole, & Busby, 1989; Spiro & Guest, 1987a). There is therefore an inherent overlap in the gene targets of FNR, FixK

and FnrN proteins. In rhizobia, this enables FixK and FnrN to regulate the same genes and thus act in a parallel and often redundant fashion.

4.2 Role of FnrN in oxygen regulation

E. coli fnr is expressed regardless of oxygen concentration but is subject to negative autoregulation under microaerobic conditions. In contrast expression of rhizobial *fnrN* homologs appears to occur only under microaerobic conditions, but autoregulation is also found in these species.

Both *fnrN* genes in Rlv UPM791 are positively and negatively autoregulated under microaerobic conditions (Colombo et al., 2000). This effect is due to the presence of two anaeroboxes in the promoters of the *fnrN* genes; a high-affinity site in the distal region (-42.5 relative to the TSS) and a low-affinity site in the proximal region (-10). FnrN binding to the distal anaerobox induces *fnrN* expression. At high FnrN concentration, the protein binds to the low-affinity site and represses *fnrN* transcription. This dual regulatory mechanism is proposed to balance FnrN concentration. Two anaeroboxes are also found in front of the *fnrN* genes of Rlv VF39, Rlv 3841 and *R. etli* CFN42, suggesting a similar auto-regulatory balancing mechanism operates in these rhizobia (Lopez et al., 2001; Shaw, 1984; Young et al., 2006). Of note, a dual anaerobox arrangement is also found in front of the *fixGHIS* operons of all three organisms. Expression of this complex may therefore also be under dual feedback control, but this has yet to be investigated.

Rlv VF39 *fnrN* is strongly induced under microaerobic conditions as expected from the presence of anaeroboxes upstream of the gene. The presence of these anaeroboxes suggests FixK or FnrN, or both, act to regulate *fnrN* expression. Conflicting results have been published on the importance of *fnrN* autoregulation in this strain. Several studies reported positive autoregulation of *fnrN* (Clark, Oresnik, & Hynes, 2001; Schlüter et al., 1992). Another study reported that hFixL was essential for microaerobic *fnrN* induction, suggesting no or limited autoregulation (Boesten & Priefer, 2004). FixK apparently did not mediate this regulation and no FxkR-binding K-box has been identified in front of *fnrN*. It is therefore unclear how hFixL regulates *fnrN* in this model. The finding that autoregulation and hFixL both play important roles may be reconciled by a model in which the positive autoregulatory feedback loop must be initiated by hFixL-dependent induction. However, this is in apparent contradiction to other results showing hFixL and FnrN are redundant regulators and that a *hfixL* mutant retains nitrogen fixation activity (Moris et al., 2004; Patschkowski

et al., 1996). Further investigations will be required to clarify the mechanisms leading to *fnrN* expression in Rlv VF39.

In *R. etli* CNPAF512, two possible regulatory mechanisms also exist with the potential to regulate microaerobic *fnrN* expression; autoregulation and control by the organism's traditional FixLJ cascade (Moris et al., 2004). Two anaeroboxes are present in front of the gene and positive autoregulation of FnrN under microaerobic conditions has been confirmed. To investigate the possibility of *fnrN* control by a FixLJ-FixK system, the gene's native *R. etli* CNPAF512 promoter was introduced into the *S. meliloti* host. However, no regulation by *S. meliloti*'s endogenous FixLJ-FixK system was found. Transcription of *fnrN* in *R. etli* CNPAF512 thus appears to rely solely on autoregulation. This implies FixK does not bind the gene's anaeroboxes to regulate it, or this binding is biologically irrelevant.

The CFN42 strain of *R. etli* has two differentially regulated copies of *fnrN*, *fnrN*d and *fnrN*chr (Lopez et al., 2001; Moris et al., 2004). Under microaerobic conditions *fnrN*d is induced through at least three mechanisms; by FnrNchr, by hFixL and by FnrNd autoregulation. Single mutants in *fnrN*chr, *hfixL* or *fnrN*d only partially reduced microaerobic induction of *fnrN*d. A single mutant in *fixK*f did significantly decrease the gene's expression, whilst *fixK*d had no effect on transcription of either *fnrN* homolog. It is believed FixKf plays such a crucial role because it is required not only for induction by the hFixL-FxkR pathway but also for FnrNd production and consequently *fnrN*d autoregulation. This supports the existence of a cooperative induction model, also theorized in Rlv VF39, with both hFixL TCS-based regulation and autoregulation of FnrN. FixKf is likely involved in both mechanisms, explaining its critical role in *fnrN*d expression. By contrast, expression of the second *fnrN* copy *fnrN*chr instead appears to be largely repressed by both FnrNd and the hFixL-FxkR pathway under microaerobic conditions. However, no expression of *fnrN*chr occurred in an *fnrN*d *hfixL* double mutant, suggesting these proteins also have a positive regulatory role. Of note, hFixL-FxkR repression of *fnrN*chr was not relieved by a mutation in *fixK*f suggesting FxkR directly inhibits transcription. However, no consensus K-box is present in the *fnrN*chr promoter, so it is unclear how FxkR regulates the gene's expression. Further, it has been proposed FnrNchr acts to induce *fixK*f, resulting in another regulatory feedback loop that may act to balance the production of FnrNchr or FixKf, or both. This multitude of interconnections between the hFixL-based cascade and the two FnrN homologs results in exquisitely complex regulation that is

evidence of oxygen's essential and finely tuned role in controlling nitrogen fixation in *R. etli* CFN42.

Besides autoregulation, genes induced by FnrN commonly include the *fixNOQP* and related *fixGHIS* operons. FnrN and FixK proteins often appear to function in a redundant fashion when they coexist in an organism, as found in Rlv VF39 and *R. etli* CFN42 (Andreas Schlüter et al., 1997; Lopez et al., 2001; Zamorano-Sánchez & Girard, 2015). The former encodes two copies of the terminal oxidase, *fixNOQP*c and *fixNOQP*d. FnrN is an important regulator of both, with the hFixL-FxkR pathway also shown to play a significant role (Boesten & Priefer, 2004; Clark et al., 2001; Schlüter et al., 1992, 1993).

Redundancy can also exist in situations where no class IC FixK homolog is present, as is the case in Rlv UPM791 (Gutiérrez et al., 1997). This organism encodes two copies of *fnrN*, a duplication which appears to compensate for the lack of a FixK homolog. Both *fnrN* homologs are individually sufficient for microaerobic induction of *fixNOQP* (Gutiérrez et al., 1998; Hernando et al., 1995). Thus FnrN and NifA together appear sufficient to regulate expression of core nitrogen fixation machinery and associated functions with no TCS involvement identified to date (Colombo et al., 2000; Ruiz-Argüeso, Palacios, & Imperial, 2001; Sánchez-Cañizares et al., 2018).

Even more redundancy is found in *R. etli* CFN42, which encodes two copies of *fnrN* (*fnrN*chr and *fnrN*d) (Lopez et al., 2001). Single *fnrN* mutants show no effect on *fixNOQP* expression but half WT levels were reported in a double *fnrN* mutant. The remainder of *fixNOQP* induction is due to FixKf regulation under control of the hFixL-FxkR pathway. Of note, the anaeroboxes of all *fnrN* and *fixNOQP* promoters are identical in this strain but differences are nevertheless observed in the regulatory roles of FnrNchr, FnrNd and FixKf. This may be due to differences in their respective anaerobox affinities, a mechanism likely to play a role in other rhizobia as well. The importance of these three regulators also varies temporally. At 32 days post-inoculation a double *fnrN* mutant had a limited effect on nitrogen fixation. At 42 days, the same *fnrN* double mutant drastically reduced activity. In this strain, the FnrN proteins therefore appear to become more important in the later stages of symbiosis. This is in line with *R. etli* CFN42's use of a highly complex oxygen regulation network to enable very finely tuned regulation, both at a spatial and temporal level. In a similar vein, in *R. etli* CNPAF512 *fnrN* is important for *fixNOQP* and *fixGHIS* expression in the early stages of symbiosis but plays a smaller role in the late stage (Moris et al., 2004).

B. japonicum's class IB FNR-like protein $FixK_1$ operates in a hierarchical cascade, in contrast to the redundancy found in other rhizobia (Nellen-Anthamatten & Rossi, 1998). The *fixK*$_1$ gene is microaerobically induced under control of the canonical FixLJ cascade via the class IC $FixK_2$ protein (Socorro Mesa et al., 2008). The $FixK_1$ regulon is far smaller than that of $FixK_2$ and microaerobic induction of both *fixNOQP* and *fixGHIS* depends exclusively on the action of $FixK_2$. There is also evidence that $FixK_1$ acts to repress genes that are activated by NifA, ensuring these remain unexpressed even at low oxygen concentration until the correct symbiotic conditions have been reached. This difference in the role of *B. japonicum* $FixK_1$ and $FixK_2$ may be because the two proteins no longer bind the same anaerobox. Similar differences may exist between FNR-like proteins across rhizobia more generally, but the structural basis for these putative divergences has yet to be investigated.

5. The NifA transcription factor

NifA is the central, essential activator of nitrogen fixation across most studied symbiotic and non-symbiotic diazotrophs (Dixon, 1984; Dixon & Kahn, 2004). In contrast to the FixLJ, hFixL-FxkR and FnrN regulatory systems, which are generally thought to act in the earlier phases of symbiotic establishment, NifA is primarily involved in the final stage of bacteroid differentiation (Hirsch & Smith, 1987). Key targets of NifA commonly include the nitrogenase (*nif*) genes found across nitrogen fixing species and the *fix* genes specific to symbiotic diazotrophs (Earl, Ronson, & Ausubel, 1987; Edgren & Nordlund, 2004; Martínez, Palacios, Imperial, & Ruiz-Argüeso, 2004; Schetgens, Hontelez, van den Bos, & van Kammen, 1985). The protein and its associated regulatory mechanisms have been best studied in non-symbiotic diazotrophs, notably *Klebsiella pneumoniae* (Postgate, 1982). In these species NifA regulation at the transcriptional and post-transcriptional level integrates multiple signals including oxygen concentration, the energy status of the cell and its nitrogen availability (Alexandre Kaminski & Elmerich, 1998; Dixon & Kahn, 2004; Martinez-Argudo, Little, Shearer, Johnson, & Dixon, 2005). In symbiotic diazotrophs, oxygen is the primary and often apparently sole NifA regulator at both the transcriptional and protein level (Fischer, 1994). Transcriptional regulation of *nifA* is more poorly conserved than that of *fixK* or *fnrN* and differs substantially at the species and strain level. Complementation relying on heterologous expression of *nifA* from its native promoter therefore often fails, but has been demonstrated in a

few cases (Baldani et al., 2011; Isabel Martinez-Argudo, Little, Shearer, Johnson, & Dixon, 2004; Yao et al., 2006). The ability of NifA to regulate targets is much better conserved, so cross-species complementation is common if the protein is successfully expressed. Expression of *B. japonicum*, *S. meliloti* and Rlv UPM791 NifA targets were activated by heterologous *K. pneumoniae* NifA (Alvarez-Morales & Hennecke, 1985; Beynon, Williams, & Cannon, 1988; Hans Martin Fischer & Hennecke, 1987; Martínez, Colombo, Palacios, Imperial, & Ruiz-Argüeso, 2008; Sundaresan, Jones, Ow, & Ausubel, 1983). Likewise, *S. meliloti* NifA activated expression of *nif* gene promoters in *K. pneumoniae* and other rhizobia (Alvarez-Morales & Hennecke, 1985; Beynon et al., 1988).

5.1 Structure of NifA

NifA is a σ^{54}-dependent enhancer binding protein (EBP) (Studholme & Dixon, 2003). The protein's flexibility has to date prevented the resolution of a full crystal structure. At the center of the protein is a highly conserved AAA + domain (ATPase associated with diverse cellular activities) (Bush & Dixon, 2012). This contains protein binding interfaces enabling NifA oligomerization and activates transcription by driving DNA unwinding and σ^{54} complex formation. It appears to be the only domain essential for transcriptional activation by NifA (Arsène, Kaminski, & Elmerich, 1996, 1999; Berger et al., 1995). Truncation of the other two domains retained activity, albeit with a loss of regulatory control.

The central domain is flanked at the C-terminus by a DNA-binding HTH domain that is also well conserved amongst members of the EBP family (Morett & Buck, 1988; Zou et al., 2008). This binds upstream activator sequence (UAS) motifs with a TGT-N_{10}-ACA consensus sequence that is conserved across free living and symbiotic diazotrophs (Avarez-Morales, Betancourt-Alvarez, Kaluza, & Hennecke, 1986; Buck, Miller, Drummond, & Dixon, 1986). In contrast to the anaerobox operator bound by FnrN and FixK proteins which is generally found within a hundred base pairs of the TSS, UAS sites can be up to 1 Kbp upstream of target genes (Gubler, 1989; Ninfa, Reitzer, & Magasanik, 1987). NifA forms a loop in the intervening sequence between the UAS and the TSS during its interaction with RNA polymerase to activate transcription (Martínez et al., 2004). Because of the range at which this mechanism can occur, functional UAS sites need not be within the traditional promoter region of a gene and can be located within the coding sequence of upstream genes (Hauser et al., 2007).

The third domain of NifA is an N-terminal GAF regulatory domain, named after three proteins in which it is found (Pflüger-Grau & Görke, 2010). Although the GAF domain appears to consistently have a regulatory function, its sequence is poorly conserved and the mechanisms and signals it responds to vary significantly across species (Ho, Burden, & Hurley, 2000; Huala & Ausubel, 1989). In non-symbiotic diazotrophs including *K. pneumoniae* and *Azotobacter vinelandii* the GAF domain plays an important role in NifA repression by NifL binding (for details see Arsène, Kaminski, & Elmerich, 1999; Isabel Martinez-Argudo et al., 2004; Ninfa & Jiang, 2005; Schmitz, Klopprogge, & Grabbe, 2002). This interaction between NifA and NifL integrates multiple regulatory signals in *K. pneumoniae*. Expression of *nifL* is activated by NtrC in response to low nitrogen availability (Beynon et al., 1988; Ray Dixon et al., 1980). Further, NifL binds an FAD cofactor allowing it to sense oxygen (Dixon & Kahn, 2004; Hill, Austin, Eydmann, Jones, & Dixon, 1996). The oxidation state of FAD responds to oxygen and induces changes in NifL conformation and its sub-cellular localization. These oxygen-regulated effects control both its binding and repression of NifA. NifL regulation of NifA thus combines a nitrogen signal at the level of *nifL* transcription and an oxygen signal at the protein level (Drummond, Contreras, & Mitchenall, 1990; Moreno et al., 1992). This repression is further regulated by the NifA GAF domain which is required for NifL binding (Arnott, Sidoti, Hill, & Merrick, 1989). In *A. vinelandii* GAF binds 2-oxoglutarate as a proxy for sensing nitrogen limiting conditions (Little & Dixon, 2003). Binding of 2-oxoglutarate prevents GAF from interacting with NifL, abolishing NifL-mediated oxygen repression. This enables *A. vinelandii* to perform nitrogen fixation under aerobic conditions, in conjunction with a host of mechanisms protecting the nitrogenase from oxygen (Drozd & Postgate, 2009; Liu et al., 1995; Sabra, Zeng, Lünsdorf, & Deckwer, 2000). Derepression of NifA and activation of fixation under nitrogen limiting conditions in *A. vinelandii* thus requires the GAF domain. In *K. pneumoniae*, the GAF domain does not bind 2-oxoglutarate and nitrogen sensing functions primarily through NifA interaction with the GlnK protein instead. There is some evidence that NifL does play a role in nitrogen regulation but it remains unclear what function the GAF domain serves in *K. pneumoniae* (Buchanan-Wollaston, Cannon, Beynon, & Cannon, 1981).

No homolog of NifL is found in rhizobial diazotrophs and the role of the GAF domain in these organisms is unknown (Dixon & Kahn, 2004; Loroch et al., 1995). Deletion of the domain in *S. meliloti* and *B. japonicum* had minimal effect on NifA activity and did not influence oxygen regulation of the

protein (Hans Martin Fischer, Bruderer, & Hennecke, 1988; Huala & Ausubel, 1989). One study reported a strong increase in *S. meliloti* NifA activity upon deletion of the GAF domain (Beynon et al., 1988). In the same species partial domain deletions abolished NifA functionality, suggesting these led to protein misfolding or instability (Huala & Ausubel, 1989). In *R. leguminosarum* bv. *trifolii*, NifA does not have a GAF domain (Iismaa & Watson, 1989). The GAF domain therefore appears not to be required for NifA activity in rhizobia, but a regulatory function remains likely. It may act to regulate NifA activity in the presence or absence of an as yet unidentified interacting protein or small molecule. *A. caulinodans* likely employs a NifA control system combining elements from free-living and symbiotic diazotrophs as it is able to fix nitrogen under both conditions (Alexandre Kaminski & Elmerich, 1998; Fischer, 1994; Michel-Reydellet & Kaminski, 1999). Uniquely amongst rhizobia studied to date, its NifA is inactive in conditions of high fixed nitrogen concentration. It has been theorized this effect is due to repression by an as-yet unidentified protein functionally similar to NifL. Deletion of GAF abolishes NifA function in *A. caulinodans*, suggesting the domain is required to relieve this repression, as shown in *K. pneumoniae*. Similarly, in *Azospirillum brasilense* the GAF domain of NifA apparently interacts directly with the P_{II} protein to mediate regulation in response to fixed nitrogen levels (Arsène et al., 1996, 1999). Sequence conservation in the GAF domain of *Herbaspirillum seropedicae* indicates a similar mechanism operates in that species (Souza, Funayama, Rigo, Yates, & Pedrosa, 1991). *A. caulinodans* NifA may therefore interact with a protein which regulates its activity in response to nitrogen availability, with the GAF domain relieving this inhibition. This putative mechanism is in contrast to most other symbiotic diazotrophs, where nitrogen levels are found not to regulate NifA (Dixon & Kahn, 2004). As discussed in more detail below, *A. caulinodans* NifA is sensitive to oxygen at the protein level so the GAF domain and NifL homolog likely play no role in oxygen control.

In symbiotic diazotrophs NifL-mediated oxygen control is replaced with direct oxygen sensing by the NifA protein (Beynon et al., 1988; Krey, Pühler, & Klipp, 1992). Oxygen sensitivity at the protein level has been demonstrated for *S. meliloti* and *B. japonicum* NifA. This is mediated by a broadly conserved cysteine-rich motif present in an inter-domain linker (IDL) region between the central AAA+ and HTH domains. All of the conserved cysteine residues in this motif were shown to be essential for the activity of *B. japonicum* NifA (Hans Martin Fischer et al., 1988). The same conserved residues are present in the *A. caulinodans*, *S. meliloti* and

R. leguminosarum NifA proteins. Several non-rhizobial species employing NifA homologs, including *A. brasilense* and *H. seropedicae*, also retain the IDL but it is not found in the NifA proteins of *K. pneumoniae* or *A. vinelandii* and appears to be mutually exclusive with the NifL-NifA regulatory system. The cysteine-rich motif is similar to metal-binding domains found in other proteins (Fischer, Fritsche, Herzog, & Hennecke, 1989). It is believed to sense the redox state of the cell as an indirect measure of oxygen concentration (Bauer, Kaspar, Fischer, & Hennecke, 1998; Klipp, Reiländer, Schlüter, Krey, & Pühler, 1989; Krey et al., 1992). A presently well supported model is that under microaerobic conditions a reduced metal ion, likely Fe^{2+}, is bound by the motif. Binding of this ion at the IDL is likely required for a conformational change that produces active NifA. Metal ion binding also appears required for protein stability. In *S. meliloti* and *B. japonicum*, NifA is rapidly degraded under aerobic conditions. Indirect oxygen control through metal ion binding by the IDL therefore strictly regulates activation by rhizobial NifA proteins.

5.2 Role of NifA in oxygen regulation

Rhizobia typically regulate *nifA* transcription in response to oxygen concentration, in contrast to free-living diazotrophs where expression responds to nitrogen levels (Dixon & Kahn, 2004; Martínez et al., 2004). Microaerobic induction is achieved through several mechanisms often acting in combination (Fig. 4). Positive NifA autoregulation resulting in a feedback loop is a very common motif in rhizobia, likely enabling rapid production of the protein upon transition to nitrogen fixation in nodules. Several species also employ a second layer of oxygen regulation at the transcriptional level, including through the FixLJ cascade.

In *A. caulinodans nifA* is transcribed from two promoters. The first is σ^{54}-dependent, and the second σ^{70}-dependent and regulated by the FixLJ system through FixK (Alexandre Kaminski & Elmerich, 1998; Nees, Stein, & Ludwig, 1988). FixK was required under symbiotic and free-living conditions, demonstrating the FixLJ system plays a role in both contexts in *A. caulinodans* (Loroch et al., 1995). The two promoters appear to be integrated in a single feedback system. Activation of the σ^{70} promoter upregulates activity from the downstream σ^{54} promoter. A putative NifA-binding motif overlaps with the σ^{70} promoter and NifA binding here likely represses transcription from both promoters. This would balance NifA production, not unlike the feedback demonstrated for *fnrN* expression in some species (see Section 4). The nitrogen status of the cell also plays a role in *nifA* regulation,

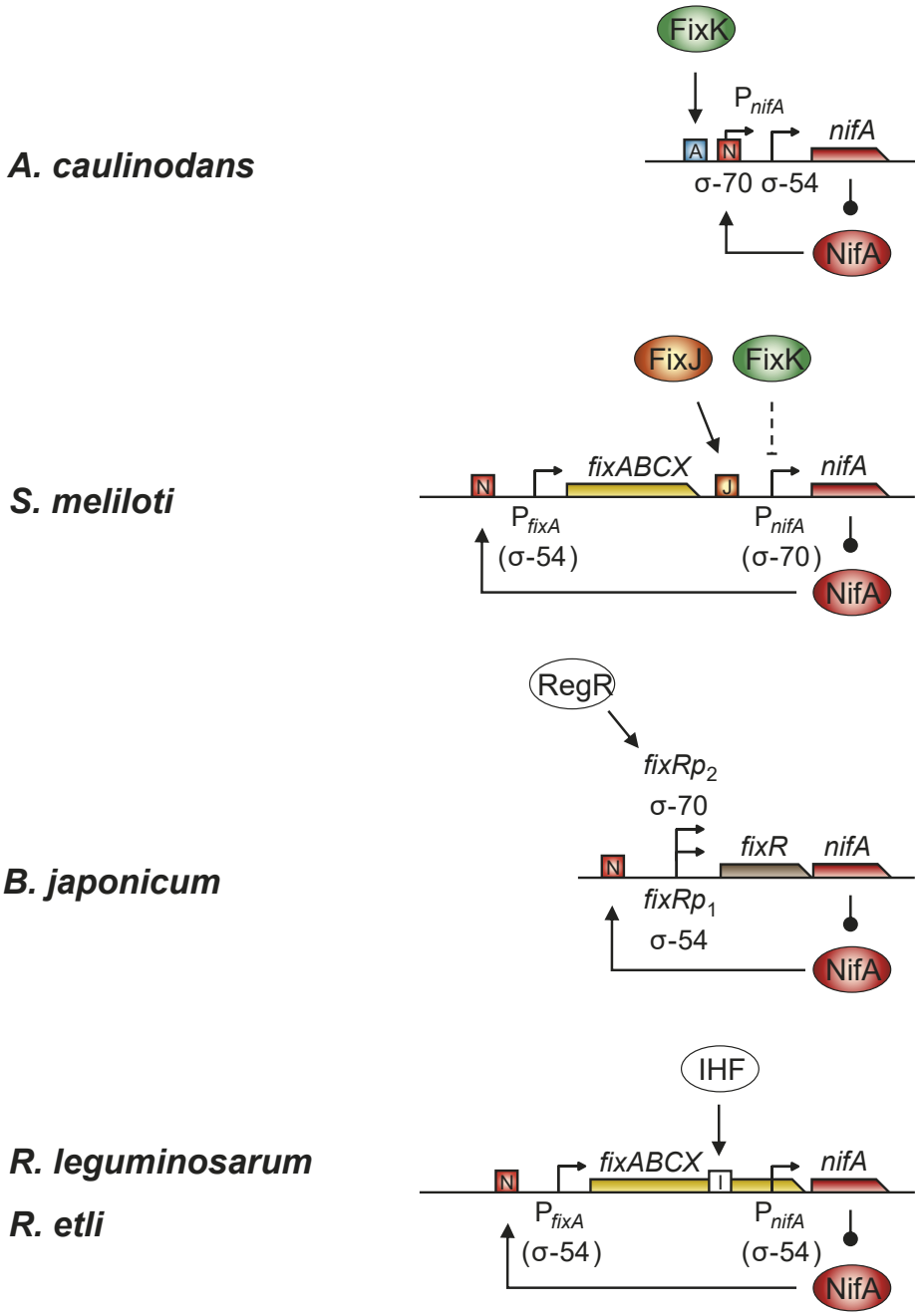


Fig. 4 ***Transcriptional regulation of nifA across rhizobia.*** Transcriptional activation and repression are shown as lines ending in arrows and bars, respectively. Squares indicate operators for NifA (N, red), FixJ (J, orange), the anaerobox bound by FixK (A, blue) and the IHF factor operator (I, white). Genes and operons are shown as pointed rectangles. Translation is shown as a line ending in a circle. Proteins are shown as ellipses. The *nifA* gene and its protein are highlighted in red, *fixABCX* in yellow and *fixR* in brown. The FixJ and FixK proteins are given in orange and green, respectively. Proteins which activate transcription but are not directly involved in oxygen regulation are coloured white. Diagrams represent typical systems, but regulation varies at the strain level; see text for details.

via the *ntrBC* and *ntrYX* systems (for details see Pawlowski, Klosse, & de Bruijn, 1991; Pawlowski, Ratet, Schell, & de Bruijn, 1987; Ratet, Pawlowski, Schell, & de Bruijn, 1989).

In *S. meliloti* a similar dual feedback system operates with two promoters, P_{fixA} (σ^{54}) and P_{nifA} (σ^{70}), regulating *nifA* (Kim, Helinski, & Ditta, 1986; Thöny, Anthamatten, & Hennecke, 1989). The gene is downstream of and co-transcribed with the *fixABCX* operon, itself a target of NifA activation. Thus *nifA* is under P_{fixA} control and auto-regulated, an effect shown to be essential for its expression in nodules (Szeto, Nixon, Ronson, & Ausubel, 1987; Thöny et al., 1989). The second promoter P_{nifA} directly upstream of *nifA* is also induced under microaerobic conditions, by FixJ (David et al., 1988; Ditta, Virts, Palomares, & Kim, 1987). Simultaneously FixJ induces *fixK*, but FixK represses transcription from P_{nifA}, thereby balancing *nifA* expression when the FixLJ system is active (Batut et al., 1989). This system does not appear widespread, and the mechanism of FixK-mediated P_{nifA} repression is unclear. *B. japonicum nifA* is not regulated by FixLJ and likewise there are no FixJ or anaerobox operators upstream of *R. leguminosarum* or *R. etli nifA* (Denise Anthamatten & Hennecke, 1991; Enrique Morett, Fischer, & Hennecke, 1991).

A similar dual control system operates in Rlv UPM791, where the adjacent *fixABCX-nifAB* operons are inside a 10-member co-transcribed gene cluster (Martínez et al., 2004). Expression of *nifA* occurs only under symbiotic conditions in this strain (Brito et al., 1997). Two σ^{54} type promoters regulate expression of the gene, P_{nifA1} transcribing the full cluster and P_{nifA2}, likely located immediately upstream of *fixA*. The promoter regulating the full cluster is auto-activated by NifA and responsible for a majority of *nifA* expression. The P_{nifA2} promoter appears to be an incomplete duplication of the first and does not contain NifA-binding UAS sites. A deletion of P_{nifA1} expressed *nifA* sufficiently from P_{nifA2} to retain limited nitrogen fixation. Sequence information suggests a similar dual promoter system operates in Rlv 3841 (Young et al., 2006). It is possible other regulators act on P_{nifA2}, but their identity has yet to be determined.

In *B. japonicum* and *R. etli* CNPAF512 *nifA* is auto-activated but regulators which do not respond to oxygen also participate in its transcription (Kim et al., 1986; Michiels, D'hooghe, Verreth, Pelemans, & Vanderleyden, 1994; Beat Thöny & Hennecke, 2006). Expression of *nifA* in these organisms occurs under aerobic conditions (Beat Thöny, Fischer, Anthamatten, Bruderer, & Hennecke, 1987; Michiels et al., 1994). This suggests the emphasis is on post-transcriptional oxygen regulation of NifA, as found in *A. brasilense*.

Auto-activation in *R. etli* CNPAF512 proceeds from a promoter upstream of *fixABCX* as in Rlv UPM791, accounting for at least half of *nifA* expression (Kim et al., 1986; Michiels et al., 1994). A second promoter exists directly upstream of *nifA* that does not respond to oxygen concentration (Benhassine, Fauvart, Vanderleyden, & Michiels, 2007). It was recently shown that this second promoter is regulated by an IHF-like protein, a mechanism also found in *H. seropedicae* (Wassem, de Souza, Yates, Pedrosa, & Buck, 2000; Wassem et al., 2002). The implications of this regulation are unclear, and it remains to be determined whether this induces or represses *nifA* transcription.

In *B. japonicum* the gene is part of a *fixR-nifA* operon (Bauer et al., 1998). Two overlapping promoters regulate expression, $fixRp_1$ and $fixRp_2$ of type σ^{54} and σ^{70} respectively (Barrios, Fischer, Hennecke, & Morett, 1995). Under microaerobic conditions $fixRp_1$ is auto-activated by NifA, resulting in five-fold induction (Thöny et al., 1989). Under aerobic conditions, *nifA* is expressed but the FixLJ-regulated *rpoN* σ^{54} factor required for NifA activity is not (Gussin, Ronson, & Ausubel, 1986; Kullik et al., 1991; Kustu, Santero, Keener, Popham, & Weiss, 1989). Thus, non-activating binding of NifA to the $fixRp_1$ promoter instead results in auto-repression when the right oxygen conditions are not present. It is likely this mechanism also represses other members of the NifA regulon. The second, housekeeping $fixRp_2$ promoter controlling *nifA* expression is regulated in part by the RegS-RegR (RegSR) TCS (Emmerich, Hennecke, & Fischer, 2000; Lindemann et al., 2007). A mutation of *regR* was found to reduce nitrogen fixation to 2% of WT levels, demonstrating the importance of this second promoter. The RegSR system has since been shown to regulate expression of other targets which play important roles in antibiotic resistance, symbiotic host specificity and denitrification (Lindemann et al., 2010; Torres et al., 2014). RegR activation of *nifA* expression is independent of oxygen, suggesting the gene can be expressed under aerobic conditions (Bauer & Wu, 2008; Emmerich, Panglungtshang, Strehler, Hennecke, & Fischer, 1999). RegS is instead believed to sense redox conditions at the cell surface in a manner analogous to the well-studied *Rhizobium capsulatus* RegBA system. Of note, a *regS* mutant only minimally reduced nitrogen fixation activity. This is in line with denitrification studies which suggest RegR is also regulated by an as yet unidentified alternative sensor kinase (Bauer et al., 1998). RegR control of *nifA* therefore likely integrates at least one additional control signal that has yet to be identified.

Beyond its well-studied *nif* and *fix* gene targets, active NifA controls a far larger regulon and many members of this have yet to be investigated (Salazar et al., 2010). It is theorized that control of some NifA targets evolved through "regulatory noise", wherein horizontally transferred genes adapt to their host through changes in their operator sequences (De Lorenzo & Pérez-Martín, 1996; Martínez et al., 2008). In some rhizobia NifA activates expression of hydrogenase genes allowing cells to use the hydrogen produced by nitrogen fixation as an energy source (Brito et al., 1997; Martínez et al., 2008; Robson & Postgate, 1980; Schubert & Evans, 1976). In *S. meliloti* and *B. japonicum* NifA also regulates a ferredoxin critical for nitrogen fixation which may function as a direct electron donor to nitrogenase (Hauser et al., 2007; Klipp et al., 1989). The NifA regulon in *B. japonicum* appears so broad that it has been suggested the protein should be thought of as a general regulator of anaerobic processes rather than one specific for nitrogen fixation. This may explain why at least part of its expression, through the σ^{70} $fixRp_2$ promoter, apparently escapes oxygen regulation. A recent study of *B. japonicum* DOA9, which like *A. caulinodans* is able to fix under free-living and symbiotic conditions, found that this strain employed two *nifA* homologs (Wongdee, Boonkerd, Teaumroong, Tittabutr, & Giraud, 2018). Both were individually sufficient for successful symbiotic nitrogen fixation, but one was specifically essential for free-living fixation. Like the functional similarity of hFixL-FxkR and FnrN, reiteration of NifA may be another rhizobial strategy to improve the robustness of oxygen regulation and create a more finely-tuned system.

As found with FixJ, the NifA regulon is probably augmented by its control of transcription factors. One of these may include the FixLJ cascade itself. In *S. meliloti*, a *nifA* mutant appeared to show increased expression of *fixLJ* (Tian et al., 2006). NifA may therefore indirectly suppress FixLJ targets during the final stage of the symbiosis. Controlling additional regulatory mechanisms would enable the protein to indirectly influence a large pool of genes beyond its direct targets. The full role of NifA thus appears to extend well beyond what is currently known.

6. Conclusions and perspectives

Rhizobia undergo a complex lifestyle transition in their partnership with legumes which requires finely tuned regulation. Nitrogen fixation is energy intensive but the nitrogenase complex must operate under microaerobic conditions, creating a conflicting demand for oxygen. These

paradoxical requirements have driven a host of evolutionary adaptations in legumes and rhizobia to establish a successful symbiosis. As rhizobia differentiate from a free-living lifestyle in soil to nitrogen fixing bacteroids in nodules, the oxygen concentration drops several orders of magnitude. Multiple oxygen-sensing mechanisms, including the FixLJ and hFixL-FxkR cascades and NifA, enable them to respond appropriately throughout development. The FixLJ and hFixL-FxkR cascades activate at a relatively high oxygen concentration to control expression early in symbiosis while NifA is activated at a lower oxygen concentration and accordingly regulates the final stages of bacteroid differentiation. The oxygen range covered by FnrN is not yet well understood, and further work is required to understand its role. They may bridge the gap between FixLJ or hFixL-FxkR and NifA regulation. In *R. etli* CFN42, regulation by hFixL-FxkR is important at the start of symbiosis but FnrN becomes critical in later stages (Lopez et al., 2001).

There is considerable inter- and intra-species variation in oxygen regulation. Much of this is due to the spread of symbiotic nitrogen fixation through horizontal gene transfer (HGT) of symbiosis islands (Masson-Boivin & Sachs, 2018). However, several regulatory themes are nevertheless broadly conserved. Rhizobial oxygen regulators consistently interact at multiple levels, be they FixK, FnrN or NifA. Although hierarchical systems have been found, it is more common for regulators to overlap and create a degree of redundancy. The *fixNOQP* and *fixGHIS* operons appear to be universally under oxygen control, often redundantly by regulators such as FixK and FnrN. Redundancy is also achieved by the reiteration of regulators within an organism, with two (or more) homologs playing subtly different roles.

Autoregulation is also common for FixK, FnrN and NifA, acting both to balance protein production and to enable a rapid increase in transcription through a self-amplifying loop. This suggests rapidly producing these proteins is important to establish symbiosis, although the reason for this remains unclear. Shifting *nifA* regulation from the transcriptional to protein-level as seen in some rhizobia may also serve to produce more responsive regulation. It has been suggested legumes decrease the pH of symbiosomes as a method to force rhizobia to metabolize succinate and produce ammonia, both of which counteract this acidification (Kannenberg & Brewin, 1989). Some of the changes rhizobia undergo during symbiosis may therefore need to occur quickly for them to survive. Another possibility is a competitive pressure between rhizobia to rapidly demonstrate their ability to fix nitrogen. Legumes can selectively sanction certain nodules if others show a higher

fixation rate (Batstone, Dutton, Wang, Yang, & Frederickson, 2017; Kiers, Rousseau, West, & Denlson, 2003; Westhoek et al., 2017). This raises the possibility that bacteria in developing nodules are selected for how quickly they begin to fix nitrogen, as well as the final rate at which they fix.

We perceive several promising areas for future work. The role of the NifA GAF domain in rhizobia remains essentially unknown. It apparently plays no role in oxygen sensing and likely integrates one or more other signals. Identifying these will provide important insights into other mechanisms regulating symbiotic establishment, and how these interact with oxygen regulation. It should be noted that the low conservation of the GAF domain suggests these signals differ between species. Broadening our understanding of the 'oxygen regulon' in rhizobia represents another promising avenue. Recent work makes it increasingly clear that the systems discussed in this review act on many targets which remain to be identified and investigated. NifA and FixK in particular have been shown to control very large regulons, directly or indirectly. One possibility is that these systems also play a role outside of symbiosis.

Other pathways involved in symbiotic oxygen regulation remain to be discovered and understood. The NtrR protein of *S. meliloti* regulates microaerobic *nif* and *fix* gene induction as well as several unrelated metabolic functions, but it is unknown what signal the protein responds to (Kelemen et al., 2004). An intermediate metabolite of purine biosynthesis represses *fixK* in *S. meliloti* through an as yet unidentified pathway (Mario Soberón, Morera, Kondorosi, Lopez, & Miranda, 2001). A purine metabolism mutant of *R. etli* was found to express the *fixNOQP* operon under free-living conditions (Soberón, Lopez, Miranda, Tabche, & Morera, 1997; Soberón et al., 2016). This and other connections likely integrate as-yet unidentified signals important in the regulation of nitrogen fixation, which may include signals from the legume host. These additional inputs no doubt play important roles in supplementing, modifying or balancing oxygen regulation, and understanding them will be essential to arrive at a holistic picture of nitrogen fixation control in rhizobia.

Acknowledgments

The authors would like to thank Dr Tim Haskett, Dr Carmen Sánchez-Cañizares and Prof Lee Sweetlove for their advice and critically reviewing the manuscript. They would also like to thank Dr Beatriz Jorrín for her input on the design of the figures. This work was supported by the Biotechnology and Biological Sciences Research Council [grant numbers BB/N003608/1 and BB/K006134/1]. PJR would also like to gratefully acknowledge his scholarship from the Biotechnology and Biological Sciences Research Council, grant number BB/M011224/1.

References

Alexandratos, N., & Bruinsma, J. (2012). World agriculture towards 2030/2050. *Land Use Policy, 20*(4), 375. https://doi.org/10.1016/S0264-8377(03)00047-4.

Allen, R. S., Tilbrook, K., Warden, A. C., Campbell, P. C., Rolland, V., Singh, S. P., et al. (2017). Expression of 16 nitrogenase proteins within the plant mitochondrial matrix. *Frontiers of Plant Science, 8*, 287. https://doi.org/10.3389/fpls.2017.00287.

Alvarez-Morales, A., & Hennecke, H. (1985). Expression of Rhizobium japonicum nifH and nifDK operons can be activated by the *Klebsiella pneumoniae* nifA protein but not by the product of ntrC. *MGG Molecular & General Genetics, 199*(2), 306–314. https://doi.org/10.1007/BF00330273.

Andrews, M. M. E., & Andrews, M. M. E. (2017). Specificity in legume-rhizobia symbioses. *International Journal of Molecular Sciences, 18*(4), 705. https://doi.org/10.3390/ijms18040705.

Anthamatten, D., & Hennecke, H. (1991). The regulatory status of the fixL- and fixJ-like genes in *Bradyrhizobium japonicum* may be different from that in *Rhizobium meliloti*. *MGG Molecular & General Genetics, 225*(1), 38–48. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF00282640.pdf.

Anthamatten, D., Scherb, B., & Hennecke, H. (1992). Characterization of a fixLJ-regulated *Bradyrhizobium japonicum* gene sharing similarity with the *Escherichia coli* fnr and *Rhizobium meliloti* fixK genes. *Journal of Bacteriology, 174*(7), 2111–2120. https://doi.org/10.1128/jb.174.7.2111-2120.1992.

Appleby, C. A. (1984). Leghemoglobin and rhizobium respiration. *Annual Review of Plant Physiology, 35*(1), 443–478. https://doi.org/10.1146/annurev.pp.35.060184.002303.

Arnott, M., Sidoti, C., Hill, S., & Merrick, M. (1989). Deletion analysis of the nitrogen fixation regulatory gene nifL of *Klebsiella pneumoniae*. *Archives of Microbiology, 151*, 180–182. Retrieved from https://link.springer.com/content/pdf/10.1007%2FBF00414436.pdf.

Arrese-Igor, C., Royuela, M., Lorenzo, C., Felipe, M. R., & Aparicio-Tejo, P. M. (1993). Effect of low rhizosphere oxygen on growth, nitrogen fixation and nodule morphology in lucerne. *Physiologia Plantarum, 89*(1), 55–63. https://doi.org/10.1111/j.1399-3054.1993.tb01786.x.

Arsène, F., Kaminski, P. A., & Elmerich, C. (1996). Modulation of NifA activity by P(II) in Azospirillum brasilense: Evidence for a regulatory role of the NifA N-terminal domain. *Journal of Bacteriology, 178*(16), 4830–4838.

Arsène, F., Kaminski, P. A., & Elmerich, C. (1999). Control of Azospirillum brasilense NifA activity by P(II): Effect of replacing Tyr residues of the NifA N-terminal domain on NifA activity. *FEMS Microbiology Letters, 179*(2), 339–343. https://doi.org/10.1016/S0378-1097(99)00426-7.

Avarez-Morales, A., Betancourt-Alvarez, M., Kaluza, K., & Hennecke, H. (1986). Activation of the *Bradyrhizobium japonicum* nifH and nifDK operons is dependent on promoter-upstream DNA sequences. *Nucleic Acids Research, 14*(10), 4207–4227. https://doi.org/10.1093/nar/14.10.4207.

Baldani, I., Lee, S., Perlova, L., De Oliveira, A., Sevilla, M., Kennedy, C., et al. (2011). Analysis of nitrogen fixation and regulatory genes in the sugar cane endophyte Acetobacter diazotrophicus. In *Nitrogen fixation with non-legumes* (pp. 11–19). Dordrecht: Springer. https://doi.org/10.1007/978-94-011-5232-7_2.

Barcellos, F. G., Menna, P., Batista, J. S. D. S., & Hungria, M. (2007). Evidence of horizontal transfer of symbiotic genes from a *Bradyrhizobium japonicum* inoculant strain to indigenous diazotrophs sinorhizobium (ensifer) fredii and Bradyrhizobium elkanii in a Brazilian Savannah soil. *Applied and Environmental Microbiology, 73*(8), 2635–2643. https://doi.org/10.1128/AEM.01823-06.

Barrios, H., Fischer, H. M., Hennecke, H., & Morett, E. (1995). Overlapping promoters for two different RNA polymerase holoenzymes control *Bradyrhizobium japonicum* nifA expression. *Journal of Bacteriology, 177*(7), 1760–1765. https://doi.org/10.1128/jb.177.7.1760-1765.1995.

Bassett, B., Goodman, R. N., & Novacky, A. (1977). Ultrastructure of soybean nodules. I: Release of rhizobia from the infection thread. *Canadian Journal of Microbiology, 23*(5), 573–582. https://doi.org/10.1139/m77-083.

Bates, D. M., Popescu, C. V., Khoroshilova, N., Vogt, K., Beinert, H., Münck, E., et al. (2000). Substitution of leucine 28 with histidine in the *Escherichia coli* transcription factor FNR results in increased stability of the [4Fe-4S]2+ cluster to oxygen. *Journal of Biological Chemistry, 275*(9), 6234–6240. https://doi.org/10.1074/jbc.275.9.6234.

Batstone, R. T., Dutton, E. M., Wang, D., Yang, M., & Frederickson, M. E. (2017). The evolution of symbiont preference traits in the model legume Medicago truncatula. *New Phytologist, 213*(4), 1850–1861. https://doi.org/10.1111/nph.14308.

Batut, J., & Boistard, P. (1994). Oxygen control in rhizobium. *Antonie Van Leeuwenhoek, 66*(1–3), 129–150. https://doi.org/10.1007/BF00871636.

Batut, J., Daveran-Mingot, M. L., David, M., Jacobs, J., Garnerone, A. M., & Kahn, D. (1989). fixK, a gene homologous with fnr and crp from *Escherichia coli*, regulates nitrogen fixation genes both positively and negatively in *Rhizobium meliloti*. *The EMBO Journal, 8*(4), 1279–1286. https://doi.org/10.1002/J.1460-2075.1989.TB03502.X.

Batut, J., de Philip, P., Reyrat, J. M., Waelkens, F., & Boistard, P. (1993). Oxygen regulation of nitrogen fixation gene expression in *Rhizobium meliloti*. In *Advances in molecular Genetics of plant-Microbe interactions* (pp. 183–191). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-0651-3_20.

Bauer, C. E., & Wu, J. (2008). RegB/RegA, a global redox-responding two-component system. In *Advances in experimental medicine and biology*. New York: Springer. https://doi.org/10.1007/978-0-387-78885-2_9.

Bauer, E., Kaspar, T., Fischer, H. M., & Hennecke, H. (1998). Expression of the fixR-nifA operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR. *Journal of Bacteriology, 180*(15), 3853–3863. Retrieved from http://jb.asm.org/.

Benhassine, T., Fauvart, M., Vanderleyden, J., & Michiels, J. (2007). Interaction of an IHF-like protein with the Rhizobium etli nifA promoter. *FEMS Microbiology Letters, 271*(1), 20–26. https://doi.org/10.1111/j.1574-6968.2007.00699.x.

Berger, D. K., Narberhaus, F., Lee, H.-S. S., Kustu, S., Berger, D. K., Narberhaus, F., et al. (1995). In vitro studies of the domains of the nitrogen fixation regulatory protein NIFA. *Journal of Bacteriology, 177*(1), 191–199. Retrieved from http://jb.asm.org/content/177/1/191.full.pdf.

Bergersen, F. J. (1994). Distribution of O2 within infected cells of soybean root nodules: A new simulation. *Protoplasma, 183*(1–4), 49–61. https://doi.org/10.1007/BF01276812.

Bergersen, F. J. (1996). Delivery of O2 to bacteroids in soybean nodule cells: Consideration of gradients of concentration of free, dissolved O2 in and near symbiosomes and beneath intercellular spaces. *Protoplasma, 191*(1–2), 9–20. https://doi.org/10.1007/BF01280821.

Bergès, H., Checroun, C., Guiral, S., Garnerone, A. M., Boistard, P., & Batut, J. (2001). A glutamine-amidotransferase-like protein modulates FixT anti-kinase activity in *Sinorhizobium meliloti*. *BMC Microbiology, 1*(6), 1–9. https://doi.org/10.1186/1471-2180-1-6.

Beynon, J. L., Williams, M. K., & Cannon, F. C. (1988). Expression and functional analysis of the *Rhizobium meliloti* nifA gene. *The EMBO Journal, 7*(1), 7–14. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16453824.

Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., et al. (1999). Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure, 7*(12), 1505–1515. https://doi.org/10.1016/S0969-2126(00)88341-0.

Bobik, C., Meilhoc, E., & Batut, J. (2006). FixJ: A major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. *Journal of Bacteriology, 188*(13), 4890–4902. https://doi.org/10.1128/JB.00251-06.

Boesten, B., & Priefer, U. B. (2004). The C-terminal receiver domain of the *Rhizobium leguminosarum* bv. viciae FixL protein is required for free-living microaerobic induction of the fnrN promoter. *Microbiology, 150*(11), 3703–3713. https://doi.org/10.1099/mic.0.27323-0.

Brito, B., Martínez, M., Fernández, D., Rey, L., Cabrera, E., Palacios, J.-M. M., et al. (1997). Hydrogenase genes from *Rhizobium leguminosarum* bv. viciae are controlled by the nitrogen fixation regulatory protein NifA. *Proceedings of the National Academy of Sciences, 94*(June), 6019–6024. https://doi.org/10.1073/pnas.94.12.6019.

Buchanan-Wollaston, V., Cannon, M. C., Beynon, J. L., & Cannon, F. C. (1981). Role of the nifA gene product in the regulation of nif expression in *Klebsiella pneumoniae*. *Nature, 294*(5843), 776–778. https://doi.org/10.1038/294776a0.

Buck, M., Miller, S., Drummond, M., & Dixon, R. (1986). Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature, 320*(6060), 374–378. https://doi.org/10.1038/320374a0.

Burén, S., López-Torrejón, G., & Rubio, L. M. (2018). Extreme bioengineering to meet the nitrogen challenge. *Proceedings of the National Academy of Sciences of the United States of America, 115*(36), 8849–8851. https://doi.org/10.1073/pnas.1812247115.

Burén, S., & Rubio, L. M. (2018). State of the art in eukaryotic nitrogenase engineering. *FEMS Microbiology Letters, 365*(2), 1–9. https://doi.org/10.1093/femsle/fnx274.

Burgess, B. K., & Lowe, D. J. (1996). Mechanism of molybdenum nitrogenase. *Chemical Reviews, 96*(7), 2983–3012. https://doi.org/10.1021/cr950055x.

Bush, M., & Dixon, R. (2012). The role of bacterial enhancer binding proteins as specialized activators of 54-dependent transcription. *Microbiology and Molecular Biology Reviews, 76*(3), 497–529. https://doi.org/10.1128/MMBR.00006-12.

Canfield, D. E., Glazer, A. N., & Falkowski, P. G. (2010). The evolution and future of Earth's nitrogen cycle. *Science, 330*(6001), 192–196. https://doi.org/10.1126/science.1186120.

Clark, S. R. D., Oresnik, I. J., & Hynes, M. F. (2001). RpoN of *Rhizobium leguminosarum* bv. viciae strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation. *MGG Molecular & General Genetics, 264*(5), 623–633. https://doi.org/10.1007/s004380000348.

Cocking, E. C., Stone, P. J., & Davey, M. R. (2005). Symbiosome-like intracellular colonization of cereals and other crop plants by nitrogen-fixing bacteria for reduced inputs of synthetic nitrogen fertilizers. *Science in China. Series C, Life Sciences, 48*(4), 888–896. https://doi.org/10.1360/062005-280.

Colombo, M. V., Gutiérrez, D., Palacios, J. M., Imperial, J., & Ruiz-Argüeso, T. (2000). A novel autoregulation mechanism of fnrN expression in *Rhizobium leguminosarum* by viciae. *Molecular Microbiology, 36*(2), 477–486. https://doi.org/10.1046/j.1365-2958.2000.01867.x.

Colonna-Romano, S., Arnold, W., Schlüter, A., Boistard, P., Pühler, A., & Priefer, U. B. (1990). An Fnr-like protein encoded in *Rhizobium leguminosarum* biovar viciae shows structural and functional homology to *Rhizobium meliloti* fixK. *MGG Molecular & General Genetics, 223*(1), 138–147. https://doi.org/10.1007/BF00315806.

Conway, G. (2000). Food for all in the 21st century. *Environment: Science and Policy for Sustainable Development, 42*(1), 8–18. https://doi.org/10.1080/00139150009604857.

Crack, J. C., Green, J., Cheesman, M. R., Le Brun, N. E., & Thomson, A. J. (2007). Superoxide-mediated amplification of the oxygen-induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR. *Proceedings of the National Academy of Sciences, 104*(7), 2092–2097. https://doi.org/10.1073/pnas.0609514104.

Crack, J., Green, J., & Thomson, A. J. (2004). Mechanism of oxygen sensing by the bacterial transcription factor fumarate-nitrate reduction (FNR). *Journal of Biological Chemistry, 279*(10), 9278–9286. https://doi.org/10.1074/jbc.M309878200.

Crack, J. C., & Le Brun, N. E. (2018). Redox-sensing iron–sulfur cluster regulators. *Antioxidants and Redox Signaling, 29*(18), 1809–1829. https://doi.org/10.1089/ars.2017.7361.

Da Re, S., Schumacher, J., Rousseau, P., Fourment, J., Ebel, C., & Kahn, D. (1999). Phosphorylation-induced dimerization of the FixJ receiver domain. *Molecular Microbiology, 34*(3), 504–511. https://doi.org/10.1046/j.1365-2958.1999.01614.x.

Dakora, F. D. (1995). A functional relationship between leghaemoglobin and nitrogenase based on novel measurements of the two proteins in legume root nodules. *Annals of Botany, 75*(1), 49–54. https://doi.org/10.1016/S0305-7364(05)80008-3.

Dakora, F., & Atkins, C. (1989). Diffusion of oxygen in relation to structure and function in legume root nodules. *Functional Plant Biology, 16*(1), 131. https://doi.org/10.1071/PP9890131.

David, M., Daveran, M.-L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., et al. (1988). Cascade regulation of nif gene expression in *Rhizobium meliloti*. *Cell, 54*(5), 671–683. https://doi.org/10.1016/S0092-8674(88)80012-6.

De Lorenzo, V., & Pérez-Martín, J. (1996). Regulatory noise in prokaryotic promoters: How bacteria learn to respond to novel environmental signals. *Molecular Microbiology, 19*(6), 1177–1184. https://doi.org/10.1111/j.1365-2958.1996.tb02463.x.

De Maagd, R. A., Yang, W. C., Goosen-de Roo, L., Mulders, I. H. M., Roest, H. P., Spaink, H. P., et al. (1994). Down-regulation of expression of the *Rhizobium leguminosarum* outer membrane protein gene ropA occurs Abruptly in interzone ll-lll of pea nodules and can Be Uncoupled from nif gene activation. *Molecular Plant-Microbe Interactions, 7*(2), 276. https://doi.org/10.1094/MPMI-7-0276.

Delgado, M. J., Bedmar, E. J., & Downie, J. A. (1998). Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation. *Advances in Microbial Physiology, 40*, 191–231. https://doi.org/10.1016/S0065-2911(08)60132-0.

Denison, R. F., & Kinraide, T. B. (1995). Oxygen-induced membrane Depolarizations in legume root nodules. *Plant Physiology, 108*(1), 235–240. https://doi.org/10.1104/pp.108.1.235.

Ditta, G., Virts, E., Palomares, A., & Kim, G. H. (1987). The nifA gene of *Rhizobium meliloti* is oxygen regulated. *Journal of Bacteriology, 169*(7), 3217–3223. https://doi.org/10.1128/jb.169.7.3217-3223.1987.

Dixon, R. (1984). The genetic Complexity of nitrogen fixation. *Microbiology, 130*(11), 2745–2755. https://doi.org/10.1099/00221287-130-11-2745.

Dixon, R., Eady, R. R., Espin, G., Hill, S., Iaccarino, M., Kahn, D., et al. (1980). Analysis of regulation of Klebsiella pneumoniae nitrogen fixation (nif) gene cluster with gene fusions. *Nature, 286*(5769), 128–132. https://doi.org/10.1038/286128a0.

Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology, 2*(8), 621–631. https://doi.org/10.1038/nrmicro954.

Dobermann, A., & Cassman, K. G. (2005). Cereal area and nitrogen use efficiency are drivers of future nitrogen fertilizer consumption. *Science in China. Series C, Life Sciences, 48*(745), 745–758. https://doi.org/10.1360/062005-268.

Downie, J. A. (2005). Legume haemoglobins: Symbiotic nitrogen fixation needs bloody nodules. *Current Biology, 15*(6), R196–R198. https://doi.org/10.1016/j.cub.2005.03.007.

Downie, J. A. (2014). Legume nodulation. *Current Biology, 24*(5), R184–R190. https://doi.org/10.1016/j.cub.2014.01.028.

Draper, J. (1986). Nitrogen Fixation and oxygen in legume root nodules. *Plant, Cell and Environment, 9*(4), 353–354. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201301759455.

Dreyfus, B. L., & Dommergues, Y. R. (1981). Stem nodules of the tropical legume, *Sesbania rostrata*. In A. H. Gibson, & W. E. Newton (Eds.), *Current Perspectives in nitrogen fixation* (p. 471). Australian Academy of Science. Retrieved from http://www.documentation.ird.fr/hor/fdi:16530.

Dreyfus, B. L., Elmerich, C., & Dommergues, Y. R. (February 1, 1983). Free-living Rhizobium strain able to grow on N2 as the sole nitrogen source. In *Applied and environmental Microbiology*. American Society for Microbiology. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16346220.

Drozd, J., & Postgate, J. R. (2009). Effects of oxygen on Acetylene reduction, cytochrome content and respiratory activity of Azotobacter chroococcum. *Journal of General Microbiology, 63*(1), 63–73. https://doi.org/10.1099/00221287-63-1-63.

Drummond, M. H., Contreras, A., & Mitchenall, L. A. (1990). The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of Klebsiella pneumoniae. *Molecular Microbiology, 4*(1), 29–37. https://doi.org/10.1111/j.1365-2958.1990.tb02012.x.

D'hooghe, I., Michiels, J., Vlassak, K., Verreth, C., Waelkens, F., & Vanderleyden, J. (1995). Structural and functional analysis of the fixLJ genes of *Rhizobium leguminosarum* biovar phaseoli CNPAF512. *MGG Molecular & General Genetics, 249*(1), 117–126. https://doi.org/10.1007/BF00290243.

Earl, C. D., Ronson, C. W., & Ausubel, F. M. (1987). Genetic and structural analysis of the *Rhizobium meliloti* fixA, fixB, fixC, and fixX genes. *Journal of Bacteriology, 169*(3), 1127–1136. https://doi.org/10.1128/jb.169.3.1127-1136.1987.

Ebert, M., Laaß, S., Thürmer, A., Roselius, L., Eckweiler, D., Daniel, R., et al. (2017). FnrL and Three Dnr regulators are used for the metabolic adaptation to low oxygen tension in Dinoroseobacter shibae. *Frontiers in Microbiology, 8*(APR), 642. https://doi.org/10.3389/fmicb.2017.00642.

Edgren, T., & Nordlund, S. (2004). The fixABCX genes in *Rhodospirillum rubrum* encode a putative membrane complex participating in electron transfer to nitrogenase. *Journal of Bacteriology, 186*(7), 2052–2060. https://doi.org/10.1128/JB.186.7.2052-2060.2004.

Emmerich, R., Hennecke, H., & Fischer, H. M. (2000). Evidence for a functional similarity between the two-component regulatory systems RegSR, ActSR, and RegBA (PrrBA) in α-proteobacteria. *Archives of Microbiology, 174*(5), 307–313. https://doi.org/10.1007/s002030000207.

Emmerich, R., Panglungtshang, K., Strehler, P., Hennecke, H., & Fischer, H. M. (1999). Phosphorylation, dephosphorylation and DNA-binding of the *Bradyrhizobium japonicum* RegSR two-component regulatory proteins. *European Journal of Biochemistry, 263*(2), 455–463. https://doi.org/10.1046/j.1432-1327.1999.00517.x.

Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z., & Winiwarter, W. (2008). How a century of ammonia synthesis changed the world. *Nature Geoscience, 1*(10), 636–639. https://doi.org/10.1038/ngeo325.

Fernández, N., Cabrera, J. J., Varadarajan, A. R., Lutz, S., Ledermann, R., Roschitzki, B., et al. (2019). An integrated systems approach unveils new aspects of microoxia-mediated regulation in Bradyrhizobium diazoefficiens. *Frontiers in Microbiology, 10*, 924. https://doi.org/10.3389/fmicb.2019.00924.

Ferrières, L., Francez-Charlot, A., Gouzy, J., Rouillé, S., & Kahn, D. (2004). FixJ-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. *Microbiology, 150*(7), 2335–2345. https://doi.org/10.1099/mic.0.27081-0.

Ferrières, L., & Kahn, D. (2002). Two distinct classes of FixJ binding sites defined by in vitro selection. *FEBS Letters, 517*(1−3), 185−189. https://doi.org/10.1016/S0014-5793(02)02618-2.

Fischer, H. M. (1994). Genetic regulation of nitrogen fixation in rhizobia. *Microbiological Reviews, 58*(3), 352−386. https://doi.org/10.1186/gb-2011-12-10-r106.

Fischer, H. M. (1996). Environmental regulation of rhizobial symbiotic nitrogen fixation genes. *Trends in Microbiology, 4*(8), 317−320. https://doi.org/10.1016/0966-842X(96)10049-4.

Fischer, H. M., Bruderer, T., & Hennecke, H. (1988). Essential and non-essential domains in the *Bradyrhizobium japonicum* nifa protein: Identification of indispensable cysteine residues potentially involved in redox reactivity and/or metal binding. *Nucleic Acids Research, 16*(5), 2207−2224. https://doi.org/10.1093/nar/16.5.2207.

Fischer, H.-M., Fritsche, S., Herzog, B., & Hennecke, H. (1989). Critical spacing between two essential cysteine residues in the interdomain linker of the *Bradyrhizobium japonicum* NifA protein. *FEBS Letters, 255*(1), 167−171. https://doi.org/10.1016/0014-5793(89)81083-X.

Fischer, H. M., & Hennecke, H. (1987). Direct response of *Bradyrhizobium japonicum* nifA-mediated nif gene regulation to cellular oxygen status. *MGG Molecular & General Genetics, 209*(3), 621−626. https://doi.org/10.1007/BF00331174.

Foussard, M., Soupène, E., Garnerone, A. M., Capela, D., Cabanes, D., Boistard, P., et al. (1998). Regulation of nitrogen fixation gene expression in rhizobia: An overview. In *Biological nitrogen fixation for the 21st century* (pp. 101−106). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-011-5159-7_33.

Gage, D. J. (2002). Analysis of infection thread development using Gfp- and DsRed-expressing *Sinorhizobium meliloti*. *Journal of Bacteriology, 184*(24), 7042−7046. https://doi.org/10.1128/JB.184.24.7042-7046.2002.

Gage, D. J., & Margolin, W. (2000). Hanging by a thread: Invasion of legume plants by rhizobia. *Current Opinion in Microbiology, 3*(6), 613−617. https://doi.org/10.1016/S1369-5274(00)00149-1.

Galinier, A., Garnerone, A. M., Reyrat, J. M., Kahn, D., Batut, J., & Boistard, P. (1994). Phosphorylation of the *Rhizobium meliloti* FixJ protein induces its binding to a compound regulatory region at the fixK promoter. *Journal of Biological Chemistry, 269*(38), 23784−23789. Retrieved from http://www.jbc.org/content/269/38/23784.full.pdf.

Gamper, M., Zimmermann, A., & Haas, D. (1991). Anaerobic regulation of transcription initiation in the arcDABC operon of *Pseudomonas aeruginosa*. *Journal of Bacteriology, 173*(15), 4742−4750. https://doi.org/10.1128/jb.173.15.4742-4750.1991.

Gao, R., & Stock, A. M. (2009). Biological insights from structures of two-component proteins. *Annual Review of Microbiology, 63*(1), 133−154. https://doi.org/10.1146/annurev.micro.091208.073214.

Garnerone, A. M., Cabanes, D., Foussard, M., Boistard, P., & Batut, J. (1999). Inhibition of the FixL sensor kinase by the FixT protein in *Sinorhizobium meliloti*. *Journal of Biological Chemistry, 274*(45), 32500−32506. https://doi.org/10.1074/jbc.274.45.32500.

Garnerone, A.-M., Foussard, M., Boistard, P., & Batut, J. (1999). Mode of action of the FixT repressor protein of *Sinorhizobium meliloti*. In *Highlights of nitrogen fixation research* (pp. 195−199). New York: Kluwer Academic/Plenum Publishers. https://doi.org/10.1007/978-1-4615-4795-2_39.

Gebhardt, C., Turner, G. L., Gibson, A. H., Dreyfus, B. L., & Bergersen, F. J. (1984). Nitrogen-fixing growth in continuous culture of a strain of rhizobium sp. isolated from stem nodules on *Sesbania rostrata*. *Journal of General Microbiology, 130*(4), 843−848. https://doi.org/10.1099/00221287-130-4-843.

Geddes, B. A., Ryu, M.-H. H., Mus, F., Garcia Costas, A., Peters, J. W., Voigt, C. A., et al. (2015). Use of plant colonizing bacteria as chassis for transfer of N_2 fixation to cereals.

Current Opinion in Biotechnology, 32, 216–222. https://doi.org/10.1016/j.copbio.2015.01.004.

Georgiadis, M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J., & Rees, D. (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science, 257*(5077), 1653–1659. https://doi.org/10.1126/science.1529353.

Girard, L., Brom, S., Dávalos, A., López, O., Soberón, M., & Romero, D. (2000). Differential regulation of fixN -reiterated genes in rhizobium etli by a novel fixL—fixK cascade. *Molecular Plant-Microbe Interactions, 13*(12), 1283–1292. https://doi.org/10.1094/MPMI.2000.13.12.1283.

Gómez-Hernández, N., Reyes-González, A., Sánchez, C., Mora, Y., Delgado, M. J., & Girard, L. (2011). Regulation and symbiotic role of nirK and norC expression in rhizobium etli. *Molecular Plant-Microbe Interactions, 24*(2), 233–245. https://doi.org/10.1094/MPMI-07-10-0173.

Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., & Chan, M. K. (1998). Structure of a biological oxygen sensor: A new mechanism for heme-driven signal transduction. *Proceedings of the National Academy of Sciences, 95*(26), 15177–15182. https://doi.org/10.1073/pnas.95.26.15177.

Gouet, P., Fabry, B., Guillet, V., Birck, C., Mourey, L., Kahn, D., et al. (1999). Structural transitions in the FixJ receiver domain. *Structure, 7*(12), 1517–1526. https://doi.org/10.1016/S0969-2126(00)88342-2.

Gourret, J.-P., & Fernandez-Arias, H. (1974). Etude ultrastructurale et cytochimique de la différenciation des bactéroïdes de Rhizobium trifolii Dangeard dans les nodules de Trifolium repens L. *Canadian Journal of Microbiology, 20*(8), 1169–1181. https://doi.org/10.1139/m74-181.

Granados-Baeza, M. J., Gómez-Hernández, N., Mora, Y., Delgado, M. J., Romero, D., & Girard, L. (2007). Novel reiterated Fnr-type proteins control the production of the symbiotic terminal oxidase cbb3 in Rhizobium etli CFN42. *Molecular Plant-Microbe Interactions: Molecular Plant-Microbe Interactions, 20*(10), 1241–1249. https://doi.org/10.1094/MPMI-20-10-1241.

Green, J., Crack, J. C., Thomson, A. J., & LeBrun, N. E. (2009). Bacterial sensors of oxygen. *Current Opinion in Microbiology, 12*(2), 145–151. https://doi.org/10.1016/j.mib.2009.01.008.

Green, J., Scott, C., & Guest, J. R. (2001). Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. In *Advances in Microbial Physiology* (Vol. 44, pp. 1–34). Academic Press. https://doi.org/10.1016/S0065-2911(01)44010-0.

Gruber, N., & Galloway, J. N. (2008). An Earth-system perspective of the global nitrogen cycle. *Nature, 451*(7176), 293–296. https://doi.org/10.1038/nature06592.

Gubler, M. (1989). Fine-tuning of nif and fix gene expression by upstream activator sequences in *Bradyrhizobium japonicum*. *Molecular Microbiology, 3*(2), 149–159. https://doi.org/10.1111/j.1365-2958.1989.tb01804.x.

Guest, J. R., Green, J., Irvine, A. S., & Spiro, S. (1996). The FNR modulon and FNR-regulated gene expression. In *Regulation of gene expression in* Escherichia coli (pp. 317–342). Boston: Springer US. https://doi.org/10.1007/978-1-4684-8601-8_16.

Guimarães, W. G., Gondim, A. C. S., da S. Costa, P. M., Gilles-Gonzalez, M. A., Lopes, L. G. F., Carepo, M. S. P., et al. (2017). Insights into signal transduction by a hybrid FixL: Denaturation study of on and off states of a multi-domain oxygen sensor. *Journal of Inorganic Biochemistry, 172*, 129–137. https://doi.org/10.1016/j.jinorgbio.2017.04.013.

Gussin, G. N., Ronson, C. W., & Ausubel, F. M. (1986). Regulation of nitrogen fixation genes. *Annual Review of Genetics, 20*(1), 567–591. https://doi.org/10.1146/annurev.ge.20.120186.003031.

Gutiérrez, D., Hernando, Y., Palacios, J. M., Imperial, J., & Ruiz-Argüeso, T. (1997). FnrN controls symbiotic nitrogen fixation and hydrogenase activities in *Rhizobium leguminosarum* biovar viciae UPM791. *Journal of Bacteriology, 179*(17), 5264–5270. https://doi.org/10.1128/jb.179.17.5264-5270.1997.

Gutiérrez, D., Hernando, Y., Palacios, J. M., Imperial, J., & Ruiz-Argüeso, T. (1998). Symbiotic expression of hydrogenase and nitrogenase activities of *Rhizobium leguminosarum* bv. Viciae are controlled by FnrN. In *Biological nitrogen fixation for the 21st century* (p. 286). Dordrecht: Springer. https://doi.org/10.1007/978-94-011-5159-7_155.

Hao, B., Isaza, C., Arndt, J., Soltis, M., & Chan, M. K. (2002). Structure-based mechanism of O2 sensing and ligand discrimination by the FixL heme domain of *Bradyrhizobium japonicum*. *Biochemistry, 41*(43), 12952–12958. https://doi.org/10.1021/bi020144l.

Hauser, F., Pessi, G., Friberg, M., Weber, C., Rusca, N., Lindemann, A., et al. (2007). Dissection of the *Bradyrhizobium japonicum* NifA+σ54 regulon, and identification of a ferredoxin gene (fdxN) for symbiotic nitrogen fixation. *Molecular Genetics and Genomics, 278*(3), 255–271. https://doi.org/10.1007/s00438-007-0246-9.

Heermann, R., & Jung, K. (2010). Stimulus perception and signaling in histidine kinases. In *Bacterial signaling* (pp. 135–161). Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA. https://doi.org/10.1002/9783527629237.ch8.

Hernando, Y., Palacios, J. M., Imperial, J., Ruiz-Argueso, T., & Ruiz-Argüeso, T. (1995). The hypBFCDE operon from *Rhizobium leguminosarum* biovar viciae is expressed from an Fnr-type promoter that escapes mutagenesis of the fnrN gene. *Journal of Bacteriology, 177*(19), 5661–5669. https://doi.org/10.1128/jb.177.19.5661-5669.1995.

Hertig, C., Li, R. Y., Louarn, A. M., Garnerone, A. M., David, M., Batut, J., et al. (1989). *Rhizobium meliloti* regulatory gene fixJ activates transcription of *R. meliloti* nifA and fixK genes in *Escherichia coli*. *Journal of Bacteriology, 171*(3), 1736–1738. https://doi.org/10.1128/jb.171.3.1736-1738.1989.

Hill, S., Austin, S., Eydmann, T., Jones, T., & Dixon, R. (1996). *Azotobacter vinelandii* NIFL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch. *Proceedings of the National Academy of Sciences, 93*(5), 2143–2148. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC39924/pdf/pnas01509-0418.pdf.

Hirsch, A. M., & Smith, C. A. (1987). Effects of *Rhizobium meliloti* nif and fix mutants on alfalfa root nodule development. *Journal of Bacteriology, 169*(3), 1137–1146. https://doi.org/10.1128/jb.169.3.1137-1146.1987.

Ho, Y.-S. S., Burden, L. M., & Hurley, J. H. (2000). Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *The EMBO Journal, 19*(20), 5288–5299. https://doi.org/10.1093/emboj/19.20.5288.

Huala, E., & Ausubel, F. M. (1989). The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti* nifH promoter. *Journal of Bacteriology, 171*(6), 3354–3365. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2722751.

Hu, Y., & Ribbe, M. W. (2013). Nitrogenase assembly. *Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1827*(8–9), 1112–1122. https://doi.org/10.1016/j.bbabio.2012.12.001.

Iismaa, S. E., & Watson, J. M. (1989). The nifA gene product from *Rhizobium leguminosarum* biovar trifolii lacks the N-terminal domain found in other NifA proteins. *Molecular Microbiology, 3*(7), 943–955. https://doi.org/10.1111/j.1365-2958.1989.tb00244.x.

Imlay, J. A. (2006). Iron-sulphur clusters and the problem with oxygen. *Molecular Microbiology, 59*(4), 1073–1082. https://doi.org/10.1111/j.1365-2958.2006.05028.x.

Jayaraman, P. S., Cole, J. A., & Busby, S. J. W. (1989). Mutational analysis of the nucleotide sequence at the FNR-dependent nirB promoter in *Escherichia coli*. *Nucleic Acids Research, 17*(1), 135–145. https://doi.org/10.1093/nar/17.1.135.

Jayaraman, P.-S., Gaston, K. L., Cole, J. A., & Busby, S. J. W. W. (1988). The nirB promoter of *Escherichia coli*: Location of nucleotide sequences essential for regulation by oxygen, the

FNR protein and nitrite. *Molecular Microbiology, 2*(4), 527–530. https://doi.org/10.1111/j.1365-2958.1988.tb00059.x.

Jervis, A. J., & Green, J. (2007). In vivo demonstration of FNR dimers in response to lower O2 availability. *Journal of Bacteriology, 189*(7), 2930–2932. https://doi.org/10.1128/JB.01921-06.

Jimenez-Zurdo, J. I. (1995). Characterization of a *Rhizobium meliloti* proline dehydrogenase mutant altered in nodulation efficiency and competitiveness on alfalfa roots. *Molecular Plant-Microbe Interactions, 8*(4), 492. https://doi.org/10.1094/MPMI-8-0492.

Jiménez-Zurdo, J. I., García-Rodríguez, F. M., & Toro, N. (1997). The *Rhizobium meliloti* putA gene: Its role in the establishment of the symbiotic interaction with alfalfa. *Molecular Microbiology, 23*(1), 85–93. https://doi.org/10.1046/j.1365-2958.1997.1861555.x.

Kahn, D., David, M., Domergue, O., Daveran, M. L., Ghai, J., Hirsch, P. R., et al. (1989). *Rhizobium meliloti* fixGHI sequence predicts involvement of a specific cation pump in symbiotic nitrogen fixation. *Journal of Bacteriology, 171*(2), 929–939. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC209684/pdf/jbacter00168-0319.pdf.

Kaminski, P. A., & Elmerich, C. (1991). Involvement of fixLJ in the regulation of nitrogen fixation in *Azorhizobium caulinodans*. *Molecular Microbiology, 5*(3), 665–673. https://doi.org/10.1111/j.1365-2958.1991.tb00738.x.

Kaminski, P. A., & Elmerich, C. (1998). The control of *Azorhizobium caulinodans* nifA expression by oxygen, ammonia and by the HF-I-like protein, NrfA. *Molecular Microbiology, 28*(3), 603–613. https://doi.org/10.1046/j.1365-2958.1998.00823.x.

Kaminski, P. A., Mandon, K., Arigoni, F., Desnoues, N., & Elmerich, C. (1991). Regulation of nitrogen fixation in *Azorhizobium caulinodans*: Identification of a fixK-like gene, a positive regulator of nifA. *Molecular Microbiology, 5*(8), 1983–1991. https://doi.org/10.1111/j.1365-2958.1991.tb00820.x.

Kannenberg, E. L., & Brewin, N. J. (1989). Expression of a cell surface antigen from *Rhizobium leguminosarum* 3841 is regulated by oxygen and pH. *Journal of Bacteriology, 171*(9), 4543–4548. https://doi.org/10.1128/jb.171.9.4543-4548.1989.

Kawashima, K., Suganuma, N., Tamaoki, M., & Kouchi, H. (2001). Two types of pea leghemoglobin genes showing different O2-binding affinities and distinct patterns of spatial expression in nodules. *Plant Physiology, 125*(2), 641–651. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11161022.

Kelemen, J. Z., Nagy, Z. B., Bodogai, M., Becker, A., Rüberg, S., Puskás, L. G., et al. (2004). Wide-range transcriptional modulating effect of ntrR under microaerobiosis in *Sinorhizobium meliloti*. *Molecular Genetics and Genomics, 272*(3), 275–289. https://doi.org/10.1007/s00438-004-1051-3.

Khoroshilova, N., Popescu, C., Munck, E., Beinert, H., & Kiley, P. J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proceedings of the National Academy of Sciences, 94*(12), 6087–6092. https://doi.org/10.1073/pnas.94.12.6087.

Kiers, E. T., Rousseau, R. A., West, S. A., & Denlson, R. F. (2003). Host sanctions and the legume-rhizobium mutualism. *Nature, 425*(6953), 78–81. https://doi.org/10.1038/nature01931.

Kim, C. H., Helinski, D. R., & Ditta, G. (1986). Overlapping transcription of the nifA regulatory gene in *Rhizobium meliloti*. *Gene, 50*(1–3), 141–148. https://doi.org/10.1016/0378-1119(86)90319-7.

King, B. J., Hunt, S., Weagle, G. E., Walsh, K. B., Pottier, R. H., Canvin, D. T., et al. (1988). Regulation of O2 concentration in soybean nodules observed by in situ spectroscopic measurement of leghemoglobin oxygenation. *Plant Physiology, 87*(2), 296–299. https://doi.org/10.1104/Pp.87.2.296.

King, N. D., Hojnacki, D., & O'Brian, M. R. (2000). The *Bradyrhizobium japonicum* proline biosynthesis gene proC is essential for symbiosis. *Applied and Environmental Microbiology, 66*(12), 5469–5471. https://doi.org/10.1128/AEM.66.12.5469-5471.2000.

Kitts, C. L., & Ludwig, R. A. (1994). *Azorhizobium caulinodans* respires with at least four terminal oxidases. *Journal of Bacteriology, 176*(3), 886–895. https://doi.org/10.1128/jb.176.3.886-895.1994.

Klipp, W., Reiländer, H., Schlüter, A., Krey, R., & Pühler, A. (1989). The *Rhizobium meliloti* fdxN gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with nifA and nifB. *MGG Molecular & General Genetics, 216*(2–3), 293–302. https://doi.org/10.1007/BF00334368.

Koch, H. G., Winterstein, C., Saribas, A. S., Alben, J. O., & Daldal, F. (2000). Roles of the ccoGHIS gene products in the biogenesis of the cbb3-type cytochrome c oxidase. *Journal of Molecular Biology, 297*(1), 49–65. https://doi.org/10.1006/jmbi.2000.3555.

Kopat, V. V., Chirak, E. R., Kimeklis, A. K., Safronova, V. I., Belimov, A. A., Kabilov, M. R., et al. (2017). Evolution of fixNOQP genes encoding cytochrome oxidase with high affinity to oxygen in rhizobia and related bacteria. *Russian Journal of Genetics, 53*(7), 1022–7954. https://doi.org/10.1134/S1022795417070067.

Körner, H., Sofia, H. J., & Zumft, W. G. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiology Reviews, 27*(5), 559–592. https://doi.org/10.1016/S0168-6445(03)00066-4.

Krey, R., Pühler, A., & Klipp, W. (1992). A defined amino acid exchange close to the putative nucleotide binding site is responsible for an oxygen-tolerant variant of the *Rhizobium meliloti* NifA protein. *MGG Molecular & General Genetics, 234*(3), 433–441. https://doi.org/10.1007/BF00538703.

Kullik, I., Fritsche, S., Knobel, H., Sanjuan, J., Hennecke, H., & Fischer, H. M. (1991). *Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the σ54 gene (rpoN). *Journal of Bacteriology, 173*(3), 1125–1138. Retrieved from http://jb.asm.org/.

Kundu, S., Trent, J. T., & Hargrove, M. S. (2003). Plants, humans and hemoglobins. *Trends in Plant Science, 8*(8), 387–393. https://doi.org/10.1016/S1360-1385(03)00163-8.

Kustu, S., Santero, E., Keener, J., Popham, D., & Weiss, D. (1989). Expression of sigma 54 (ntrA)-dependent genes is probably united by a common mechanism. *Microbiological Reviews, 53*(3), 367–376. Retrieved from http://mmbr.asm.org/.

Kuzma, M. M., Hunt, S., & Layzell, D. B. (1993). Role of oxygen in the limitation and inhibition of nitrogenase activity and respiration rate in individual soybean nodules. *Plant Physiology, 101*(1), 161–169. pii:101/1/161.

Layzell, D. B., Diaz del Castillo, L., Hunt, S., Kuzma, M., Van Cauwenberghe, O., & Oresnik, I. (1993). *The regulation of oxygen and its role in regulating nodule metabolism* (pp. 393–398). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-2416-6_39.

Layzell, D. B., & Hunt, S. (1990). Oxygen and the regulation of nitrogen fixation in legume nodules. *Physiologia Plantarum, 80*(2), 322–327. https://doi.org/10.1111/j.1399-3054.1990.tb04414.x.

Lazazzera, B. A., Bates, D. M., & Kiley, P. J. (1993). The activity of the *Escherichia coli* transcription factor FNR is regulated by a change in oligomeric state. *Genes & Development, 7*(10), 1993–2005. https://doi.org/10.1101/gad.7.10.1993.

Ledbetter, R. N., Garcia Costas, A. M., Lubner, C. E., Mulder, D. W., Tokmina-Lukaszewska, M., Artz, J. H., et al. (2017). The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii*: Generation of low-potential reducing equivalents for nitrogenase catalysis. *Biochemistry, 56*(32), 4177–4190. https://doi.org/10.1021/acs.biochem.7b00389.

Lee, K.-B., De Backer, P., Aono, T., Liu, C.-T., Suzuki, S., Suzuki, T., et al. (2008). The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. *BMC Genomics, 9*(1), 271. https://doi.org/10.1186/1471-2164-9-271.

Li, H., Xu, F., Ren, X., & Chen, S. (2010). Functional analysis of the fixL/fixJ and fixK genes in Azospirillum brasilense Sp7. *Annals of Microbiology, 60*(3), 469–480. https://doi.org/10.1007/s13213-010-0065-9.

Lindemann, A., Balsiger, S., Müller, A. J., Pessi, G., Fischer, H.-M., Koch, M., et al. (2010). Host-specific symbiotic requirement of BdeAB, a RegR-controlled RND-type efflux system in *Bradyrhizobium japonicum*. *FEMS Microbiology Letters, 312*(2), 184–191. https://doi.org/10.1111/j.1574-6968.2010.02115.x.

Lindemann, A., Moser, A., Pessi, G., Hauser, F., Friberg, M., Hennecke, H., et al. (2007). New target genes controlled by the *Bradyrhizobium japonicum* two-component regulatory system RegSR. *Journal of Bacteriology, 189*(24), 8928–8943. https://doi.org/10.1128/JB.01088-07.

Little, R., & Dixon, R. (2003). The amino-terminal GAF domain of *Azotobacter vinelandii* NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions. *Journal of Biological Chemistry, 278*(31), 28711–28718. https://doi.org/10.1074/jbc.M301992200.

Liu, J., Lee, F., Lin, C., Yao, X., Davenport, J. W., & Wong, T. (1995). Alternative function of the electron transport system in *Azotobacter vinelandii*: Removal of excess reductant by the cytochrome d pathway. *Applied and Environmental Microbiology, 61*(11), 3998–4003. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16535163.

Lois, A. F., Weinstein, M., Ditta, G. S., & Helinski, D. R. (1993). Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of *Rhizobium meliloti* are coordinately regulated by oxygen. *Journal of Biological Chemistry, 268*(6), 4370–4375. Retrieved from http://www.jbc.org/content/268/6/4370.short.

López-Torrejón, G., Jiménez-Vicente, E., Buesa, J. M., Hernandez, J. A., Verma, H. K., & Rubio, L. M. (2016). Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nature Communications, 7*, 11426. https://doi.org/10.1038/ncomms11426.

Lopez, O., Morera, C., Miranda-Rios, J., Girard, L., Romero, D., & Soberon, M. (2001). Regulation of gene expression in response to oxygen in *Rhizobium etli*: Role of FnrN in fixNOQP expression and in symbiotic nitrogen fixation. *Journal of Bacteriology, 183*(24), 6999–7006. https://doi.org/10.1128/JB.183.24.6999-7006.2001.

Loroch, A. I., Nguyen, B. G., & Ludwig, R. A. (1995). Interactive regulation of *Azorhizobium* nifA transcription via overlapping promoters. *Journal of Bacteriology, 177*(24), 7210–7221. https://doi.org/10.1128/jb.177.24.7210-7221.1995.

Mandon, K., Kaminski, P. A., & Elmerich, C. (1994). Functional analysis of the fixNOQP region of *Azorhizobium caulinodans*. *Journal of Bacteriology, 176*(9), 2560–2568. https://doi.org/10.1128/jb.176.9.2560-2568.1994.

Mandon, K., Kaminski, P. A., Mougel, C., Desnoues, N., Dreyfus, B., & Elmerich, C. (1993). Role of the fixGHI region of *Azorhizobium caulinodans* in free-living and symbiotic nitrogen fixation. *FEMS Microbiology Letters, 114*(2), 185–189. pii:0378-1097(93)90517-6.

Marchal, K., & Vanderleyden, J. (2000). The "oxygen paradox" of dinitrogen-fixing bacteria. *Biology and Fertility of Soils, 30*(5–6), 363–373. https://doi.org/10.1007/s003740050017.

Martin, F. M., Uroz, S., & Barker, D. G. (2017). Ancestral alliances: Plant mutualistic symbioses with fungi and bacteria. *Science, 356*(6340), 1–9. https://doi.org/10.1126/science.aad4501.

Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., & Dixon, R. (2004). The NifL-NIFA system: A multidomain transcriptional regulatory complex that integrates

environmental signals. *Journal of Bacteriology, 186*(3), 601–610. https://doi.org/10.1128/JB.186.3.601-610.2004.

Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., & Dixon, R. (2005). Nitrogen fixation: Key genetic regulatory mechanisms. *Biochemical Society Transactions, 33*(1), 152–156. https://doi.org/10.1042/BST0330152.

Martínez-Romero, E. (2009). Coevolution in rhizobium-legume symbiosis? *DNA and Cell Biology, 28*(8), 361–370. https://doi.org/10.1089/dna.2009.0863.

Martínez, M., Colombo, M.-V. V., Palacios, J.-M. M., Imperial, J., & Ruiz-Argüeso, T. (2008). Novel arrangement of enhancer sequences for NifA-dependent activation of the hydrogenase gene promoter in *Rhizobium leguminosarum* bv. viciae. *Journal of Bacteriology, 190*(9), 3185–3191. https://doi.org/10.1128/JB.00107-08.

Martínez, M., Palacios, J. M., Imperial, J., & Ruiz-Argüeso, T. (2004). Symbiotic autoregulation of nifA expression in *Rhizobium leguminosarum* bv. viciae. *Journal of Bacteriology, 186*(19), 6586–6594. https://doi.org/10.1128/JB.186.19.6586-6594.2004.

Masson-Boivin, C., Giraud, E., Perret, X., & Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: How many rhizobium recipes? *Trends in Microbiology, 17*(10), 458–466. https://doi.org/10.1016/j.tim.2009.07.004.

Masson-Boivin, C., & Sachs, J. L. (2018). Symbiotic nitrogen fixation by rhizobia — the roots of a success story. *Current Opinion in Plant Biology, 44*, 7–15. https://doi.org/10.1016/j.pbi.2017.12.001.

McGinnis, S. D., & O'Brian, M. R. (1995). The rhizobial hemA gene is required for symbiosis in species with deficient delta-aminolevulinic acid uptake activity. *Plant Physiology, 108*(4), 1547–1552. https://doi.org/10.1104/pp.108.4.1547.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis. *Proceedings of the National Academy of Sciences, 103*(13), 5230–5235. https://doi.org/10.1073/pnas.0600912103.

Mesa, S., Bedmar, E. J., Chanfon, A., Hennecke, H., & Fischer, H. M. (2003). *Bradyrhizobium japonicum* NnrR, a denitrification regulator, expands the FixLJ-FixK2 regulatory cascade. *Journal of Bacteriology, 185*(13), 3978–3982. https://doi.org/10.1128/JB.185.13.3978-3982.2003.

Mesa, S., Hauser, F., Friberg, M., Malaguti, E., Fischer, H.-M., & Hennecke, H. (2008). Comprehensive assessment of the regulons controlled by the FixLJ-FixK2-FixK1 cascade in *Bradyrhizobium japonicum*. *Journal of Bacteriology, 190*(20), 6568–6579. https://doi.org/10.1128/JB.00748-08.

Mesa, S., Hennecke, H., & Fischer, H.-M. (2006). A multitude of CRP/FNR-like transcription proteins in *Bradyrhizobium japonicum*. *Biochemical Society Transactions, 34*(1), 156–159. https://doi.org/10.1042/bst0340156.

Mesa, S., Reutimann, L., Fischer, H.-M., & Hennecke, H. (2009). Posttranslational control of transcription factor FixK2, a key regulator for the *Bradyrhizobium japonicum*-soybean symbiosis. *Proceedings of the National Academy of Sciences, 106*(51), 21860–21865. https://doi.org/10.1073/pnas.0908097106.

Mesa, S., Ucurum, Z., Hennecke, H., & Fischer, H. M. (2005). Transcription activation in vitro by the *Bradyrhizobium japonicum* regulatory protein FixK2. *Journal of Bacteriology, 187*(10), 3329–3338. https://doi.org/10.1128/JB.187.10.3329-3338.2005.

Michel-Reydellet, N., & Kaminski, P. A. (1999). *Azorhizobium caulinodans* P(II) and GlnK proteins control nitrogen fixation and ammonia assimilation. *Journal of Bacteriology, 181*(8), 2655–2658. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10198037.

Michiels, J., D'hooghe, I., Verreth, C., Pelemans, H., & Vanderleyden, J. (1994). Characterization of the *Rhizobium leguminosarum* biovar phaseoli nifA gene, a positive regulator of

nif gene expression. *Archives of Microbiology, 161*(5), 404–408. https://doi.org/10.1007/BF00288950.

Michiels, J., Van Soom, T., D'hooghe, I., Dombrecht, B., Benhassine, T., de Wilde, P., et al. (1998). The rhizobium etli rpoN locus: DNA sequence analysis and phenotypical characterization of rpoN, ptsN, and ptsA mutants. *Journal of Bacteriology, 180*(7), 1729–1740. Retrieved from http://jb.asm.org/.

Millar, A. H., Day, D. A., & Bergersen, F. J. (1995). Microaerobic respiration and oxidative phosphorylation by soybean nodule mitochondria: Implications for nitrogen fixation. *Plant, Cell and Environment, 18*(7), 715–726. https://doi.org/10.1111/j.1365-3040.1995.tb00574.x.

Minchin, F. R. (1997). Regulation of oxygen diffusion in legume nodules. *Soil Biology and Biochemistry, 29*(516), 88–89. https://doi.org/10.1016/S0038-0717(96)00204-0.

Minchin, F. R., James, E. K., & Becana, M. (2008). Oxygen diffusion, production of reactive oxygen and nitrogen species, and antioxidants in legume nodules. In *Nitrogen-fixing leguminous symbioses* (pp. 321–362). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-3548-7_11.

Minchin, F. R., Sheehy, J. E., & Witty, J. F. (1985). Factors limiting N2 fixation by the legume-rhizobium symbiosis. In *Nitrogen fixation research progress* (pp. 285–291). Dordrecht: Martinus Nijhoff Publishers. https://doi.org/10.1007/978-94-009-5175-4_40.

Montiel, J., Downie, J. A., Farkas, A., Bihari, P., Herczeg, R., Bálint, B., et al. (2017). Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides. *Proceedings of the National Academy of Sciences, 114*(19), 5041–5046. https://doi.org/10.1073/pnas.1704217114.

Moore, L. J., & Kiley, P. J. (2001). Characterization of the dimerization domain in the FNR transcription factor. *Journal of Biological Chemistry, 276*(49), 45744–45750. https://doi.org/10.1074/jbc.M106569200.

Moore, L. J., Mettert, E. L., & Kiley, P. J. (2006). Regulation of FNR dimerization by subunit charge repulsion. *Journal of Biological Chemistry, 281*(44), 33268–33275. https://doi.org/10.1074/jbc.M608331200.

Moreno, S., Patriarca, E. J., Chiurazzi, M., Meza, R., Defez, R., Lamberti, A., et al. (1992). Phenotype of a *Rhizobium leguminosarum* ntrC mutant. *Research in Microbiology, 143*(2), 161–171. https://doi.org/10.1016/0923-2508(92)90005-9.

Morett, E., & Buck, M. (1988). NifA-dependent in vivo protection demonstrates that the upstream activator sequence of nif promoters is a protein binding site. *Proceedings of the National Academy of Sciences, 85*(24), 9401–9405. https://doi.org/10.1073/pnas.85.24.9401.

Morett, E., Fischer, H. M., & Hennecke, H. (1991). Influence of oxygen on DNA binding, positive control, and stability of the *Bradyrhizobium japonicum* NifA regulatory protein. *Journal of Bacteriology, 173*(11), 3478–3487. https://doi.org/10.1128/JB.173.11.3478-3487.1991.

Moris, M., Dombrecht, B., Xi, C., Vanderleyden, J., & Michiels, J. (2004). Regulatory role of rhizobium etli CNPAF512 fnrN during symbiosis. *Applied and Environmental Microbiology, 70*(3), 1287–1296. https://doi.org/10.1128/AEM.70.3.1287-1296.2004.

Murray, J. D. (2011). Invasion by invitation: *Rhizobial* infection in legumes. *Molecular Plant-Microbe Interactions, 24*(6), 631–639. https://doi.org/10.1094/mpmi-08-10-0181.

Mus, F., Crook, M. B., Garcia, K., Costas, A. G., Geddes, B. A., Kouri, E. D., et al. (2016). Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Applied and Environmental Microbiology, 82*(13), 3698–3710. https://doi.org/10.1128/AEM.01055-16.

Mylona, P., Pawlowski, K., & Bisseling, T. (1995). Symbiotic nitrogen fixation. *The Plant Cell, 7*(7), 869–885. https://doi.org/10.1105/tpc.7.7.869.

Nees, D. W., Stein, P. A., & Ludwig, R. A. (1988). The *Azorhizobium caulinodans* nifA gene: Identification of upstream-activating sequences including a new element, the "anaerobox. *Nucleic Acids Research, 16*(20), 9839–9853. https://doi.org/10.1093/nar/16.20.9839.

Nellen-Anthamatten, D., & Rossi, P. (1998). *Bradyrhizobium japonicum* FixK2, a crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels. *Journal of Bacteriology, 180*(19), 5251–5255. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9748464.

Neo, H. H., & Layzell, D. B. (1997). Phloem glutamine and the regulation of O2 diffusion in legume nodules. *Plant Physiology, 113*(1), 259–267. https://doi.org/10.1104/pp.113.1.259.

Ninfa, A. J., & Jiang, P. (2005). PII signal transduction proteins: Sensors of α-ketoglutarate that regulate nitrogen metabolism. *Current Opinion in Microbiology, 8*(2), 168–173. https://doi.org/10.1016/j.mib.2005.02.011.

Ninfa, A. J., Reitzer, L. J., & Magasanik, B. (1987). Initiation of transcription at the bacterial glnAp2 promoter by purified *E. coli* components is facilitated by enhancers. *Cell, 50*(7), 1039–1046. https://doi.org/10.1016/0092-8674(87)90170-X.

Oldroyd, G. E. D., & Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. *Current Opinion in Biotechnology, 26*, 19–24. https://doi.org/10.1016/j.copbio.2013.08.006.

Oldroyd, G. E. D., & Downie, J. A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annual Review of Plant Biology, 59*(1), 519–546. https://doi.org/10.1146/annurev.arplant.59.032607.092839.

Oldroyd, G. E. D., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annual Review of Genetics, 45*(1), 119–144. https://doi.org/10.1146/annurev-genet-110410-132549.

Oono, R., Denison, R. F., & Kiers, E. T. (2009). Controlling the reproductive fate of rhizobia: How universal are legume sanctions? *New Phytologist, 183*(4), 967–979. https://doi.org/10.1111/j.1469-8137.2009.02941.x.

Oresnik, I. J., Atkins, C. A., & Layzell, D. B. (1995). The legume symbiosis: C-limited bacteria living within O2 limited plant cells?. In I. A. Tikhonovich, N. A. Provorov, V. I. Romanov, & W. E. Newton (Eds.), *Nitrogen fixation: Fundamentals and applications. Proceed. 10th Intl. Cong. Nitrogen fix., St. Petersburg* (Vol. 27, p. 601) Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-0379-4.

Ott, T., Van Dongen, J. T., Günther, C., Krusell, L., Desbrosses, G., Vigeolas, H., et al. (2005). Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development. *Current Biology, 15*(6), 531–535. https://doi.org/10.1016/j.cub.2005.01.042.

Page, K. M., & Guerinot, M. L. (1995). Oxygen control of the *Bradyrhizobium japonicum* hemA gene. *Journal of Bacteriology, 177*(14), 3979–3984. https://doi.org/10.1128/jb.177.14.3979-3984.1995.

Patschkowski, T., Schlüter, A., & Priefer, U. B. (1996). *Rhizobium leguminosarum* bv. viciae contains a second fnr/fixK-like gene and an unusual fixL homologue. *Molecular Microbiology, 21*(2), 267–280. https://doi.org/10.1046/j.1365-2958.1996.6321348.x.

Pawlowski, K., Klosse, U., & de Bruijn, F. J. (1991). Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism. *MGG Molecular & General Genetics, 231*(1), 124–138. https://doi.org/10.1007/BF00293830.

Pawlowski, K., Ratet, P., Schell, J., & de Bruijn, F. J. (1987). Cloning and characterization of nifA and ntrC genes of the stem nodulating bacterium ORS571, the nitrogen fixing symbiont of *Sesbania rostrata*: Regulation of nitrogen fixation (nif) genes in the free living

versus symbiotic state. *MGG Molecular & General Genetics, 206*(2), 207–219. https://doi.org/10.1007/BF00333576.

Pflüger-Grau, K., & Görke, B. (2010). Regulatory roles of the bacterial nitrogen-related phosphotransferase system. *Trends in Microbiology, 18*(5), 205–214. https://doi.org/10.1016/j.tim.2010.02.003.

Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: From saprophytes to endosymbionts. *Nature Reviews Microbiology, 16*(5), 291–303. https://doi.org/10.1038/nrmicro.2017.171.

Postgate, J. R. (1974). New advances and future potential in biological nitrogen fixation. *Journal of Applied Bacteriology, 37*(2), 185–202. https://doi.org/10.1111/j.1365-2672.1974.tb00431.x.

Postgate, J. R. (1982). Biology nitrogen fixation: Fundamentals. *Philosophical Transactions of the Royal Society B: Biological Sciences, 296*(1082), 375–385. https://doi.org/10.1098/rstb.1982.0013.

Preisig, O., Anthamatten, D., & Hennecke, H. (1993). Genes for a microaerobically induced oxidase complex in *Bradyrhizobium japonicum* are essential for a nitrogen-fixing endosymbiosis. *Proceedings of the National Academy of Sciences of the United States of America, 90*(8), 3309–3313. https://doi.org/10.1073/pnas.90.8.3309.

Preisig, O., Zufferey, R., Thöny-Meyer, L., Appleby, C. A., & Hennecke, H. (1996). A high-affinity cbb3-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. *Journal of Bacteriology, 178*(6), 1532–1538. https://doi.org/10.1128/jb.178.6.1532-1538.1996.

Prell, J., White, J. P., Bourdes, A., Bunnewell, S., Bongaerts, R. J., & Poole, P. S. (2009). Legumes regulate Rhizobium bacteroid development and persistence by the supply of branched-chain amino acids. *Proceedings of the National Academy of Sciences, 106*(30), 12477–12482. https://doi.org/10.1073/pnas.0903653106.

Ratet, P., Pawlowski, K., Schell, J., & de Bruijn, F. J. (1989). The *Azorhizobium caulinodans* nitrogen-fixation regulatory gene, nifA, is controlled by the cellular nitrogen and oxygen status. *Molecular Microbiology, 3*(6), 825–838. https://doi.org/10.1111/j.1365-2958.1989.tb00231.x.

Ray, A., & Williams, H. D. (2006). The effects of mutation of the anr gene on the aerobic respiratory chain of *Pseudomonas aeruginosa*. *FEMS Microbiology Letters, 156*(2), 227–232. https://doi.org/10.1111/j.1574-6968.1997.tb12732.x.

Re, S. D., Bertagnoli, S., Fourment, J., Reyrat, J. marc, & Kahn, D. (1994). Intramolecular signal transduction within the FixJ transcriptional activator: In vitro evidence for the inhibitory effect of the phosphorylatable regulatory domain. *Nucleic Acids Research, 22*(9), 1555–1561. https://doi.org/10.1093/nar/22.9.1555.

Renalier, M. H., Batut, J., Ghai, J., Terzaghi, B., Gherardi, M., David, M., et al. (1987). A new symbiotic cluster on the pSym megaplasmid of *Rhizobium meliloti* 2011 carries a functional fix gene repeat and a nod locus. *Journal of Bacteriology, 169*(5), 2231–2238. https://doi.org/10.1128/jb.169.5.2231-2238.1987.

Reutimann, L., Mesa, S., & Hennecke, H. (2010). Autoregulation of fixK 2 gene expression in *Bradyrhizobium japonicum*. *Molecular Genetics and Genomics, 284*(1), 25–32. https://doi.org/10.1007/s00438-010-0547-2.

Reyes-González, A., Talbi, C., Rodríguez, S., Rivera, P., Zamorano-Sánchez, D., & Girard, L. (2016). Expanding the regulatory network that controls nitrogen fixation in *sinorhizobium meliloti*: Elucidating the role of the two-component system hFixL-FxkR. *Microbiology, 162*(6), 979–988. https://doi.org/10.1099/mic.0.000284.

Reyrat, J.-M. M., David, M., Blonski, C., Boistard, P., Batut, J., Reyrat, J.-M. M., et al. (1993). Oxygen-regulated in vitro transcription of *Rhizobium meliloti* nifA and fixK genes. *Journal of Bacteriology, 175*(21), 6867–6872. Retrieved from http://jb.asm.org/content/175/21/6867.full.pdf.

Robson, R. L., & Postgate, J. R. (1980). Oxygen and hydrogen in biological nitrogen fixation. *Annual Review of Marine Science, 34*, 183–207. Retrieved from http://www.annualreviews.org/doi/pdf/10.1146/annurev.mi.34.100180.001151.

Roche, P., Mouawad, L., Perahia, D., Samama, J.-P., & Kahn, D. (2002). Molecular dynamics of the FixJ receiver domain: Movement of the beta4-alpha4 loop correlates with the in and out flip of Phe101. *Protein Science, 11*(11), 2622–2630. https://doi.org/10.1110/ps.0218802.

Rodgers, K. R., & Lukat-Rodgers, G. S. (2005). Insights into heme-based O2 sensing from structure-function relationships in the FixL proteins. *Journal of Inorganic Biochemistry, 99*(4), 963–977. https://doi.org/10.1016/j.jinorgbio.2005.02.016.

Rogel, M. A., Hernández-Lucas, I., Kuykendall, L. D., Balkwill, D. L., & Martinez-Romero, E. (2001). Nitrogen-fixing nodules with ensifer adhaerens harboring rhizobium tropici symbiotic plasmids. *Applied and Environmental Microbiology, 67*(7), 3264–3268. https://doi.org/10.1128/AEM.67.7.3264-3268.2001.

Rogers, C., & Oldroyd, G. E. D. (2014). Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. *Journal of Experimental Botany, 65*(8), 1939–1946. https://doi.org/10.1093/jxb/eru098.

Romanov, V. I., Gordon, A. J., Minchin, F. R., Witty, J. F., Skøt, L., James, C. L., et al. (1995). Anatomy, physiology and biochemistry of root nodules of sprint-2 Fix, a symbiotically defective mutant of pea (Pisum sativum L.). *Journal of Experimental Botany, 46*(12), 1809–1816. https://doi.org/10.1093/jxb/46.12.1809.

Roth, L. E., & Stacey, G. (1989). Bacterium release into host cells of nitrogen-fixing soybean nodules: The symbiosome membrane comes from three sources. *European Journal of Cell Biology, 49*(1), 13–23. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2759097.

Rubio, L. M., & Ludden, P. W. (2005). Maturation of nitrogenase: A biochemical puzzle. *Journal of Bacteriology, 187*(2), 405–414. https://doi.org/10.1128/JB.187.2.405-414.2005.

Ruiz-Argüeso, T., Palacios, J. M., & Imperial, J. (2001). Regulation of the hydrogenase system in *Rhizobium leguminosarum*. *Plant and Soil, 230*(1), 49–57. https://doi.org/10.1023/A:1004578324977.

Sabra, W., Zeng, A.-P., Lünsdorf, H., & Deckwer, W.-D. (2000). Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. *Applied and Environmental Microbiology, 66*(9), 4037–4044. https://doi.org/10.1128/AEM.66.9.4037-4044.2000.

Salazar, E., Javier Díaz-Mejía, J., Moreno-Hagelsieb, G., Martínez-Batallar, G., Mora, Y., Mora, J., et al. (2010). Characterization of the Nif A-RpoN regulon in rhizobium etli in free life and in symbiosis with phaseolus vulgaris. *Applied and Environmental Microbiology, 76*(13), 4510–4520. https://doi.org/10.1128/AEM.02007-09.

Sánchez-Cañizares, C., Jorrín, B., Durán, D., Nadendla, S., Albareda, M., Rubio-Sanz, L., et al. (2018). Genomic diversity in the endosymbiotic bacterium *rhizobium leguminosarum*. *Genes, 9*(2), 60. https://doi.org/10.3390/genes9020060.

Sawers, R. G. (1991). Identification and molecular characterization of a transcriptional regulator from *Pseudomonas aeruginosa* PAO1 exhibiting structural and functional similarity to the FNR protein of *Escherichia coli*. *Molecular Microbiology, 5*(6), 1469–1481. https://doi.org/10.1111/j.1365-2958.1991.tb00793.x.

Schetgens, R. M. P., Hontelez, J. G. J., van den Bos, R. C., & van Kammen, A. (1985). Identification and phenotypical characterization of a cluster of fix genes, including a nif regulatory gene, from *Rhizobium leguminosarum* PRE. *MGG Molecular & General Genetics, 200*(3), 368–374. https://doi.org/10.1007/BF00425719.

Schlüter, A., Patschkowski, T., Quandt, J., Selinger, L. B., Weidner, S., Krämer, M., et al. (1997). Functional and regulatory analysis of the two copies of the fixNOQP operon

of *rhizobium leguminosarum* strain VF39. *Molecular Plant-Microbe Interactions, 10*(5), 605–616. https://doi.org/10.1094/MPMI.1997.10.5.605.

Schlüter, A., Patschkowski, T., Unden, G., & Priefer, U. B. (1992). The *Rhizobium leguminosarum* FnrN protein is functionally similar to *Escherichia coli* Fnr and promotes heterologous oxygen-dependent activation of transcription. *Molecular Microbiology, 6*(22), 3395–3404. https://doi.org/10.1111/j.1365-2958.1992.tb02207.x.

Schlüter, A., Patschkowski, T., Weidner, S., Unden, G., Hynes, M. F., & Priefer, U. B. (1993). Function and regulatory characteristics of FnrN, an oxygen-responsive transcriptional activator in *Rhizobium leguminosarum* bv. viciae. In *New Horizons in nitrogen fixation* (p. 493).

Schmitz, R. A., Klopprogge, K., & Grabbe, R. (2002). Regulation of nitrogen fixation in *Klebsiella pneumoniae* and *Azotobacter vinelandii*: NifL, transducing two environmental signals to the nif transcriptional activator NifA. *Journal of Molecular Microbiology and Biotechnology, 4*(3), 235–242.

Schubert, K. R., & Evans, H. J. (1976). Hydrogen evolution: A major factor affecting the efficiency of nitrogen fixation in nodulated symbionts. *Proceedings of the National Academy of Sciences, 73*(4), 1207–1211. https://doi.org/10.1073/pnas.73.4.1207.

Schulze, J. (2004). How are nitrogen fixation rates regulated in legumes? *Journal of Plant Nutrition and Soil Science, 167*(2), 125–137. https://doi.org/10.1002/jpln.200320358.

Schumacher, J., Da Re, S., Fourment, J., Roche, P., Rousseau, P., Ton-Hoang, B., et al. (2000). Structural basis for signal transduction within the FixJ transcriptional activator. In *Nitrogen fixation: From molecules to crop Productivity* (Vol. 408, pp. 99–100). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47615-0_36.

Sciotti, M. A., Chanfon, A., Hennecke, H., & Fischer, H. M. (2003). Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*. *Journal of Bacteriology, 185*(18), 5639–5642. https://doi.org/10.1128/JB.185.18.5639-5642.2003.

Seefeldt, L. C., Hoffman, B. M., & Dean, D. R. (2009). Mechanism of Mo-dependent nitrogenase. *Annual Review of Biochemistry, 78*(1), 701–722. https://doi.org/10.1146/annurev.biochem.78.070907.103812.

Shah, V. K., & Brill, W. J. (1977). Isolation of an iron-molybdenum cofactor from nitrogenase★. *Proceedings of the National Academy of Sciences, 74*(8), 3249–3253. https://doi.org/10.1073/pnas.74.8.3249.

Shaw, B. D. (1984). Oxygen control mechanisms in nitrogen-fixing systems. In *Current Developments in biological nitrogen fixation* (pp. 111–135). Cambridge: Cambridge University Press. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201302645498.

Sheehy, J. E., Minchin, F. R., & Witty, J. F. (1985). Control of nitrogen fixation in a legume nodule: An analysis of the role of oxygen diffusion in relation to nodule structure. *Annals of Botany, 55*(4), 549–562. https://doi.org/10.1093/oxfordjournals.aob.a086930.

Smil, V. (2001). *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of world Food production*. Boston: MIT Press. Retrieved from https://books.google.co.uk/books?hl=en&lr=&id=G9FljcEASycC&oi=fnd&pg=PR11&dq=Enriching+the+earth:+Fritz+Haber,+Carl+Bosch,+and+the+transformation+of+world+food+production&ots=qPUYMJ0LVI&sig=Jl_716yGeKS-5zw0n7lv1s9RVww#v=onepage&q=Enriching the earth%3A Fr.

Soberón, M., Dávalos, A., Encarnación, S., Taboada, H., Morera, C., Mora, J., et al. (2016). Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic terminal oxidase cbb3 in Rhizobium etli. *Journal of Bacteriology, 179*(22), 6887–6893. https://doi.org/10.1128/jb.179.22.6887-6893.1997.

Soberón, M., Lopez, O., Miranda, J., Tabche, M. L., & Morera, C. (1997). Genetic evidence for 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) as a negative effector of

cytochrome terminal oxidase cbb 3 production in Rhizobium etli. *MGG Molecular & General Genetics, 254*(6), 665–673. https://doi.org/10.1007/s004380050464.

Soberón, M., Morera, C., Kondorosi, A., Lopez, O., & Miranda, J. (2001). A purine-related metabolite negatively regulates fixNOQP expression in *Sinorhizobium meliloti* by modulation of fixK expression. *Molecular Plant-Microbe Interactions, 14*(4), 572–576. https://doi.org/10.1094/MPMI.2001.14.4.572.

Soupène, E., Foussard, M., Boistard, P., Truchet, G., & Batut, J. (1995). Oxygen as a key developmental regulator of *Rhizobium meliloti* N2-fixation gene expression within the alfalfa root nodule. *Proceedings of the National Academy of Sciences, 92*(9), 3759–3763. https://doi.org/10.1073/pnas.92.9.3759.

Sousa, E. H. S. S., Tuckerman, J. R., Gondim, A. C. S. S., Gonzalez, G., & Gilles-Gonzalez, M.-A. A. (2013). Signal transduction and phosphoryl transfer by a FixL hybrid kinase with low oxygen affinity: Importance of the vicinal PAS domain and receiver aspartate. *Biochemistry, 52*(3), 456–465. https://doi.org/10.1021/bi300991r.

Sousa, E. H. S., Tuckerman, J. R., Gonzalez, G., & Gilles-Gonzalez, M. A. (2007). A memory of oxygen binding explains the dose response of the heme-based sensor FixL. *Biochemistry, 46*(21), 6249–6257. https://doi.org/10.1021/bi7003334.

Souza, E. M., Funayama, S., Rigo, L. U., Yates, M. G., & Pedrosa, F. O. (1991). Sequence and structural organization of a nifA-like gene and part of a nifB-like gene of Herbaspirillum seropedicae strain Z78. *Journal of General Microbiology, 137*(7), 1511–1522. https://doi.org/10.1099/00221287-137-7-1511.

Spiro, S. (1994). The FNR family of transcriptional regulators. *Antonie Van Leeuwenhoek, 66*(1–3), 23–36. https://doi.org/10.1007/BF00871630.

Spiro, S., & Guest, J. R. (1987a). Activation of the lac Operon of *Escherichia coli* by a mutant FNR protein. *Molecular Microbiology, 1*(3), 53–58. https://doi.org/10.1111/j.1365-2958.1987.tb00526.x.

Spiro, S., & Guest, J. R. (1987b). Regulation and over-expression of the fnr gene of *Escherichia coli*. *Journal of General Microbiology, 133*(12), 3279–3288. https://doi.org/10.1099/00221287-133-12-3279.

Spiro, S., & Guest, J. R. (1990). FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiology Letters, 75*(4), 399–428. https://doi.org/10.1016/0378-1097(90)90690-R.

Stigter, J. (1994). *Regulation of* Azorhizobium caulinodans *ORS571 nitrogen fixation (NIF/FIX) genes*. Wageningen: Landbouwuniversiteit te Wageningen. Retrieved from http://edepot.wur.nl/206538.

Stock, J. B., Park, P., Surette, M. G., & Levit, M. (1995). Two-component signal transduction systems: Structure-function relationships and mechanisms of catalysis. In *Two-component signal transduction* (pp. 25–51). Washington, D.C: American Society of Microbiology. https://doi.org/10.1128/9781555818319.ch3.

Studholme, D. J., & Dixon, R. (2003). Domain architectures of 54-dependent transcriptional activators. *Journal of Bacteriology, 185*(6), 1757–1767. https://doi.org/10.1128/JB.185.6.1757-1767.2003.

Sundaresan, V., Jones, J. D. G., Ow, D. W., & Ausubel, F. M. (1983). *Klebsiella pneumoniae* nifA product activates the *Rhizobium meliloti* nitrogenase promoter. *Nature, 301*(5902), 728–732. https://doi.org/10.1038/301728a0.

Sutton, V. R., Mettert, E. L., Beinert, H., & Kiley, P. J. (2004). Kinetic analysis of the oxidative conversion of the [4Fe-4S]2+ cluster of FNR to a [2Fe-2S]2+ cluster. *Journal of Bacteriology, 186*(23), 8018–8025. https://doi.org/10.1128/JB.186.23.8018-8025.2004.

Sutton, V. R., Stubna, A., Patschkowski, T., Münck, E., Beinert, H., & Kiley, P. J. (2004). Superoxide destroys the [2Fe-2S]2+ cluster of FNR from *Escherichia coli*. *Biochemistry, 43*(3), 791–798. https://doi.org/10.1021/bi0357053.

Szeto, W. W., Nixon, B. T., Ronson, C. W., & Ausubel, F. M. (1987). Identification and characterization of the *Rhizobium meliloti* ntrC gene: *R. Meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells. *Journal of Bacteriology, 169*(4), 1423–1432. https://doi.org/10.1128/jb.169.4.1423-1432.1987.

Terpolilli, J. J., Hood, G. A., & Poole, P. S. (2012). What determines the efficiency of N2-fixing rhizobium-legume symbioses?. In *Advances in Microbial Physiology* (Vol. 60, pp. 325–389) Elsevier/Academic Press. https://doi.org/10.1016/B978-0-12-398264-3.00005-X.

Thöny, B., Anthamatten, D., & Hennecke, H. (1989). Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon fixR nifA: Analysis of cis- and trans-acting elements. *Journal of Bacteriology, 171*(8), 4162–4169. https://doi.org/10.1128/jb.171.8.4162-4169.1989.

Thöny, B., Fischer, H.-M., Anthamatten, D., Bruderer, T., & Hennecke, H. (1987). The symbiotic nitrogen fixation regulatory operon (fixRnifA) of *Bradyrhizobium japonicum* is expressed aerobically and is subject to a novel, nifA -independent type of activation. *Nucleic Acids Research, 15*(20), 8479–8499. https://doi.org/10.1093/nar/15.20.8479.

Thöny, B., & Hennecke, H. (2006). The -24/-12 promoter comes of age. *FEMS Microbiology Letters, 63*(4), 341–357. https://doi.org/10.1111/j.1574-6968.1989.tb03404.x.

Thumfort, P. P., Atkins, C. A., & Layzell, D. B. (1994). A Re-evaluation of the role of the infected cell in the control of O2 diffusion in legume nodules. *Plant Physiology, 105*(4), 1321–1333. https://doi.org/10.1104/pp.105.4.1321.

Tian, Z., Zou, H., Li, J., Zhang, Y., Liu, Y., Yu, G., et al. (2006). Transcriptome analysis of *Sinorhizobium meliloti* nodule bacteria in nifA mutant background. *Chinese Science Bulletin, 51*(17), 2079–2086. https://doi.org/10.1007/s11434-006-2092-2.

Timmers, A. C. J., Soupène, E., Auriac, M.-C., de Billy, F., Vasse, J., Boistard, P., et al. (2000). Saprophytic intracellular rhizobia in alfalfa nodules. *Molecular Plant-Microbe Interactions, 13*(11), 1204–1213. https://doi.org/10.1094/MPMI.2000.13.11.1204.

Ton-Hoang, B., Salhi, M., Schumacher, J., Da Re, S., & Kahn, D. (2001). Promoter-specific involvement of the FixJ receiver domain in transcriptional activation. *Journal of Molecular Biology, 312*(4), 583–589. https://doi.org/10.1006/jmbi.2001.5014.

Torres, M. J., Argandoñ, M., Vargas, C., Bedmar, E. J., Fischer, H.-M., Mesa, S., et al. (2014). The global response regulator RegR controls expression of denitrification genes in *Bradyrhizobium japonicum*. *PLoS One, 9*(6), e99011. https://doi.org/10.1371/journal.pone.0099011.

Trageser, M., & Unden, G. (1989). Role of cysteine residues and of metal ions in the regulatory functioning of FNR, the transcriptional regulator of anaerobic respiration in *Escherichia coli*. *Molecular Microbiology, 3*(5), 593–599. https://doi.org/10.1111/j.1365-2958.1989.tb00206.x.

Trzebiatowski, J. R., Ragatz, D. M., & De Bruijn, F. J. (2001). Isolation and regulation of *Sinorhizobium meliloti* 1021 Loci induced by oxygen limitation. *Applied and Environmental Microbiology, 67*(8), 3728–3731. https://doi.org/10.1128/AEM.67.8.3728-3731.2001.

Turner, S. L., & Young, J. P. W. (2000). The glutamine synthetases of rhizobia: Phylogenetics and evolutionary implications. *Molecular Biology and Evolution, 17*(2), 309–319. https://doi.org/10.1093/oxfordjournals.molbev.a026311.

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annual Review of Plant Biology, 64*(1), 781–805. https://doi.org/10.1146/annurev-arplant-050312-120235.

Unden, G., & Trageser, M. (1991). Oxygen regulated gene expression in *Escherichia coli*: Control of anaerobic respiration by the FNR protein. *Antonie van Leeuwenhoek, 59*(2), 65–76. https://doi.org/10.1007/BF00445650.

Vance, C. P. (1983). Rhizobium infection and nodulation: A beneficial plant disease? *Annual Review of Microbiology, 37*(1), 399–424. https://doi.org/10.1146/annurev.mi.37.100183.002151.

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., et al. (2010). Plant peptides Govern terminal differentiation of bacteria in symbiosis. *Science, 327*(5969), 1122–1126. https://doi.org/10.1126/science.1184057.

Vasse, J., De Billy, F., Camut, S., & Truchet, G. (1990). Correlation between ultrastructural differentiation of bacteriods and nitrogen fixation in alfalfa nodules. *Journal of Bacteriology, 172*(8), 4295–4306. Retrieved from http://jb.asm.org/.

Vicente, E. J., & Dean, D. R. (2017). Keeping the nitrogen-fixation dream alive. *Proceedings of the National Academy of Sciences, 114*(12), 3009–3011. https://doi.org/10.1073/pnas.1701560114.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., et al. (1997). Human alteration of the global nitrogen cycle: Sources and consequences. *Source: Ecological Applications Ecological Applications Ecological Applications, 7*(3), 737–750. https://doi.org/10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2.

Vitousek, P. M., Menge, D. N. L., Reed, S. C., & Cleveland, C. C. (2013). Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences, 368*(1621), 1–9. https://doi.org/10.1098/rstb.2013.0119.

Vollack, K. U., Härtig, E., Körner, H., & Zumft, W. G. (1999). Multiple transcription factors of the FNR family in denitrifying *Pseudomonas stutzeri*: Characterization of four fnr-like genes, regulatory responses and cognate metabolic processes. *Molecular Microbiology, 31*(6), 1681–1694. https://doi.org/10.1046/j.1365-2958.1999.01302.x.

Waelkens, F., Foglia, A., Morel, J.-B., Fourment, J., Batut, J., & Boistard, P. (1992). Molecular genetic analysis of the *Rhizobium meliloti* fixK promoter: Identification of sequences involved in positive and negative regulation. *Molecular Microbiology, 6*(11), 1447–1456. https://doi.org/10.1111/j.1365-2958.1992.tb00865.x.

Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., et al. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science, 327*(5969), 1126–1129. https://doi.org/10.1126/science.1184096.

Wassem, R., de Souza, E. M., Yates, M. G., Pedrosa, F. de O., & Buck, M. (2000). Two roles for integration host factor at an enhancer-dependent nifA promoter. *Molecular Microbiology, 35*(4), 756–764. https://doi.org/10.1046/j.1365-2958.2000.01746.x.

Wassem, R., Pedrosa, F. O., Yates, M. G., Rego, F. G., Chubatsu, L. S., Rigo, L. U., et al. (2002). Control of autogenous activation of Herbaspirillum seropedicae nifA promoter by the IHF protein. *FEMS Microbiology Letters, 212*(2), 177–192. https://doi.org/10.1111/j.1574-6968.2002.tb11263.x.

Weidner, S., Baumgarth, B., Gottfert, M., Jaenicke, S., Puhler, A., Schneiker-Bekel, S., et al. (2013). Genome sequence of *Sinorhizobium meliloti* Rm41. *Genome Announcements, 1*(1). https://doi.org/10.1128/genomeA.00013-12. e00013-12.

Wei, H., & Layzell, D. B. (2006). Adenylate-coupled ion movement. A mechanism for the control of nodule permeability to O2 diffusion. *Plant Physiology, 141*(1), 280–287. https://doi.org/10.1104/pp.106.077552.

Weisbach, C., Walther, P., Hartwig, U. A., & Nosberger, J. (1999). Electron microscopic investigation of water occlusions in intercellular spaces in the inner cortex of lucerne nodules. *Journal of Structural Biology, 126*(1), 59–71. https://doi.org/10.1006/jsbi.1999.4100.

Westhoek, A., Field, E., Rehling, F., Mulley, G., Webb, I., Poole, P. S., et al. (2017). Policing the legume-rhizobium symbiosis: A critical test of partner choice. *Scientific Reports, 7*(1), 1419. https://doi.org/10.1038/s41598-017-01634-2.

West, A. H., & Stock, A. M. (2001). Histidine kinases and response regulator proteins in two-component signaling systems. *Trends in Biochemical Sciences, 26*(6), 369–376. https://doi.org/10.1016/S0968-0004(01)01852-7.

Wheatley, R. M., & Poole, P. S. (2018). Mechanisms of bacterial attachment to roots. *FEMS Microbiology Reviews, 42*(4), 448–461. https://doi.org/10.1093/femsre/fuy014.

van de Wiel, C., Scheres, B., Franssen, H., van Lierop, M. J., van Lammeren, A., van Kammen, A., et al. (1990). The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules. *The EMBO Journal, 9*(1), 1–7. https://doi.org/10.1002/j.1460-2075.1990.tb08073.x.

Wittenberg, J. B., Bergersen, F. J., Appleby, C. A., & Turner, G. L. (1974). Facilitated oxygen diffusion: The role of leghemoglobin in nitrogen fixation by bacteroids isolated from soybean root nodules. *Journal of Biological Chemistry, 249*(13), 4057–4066. Retrieved from http://www.jbc.org/.

Wittenberg, J. B., & Wittenberg, B. A. (2011). *Facilitated oxygen diffusion by oxygen carriers*. New York: Springer. https://doi.org/10.1007/978-1-4612-5890-2_9.

Witty, J. F., & Minchin, F. R. (1990). Oxygen diffusion in the legume root nodule. In *Nitrogen fixation* (pp. 285–292). Boston: Springer. https://doi.org/10.1007/978-1-4684-6432-0_29.

Witty, J. F., Minchin, F. R., Skot, L., & Sheehy, J. E. (1986). Nitrogen fixation and oxygen in legume root nodules. In *Oxford Surveys of plant molecular & cell Biology* (Vol. 3, pp. 275–314). Oxford: Oxford University Press. Retrieved from file http://home/ford/Documents/PDF/rfd850.pdf.

Witty, J. F., Skøt, L., & Revsbech, N. P. (1987). Direct evidence for changes in the resistance of legume root nodules to O2 diffusion. *Journal of Experimental Botany, 38*(7), 1129–1140. https://doi.org/10.1093/jxb/38.7.1129.

Wongdee, J., Boonkerd, N., Teaumroong, N., Tittabutr, P., & Giraud, E. (2018). Regulation of nitrogen fixation in Bradyrhizobium sp. Strain DOA9 involves two distinct NifA regulatory proteins that are functionally redundant during symbiosis but not during free-living growth. *Frontiers in Microbiology, 9*(JUL), 1–11. https://doi.org/10.3389/fmicb.2018.01644.

Wright, G. S. A., Saeki, A., Hikima, T., Nishizono, Y., Hisano, T., Kamaya, M., et al. (2018). Architecture of the complete oxygen-sensing FixL-FixJ two-component signal transduction system. *Science Signaling, 11*(525), 1–12. https://doi.org/10.1126/scisignal.aaq0825.

Wycoff, K. L., Hunt, S., Gonzales, M. B., VandenBosch, K. A., Layzell, D. B., & Hirsch, A. M. (1998). Effects of oxygen on nodule physiology and expression of nodulins in alfalfa. *Plant Physiology, 117*(2), 385–395. https://doi.org/10.1104/pp.117.2.385.

Yamada, S., Sugimoto, H., Kobayashi, M., Ohno, A., Nakamura, H., & Shiro, Y. (2009). Structure of PAS-linked histidine kinase and the response regulator complex. *Structure, 17*(10), 1333–1344. https://doi.org/10.1016/j.str.2009.07.016.

Yao, Z., Tian, Z., Dai, X., Becker, A., Li, J., Yan, H., et al. (2006). Complementation analyses of *Sinorhizobium meliloti* nifA mutant with different originated nifA genes. *Chinese Science Bulletin, 51*(22), 2748–2754. https://doi.org/10.1007/s11434-006-2203-0.

Young, J. P. W., Crossman, L. C., Johnston, A. W. B., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., et al. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biology, 7*(4), R34. https://doi.org/10.1186/gb-2006-7-4-r34.

Zamorano-Sánchez, D., & Girard, L. (2015). FNR-like proteins in rhizobia: Past and future. In *Biological nitrogen fixation* (Vol. 1, pp. 155–166). Hoboken: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119053095.ch15.

Zamorano-Sánchez, D., Reyes-González, A., Gómez-Hernández, N., Rivera, P., Georgellis, D., & Girard, L. (2012). FxkR provides the missing link in the fixL-fixK

signal transduction cascade in rhizobium etli CFN42. *Molecular Plant-Microbe Interactions, 25*(11), 1506–1517. https://doi.org/10.1094/MPMI-05-12-0136-R.

Zou, X., Zhu, Y. Y., Pohlmann, E. L., Li, J., Zhang, Y., & Roberts, G. P. (2008). Identification and functional characterization of NifA variants that are independent of GlnB activation in the photosynthetic bacterium *Rhodospirillum rubrum*. *Microbiology, 154*(9), 2689–2699. https://doi.org/10.1099/mic.0.2008/019406-0.

A.2 Appendix 2: Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a *Rhizobium*-legume symbiosis

PLOS GENETICS

RESEARCH ARTICLE

Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a *Rhizobium*-legume symbiosis

Paul J. Rutten[1], Harrison Steel[2], Graham A. Hood[3], Vinoy K. Ramachandran[1], Lucie McMurtry[1], Barney Geddes[1], Antonis Papachristodoulou[2], Philip S. Poole[1]*

1 Department of Plant Sciences, University of Oxford, Oxford, United Kingdom, **2** Department of Engineering Science, University of Oxford, Oxford, United Kingdom, **3** Department of Molecular Microbiology, John Innes Centre, Norwich, United Kingdom

* philip.poole@plants.ox.ac.uk

OPEN ACCESS

Citation: Rutten PJ, Steel H, Hood GA, Ramachandran VK, McMurtry L, Geddes B, et al. (2021) Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a *Rhizobium*-legume symbiosis. PLoS Genet 17(2): e1009099. https://doi.org/10.1371/journal.pgen.1009099

Editor: Sean Crosson, Michigan State University, UNITED STATES

Received: September 4, 2020

Accepted: December 4, 2020

Published: February 4, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pgen.1009099

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Abstract

Regulation by oxygen (O_2) in rhizobia is essential for their symbioses with plants and involves multiple O_2 sensing proteins. Three sensors exist in the pea microsymbiont *Rhizobium leguminosarum* Rlv3841: hFixL, FnrN and NifA. At low O_2 concentrations (1%) hFixL signals via FxkR to induce expression of the FixK transcription factor, which activates transcription of downstream genes. These include *fixNOQP*, encoding the high-affinity cbb_3-type terminal oxidase used in symbiosis. In free-living Rlv3841, the hFixL-FxkR-FixK pathway was active at 1% O_2, and confocal microscopy showed hFixL-FxkR-FixK activity in the earliest stages of Rlv3841 differentiation in nodules (zones I and II). Work on Rlv3841 inside and outside nodules showed that the hFixL-FxkR-FixK pathway also induces transcription of *fnrN* at 1% O_2 and in the earliest stages of Rlv3841 differentiation in nodules. We confirmed past findings suggesting a role for FnrN in *fixNOQP* expression. However, unlike hFixL-FxkR-FixK, Rlv3841 FnrN was only active in the near-anaerobic zones III and IV of pea nodules. Quantification of *fixNOQP* expression in nodules showed this was driven primarily by FnrN, with minimal direct hFixL-FxkR-FixK induction. Thus, FnrN is key for full symbiotic expression of *fixNOQP*. Without FnrN, nitrogen fixation was reduced by 85% in Rlv3841, while eliminating hFixL only reduced fixation by 25%. The hFixL-FxkR-FixK pathway effectively primes the O_2 response by increasing *fnrN* expression in early differentiation (zones I-II). In zone III of mature nodules, near-anaerobic conditions activate FnrN, which induces *fixNOQP* transcription to the level required for wild-type nitrogen fixation activity. Modelling and transcriptional analysis indicates that the different O_2 sensitivities of hFixL and FnrN lead to a nuanced spatiotemporal pattern of gene regulation in different nodule zones in response to changing O_2 concentration. Multi-sensor O_2 regulation is prevalent in rhizobia, suggesting the fine-tuned control this enables is common and maximizes the effectiveness of the symbioses.

Funding: This work was supported by the Biotechnology and Biological Sciences Research Council [grant numbers BB/L011484/1 and BB/M011224/1]. AP and HS were supported by the Engineering and Physical Sciences Research Council [grant number EP/M002454/1]. The funders played no role in the study design, data collection, decision to publish or preparation of the manuscript. https://bbsrc.ukri.org.

Competing interests: No competing interests.

Author summary

Rhizobia are soil bacteria that form a symbiosis with legume plants. In exchange for shelter from the plant, rhizobia provide nitrogen fertilizer, produced by nitrogen fixation. Fixation is catalysed by the nitrogenase enzyme, which is inactivated by oxygen. To prevent this, plants house rhizobia in root nodules, which create a low oxygen environment. However, rhizobia need oxygen, and must adapt to survive the low oxygen concentration in the nodule. Key to this is regulating their genes based on oxygen concentration. We studied one *Rhizobium* species which uses three different protein sensors of oxygen, each turning on at a different oxygen concentration. As the bacteria get deeper inside the plant nodule and the oxygen concentration drops, each sensor switches on in turn. Our results also show that the first sensor to turn on, hFixL, primes the second sensor, FnrN. This prepares the rhizobia for the core region of the nodule where oxygen concentration is lowest and most nitrogen fixation takes place. If both sensors are removed, the bacteria cannot fix nitrogen. Many rhizobia have several oxygen sensing proteins, so using multiple sensors is likely a common strategy enabling rhizobia to adapt to low oxygen precisely and in stages during symbiosis.

Introduction

Rhizobia are alpha-proteobacteria that engage in symbiosis with legume plants [1]. The bacteria convert inert atmospheric N_2 into biologically accessible ammonia and provide it to their plant host in a process called nitrogen fixation [2,3]. All biological fixation is catalysed by the nitrogenase enzyme complex that evolved before the Great Oxygenation Event and requires near-anoxic conditions to function [4–6]. However, rhizobia are obligate aerobes and must respire to meet the high energy demands of nitrogen fixation [7,8]. These competing requirements create an 'oxygen paradox' in symbiotic nitrogen fixation [9,10]. To overcome this paradox, intricate cooperation between rhizobia and their plant partners has evolved (reviewed in [11,12]). Legume plants host rhizobia in dedicated root nodules which form where bacteria have entered the plant root, usually via infection threads (reviewed in [13,14]). Nodules create a near-anoxic internal environment suitable for nitrogenase activity [15–17]. To produce this environment, oxygen (O_2) is captured and shuttled to bacteroids by plant leghaemoglobins [18–21]. The concentration of remaining free O_2 in the core nitrogen fixation zone of nodules is as low as 20–50 nM [22,23]. Rhizobia undergo a radical lifestyle change after nodule entry to survive and fix nitrogen in these conditions (reviewed in [24,25]). In indeterminate nodules, such as those produced by *Pisum sativum* (pea), rhizobia are initially free-living upon entry [26,27]. They then undergo irreversible lifestyle changes as they move from the nodule tip to its core [28,29]. Beginning in zone II and accelerating in the II-III interzone, rhizobia terminally differentiate into quasi-organelle bacteroids specialized for nitrogen fixation [30,31]. Zone III of indeterminate nodules contains differentiated bacteroids which are actively fixing nitrogen [32]. Rhizobial regulatory mechanisms sensitive to O_2 tension are essential for successful differentiation into bacteroids and the establishment of a productive symbiosis [33–35].

Multiple O_2 sensors have evolved in rhizobia, three of which are widespread and often co-exist within the same organism [11,36]. The first is the membrane-bound FixL protein, which forms a two-component system (TCS) with the FixJ receiver protein (reviewed in [37,38]). Under microaerobic conditions, FixL phosphorylates FixJ, which in turn induces expression of the *fixK* transcription factor [39–41]. FixK induces expression of downstream genes by

binding as a dimer to an 'anaerobox' motif (TTGAT-N_4-ATCAA) upstream of their promoters [42,43].

The second common O_2 sensor is a variant of FixL called hybrid FixL (hFixL) [44,45]. This forms an alternative TCS with FxkR acting as the receiver protein. FxkR is not a FixJ homolog but similarly induces expression of *fixK*, by binding to an upstream 'K-box' motif (GTTA-CA-N_4-GTTACA) [46]. The third O_2 sensor is the FnrN transcription factor. Like FixK, FnrN binds the anaerobox motif as a dimer and both are close homologs of the *E. coli* anaerobiosis regulator FNR [47–49]. Unlike FixK but like FNR, FnrN contains an N-terminal cysteine-rich cluster that makes the protein a direct sensor of O_2 [50–53]. The FixL and hFixL sensors are known to become active at relatively mildly microaerobic conditions, including in free-living rhizobia [54–56]. FnrN is likely to be far less O_2 tolerant. The O_2 sensitivity of FnrN has not been determined, but the *E. coli* FNR homolog is active only under anaerobic conditions [57–59]. All symbiotic rhizobia studied to date employ at least one of these three sensors [11]. It is common for these sensors to coexist, notably in *Rhizobium leguminosarum* biovar *viciae* VF39, multiple strains of *Ensifer meliloti* (previously *Sinorhizobium meliloti*) and *Rhizobium etli* CFN42 [44,45,60–62].

Further emphasizing the importance of O_2 regulation in symbiotic nitrogen fixation, rhizobia also employ the O_2 sensing NifA transcription factor to regulate their final differentiation into nitrogen fixing bacteroids (for reviews see [38,63]). NifA oxygen sensitivity is thought to derive from a metal-binding cysteine-rich motif in an inter-domain linker of the protein [64–66]. The protein has a large regulon, notably including nitrogenase components such as *nifH* [67–70]. Expression of *nifA* is typically auto-regulated in rhizobia, often via read-through from an upstream gene or operon that is NifA regulated, in many cases *fixABCX* [69,71–73]. In Rlv3841, a *fixABCX* operon is found directly upstream of *nifA*, suggesting such a read-through NifA auto-activation mechanism. Usually, neither expression of *nifA* nor the activity of the protein is directly regulated by the three O_2 sensors described above [11]. One notable exception is *E. meliloti*, where *nifA* is regulated by the FixLJ system [74–76]. There is no evidence that FixK or FnrN directly regulates *nifA* expression in Rlv3841.

There appears to be a spectrum among rhizobia, with some species segregating oxygen sensors into separate pathways, whilst in other species these sensors have partially or completely merged into a combined hierarchical pathway [77,78]. Where oxygen sensors are in separated pathways, redundancy often exists. In these situations the loss of one oxygen sensor does not abolish nitrogen fixation activity [79,80]. By contrast, where sensors have been merged into a single regulatory pathway, some components are individually essential [11]. Thus, loss of one oxygen sensor can severely impair nitrogen fixation even if other sensors remain.

In *R. leguminosarum* bv. VF39, knocking out FnrN or hFixL reduced nitrogen fixation to 30% or 50% of WT respectively, suggesting a non-hierarchical, redundant arrangement [60]. The hFixL-FxkR-FixK pathway of *R. etli* CFN42 is dispensable as a double *fixK* mutant had no effect on nitrogen fixation, whilst a double *fnrN* mutant reduced fixation to 20% of WT levels. By contrast, *R. etli* CFN42 appears to employ a complex hierarchical pathway, with multiple homologs of FixK and FnrN regulating each other's expression [61,62]. Species encoding homologs of only hFixL or FnrN have also been found. *Rhizobium leguminosarum* biovar *viciae* UPM791 contains two FnrN homologs but neither FixL nor hFixL [81]. It is unknown whether the two FnrN proteins respond to different O_2 concentrations or act in a redundant fashion. *E. meliloti* 1021 contains no FnrN homolog but a well-studied FixLJ system and appears to have homologs of hFixL and FxkR [82–84].

To examine the relationship between hFixL and FnrN, we studied the model organism *Rhizobium leguminosarum* biovar *viciae* 3841 (Rlv3841) which employs both sensors (Fig 1) [85,86]. Rlv3841 has a single chromosome whose gene names start with RL, and six

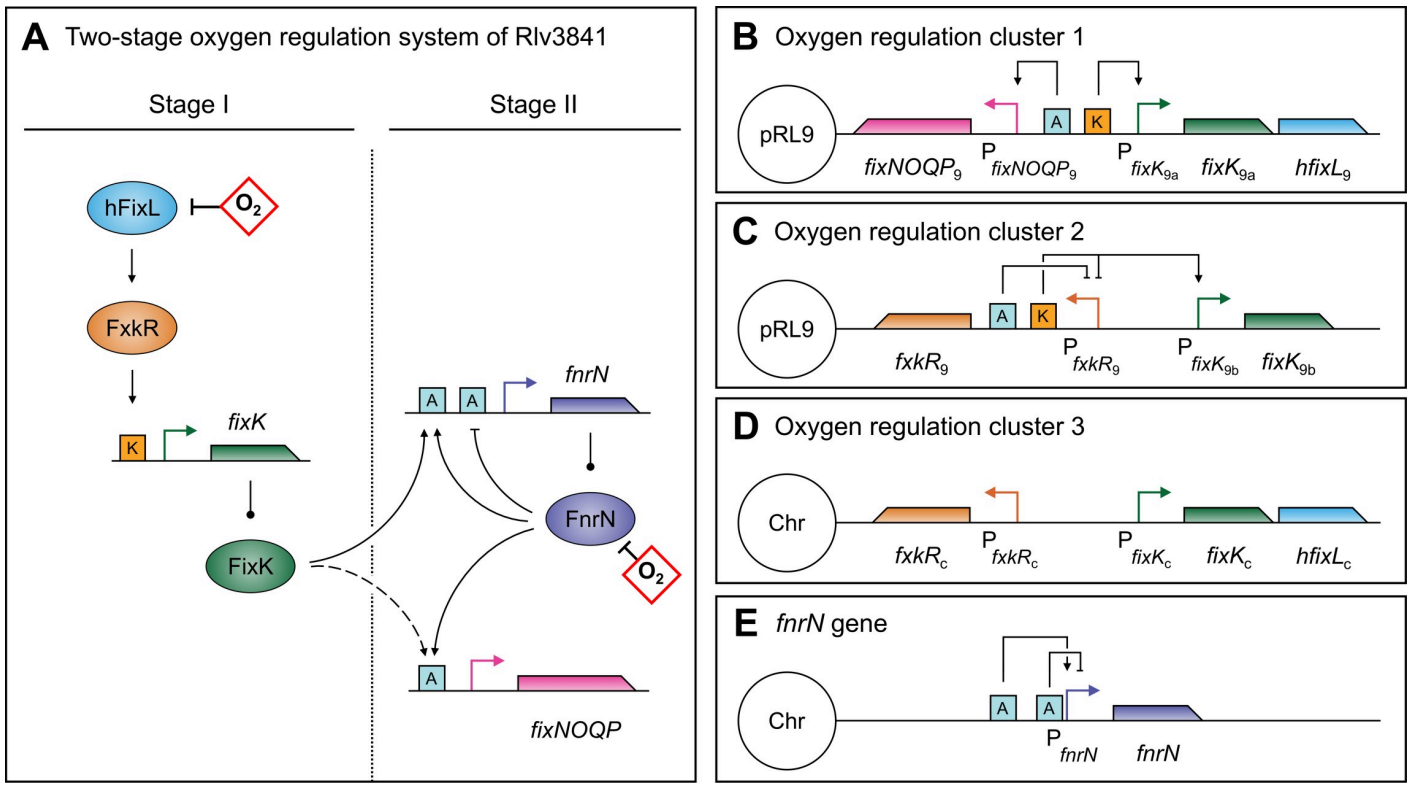


Fig 1. The integrated hFixL-FxkR-FixK and FnrN oxygen regulation systems of Rlv3841 form a single pathway and are genetically clustered. Oxygen is shown in red diamonds. Proteins are shown as ovals, operator sites as squares and genes as pointed rectangles. Transcription start sites are shown as right-angled arrows. Line endings indicate activation (arrows), inhibition (blunt end) and translation (circle). (**A**) The single pathway formed by the two sensors acts in two stages. Stage I starts under microaerobic conditions and can function outside the nodule. In this stage, hFixL is active but FnrN is not. hFixL activates FxkR, which binds to the K-box operator (orange "K" squares) to induce expression of *fixK*. FixK binds to anaerobox operators (blue "A" squares) to induce expression, including upstream of *fixNOQP* (dashed line) and *fnrN*. Once oxygen in the bacteria reaches near-anaerobic levels, FnrN becomes active and stage II begins. Like FixK, FnrN binds anaeroboxes. It auto-regulates *fnrN* both positively and negatively and induces *fixNOQP* expression. (**B**) Rlv3841 has multiple copies of several oxygen regulation genes and many are arranged in clusters. On megaplasmid pRL9, $fixK_{9a}$ forms an operon with $hfixL_9$, regulated by a K-box. This operon is adjacent to $fixNOQP_9$, regulated by an anaerobox. (**C**) $fixK_{9b}$ and $fxkR_9$ are adjacent, with an anaerobox and a K-box in their intergenic region. (**D**) The Rlv3841 chromosome also has a cluster, containing $fxkR_c$, $fixK_c$ and $hfixL_c$. Unlike the similar clusters on pRL9, the intergenic region of this cluster contains no anaerobox or K-box operators. (**E**) The *fnrN* gene is not part of a cluster and is positively and negatively regulated by a distal and proximal anaerobox, respectively. Details of transcription start site, anaerobox and K-box locations can be found in Table 1.

https://doi.org/10.1371/journal.pgen.1009099.g001

megaplasmids pRL7-12 whose gene names start with e.g. pRL9. The main symbiotic plasmid is pRL10, but many symbiotic genes are also found on pRL9 including a copy of the *fixNOQP* and *fixGHIS* operon. Rlv3841 encodes two copies of *hfixL*, which we named $hfixL_9$ (pRL90020 on pRL9) and $hfixL_c$ (RL1879 on the chromosome), with 54.9% identity at the protein level. The strain also contains two homologs (58% identity) of *fxkR*, $fxkR_9$ (pRL90026) and $fxkR_c$ (RL1881). It has three putative *fixK* genes, which we designated $fixK_{9a}$ (pRL90019), $fixK_{9b}$ (pRL90025) and $fixK_c$ (RL1880). The $fixK_{9a}$ and $fixK_{9b}$ sequences have 53% amino acid identity, whilst $fixK_c$ shares 38% and 47% identity with these proteins, respectively. Both $fixK_{9a}$-$hfixL_9$ and $fixK_c$-$hfixL_c$ appear to form operons (Fig 1B and 1D). Rlv3841 has one copy of *fnrN* (RL2818), regulated by two anaeroboxes. A similar dual-anaerobox arrangement exists in Rlv UPM791, where FnrN positively and negatively auto-regulates its own expression [87]. Binding of FnrN to the distal anaerobox induces *fnrN* transcription and binding to the proximal anaerobox represses it. Auto-activation of FnrN has also been reported in *Rhizobium etli* CNPAF512 [88]. FixK regulation of *fnrN* expression is likely as it also binds anaeroboxes, but this had not been investigated.

A study in *R. leguminosarum* VF39 found that microaerobic expression of *fnrN* also requires RpoN [89]. This finding has not been replicated elsewhere, and its significance remains unclear. Work in *R. etli* CNPAF512 showed that *fnrN* is not controlled by RpoN in that organism [88]. Rlv3841 encodes one putative *rpoN* gene (RL0422), but we found no RpoN binding sites upstream of the Rlv3841 *fnrN* transcription start site. RpoN therefore does not appear to be required for *fnrN* expression in Rlv3841.

A parallel arrangement of hFixL-FxkR-FixK and FnrN in Rlv3841 would produce redundancy, whereas an arrangement in series would create hierarchy between the two regulators. Our goal is therefore to understand how the two sensors interact in Rlv3841 and to provide insight into why they coexist.

Results

Expression of *fnrN* is auto-regulated and controlled by the hFixL-FxkR-FixK pathway

The hFixL-FxkR-FixK pathway is known to be active at relatively high O_2 concentrations, including in free-living rhizobia under microaerobic conditions [45,90]. The role of FnrN is less well understood and we began by investigating this sensor. The *fnrN* gene contains two anaeroboxes upstream of its promoter, a proximal site at -2 relative to the transcription start site (TSS) and a distal site at -34 (Fig 1E and Table 1). Binding at the proximal site inhibits transcription through steric hindrance, whilst the distal site activates it [87]. As expected from the presence of the distal site, *fnrN* was induced under microaerobic conditions in free-living Rlv3841 (Fig 2). Both FixK and FnrN bind to anaerobox operators and induce gene expression under microaerobic conditions [91,92]. Therefore, microaerobic induction of *fnrN* could be due to FnrN auto-activation and/or induction by the hFixL-FxkR-FixK pathway. To determine their respective importance, expression of *fnrN* was studied in Rlv3841 mutants defective in either hFixL or FnrN.

In a double *hfixL* mutant (LMB496; $hfixL_9$::ΩSpec $hfixL_c$:pK19 single recombination), free-living microaerobic expression of *fnrN* was reduced to 25% of its wild-type (WT) level (Fig 3).

Table 1. Location of transcription start sites, anaeroboxes and K-boxes for select oxygen regulation genes in Rlv3841.

Gene name	TSS coordinate	Anaerobox 1 location	Anaerobox 2 location	K-box location
fnrN (RL2818)	2978390	-34	-2	Not present
$fixK_{9a}$ (pRL90019)	19878	Not present	Not present	-62
$fixK_{9b}$ (pRL90025)	27090	Not present	Not present	-62
$fixK_c$ (RL1880)	1977111	Not present	Not present	Not present
$fxkR_9$ (pRL90026)	27146	Not present	+38	+6
$fxkR_c$ (RL1881)	1977167	Not present	Not present	Not present
$fixNOQP_9$ (pRL90018-16)	19721	-33	Not present	Not present
$fixNOQP_{10}$ (pRL100205-207)	206214	-34	Not present	Not present
$fixGHIS_9$ (pRL90015-12A)	15908	-32	Not present	Not present
$fixGHIS_{10}$ (pRL100208)	210004	-32	Not present	Not present

Locations of anaeroboxes and K-boxes are given relative to the TSS of each gene respectively. Anaeroboxes more than 90 bp upstream of transcription start sites are not included. The location of activating anaeroboxes is well conserved across the genes shown above (between -32 and -34 relative to the TSS). Likewise, both megaplasmid-encoded *fixK* copies have their respective K-boxes in the same position relative to the TSS (-62). The anaerobox downstream (+38) of the $fxkR_9$ transcription start site likely represses it when bound. A single K-box is shared between $fxkR_9$ and $fixK_{9b}$ (see Fig 1). It is correctly located (-62) to induce $fixK_{9b}$ when bound, and its location suggests it simultaneously represses $fxkR_9$ (+6). The operators downstream of the $fxkR_9$ TSS may create a negative feedback loop in the Rlv3841 hFixL-FxkR-FixK pathway. The downstream anaerobox also suggests FnrN repression of hFixL-FxkR-FixK via repression of $fxkR_9$ transcription, but this requires further study.

https://doi.org/10.1371/journal.pgen.1009099.t001

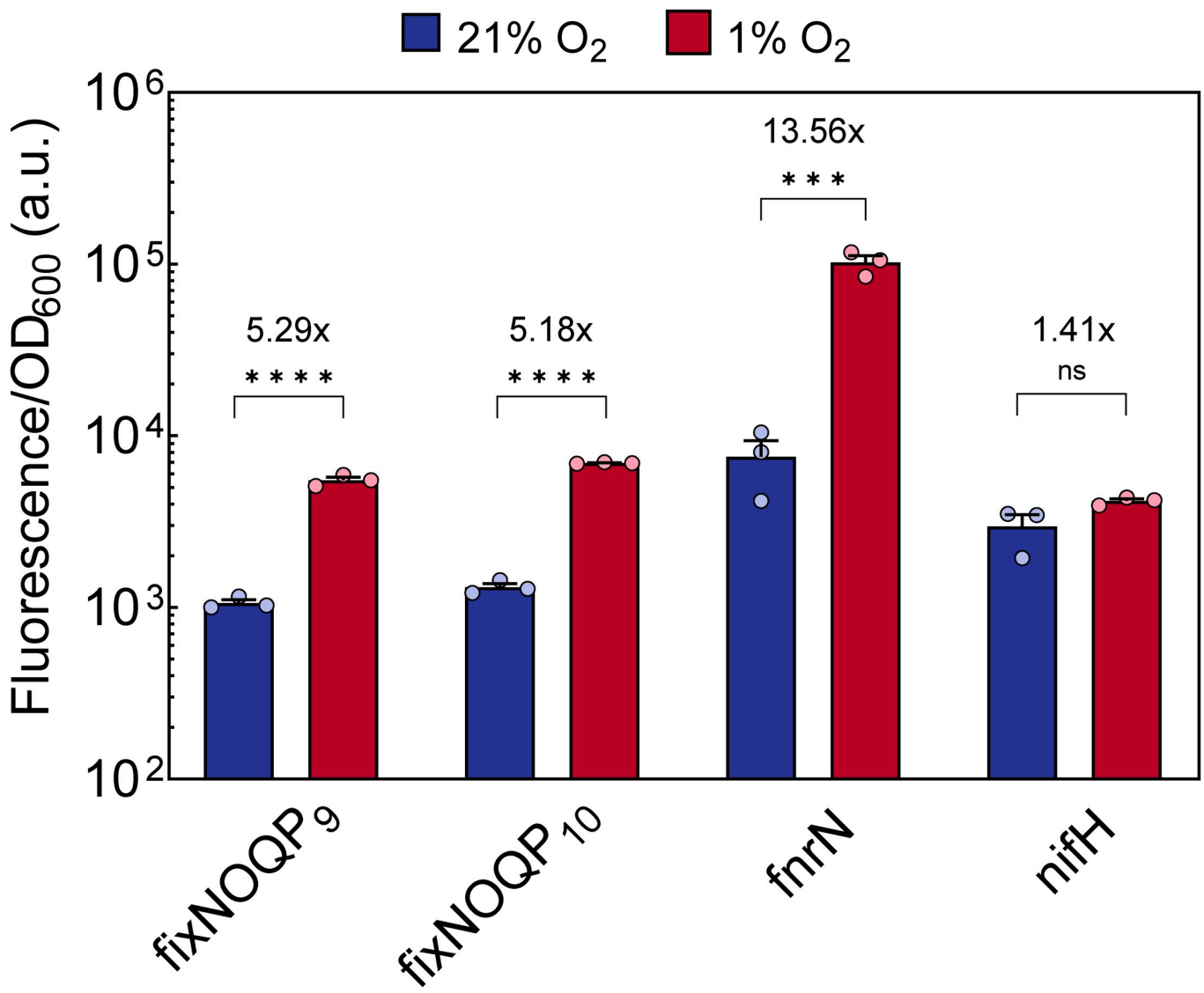


Fig 2. Microaerobiosis induces *fixNOQP* and *fnrN* genes in free-living Rlv3841. Promoter fusions of *fixNOQP*$_9$ (OPS1267), *fixNOQP*$_{10}$ (OPS1287) and *fnrN* (OPS1296) were used to measure the activity of these promoters in free-living Rlv3841 at 1% O_2 (red bars) relative to 21% O_2 (blue bars). Activity of all three promoters, measured as fluorescence normalised by OD_{600}, increased under microaerobic conditions. A positive control (OPS1294) showed no impact on OD-normalised fluorescence due to the microaerobic environment. A similar fold induction of ~5 was recorded for both *fixNOQP* operons, but *fnrN* showed more than double this fold change indicating stronger induction. No effect of O_2 concentration on *nifH* expression (OPS1268) was observed, indicating no NifA activity. Values are plotted on a logarithmic scale. Data are averages (±SEM) from three biological replicates, ns (not significant) $P \geq 0.05$; $^{***}P < 0.001$; $^{****}P < 0.0001$; by Student's t test.

https://doi.org/10.1371/journal.pgen.1009099.g002

The single mutant of *hfixL*$_9$ (LMB495; *hfixL*$_9$::ΩSpec) individually reproduced most of this reduction whilst the single mutant of *hfixL*$_c$ (LMB403; *hfixL*$_c$:pK19 single recombination) did not reduce *fnrN* expression. This suggests hFixL$_9$ is the critical hFixL protein under free-living microaerobic conditions, with hFixL$_c$ playing little to no role.

hFixL acts through the FxkR intermediary in rhizobia and Rlv3841 contains two FxkR homologs [44]. *fxkR*$_9$ forms an O_2 regulation cluster with *fixK*$_{9b}$ which we labelled cluster 2 (Fig 1C) and *fxkR*$_c$ forms a cluster with *fixK*$_c$-*hfixL*$_c$, which we labelled cluster 3 (Fig 1D). Cluster 2 contains an anaerobox and a K-box. The anaerobox is relatively far from the *fixK*$_{9b}$ TSS

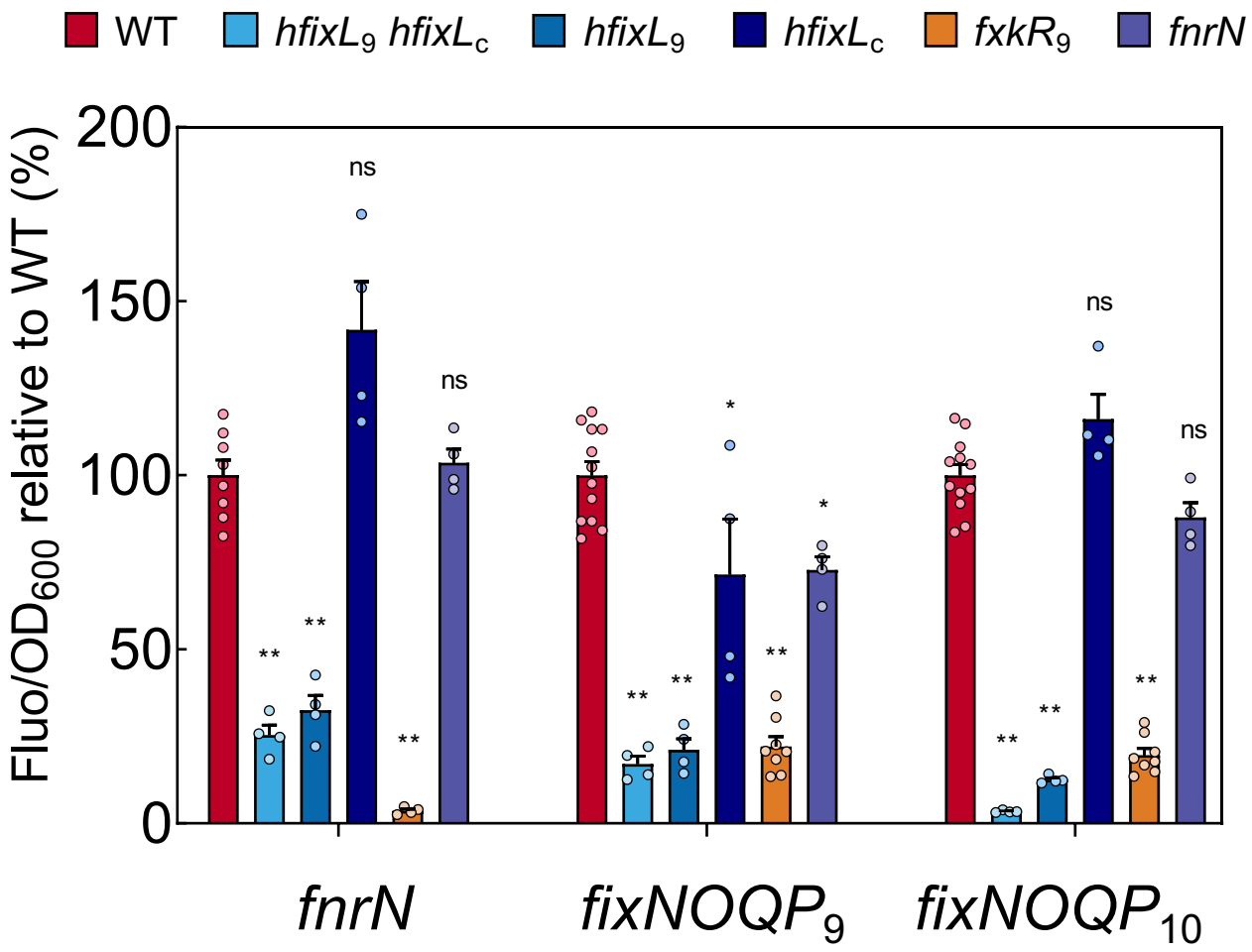


Fig 3. Under microaerobic (1% O_2) conditions in free living cells, the hFixL-FxkR-FixK pathway and not FnrN is a key activator of anaerobox controlled genes. Plasmids with promoter fusions to *syfp2* for *fnrN* (pOPS0980), *fixNOQP*$_9$ (pOPS0978) and *fixNOQP*$_{10}$ (pOPS0977) were conjugated into Rlv3841 WT and O_2 regulation mutants (*hfixL*$_9$ *hfixL*$_c$, LMB496; *hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403; *fxkR*$_9$, OPS1808; *fnrN*, LMB648). Fluorescence and OD_{600} measurements were taken after cells were grown under microaerobic conditions (1% O_2). Individual values (Fluo/OD_{600}) are normalised such that the WT average is 100% for each reporter group. Activity from all three promoters was critically reduced or nearly abolished in the double *hfixL* and *fxkR* mutant backgrounds. The *hfixL*$_9$ homolog had a far more pronounced effect on expression of all three genes than did the *hfixL*$_c$ homolog. Little or no reduction in expression was observed when *fnrN* was mutated. Data are averages (±SEM) from at least four biological replicates. Statistical tests are differences relative to WT expression; ns (no significant decrease) $P \geq 0.05$; $^{*}P < 0.001$; $^{**}P < 0.0001$; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

https://doi.org/10.1371/journal.pgen.1009099.g003

(-94), but downstream of the *fxkR*$_9$ TSS, suggesting its main role is to repress *fxkR*$_9$ transcription (Table 1). Its effect on *fixK*$_{9b}$ transcription may be minimal, if any. The K-box is also downstream of the *fxkR*$_9$ TSS, suggesting it also represses transcription. Simultaneously, this K-box also likely acts to induce *fixK*$_{9b}$, as its location upstream of the *fixK*$_{9b}$ TSS (-62) is identical to the relative position of the upstream *fixK*$_{9a}$ K-box in cluster 1 (Fig 1B and Table 1). Both K-boxes therefore appear functional for *fixK* induction, and the cluster 2 K-box appears to have a dual function, repressing *fxkR*$_9$ and inducing *fixK*$_{9b}$. The second cluster, containing *fxkR*$_c$, has no anaeroboxes or K-boxes (Fig 1D). This could imply minimal expression from the genes in this cluster, or constitutive expression that is not O_2 regulated. Neither *fixK*$_c$ (1.6-fold upregulated, p = 0.140) nor *fxkR*$_c$ (1.3-fold upregulated, p = 0.145) was significantly upregulated in 21 day old bacteroids compared to free-living Rlv3841 [93]. Based on these findings,

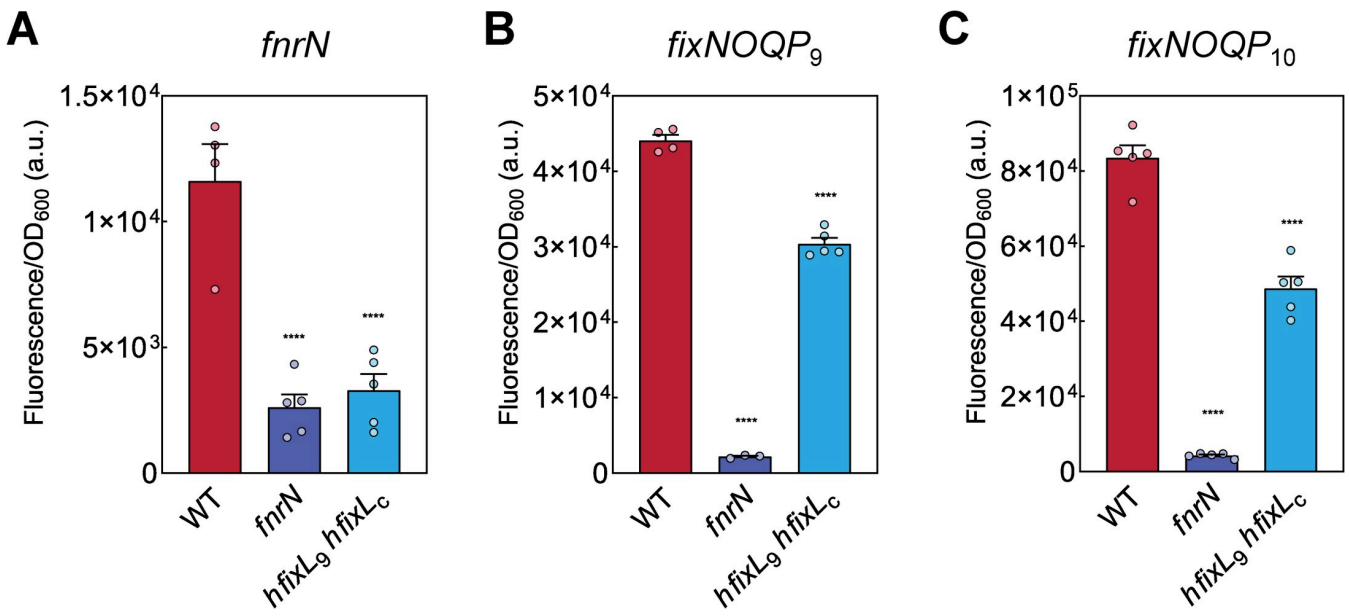


Fig 4. *In planta*, *fnrN* is both auto-regulated and controlled by the hFixL-FxkR-FixK pathway, whilst the *fixNOQP* operons are primarily controlled by FnrN. Rlv3841 WT and mutant strains (*fnrN*, LMB648; *hfixL*$_9$ *hfixL*$_c$, LMB496) containing promoter fusions to *syfp2* for *fnrN* (pOPS0980), *fixNOQP*$_9$ (pOPS0978) and *fixNOQP*$_{10}$ (pOPS0977) were inoculated on plants and bacteroids isolated for measurements. Expression in bacteroids of *fnrN* (**A**) is impaired in the *fnrN* background where auto-activation cannot take place. Expression of *fnrN* is similarly impaired in the double *hfixL* mutant, indicating the hFixL-FxkR-FixK pathway also plays an important role in symbiotic *fnrN* induction. Expression of *fixNOQP*$_9$ (**B**) and *fixNOQP*$_{10}$ (**C**) is significantly reduced in the double *hfixL* mutant and almost abolished in the *fnrN* mutant. Thus both FnrN and the hFixL-FxkR-FixK pathway play an important role in the expression of all three genes. Data are averages (±SEM) from at least three plants, ****P < 0.0001; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

https://doi.org/10.1371/journal.pgen.1009099.g004

we speculated that FxkR$_9$ is the main FxkR protein and *fxkR*$_9$ was deleted to produce strain OPS1808 (Δ*fxkR*$_9$). This mutant reproduced the reduced induction of *fnrN* under free-living microaerobic conditions observed in the double *hfixL* mutant (Fig 3). This finding supports the role of FxkR$_9$ as the mediator of hFixL O_2 regulation in Rlv3841, in agreement with studies in other rhizobia [44,45]. Studying the role of the hFixL-FxkR-FixK pathway in *fnrN* expression *in planta*, we observed that the double *hfixL* mutant reduced *fnrN* expression to 28% of WT levels (Fig 4A). This indicates the pathway also plays an important role in inducing *fnrN* during symbiosis.

We next studied the role of FnrN auto-regulation. A mutant of *fnrN* (LMB648, *fnrN*::ΩTet) had no effect on expression of *fnrN* at 1% O_2 (Fig 3), indicating FnrN auto-activation does not occur under free-living microaerobic conditions. By contrast, *in planta* the *fnrN* mutant reduced *fnrN* expression to 22% of WT levels (Fig 4A), similar to the reduction observed in the *hfixL* double mutant. FnrN auto-activation is thus an important regulatory effect during symbiosis but not under microaerobic conditions. During symbiosis, expression of *fnrN* is driven both by auto-activation and the hFixL-FxkR-FixK pathway, and both are required to attain full WT-level expression of the gene.

FnrN is critical for symbiotic gene expression but hFixL also plays an important role

To understand the respective importance of FnrN and hFixL as regulators of anaerobox controlled genes during symbiosis, their role in *fixNOQP* expression was studied. The *fixNOQP* operon encodes a high-affinity *cbb*$_3$-type terminal oxidase required for respiration during symbiosis [94–96]. It is typically regulated by an anaerobox [97]. Some rhizobia

encode multiple redundant terminal oxidases controlled by different regulators, but no alternatives appear to be encoded by Rlv3841 [98,99]. The strain therefore likely relies entirely on *fixNOQP* for respiration during symbiosis. Three putative homologs of *fixNOQP* exist in Rlv3841, which we labelled $fixNOQP_9$ (encoded on pRL9), $fixNOQP_{10}$ (encoded on pRL10) and $fixNOQP_c$ (encoded on the chromosome) [85]. Rlv VF39 encodes two *fixNOQP* operons, and either was able to sustain nitrogen fixation activity [56]. In Rlv3841, the plasmid-encoded operons are near-identical (>90% protein identity) but diverge from the $fixNOQP_c$ operon, with which they share approximately 50% identity. Only the plasmid-encoded $fixNOQP_9$ and $fixNOQP_{10}$ operons contain an upstream anaerobox, at -33 and -34 relative to their TSS respectively (Table 1). Past microarray work in our group found no significant upregulation of $fixN_c$ expression (1.7-fold upregulated, p-value 0.101) in 21 day old bacteroids compared to free-living Rlv3841 [93]. This contrasts sharply with $fixN_9$ (38.1-fold up, p = 0.010) and $fixN_{10}$ (119.6-fold up, p = 0.003), both highly upregulated. Taken together, these findings suggest the two plasmid-encoded *fixNOQP* operons of Rlv3841 are functional and anaerobox-regulated, whilst $fixNOQP_c$ is not. Rlv3841 also has two homologs of the *fixGHIS* operon (>90% protein identity), encoding the assembly machine for the *fixNOQP* terminal oxidase [51,100]. Like *fixNOQP*, *fixGHIS* operons are typically anaerobox regulated [47,62,101]. Both *fixGHIS* operons in Rlv3841 have a single upstream anaerobox, in a near-identical position to those of the *fixNOQP* operons [85] (Table 1). In line with findings in other rhizobia, the *fixGHIS* and *fixNOQP* operons are therefore likely regulated by oxygen in a similar way [89,102,103].

Both *fixNOQP* operons were induced in cultured cells under microaerobic conditions, confirming their regulation by O_2 (Fig 2). In culture, the double *hfixL* mutant severely reduced this microaerobic induction, resulting in minimal expression of $fixNOQP_{10}$ and 17% of WT $fixNOQP_9$ expression (Fig 3). The single $hfixL_9$ mutant significantly reduced expression of both operons whilst the $hfixL_c$ mutant only reduced expression of $fixNOQP_9$, to 71% of WT. These results indicate $hFixL_9$ is the dominant protein and $hFixL_c$ plays only a minor role in *fixNOQP* expression, matching their respective importance for microaerobic *fnrN* expression (Fig 3). In the Rlv3841 $fxkR_9$ mutant, expression of both *fixNOQP* operons was reduced to less than 25% of WT, indicating that the protein is required for hFixL regulation of *fixK* and hence *fixNOQP*, as found in other rhizobia [44]. The hFixL-FxkR-FixK pathway is thus a key regulator of *fixNOQP* expression under free-living microaerobic conditions. The remaining expression of $fixNOQP_9$ and $fixNOQP_{10}$ in the $fxkR_9$ mutant may be due to redundancy via the $fxkR_c$ homolog, or the result of background FixK or FnrN activity. By contrast, in these conditions the *fnrN* mutant minimally affected *fixNOQP* expression, with only $fixNOQP_9$ showing a small albeit statistically significant reduction (73% of WT). In line with our study of *fnrN* expression, the hFixL-FxkR-FixK pathway is crucial for *fixNOQP* expression under free-living microaerobic conditions whilst FnrN plays a minimal role. It is likely that the FnrN protein remains mostly inactive at the O_2 concentration (1% O_2) used in our free-living experiments.

We also checked the activity of NifA, a central activator of nitrogen fixation genes, in free-living microaerobic conditions [64,65,104]. NifA is O_2 sensitive and in most rhizobia is active only in the near-anoxic core of nodules [32,55,105,106]. Work in *E. meliloti* has however suggested the protein may already be active in the early stages of nodule development [107]. Recent work has suggested some rhizobial *nifA* variants can be active outside of the nodule [108,109]. We checked the NifA dependant induction of *nifH*, a component of the nitrogenase complex [70,73,110]. As expected, *nifH* expression did not increase under microaerobic conditions (Fig 2), indicating Rlv3841 NifA is not expressed or is inactive under these conditions.

Next, the role of FnrN and hFixL on expression of *fixNOQP in planta* during symbiosis was studied. We found that nodules formed by the *fnrN* mutant expressed both *fixNOQP* operons at only 5% of WT (Fig 4B and 4C). The FnrN sensor is thus critical for *fixNOQP* expression

inside the nodule. In nodules infected by the double *hfixL* mutant, expression of $fixNOQP_9$ and $fixNOQP_{10}$ was reduced to 68% and 58% of WT, respectively. Expression of $fixK_{9a}$ was abolished (S1 Fig), suggesting minimal FixK production in the absence of hFixL-FxkR TCS activity. Taken together, our results indicate that FnrN is critical for *fixNOQP* expression during symbiosis but the hFixL-FxkR-FixK pathway also plays a significant role.

To assess the impact of FnrN and the hFixL-FxkR-FixK pathway on symbiotic nitrogen fixation, acetylene reduction assays were performed on pea plants inoculated with O_2 regulation mutants. In line with its poor expression of *fixNOQP*, the *fnrN* mutant was critically impaired in nitrogen fixation, reducing acetylene at only 15% of the WT level (Fig 5A). Plants inoculated with this mutant produced only small and unelongated pale or brown nodules indicative of poor development and low leghaemoglobin production (Fig 5D). Thus, FnrN is critical for effective nitrogen fixation by Rlv3841. Complementation restored 88% of WT acetylene reduction activity and produced nodules indistinguishable from WT (S2 Fig).

Plants inoculated with either individual *hfixL* mutants or the double mutant were also impaired in nitrogen fixation but retained approximately 75% of WT acetylene reduction activity (Fig 5A). No morphological changes were observed in these nodules (Fig 5C). Thus, the hFixL-FxkR-FixK pathway is also an important contributor to symbiotic fixation activity and is required to attain a WT level of fixation. Complementation of the double *hfixL* mutant was attempted but the gene was found to be toxic in *E. coli* (see Materials and Methods for details). The $fxkR_9$ mutant impaired acetylene reduction rates but the decrease was insufficient to be significant ($p = 0.0584$). The $FxkR_c$ homolog is likely at least partially active and sufficiently produced to rescue hFixL regulation in the absence of $FxkR_9$. In the triple *fnrN* $hfixL_9$ $hfixL_c$ mutant (LMB673; *fnrN*::ΩTet $hfixL_9$::ΩSpec $hfixL_c$:pK19) only negligible levels of fixation were recorded. This reinforces the importance of the contribution from both the hFixL-FxkR-FixK pathway and FnrN, suggesting no additional regulators exist which induce these anaerobox controlled genes in Rlv3841 during symbiosis.

hFixL and FnrN are active in spatially distinct nodule zones during symbiosis

Legume nodules create a large internal O_2 gradient, with semi-aerobic conditions at their tip and near-anoxic conditions as low as 20 nM O_2 in the central nitrogen fixing zone [26,111]. This gradient is typically split into four zones (Fig 6A) containing different O_2 concentrations and rhizobia in different stages of differentiation (for reviews, see [26,31,112]). To understand how hFixL-FxkR-FixK and FnrN operate in this context, we used confocal microscopy to map the spatial expression of *fnrN* and *fixNOQP* in nodules.

Expression of *fnrN* in nodules infected with WT Rlv3841 (Fig 6B) was visible throughout all nodule zones. This included expression in infection threads in zone I, indicating that low O_2 induction of *fnrN* begins when Rlv3841 first enters the nodule and before the bacteria have differentiated into bacteroids. By contrast, *fnrN* expression in zone I was greatly reduced in nodules infected with the double *hfixL* mutant (Fig 6B). This suggests the O_2 concentration in the relatively aerobic environment of zone I is sufficiently low to activate the hFixL-FxkR-FixK pathway. In the absence of this pathway, some *fnrN* expression was retained in zone II and interzone II-III, but this was weaker than WT. Minimal *fnrN* expression was observed in zone III in the *hfixL* double mutant.

In the *fnrN* mutant, expression of *fnrN* appeared to be localized primarily in infection threads, around the entire periphery of the nodule (Fig 6B). Nodules infected by this mutant were severely impaired in their development, failed to elongate and contained little to no leghaemoglobin (Fig 5D). Free O_2 concentration is unlikely to drop as much in these nodules

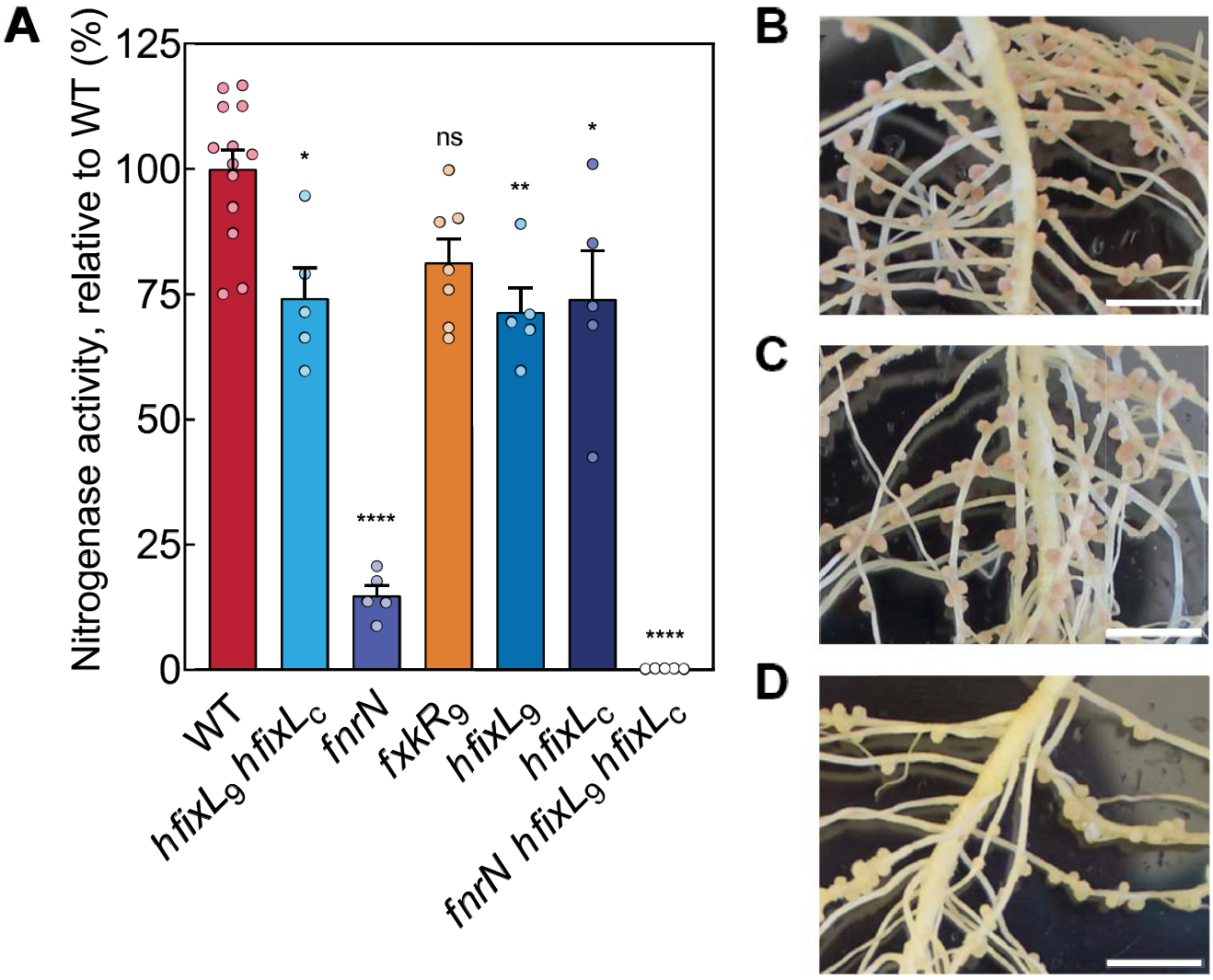


Fig 5. Effect of oxygen regulation mutants on nodule morphology and acetylene reduction rates. (A) Acetylene reduction rates of Rlv3841 mutant strains, normalised by WT activity (5.8 μmoles ethylene plant^{-1} hr^{-1}, 16.8×10^{-3} μmoles ethylene mg^{-1} of nodules hr^{-1}). Knocking out the *hfixL* genes individually (*hfixL*$_9$, LMB495; *hfixL*$_c$, LMB403) and in combination (*hfixL*$_9$ *hfixL*$_c$, LMB496) only slightly reduced fixation. The *fnrN* mutant (LMB648) critically reduced fixation. The single *fxkR*$_9$ mutant (OPS1808) did not significantly reduce fixation (p = 0.0584), possibly because of redundancy through the *fxkR*$_c$ homolog. The mutant lacking both FnrN and hFixL-FxkR-FixK function (LMB673) fixed at only a negligible rate. Rates are normalised per plant to total mass of nodules. Data are averages (±SEM) from at least five plants, ns (not significant) $P \geq 0.05$; $^{*}P < 0.05$; $^{**}P < 0.01$, $^{****}P < 0.0001$ by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons. Photos of nodules colonized by WT **(B)**, the double *hfixL* knockout **(C)** and the *fnrN* knockout **(D)**. Scale bar, 1 cm.

https://doi.org/10.1371/journal.pgen.1009099.g005

as it does in fully developed nodules. It is therefore noteworthy that the hFixL-FxkR-FixK pathway is nevertheless active, suggesting even poorly developed nodules produce a sufficiently low O_2 concentration to activate the pathway.

Expression patterns of *fixNOQP*$_9$ (Fig 7A) and *fixNOQP*$_{10}$ (Fig 7B) were similar. In WT Rlv3841, expression of both started abruptly in the II-III interzone of nodules, in agreement with past studies [55,56,113,114]. This abrupt start was absent in nodules infected with the double *hfixL* mutant, indicating it requires the hFixL-FxkR-FixK pathway (Fig 7C and 7D). Without hFixL-FxkR-FixK, expression of *fixNOQP*$_9$ and *fixNOQP*$_{10}$ started gradually after the II-III interzone, presumably driven by FnrN. Expression was also weaker than in the WT. In

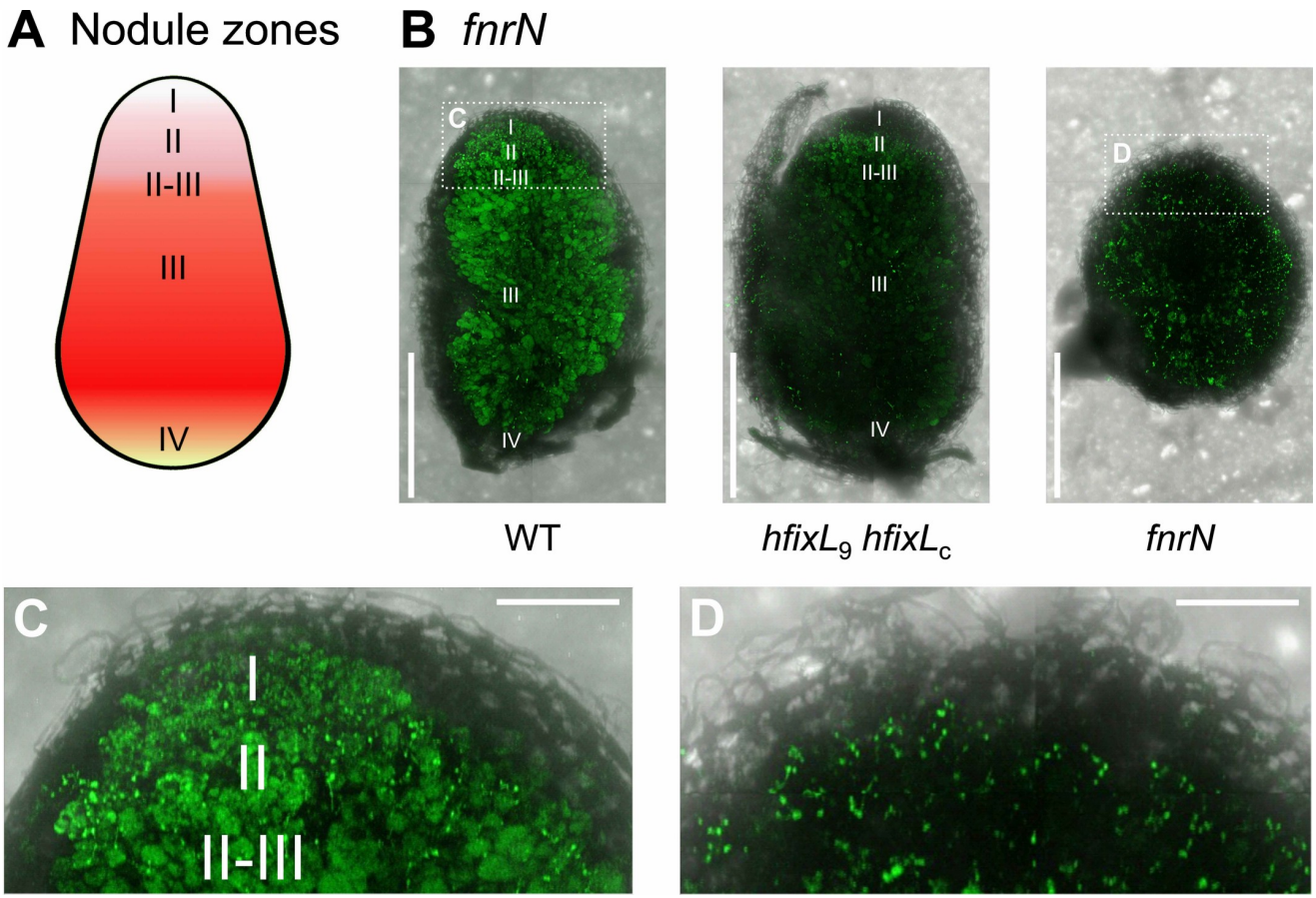


Fig 6. Spatial expression pattern of *fnrN* in nodules infected with Rlv3841 WT and mutants. (**A**) Schematic representation of an indeterminate nodule formed by *P. sativum*. Zone I contains undifferentiated rhizobia in infection threads. Rhizobia enter plant cells in zone II and in the II-III interzone undergo substantial differentiation towards becoming bacteroids. Zone III is the main nitrogen fixing zone. Zone IV contains bacteroids which are beginning to senesce. (**B**) Nodule cross sections showing expression of *fnrN* when inoculated with strains of Rlv3841 (Tn7 integrated *syfp2* promoter fusion: WT, OPS2429; $hfixL_9$ $hfixL_c$ double mutant, OPS2435; *fnrN* mutant, OPS2432). Expression begins immediately in zone I in nodules inoculated with WT; see C for a close-up of the region highlighted in white. A similar level of expression is present across all zones. When inoculated with the double *hfixL* mutant, expression began in zone II and was highest in this zone. In nodules inoculated with the *fnrN* mutant, expression was observed in infection threads around the periphery of the nodule; see D for a close up of the region highlighted in white. This mutant does not form mature nodules, and the normal zones are therefore unlikely to be fully developed. (**C**) Magnified view of *fnrN* expression in WT bacteria in zone I of the nodule. (**D**) Magnified view of *fnrN* expression in the infection threads of a nodule inoculated with the *fnrN* mutant. Scale bar; 1 mm (B), 0.25 mm (C and D). All images were captured and processed using identical parameters; see Materials and Methods for details.

https://doi.org/10.1371/journal.pgen.1009099.g006

the *fnrN* mutant, we observed minimal expression of $fixNOQP_9$. This may be due in part to the poor development of these nodules, but the expression of *fnrN* in the *fnrN* mutant (Fig 6B) indicates that the hFixL-FxkR-FixK pathway is still relatively active in these underdeveloped nodules. The lack of *fixNOQP* expression in the *fnrN* mutant therefore indicates that hFixL-FxkR-FixK cannot directly induce much *fixNOQP* expression in zone III of mature nodules. Although FnrN is the main driver of *fixNOQP* expression, the hFixL-FxkR-FixK pathway is required for full *fnrN* expression and also plays an important role, albeit indirectly.

Integration of FnrN and hFixL improves the O_2 response

To study the dynamics of an integrated cascade containing both hFixL and FnrN, we constructed an ordinary differential equation model of their combined pathway based on past

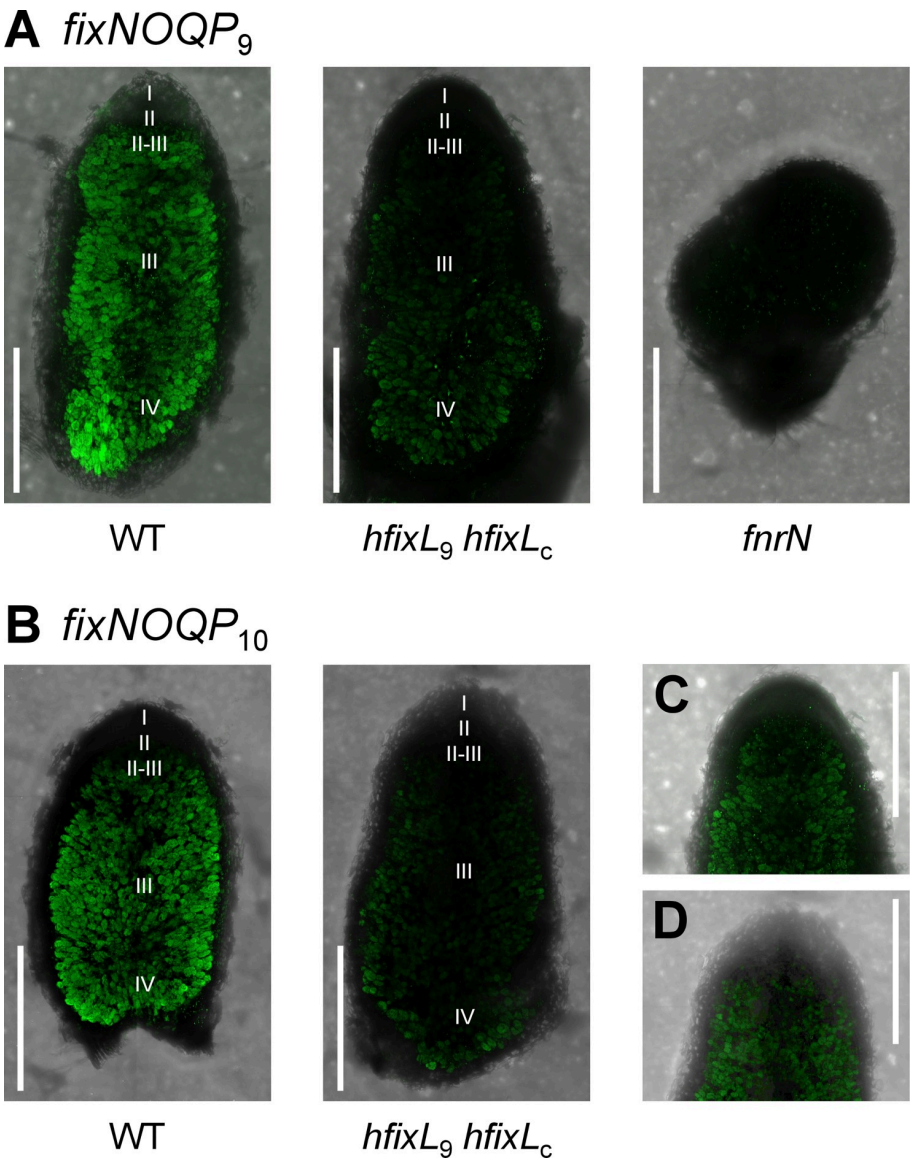


Fig 7. Spatial expression pattern of the *fixNOQP* operons in nodules infected with Rlv3841 WT and mutants. (**A**) Expression of $fixNOQP_9$ in strains of Rlv3841 (Tn7 integrated *syfp2* promoter fusion: WT, OPS2428; $hfixL_9$ $hfixL_c$ double mutant, OPS2434; *fnrN* mutant, OPS2431). In nodules inoculated with WT, expression starts abruptly at the II-III interzone. In the double *hfixL* mutant, expression is reduced and begins gradually and at a point more proximal to the root. Almost no expression is found in *fnrN* mutant nodules. (**B**) Expression of $fixNOQP_{10}$ in WT Rlv3841 and the double *hfixL* mutant followed a similar pattern as $fixNOQP_9$ (pJP2 reporter plasmid *syfp2* promoter fusion: WT, OPS2468; $hfixL_9$ $hfixL_c$, OPS2469). Expression begins at the II-III interzone. In the double *hfixL* mutant, expression again begins at a point more proximal to the root, in zone III of the nodule, and is reduced. (**C**) and (**D**) are areas of the *hfixL* double mutant reporter images (for $fixNOQP_9$ and $fixNOQP_{10}$ respectively) with their brightness and contrast altered to better display the distribution of fluorescence. Images within a set were captured and processed using identical parameters. Fluorescence intensity between the A and B image sets, and across C and D, should not be compared as intensity was normalised.

https://doi.org/10.1371/journal.pgen.1009099.g007

literature (S1 Text). A map of the regulatory connections incorporated in the model is given in S3 Fig. One each of hFixL, FxkR, FixK, FnrN and FixNOQP is considered in the model. hFixL was assumed to become active near a headspace concentration of 1% O_2 and FnrN near 0.01%

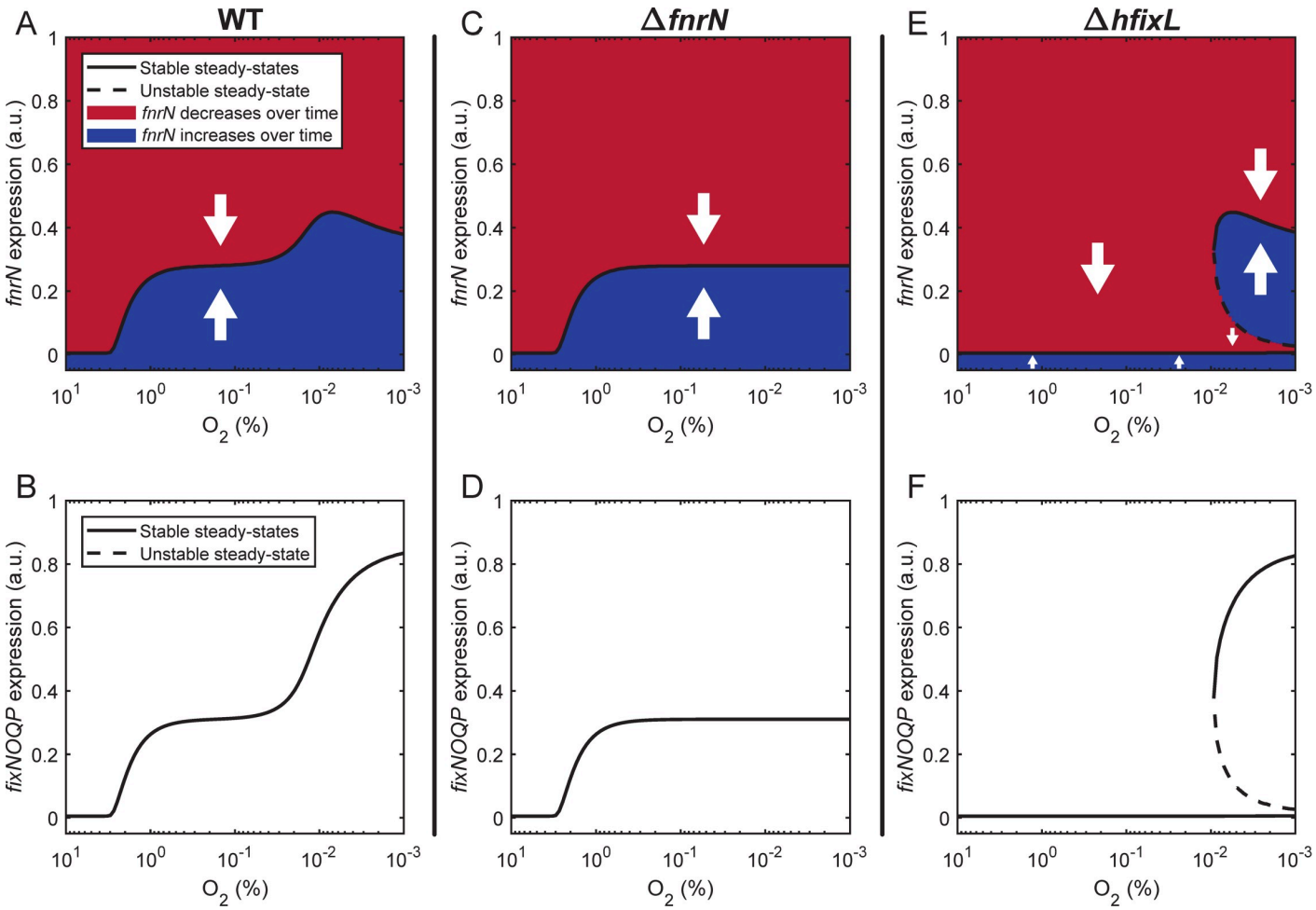


Fig 8. Modelling predicts the biphasic response of *fnrN* and *fixNOQP* expression controlled by a cascade integrating the hFixL and FnrN oxygen sensors. (**A, C, E**) Expression of *fnrN*. (**B, D, F**) Expression of *fixNOQP*. Black lines indicate stable expression steady-states, dashed lines indicate unstable steady states. In cells whose state is in a red shaded area, expression is expected to decrease; in a blue shaded area, it is expected to increase (white arrows indicate expected direction of change). (**A**) In a WT system with both sensors, expression of *fnrN* initially begins under microaerobic conditions. Expression increases slightly due to FnrN auto-activation as O_2 concentration drops, but is stabilized by auto-repression. (**B**) Expression of *fixNOQP* begins under microaerobic conditions, then increases when FnrN becomes active as O_2 drops further. (**C, D**) In the absence of FnrN, expression of *fnrN* and *fixNOQP* begins under microaerobic conditions driven by the hFixL-FxkR-FixK pathway, but does not subsequently increase. (**E, F**) In cells where only the FnrN system is present, expression of the *fnrN* gene and *fixNOQP* operon does not occur until O_2 concentration drops sufficiently for FnrN to be active. Once FnrN is active, the system exhibits bistability, with stable states of both near-zero and high levels of *fnrN* and *fixNOQP* expression. As the O_2 concentration continues to drop, an increasing proportion of cells are expected to transition to the high expression state due to stochastic variations in expression.

https://doi.org/10.1371/journal.pgen.1009099.g008

O_2, corresponding to dissolved O_2 concentrations of ~12 μM and ~120 nM respectively at equilibrium. Past studies have suggested that hFixL binds O_2 cooperatively, and that FnrN and FixK bind DNA as dimers [87,115–118]. Consequently all three of these binding processes were modelled using Hill functions with a Hill coefficient of 2 [119,120]. FixK and FnrN were assumed to have an identical induction effect on transcription when bound to the anaerobox motif. However, based on the critical role that FnrN but not FixK plays in regulation, FnrN was assumed to have a greater binding affinity for anaerobox motifs than FixK.

Our model reproduced the biphasic response of the integrated hFixL-FnrN cascade observed in WT Rlv3841 as O_2 dropped from atmospheric (~21%) to near-anoxic concentrations (0.001%) (Fig 8A and 8B). Expression of *fnrN* and *fixNOQP* first began in microaerobic

conditions (10–1% O_2) under the action of the hFixL-FxkR-FixK pathway. Subsequent activation of FnrN around 0.01% O_2 led to a further increase in *fixNOQP* expression, mirroring our findings *in planta*. Expression of *fnrN* near 0.01% O_2 initially increased due to auto-activation. Subsequently, as the O_2 concentration continued to drop, *fnrN* expression decreased due to auto-repression. Thus, the model correctly predicts a generally homogenous level of *fnrN* expression throughout the nodule, and increased *fixNOQP* expression in the core of the nodule relative to the tip.

In the absence of FnrN (Fig 8C and 8D), the model shows initial expression of *fnrN* and *fixNOQP* under microaerobic conditions due to the hFixL-FxkR-FixK pathway. There is however no further induction as O_2 continues to drop. This agrees with our finding that some *fnrN* expression takes place in the Rlv3841 *fnrN* mutant, albeit at a reduced level relative to WT. The model also predicts a lower level of *fixNOQP* expression in the *fnrN* mutant, consistent with our confocal microscopy results for $fixNOQP_9$ expression (Fig 7A).

In the absence of *hfixL* (Fig 8E and 8F), an important new behaviour of the pathway is predicted by our model. As expected, no induction of *fnrN* or *fixNOQP* takes place under microaerobic conditions, in line with our experimental findings. However, once O_2 drops below 0.01%, our model suggests that the pathway may be bistable, with possible steady states at either high *fnrN* expression or near-zero expression. As the O_2 concentration continues to decrease, the disturbance needed to move from minimal expression to the high expression steady state becomes smaller. Thus, as the bacteria experience increasingly anaerobic conditions moving to areas of the nodule more proximal to the root, the model predicts that an increasing proportion of cells will transition from near-zero to high *fnrN* expression due to stochastic variations in expression of the gene. This agrees with the gradual increase in *fixNOQP* expression observed in the *hfixL* double mutant in Rlv3841 (Fig 7A and 7B).

Discussion

O_2 regulation is essential for rhizobia to establish a successful symbiosis with their legume partners. The model *Rhizobium* Rlv3841 employs three O_2 sensors in symbiosis: hFixL, FnrN and NifA. In the present study, we examined this multi-sensor arrangement through a combination of *in vitro*, *in planta* and *in silico* approaches. The hFixL-FxkR-FixK pathway is active in the earliest stages of symbiosis, followed by FnrN as the bacteria move to areas of the nodule more proximal to the root. Both regulate genes required for symbiotic survival, such as *fixNOQP*. NifA is active at a later stage, in zone III of nodules, and regulates activation of core nitrogen fixation machinery. The hFixL-FxkR-FixK pathway is the most O_2 tolerant of the three sensors, active in free-living bacteria under microaerobic conditions and *in planta* beginning in zone I of nodules. The FnrN protein is inactive under free-living microaerobic conditions and only becomes active from the II-III interzone onwards. FnrN is critical for expression of *fixNOQP* and nitrogen fixation activity. Indirectly, the hFixL-FxkR-FixK pathway also plays an important role by inducing *fnrN* expression under microaerobic conditions, priming it for auto-activation in the central nitrogen fixing zone. Our modelling results suggest the hFixL-FxkR-FixK pathway also prevents bistability in the low O_2 response, thereby ensuring all cells commit to *fixNOQP* expression in the central nitrogen fixing zone. Thus hFixL-FxkR-FixK and FnrN act as a single regulation pathway which integrates both O_2 sensors.

Rhizobia experience a drop in O_2 concentration of at least three orders of magnitude as they transition from a free-living lifestyle in soil to terminally differentiated bacteroids in nodules. Like Rlv3841, it is common for other rhizobia to employ multiple O_2 sensors during this transition. These multiple sensors may be used to create redundancy, a feature often found in

key regulatory pathways to improve their robustness. Elements of this redundancy are present in Rlv3841, including the multiple hFixL homologs and the overlap between their role and that of FnrN. However, our results also demonstrate that each sensor plays an important distinct role. Thus, integrating sensors into a single cascade in Rlv3841 also improves the responsiveness of regulation and allows the bacteria to respond appropriately across the entire range of O_2 concentrations experienced during symbiosis. A similar dual-sensor arrangement has also previously been described in *Rhodopseudomonas palustris*, which combines a FixLJ-FixK pathway with the FnrN homolog AadR [121–123]. *R. palustris* is not symbiotic but is noted for its ability to grow under both aerobic and anaerobic conditions [124]. The combined pathway in *R. palustris* was shown to provide fine-tuned regulation for adapting to the large range of O_2 concentrations it experiences.

The prevalence of multi-sensor O_2 regulation arrangements in rhizobia may also have arisen in response to competitive fitness pressures. Legume plants can sanction rhizobia based on their nitrogen fixation activity [125–127]. We speculate the bacteria may also be selected based on the speed with which they are able to adapt to life inside nodules and begin productively fixing nitrogen. This would create pressure for strains to rapidly demonstrate their effectiveness to their legume host. Past work has suggested that one of the benefits of FnrN compared to FixLJ-FixK is that it is more responsive to O_2 concentration, providing more flexible regulation [86,87,128]. By enabling more fine-tuned control, integrated multi-sensor O_2 regulatory pathways may speed up the symbiotic transition, providing a competitive advantage.

Materials and methods

Bacterial strains and growth conditions

E. coli strains were grown in liquid or solid LB medium[129] at 37˚C supplemented with appropriate antibiotics (μg mL^{-1}): ampicillin 100, kanamycin 20, spectinomycin 50 and gentamicin 10. Rlv3841 strains were grown at 28˚C in Tryptone-Yeast (TY) extract[130] or Universal Minimal Salts (UMS)[131] with glucose and ammonium chloride at 10 mM each. Antibiotics for Rlv3841 were used at the following concentrations (μg mL^{-1}): gentamicin 20, kanamycin 50, spectinomycin 100, streptomycin 500, tetracycline 2, neomycin 80 and nitrofurantoin 20. A list of the strains used is given in S2 Text.

Cloning, colony PCRs and conjugations

All routine DNA analyses were done using standard protocols [129]. PCR reactions for cloning were carried out according to the manufacturer's instructions with Q5 High-Fidelity DNA Polymerase (New England Biolabs). Colony PCRs used OneTaq DNA Polymerase (NEB). Restriction enzymes (NEB) were used according to the manufacturer's instructions. Sanger sequencing was carried out by Eurofins Genomics. Assemblies using BD In-Fusion cloning (Takara Bio) were performed according to the manufacturer's instructions. Triparental conjugations, and transductions with bacteriophage RL38, were performed as previously described [132,133]. Tn7 integrations were performed according to the method described by Choi and colleagues [134,135]. A list of the plasmids and primers used is given in S2 Text.

Mutant generation and complementation

Rlv3841 *hfixL*$_c$ (RL1879) mutant, LMB403. A 1 Kb internal fragment of *hfixL*$_c$ was PCR amplified from Rlv3841 with primers pr0988/0989, adding XbaI sites at the 5' and 3' ends. This fragment was cloned into pK19mob digested with XbaI, using BD In-Fusion cloning, to

produce plasmid pLMB441. Triparental filter conjugation of pLMB441 into WT Rlv3841 was then performed using kanamycin selection. Colonies were screened by colony PCR using primers pr0482 and pK19A, which bind upstream of $hfixL_c$ and inside the integrated pK19 backbone respectively. This gave mutant strain LMB403.

Rlv3841 $hfixL_9$ (pRL90020) mutant, LMB495. A 1 Kb region containing $hfixL_9$ was PCR amplified from Rlv3841 with primers pr1270/1271. This region was subcloned into pJET1.2/blunt to produce plasmid pLMB581. Plasmid pLMB581 was then digested with XbaI/XhoI and the $hfixL_9$ region cloned into pJQ200SK using BD In-Fusion, digested with the same enzymes, to give plasmid pLMB585. A spectinomycin resistance cassette was digested out of the pHP45ΩSpc plasmid with SmaI and cloned into pLMB585 at a unique StuI site blunted using the Klenow fragment to give plasmid pLMB590. Triparental filter conjugation of pLMB590 into WT Rlv3841 was then performed using spectinomycin selection. Colonies were screened by colony PCR using primers pr1272/1273. This gave mutant strain LMB495.

Rlv3841 double $hfixL_c$ $hfixL_9$ mutant, LMB496. An Rlv3841 mutant in both *hfixL* genes was generated by triparental filter conjugation of pLMB441 into strain LMB495, producing double mutant strain LMB496.

Rlv3841 *fnrN* (RL2818) mutant, LMB648. A 2.5 Kb region containing *fnrN* was PCR amplified from Rlv3841 with primers pr1381/1382. This fragment was digested with XbaI/XhoI and cloned using BD In-Fusion into pJQ200SK linearized with digestion by the same enzymes to make plasmid pLMB732. A tetracycline resistance cassette was then digested out of the pHP45ΩTet plasmid with EcoRI and cloned into pLMB732 at a unique MfeI site to give plasmid pLMB733. Triparental filter conjugation of pLMB733 into WT Rlv3841 was then performed using tetracycline selection. Colonies were screened by colony PCR using primers pr1432/1433. This gave mutant strain LMB648.

Rlv3841 triple $hfixL_c$ $hfixL_9$ *fnrN* mutant, LMB673. A triple Rlv3841 mutant, in both *hfixL* genes and the *fnrN* gene, was generated by transducing *fnrN*::ΩTet from LMB648 into LMB496 to produce strain LMB673.

Rlv3841 $fxkR_9$ (pRL90026) mutant, OPS1808. Two 1 Kb regions, one upstream and one downstream of $fxkR_9$, were PCR amplified from Rlv3841 with primer pairs oxp2874/2875 and oxp2876/2877 respectively. These were cloned with BD In-Fusion into pK19mobSacB digested with PstI and EcoRI to produce plasmid pOPS1199. Triparental filter conjugation of pOPS1199 into WT Rlv3841 was then performed using kanamycin selection. Colonies were screened by colony PCR using primers oxp3155 and pK19A. Colonies with correct integration were subsequently subjected to sucrose selection to remove plasmid pK19mobSacB as previously described [136]. Colonies were then screened for loss of kanamycin resistance and using colony PCR with primers oxp3155/3156 to isolate mutant strain OPS1808.

Complemented Rlv3841 *fnrN* mutant (OPS2260). The *fnrN* gene with its native promoter was amplified from Rlv3841 with primers oxp4115/4116 and cloned into BsaI-digested pOGG280 using BD In-Fusion. This plasmid was then genomically integrated with kanamycin selection and colonies screened with primers oxp2327/2328 and confirmed with sequencing. This produced strain OPS2260.

Attempts at complementing the Rlv3841 $hfixL_9$ mutant. Our $hfixL_9$ (LMB495) mutant showed the largest phenotypic effect out of the two single *hfixL* mutants, and we therefore attempted to complement this strain. We first attempted complementation via Tn7 integration using the pOGG280 backbone in which $hfixL_9$ was under P*lac* control. This construct assembled in *E. coli* but could not be successfully conjugated into Rlv3841. We theorized the protein was being produced to a toxic level in Rlv3841 due to poor LacI repression. We next attempted to rectify this problem by driving $hfixL_9$ from the native $fixK_9$ promoter and RBS instead ($fixK_9$ and $hfixL_9$ likely form an operon). However, this construct could not be transformed

into *E. coli*, suggesting P*fixK*$_{9a}$ was causing toxic levels of *hfixL*$_9$ production. Finally, we sought to strike a balance between these two approaches by using the P*lac* promoter but the native *hfixL*$_9$ RBS, in a pOGG250 backbone. This construct assembled in *E. coli* and could be conjugated into Rlv3841 but failed to complement the mutant. It is likely that this promoter-RBS combination avoided toxicity by reducing *hfixL*$_9$ production but produced insufficient protein to achieve complementation.

Microaerobic induction measurements in cultured cells

Rlv3841 strains were first grown on TY slopes with appropriate antibiotics for three days. Cells were resuspended and washed three times by centrifugation at 5,000 RCF for 10 minutes. Washed cells were used to inoculate 10 mL liquid UMS cultures to OD_{600} 0.01 and grown overnight without antibiotics. Cultures were then diluted to OD 0.1 in 400 μL UMS per well in a 24-well microtiter plate (4titude). A gas-permeable membrane (4titude) was applied to microtiter plates. Plates were then incubated in a FLUOstar Omega plate reader equipped with an Atmospheric Control Unit (both produced by BMG) to adjust O_2 concentration to 1% and CO_2 concentration to 0.1%. Readings were taken every 30 minutes and plates shaken at 700 rpm in double orbital mode between readings. Induction was measured at 18 hrs post-inoculation, when all cultures had reached stationary phase.

Plate-based measurements

sYFP2 measurements were made on a BMG FLUOstar Omega plate reader using the bottom optic with a gain setting of 2,000 and orbital averaging enabled (53 readings, 6 mm radius). Measurements were filter-based with excitation at 520 nm and emission recorded at 540 nm. Luminescence measurements were made on a Promega GloMax plate reader using the manufacturer's protocol. Luminescence was used for S1 Fig as this reporter construct was already available.

Plant growth and acetylene reduction

Pisum sativum cv. Avola seeds were surface sterilized using 95% ethanol and 2% sodium hypochlorite before sowing. Plants were inoculated with 1×10^7 cells of the appropriate rhizobial strain and grown in 1 L beakers filled with sterile medium-grade vermiculite and nitrogen-free nutrient solution as previously described in a growth room (16h light/8h dark) [137]. Harvesting was 21 days later and acetylene reduction rate was determined as previously described [138]. All nodules for each plant were counted and their combined mass weighed; acetylene reduction rates were normalised by total nodule weight.

Bacteroid isolation

Bacteroids were isolated from root nodules after 21 days of plant growth following a differential spin protocol adapted from Tsukada et al. 2009 [139]. Approximately 100 mg of nodules were picked per plant. Nodules were immersed in 1 mL of sterile isolation buffer (1 M K_2HPO_4, 1 M KH_2PO_4, 300 mM sucrose, 2 mM $MgCl_2$) and macerated. The mixture was spun down at 200 RCF for 5 minutes to remove plant debris. The supernatant was transferred to a fresh tube and spun down at 3,500 RCF for 5 minutes. The supernatant from this second spin was discarded and the pelleted fraction, containing the isolated bacteroids, was resuspended in isolation buffer and used for microtiter plate measurements.

Transcription start site mapping

The TSS data set referenced in this paper can be found in full on the NCBI SRA database, BioProject number PRJNA667846. A publication discussing the data in full is forthcoming. Protocol details can be found in S3 Text.

Statistical analysis

All analyses were performed using GraphPad Prism 8 (GraphPad Software). Significant differences were determined by Student's t-test or one-way ANOVA followed by Dunnett's multiple comparisons post-hoc test correction. A p-value less than 0.05 was considered statistically significant.

Confocal microscopy

Reporters were constructed by transcriptional fusion of promoters to an ORF of the sYFP2 fluorescent protein. Reporters were subsequently genomically integrated into Rlv3841 strains using the mini-Tn7 system [134]. Plants were inoculated with marked strains and grown as described above. After 21 days, nodules were picked and immersed in water then cut in half longitudinally. Images were taken with an LSM 880 confocal laser-scanning microscope equipped with the Axio Imager.Z2 (Zeiss), using the manufacturer's ZEN Black software. A Plan-Apochromat 10×/0.45 M27 objective (Zeiss) was used. Excitation was at 514 nm with an Argon laser and emission measurements filtered to a range of 519–572 nm. Acquisitions were tile scans with 2×3 tiles per image. 31 Z-stack slices were taken for each tile, separated by a height of 10 μm. Images shown in this publication are maximum-intensity orthogonal projections produced with the ZEN Blue software (Zeiss).

Figure data

Data for main text Figs 2–5 and S1 and S2 Figs are given in S4 Text. Data for main text Fig 8 are given in S1 Spreadsheet.

Supporting information

S1 Fig. hFixL is required for *in planta fixK*$_{9a}$ expression in Rlv3841. A reporter (pOPS0136) was built with the *luxCDABE* reporter operon fused to the *fixK*$_{9a}$ promoter. The promoter was active in isolated WT Rlv3841 bacteroids (OPS0376), but no luminescence above no-reporter background was recorded in double *hfixL* mutant bacteroids (OPS0528). Data are averages (±SEM) from at least four plants, $^{*}P < 0.05$; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.
(EPS)

S2 Fig. Complementation of the Rlv3841 *fnrN* mutant. (A) Acetylene reduction rates; the activity of the *fnrN* mutant (LMB648) was 20% of WT Rlv3841. The complemented strain (OPS2260) fixed at 88% of WT. Nodules colonized by Rlv3841 **(B)** WT, **(C)** the complemented *fnrN* mutant and **(D)** the *fnrN* mutant. Acetylene reduction rates are normalised to total weight of nodules per plant. Data are averages (±SEM) from seven plants, ns (not significant) $P \geq 0.05$; $^{****}P < 0.0001$; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons. Complementation also restored nodule morphology.
(EPS)

S3 Fig. Simplified map of the Rlv3841 dual sensor oxygen cascade used for modelling. Only one copy each is included of *hfixL*, *fixK*, *fxkR* and *fixNOQP*. Both FnrN and FixK can

positively and negatively regulate expression of *fnrN*. FxkR negative auto-regulation is not included in the model. Regulation is indicated with lines ending in arrows (positive regulation) and ending in blunt ends (negative regulation). Translation is shown as lines ending in circles.
(EPS)

S1 Text. Modelling oxygen regulation in Rlv3841.
(PDF)

S2 Text. Strains, plasmids and primers used in the study.
(PDF)

S3 Text. Further materials and methods details for transcription start site mapping.
(PDF)

S4 Text. Data tables for main text Figs 2–5 and S1 and S2 Figs.
(PDF)

S1 Spreadsheet. Data table for main text Fig 8.
(XLSX)

Acknowledgments

The authors would like to thank Dr Tim Haskett, Dr Carmen Sánchez-Cañizares and Prof Lee Sweetlove for their advice and critically reviewing the manuscript. They would also like to thank Dr Niloufer Irani for her help with confocal microscopy, and Dr Beatriz Jorrín for her help with the Tn7-based reporters used in this study.

Author Contributions

Conceptualization: Paul J. Rutten, Harrison Steel, Graham A. Hood, Antonis Papachristodoulou, Philip S. Poole.

Data curation: Paul J. Rutten, Harrison Steel, Graham A. Hood, Vinoy K. Ramachandran.

Formal analysis: Paul J. Rutten, Harrison Steel, Vinoy K. Ramachandran.

Funding acquisition: Antonis Papachristodoulou, Philip S. Poole.

Investigation: Paul J. Rutten, Graham A. Hood, Vinoy K. Ramachandran, Lucie McMurtry, Barney Geddes.

Methodology: Paul J. Rutten, Harrison Steel, Graham A. Hood, Vinoy K. Ramachandran, Antonis Papachristodoulou, Philip S. Poole.

Project administration: Paul J. Rutten, Antonis Papachristodoulou, Philip S. Poole.

Resources: Harrison Steel, Graham A. Hood, Barney Geddes, Antonis Papachristodoulou, Philip S. Poole.

Software: Harrison Steel, Antonis Papachristodoulou.

Supervision: Barney Geddes, Antonis Papachristodoulou, Philip S. Poole.

Validation: Paul J. Rutten, Vinoy K. Ramachandran, Lucie McMurtry, Barney Geddes.

Visualization: Paul J. Rutten, Harrison Steel.

Writing – original draft: Paul J. Rutten, Harrison Steel.

Writing – review & editing: Paul J. Rutten, Harrison Steel, Graham A. Hood, Vinoy K. Ramachandran, Lucie McMurtry, Barney Geddes, Antonis Papachristodoulou, Philip S. Poole.

References

1. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. Nat Rev Microbiol. 2018; 16: 291–303. https://doi.org/10.1038/nrmicro.2017.171 PMID: 29379215
2. Mylona P, Pawlowski K, Bisseling T. Symbiotic nitrogen fixation. Plant Cell. 1995; 7: 869–885. https://doi.org/10.1105/tpc.7.7.869 PMID: 12242391
3. Oldroyd GED, Dixon R. Biotechnological solutions to the nitrogen problem. Curr Opin Biotechnol. 2014; 26: 19–24. https://doi.org/10.1016/j.copbio.2013.08.006 PMID: 24679253
4. Seefeldt LC, Hoffman BM, Dean DR. Mechanism of Mo-dependent nitrogenase. Annu Rev Biochem. 2009; 78: 701–722. https://doi.org/10.1146/annurev.biochem.78.070907.103812 PMID: 19489731
5. Leister D. Thawing out frozen metabolic accidents. BMC Biol. 2019; 17: 8. https://doi.org/10.1186/s12915-018-0621-5 PMID: 30700284
6. Layzell DB, Diaz del Castillo L, Hunt S, Kuzma M, Van Cauwenberghe O, Oresnik I. The regulation of oxygen and its role in regulating nodule metabolism. Dordrecht: Springer; 1993. pp. 393–398. https://doi.org/10.1007/978-94-017-2416-6_39
7. Millar AH, Day DA, Bergersen FJ. Microaerobic respiration and oxidative phosphorylation by soybean nodule mitochondria: implications for nitrogen fixation. Plant, Cell Environ. 1995; 18: 715–726. https://doi.org/10.1111/j.1365-3040.1995.tb00574.x
8. Talbi C, Sanchez C, Hidalgo-Garcia A, Gonzalez EM, Arrese-Igor C, Girard L, et al. Enhanced expression of *Rhizobium etli cbb*3 oxidase improves drought tolerance of common bean symbiotic nitrogen fixation. J Exp Bot. 2012; 63: 5035–5043. https://doi.org/10.1093/jxb/ers101 PMID: 22511804
9. Marchal K, Vanderleyden J. The "oxygen paradox" of dinitrogen-fixing bacteria. Biol Fertil Soils. 2000; 30: 363–373. https://doi.org/10.1007/s003740050017
10. Schulze J. How are nitrogen fixation rates regulated in legumes? J Plant Nutr Soil Sci. 2004; 167: 125–137. https://doi.org/10.1002/jpln.200320358
11. Rutten PJ, Poole PS. Oxygen regulatory mechanisms of nitrogen fixation in rhizobia. In: Poole RK, editor. Advances in Microbial Physiology. Academic Press; 2019. pp. 325–389. https://doi.org/10.1016/bs.ampbs.2019.08.001 PMID: 31655741
12. Masson-Boivin C, Giraud E, Perret X, Batut J. Establishing nitrogen-fixing symbiosis with legumes: how many *rhizobium* recipes? Trends Microbiol. 2009; 17: 458–466. https://doi.org/10.1016/j.tim.2009.07.004 PMID: 19766492
13. Gage DJ, Margolin W. Hanging by a thread: Invasion of legume plants by rhizobia. Curr Opin Microbiol. 2000; 3: 613–617. https://doi.org/10.1016/s1369-5274(00)00149-1 PMID: 11121782
14. Ferguson BJ, Mens C, Hastwell AH, Zhang M, Su H, Jones CH, et al. Legume nodulation: the host controls the party. Plant Cell Environ. 2019; 42: 41–51. https://doi.org/10.1111/pce.13348 PMID: 29808564
15. Martin FM, Uroz S, Barker DG. Ancestral alliances: plant mutualistic symbioses with fungi and bacteria. Science (80-). 2017; 356: 1–9. https://doi.org/10.1126/science.aad4501 PMID: 28546156
16. Witty JF, Minchin FR. Oxygen diffusion in the legume root nodule. Nitrogen Fixation. Boston: Springer; 1990. pp. 285–292. https://doi.org/10.1007/978-1-4684-6432-0_29
17. Oldroyd GED, Downie JA. Coordinating nodule morphogenesis with rhizobial infection in legumes. Annu Rev Plant Biol. 2008; 59: 519–546. https://doi.org/10.1146/annurev.arplant.59.032607.092839 PMID: 18444906
18. Downie JA. Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules. Curr Biol. 2005; 15: R196–R198. https://doi.org/10.1016/j.cub.2005.03.007 PMID: 15797009
19. Kawashima K, Suganuma N, Tamaoki M, Kouchi H. Two types of pea leghemoglobin genes showing different O2-binding affinities and distinct patterns of spatial expression in nodules. Plant Physiol. 2001; 125: 641–51. https://doi.org/10.1104/pp.125.2.641 PMID: 11161022
20. Bergersen FJ. Delivery of O2 to bacteroids in soybean nodule cells: consideration of gradients of concentration of free, dissolved O2 in and near symbiosomes and beneath intercellular spaces. Protoplasma. 1996; 191: 9–20. https://doi.org/10.1007/BF01280821
21. Appleby CA. Leghemoglobin and *Rhizobium* respiration. Annu Rev Plant Physiol. 1984; 35: 443–478. https://doi.org/10.1146/annurev.pp.35.060184.002303

22. King BJ, Hunt S, Weagle GE, Walsh KB, Pottier RH, Canvin DT, et al. Regulation of O2 concentration in soybean nodules observed by *in situ* spectroscopic measurement of leghemoglobin oxygenation. Plant Physiol. 1988; 87: 296–9. https://doi.org/10.1104/pp.87.2.296 PMID: 16666136

23. Minchin FR. Regulation of oxygen diffusion in legume nodules. Soil Biol Biochem. 1997; 29: 88–89. https://doi.org/10.1016/S0038-0717(96)00204-0

24. Downie JA. Legume nodulation. Curr Biol. 2014; 24: R184–R190. https://doi.org/10.1016/j.cub.2014.01.028 PMID: 24602880

25. Prell J, White JP, Bourdes A, Bunnewell S, Bongaerts RJ, Poole PS. Legumes regulate *Rhizobium* bacteroid development and persistence by the supply of branched-chain amino acids. Proc Natl Acad Sci. 2009; 106: 12477–12482. https://doi.org/10.1073/pnas.0903653106 PMID: 19597156

26. Vasse J, De Billy F, Camut S, Truchet G. Correlation between ultrastructural differentiation of bacteriods and nitrogen fixation in alfalfa nodules. J Bacteriol. 1990; 172: 4295–4306. https://doi.org/10.1128/jb.172.8.4295-4306.1990 PMID: 2376562

27. Long SR. *Rhizobium*-legume nodulation: life together in the underground. Cell. 1989; 56: 203–214. https://doi.org/10.1016/0092-8674(89)90893-3 PMID: 2643474

28. Van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, et al. Plant peptides govern terminal differentiation of bacteria in symbiosis. Science (80-). 2010; 327: 1122–1126. https://doi.org/10.1126/science.1184057 PMID: 20185722

29. Mergaert P, Uchiumi T, Alunni B, Evanno G, Cheron A, Catrice O, et al. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. Proc Natl Acad Sci. 2006; 103: 5230–5235. https://doi.org/10.1073/pnas.0600912103 PMID: 16547129

30. Łotocka B, Kopcińska J, Skalniak M. Review article: the meristem in indeterminate root nodules of Faboideae. Symbiosis. 2012; 58: 63–72. https://doi.org/10.1007/s13199-013-0225-3 PMID: 23482442

31. Popp C, Ott T. Regulation of signal transduction and bacterial infection during root nodule symbiosis. Curr Opin Plant Biol. 2011; 14: 458–467. https://doi.org/10.1016/j.pbi.2011.03.016 PMID: 21489860

32. Gavrin A, Kaiser BN, Geiger D, Tyerman SD, Wen Z, Bisseling T, et al. Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalization, HOPS suppression, and TIP1g retargeting in *Medicago*. Plant Cell. 2014; 26: 3809–3822. https://doi.org/10.1105/tpc.114.128736 PMID: 25217511

33. Layzell DB, Hunt S. Oxygen and the regulation of nitrogen fixation in legume nodules. Physiol Plant. 1990; 80: 322–327. https://doi.org/10.1111/j.1399-3054.1990.tb04414.x

34. Neo HH, Layzell DB. Phloem glutamine and the regulation of O2 diffusion in legume nodules. Plant Physiol. 1997; 113: 259–267. https://doi.org/10.1104/pp.113.1.259 PMID: 12223605

35. Oresnik IJ, Atkins CA, Layzell DB. The legume symbiosis: C-limited bacteria living within O2 limited plant cells? In: Tikhonovich IA, Provorov NA, Romanov VI, Newton WE, editors. Nitrogen fixation: fundamentals and applications Proceed 10th Intl Cong Nitrogen Fix, St Petersburg. Dordrecht: Springer Netherlands; 1995. p. 601. https://doi.org/10.1007/978-94-011-0379-4

36. Batut J, Boistard P. Oxygen control in *Rhizobium*. Antonie Van Leeuwenhoek. 1994; 66: 129–150. https://doi.org/10.1007/BF00871636 PMID: 7747928

37. West AH, Stock AM. Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem Sci. 2001; 26: 369–376. https://doi.org/10.1016/s0968-0004(01)01852-7 PMID: 11406410

38. Dixon R, Kahn D. Genetic regulation of biological nitrogen fixation. Nat Rev Microbiol. 2004; 2: 621–631. https://doi.org/10.1038/nrmicro954 PMID: 15263897

39. Da Re S, Schumacher J, Rousseau P, Fourment J, Ebel C, Kahn D. Phosphorylation-induced dimerization of the FixJ receiver domain. Mol Microbiol. 1999; 34: 504–511. https://doi.org/10.1046/j.1365-2958.1999.01614.x PMID: 10564492

40. Birck C, Mourey L, Gouet P, Fabry B, Schumacher J, Rousseau P, et al. Conformational changes induced by phosphorylation of the FixJ receiver domain. Structure. 1999; 7: 1505–1515. https://doi.org/10.1016/s0969-2126(00)88341-0 PMID: 10647181

41. Wright GSA, Saeki A, Hikima T, Nishizono Y, Hisano T, Kamaya M, et al. Architecture of the complete oxygen-sensing FixL-FixJ two-component signal transduction system. Sci Signal. 2018; 11: 1–12. https://doi.org/10.1126/scisignal.aaq0825 PMID: 29636388

42. Mesa S, Reutimann L, Fischer H-M, Hennecke H. Posttranslational control of transcription factor FixK2, a key regulator for the *Bradyrhizobium japonicum*-soybean symbiosis. Proc Natl Acad Sci. 2009; 106: 21860–21865. https://doi.org/10.1073/pnas.0908097106 PMID: 19955406

43. Sawers RG. Identification and molecular characterization of a transcriptional regulator from *Pseudomonas aeruginosa* PAO1 exhibiting structural and functional similarity to the FNR protein of

Escherichia coli. Mol Microbiol. 1991; 5: 1469–1481. https://doi.org/10.1111/j.1365-2958.1991.tb00793.x PMID: 1787797

44. Zamorano-Sánchez D, Reyes-González A, Gómez-Hernández N, Rivera P, Georgellis D, Girard L. FxkR provides the missing link in the *fixL-fixK* signal transduction cascade in *Rhizobium etli* CFN42. Mol Plant-Microbe Interact. 2012; 25: 1506–1517. https://doi.org/10.1094/MPMI-05-12-0136-R PMID: 22809273

45. Reyes-González A, Talbi C, Rodríguez S, Rivera P, Zamorano-Sánchez D, Girard L. Expanding the regulatory network that controls nitrogen fixation in *Sinorhizobium meliloti*: elucidating the role of the two-component system hFixL-FxkR. Microbiology. 2016; 162: 979–988. https://doi.org/10.1099/mic.0.000284 PMID: 27010660

46. Tsoy O V., Ravcheev DA, Čuklina J, Gelfand MS. Nitrogen fixation and molecular oxygen: comparative genomic reconstruction of transcription regulation in *Alphaproteobacteria*. Front Microbiol. 2016; 7: 1343. https://doi.org/10.3389/fmicb.2016.01343 PMID: 27617010

47. Zamorano-Sánchez D, Girard L. FNR-like proteins in rhizobia: past and future. Biological Nitrogen Fixation. Hoboken: John Wiley & Sons, Inc; 2015. pp. 155–166. https://doi.org/10.1002/9781119053095.ch15

48. Schlüter A, Patschkowski T, Weidner S, Unden G, Hynes MF, Priefer UB. Function and regulatory characteristics of FnrN, an oxygen-responsive transcriptional activator in *Rhizobium leguminosarum* bv. *viciae*. New Horizons Nitrogen Fixat. 1993; 493.

49. Dufour YS, Kiley PJ, Donohue TJ. Reconstruction of the core and extended regulons of global transcription factors. PLoS Genet. 2010; 6: 1–20. https://doi.org/10.1371/journal.pgen.1001027 PMID: 20661434

50. Gamper M, Zimmermann A, Haas D. Anaerobic regulation of transcription initiation in the *arcDABC* operon of *Pseudomonas aeruginosa*. J Bacteriol. 1991; 173: 4742–4750. https://doi.org/10.1128/jb.173.15.4742-4750.1991 PMID: 1906871

51. Kahn D, David M, Domergue O, Daveran ML, Ghai J, Hirsch PR, et al. *Rhizobium meliloti fixGHI* sequence predicts involvement of a specific cation pump in symbiotic nitrogen fixation. J Bacteriol. 1989; 171: 929–39. https://doi.org/10.1128/jb.171.2.929-939.1989 PMID: 2536685

52. Guest JR, Green J, Irvine AS, Spiro S. The FNR modulon and FNR-regulated gene expression. Regulation of gene expression in *Escherichia coli*. Boston: Springer US; 1996. pp. 317–342. https://doi.org/10.1007/978-1-4684-8601-8_16

53. Spiro S. The FNR family of transcriptional regulators. Antonie Van Leeuwenhoek. 1994; 66: 23–36. https://doi.org/10.1007/BF00871630 PMID: 7747934

54. Sciotti MA, Chanfon A, Hennecke H, Fischer HM. Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium japonicum*. J Bacteriol. 2003; 185: 5639–5642. https://doi.org/10.1128/jb.185.18.5639-5642.2003 PMID: 12949117

55. Soupène E, Foussard M, Boistard P, Truchet G, Batut J. Oxygen as a key developmental regulator of *Rhizobium meliloti* N2-fixation gene expression within the alfalfa root nodule. Proc Natl Acad Sci. 1995; 92: 3759–63. https://doi.org/10.1073/pnas.92.9.3759 PMID: 7731979

56. Schlüter A, Patschkowski T, Quandt J, Selinger LB, Weidner S, Krämer M, et al. Functional and regulatory analysis of the two copies of the *fixNOQP* operon of *Rhizobium leguminosarum* strain VF39. Mol Plant-Microbe Interact. 1997; 10: 605–16. https://doi.org/10.1094/MPMI.1997.10.5.605 PMID: 9204566

57. Spiro S, Guest JR. Regulation and over-expression of the *fnr* gene of *Escherichia coli*. J Gen Microbiol. 1987; 133: 3279–3288. https://doi.org/10.1099/00221287-133-12-3279 PMID: 2846747

58. Unden G, Trageser M. Oxygen regulated gene expression in *Escherichia coli*: control of anaerobic respiration by the FNR protein. Antonie Van Leeuwenhoek. 1991; 59: 65–76. https://doi.org/10.1007/BF00445650 PMID: 1854188

59. Jervis AJ, Green J. In vivo demonstration of FNR dimers in response to lower O2 availability. J Bacteriol. 2007; 189: 2930–2932. https://doi.org/10.1128/JB.01921-06 PMID: 17277055

60. Patschkowski T, Schlüter A, Priefer UB. *Rhizobium leguminosarum* bv. *viciae* contains a second *fnr/fixK*-like gene and an unusual *fixL* homologue. Mol Microbiol. 1996; 21: 267–280. https://doi.org/10.1046/j.1365-2958.1996.6321348.x PMID: 8858582

61. Girard L, Brom S, Dávalos A, López O, Soberón M, Romero D. Differential regulation of *fixN*-reiterated genes in *Rhizobium etli* by a novel *fixL—fixK* cascade. Mol Plant-Microbe Interact. 2000; 13: 1283–1292. https://doi.org/10.1094/MPMI.2000.13.12.1283 PMID: 11106020

62. Granados-Baeza MJ, Gómez-Hernández N, Mora Y, Delgado MJ, Romero D, Girard L. Novel reiterated Fnr-type proteins control the production of the symbiotic terminal oxidase *cbb*3 in *Rhizobium etli*

CFN42. Mol Plant-Microbe Interact. 2007; 20: 1241–1249. https://doi.org/10.1094/MPMI-20-10-1241 PMID: 17918626

63. Martinez-Argudo I, Little R, Shearer N, Johnson P, Dixon R. Nitrogen fixation: key genetic regulatory mechanisms. Biochem Soc Trans. 2005; 33: 152–156. https://doi.org/10.1042/BST0330152 PMID: 15667291

64. Krey R, Pühler A, Klipp W. A defined amino acid exchange close to the putative nucleotide binding site is responsible for an oxygen-tolerant variant of the *Rhizobium meliloti* NifA protein. MGG Mol Gen Genet. 1992; 234: 433–41. https://doi.org/10.1007/BF00538703 PMID: 1406589

65. Fischer H-M, Fritsche S, Herzog B, Hennecke H. Critical spacing between two essential cysteine residues in the interdomain linker of the *Bradyrhizobium japonicum* NifA protein. FEBS Lett. 1989; 255: 167–171. https://doi.org/10.1016/0014-5793(89)81083-x PMID: 2792368

66. Beynon JL, Williams MK, Cannon FC. Expression and functional analysis of the *Rhizobium meliloti nifA* gene. EMBO J. 1988; 7: 7–14. https://doi.org/10.1002/j.1460-2075.1988.tb02777.x PMID: 16453824

67. Burgess BK, Lowe DJ. Mechanism of molybdenum nitrogenase. Chem Rev. 1996; 96: 2983–3012. https://doi.org/10.1021/cr950055x PMID: 11848849

68. Earl CD, Ronson CW, Ausubel FM. Genetic and structural analysis of the *Rhizobium meliloti fixA*, *fixB*, *fixC*, and *fixX* genes. J Bacteriol. 1987; 169: 1127–1136. https://doi.org/10.1128/jb.169.3.1127-1136.1987 PMID: 3029021

69. Martínez M, Palacios JM, Imperial J, Ruiz-Argüeso T. Symbiotic autoregulation of *nifA* expression in *Rhizobium leguminosarum* bv. *viciae*. J Bacteriol. 2004; 186: 6586–6594. https://doi.org/10.1128/JB.186.19.6586-6594.2004 PMID: 15375140

70. Salazar E, Javier Díaz-Mejía J, Moreno-Hagelsieb G, Martínez-Batallar G, Mora Y, Mora J, et al. Characterization of the NifA-RpoN regulon in *Rhizobium etli* in free life and in symbiosis with *Phaseolus vulgaris*. Appl Environ Microbiol. 2010; 76: 4510–4520. https://doi.org/10.1128/AEM.02007-09 PMID: 20453139

71. Thöny B, Anthamatten D, Hennecke H. Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon *fixR nifA*: analysis of *cis*- and *trans*-acting elements. J Bacteriol. 1989; 171: 4162–4169. https://doi.org/10.1128/jb.171.8.4162-4169.1989 PMID: 2753853

72. Thöny B, Hennecke H. The -24/-12 promoter comes of age. FEMS Microbiol Lett. 2006; 63: 341–357. https://doi.org/10.1111/j.1574-6968.1989.tb03404.x

73. Michiels J, D'hooghe I, Verreth C, Pelemans H, Vanderleyden J. Characterization of the *Rhizobium leguminosarum* biovar *phaseoli nifA* gene, a positive regulator of *nif* gene expression. Arch Microbiol. 1994; 161: 404–408. https://doi.org/10.1007/BF00288950 PMID: 8042903

74. Hertig C, Li RY, Louarn AM, Garnerone AM, David M, Batut J, et al. *Rhizobium meliloti* regulatory gene *fixJ* activates transcription of *R. meliloti nifA* and *fixK* genes in *Escherichia coli*. J Bacteriol. 1989; 171: 1736–1738. https://doi.org/10.1128/jb.171.3.1736-1738.1989 PMID: 2646295

75. David M, Daveran M-L, Batut J, Dedieu A, Domergue O, Ghai J, et al. Cascade regulation of *nif* gene expression in *Rhizobium meliloti*. Cell. 1988; 54: 671–683. https://doi.org/10.1016/s0092-8674(88)80012-6 PMID: 2842062

76. Ditta G, Virts E, Palomares A, Kim GH. The *nifA* gene of *Rhizobium meliloti* is oxygen regulated. J Bacteriol. 1987; 169: 3217–3223. https://doi.org/10.1128/jb.169.7.3217-3223.1987 PMID: 2439489

77. Colonna-Romano S, Arnold W, Schlüter A, Boistard P, Pühler A, Priefer UB. An Fnr-like protein encoded in *Rhizobium leguminosarum* biovar *viciae* shows structural and functional homology to *Rhizobium meliloti fixK*. MGG Mol Gen Genet. 1990; 223: 138–147. https://doi.org/10.1007/BF00315806 PMID: 2175385

78. Lopez O, Morera C, Miranda-Rios J, Girard L, Romero D, Soberon M. Regulation of gene expression in response to oxygen in *Rhizobium etli*: role of FnrN in *fixNOQP* expression and in symbiotic nitrogen fixation. J Bacteriol. 2001; 183: 6999–7006. https://doi.org/10.1128/JB.183.24.6999-7006.2001 PMID: 11717256

79. Láruson ÁJ, Yeaman S, Lotterhos KE. The importance of genetic redundancy in evolution. Trends in Ecology and Evolution. Elsevier Ltd; 2020. pp. 809–822. https://doi.org/10.1016/j.tree.2020.04.009 PMID: 32439075

80. Tononi G, Sporns O, Edelman GM. Measures of degeneracy and redundancy in biological networks. Proc Natl Acad Sci U S A. 1999; 96: 3257–3262. https://doi.org/10.1073/pnas.96.6.3257 PMID: 10077671

81. Gutiérrez D, Hernando Y, Palacios JM, Imperial J, Ruiz-Argüeso T. FnrN controls symbiotic nitrogen fixation and hydrogenase activities in *Rhizobium leguminosarum* biovar *viciae* UPM791. J Bacteriol. 1997; 179: 5264–5270. https://doi.org/10.1128/jb.179.17.5264-5270.1997 PMID: 9286975

82. Mesa S, Hauser F, Friberg M, Malaguti E, Fischer H-M, Hennecke H. Comprehensive assessment of the regulons controlled by the FixLJ-FixK2-FixK1 cascade in *Bradyrhizobium japonicum*. J Bacteriol. 2008; 190: 6568–6579. https://doi.org/10.1128/JB.00748-08 PMID: 18689489

83. Galibert F, Finan TM, Long SR, Pühler A, Abola P, Ampe F, et al. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science (80-). 2001; 293: 668–672. https://doi.org/10.1126/science.1060966 PMID: 11474104

84. Ferrières L, Francez-Charlot A, Gouzy J, Rouillé S, Kahn D. FixJ-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. Microbiology. 2004; 150: 2335–2345. https://doi.org/10.1099/mic.0.27081-0 PMID: 15256575

85. Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull KH, et al. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol. 2006; 7: R34. https://doi.org/10.1186/gb-2006-7-4-r34 PMID: 16640791

86. Sánchez-Cañizares C, Jorrín B, Durán D, Nadendla S, Albareda M, Rubio-Sanz L, et al. Genomic diversity in the endosymbiotic bacterium *Rhizobium leguminosarum*. Genes (Basel). 2018; 9: 60. https://doi.org/10.3390/genes9020060 PMID: 29364862

87. Colombo MV, Gutiérrez D, Palacios JM, Imperial J, Ruiz-Argüeso T. A novel autoregulation mechanism of *fnrN* expression in *Rhizobium leguminosarum* bv *viciae*. Mol Microbiol. 2000; 36: 477–486. https://doi.org/10.1046/j.1365-2958.2000.01867.x PMID: 10792733

88. Moris M, Dombrecht B, Xi C, Vanderleyden J, Michiels J. Regulatory role of *Rhizobium etli* CNPAF512 *fnrN* during symbiosis. Appl Environ Microbiol. 2004; 70: 1287–1296. https://doi.org/10.1128/aem.70.3.1287-1296.2004 PMID: 15006745

89. Clark SRD, Oresnik IJ, Hynes MF. RpoN of *Rhizobium leguminosarum* bv. *viciae* strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation. MGG Mol Gen Genet. 2001; 264: 623–633. https://doi.org/10.1007/s004380000348 PMID: 11212917

90. Silva Sousa EH, Gonzalez G, Gilles-Gonzalez MA. Oxygen blocks the reaction of the FixL-FixJ complex with ATP but does not influence binding of FixJ or ATP to FixL. Biochemistry. 2005; 44: 15359–15365. https://doi.org/10.1021/bi051661h PMID: 16285740

91. Schlüter A, Patschkowski T, Unden G, Priefer UB. The *Rhizobium leguminosarum* FnrN protein is functionally similar to *Escherichia coli* Fnr and promotes heterologous oxygen-dependent activation of transcription. Mol Microbiol. 1992; 6: 3395–3404. https://doi.org/10.1111/j.1365-2958.1992.tb02207.x PMID: 1484491

92. Kaminski PA, Mandon K, Arigoni F, Desnoues N, Elmerich C. Regulation of nitrogen fixation in *Azorhizobium caulinodans*: identification of a *fixK*-like gene, a positive regulator of *nifA*. Mol Microbiol. 1991; 5: 1983–1991. https://doi.org/10.1111/j.1365-2958.1991.tb00820.x PMID: 1766374

93. Karunakaran R, Ramachandran VK, Seaman JC, East AK, Mouhsine B, Mauchline TH, et al. Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J Bacteriol. 2009; 191: 4002–4014. https://doi.org/10.1128/JB.00165-09 PMID: 19376875

94. Delgado MJ, Bedmar EJ, Downie JA. Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation. Adv Microb Physiol. 1998; 40: 191–231. https://doi.org/10.1016/s0065-2911(08)60132-0 PMID: 9889979

95. Preisig O, Zufferey R, Thöny-Meyer L, Appleby CA, Hennecke H. A high-affinity *cbb*3-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of *Bradyrhizobium japonicum*. J Bacteriol. 1996; 178: 1532–1538. https://doi.org/10.1128/jb.178.6.1532-1538.1996 PMID: 8626278

96. Kopat VV, Chirak ER, Kimeklis AK, Safronova VI, Belimov AA, Kabilov MR, et al. Evolution of *fixNOQP* genes encoding cytochrome oxidase with high affinity to oxygen in rhizobia and related bacteria. Russ J Genet. 2017; 53: 766–774. https://doi.org/10.1134/S1022795417070067

97. Fischer HM. Genetic regulation of nitrogen fixation in rhizobia. Microbiol Rev. 1994; 58: 352–386. https://doi.org/10.1186/gb-2011-12-10-r106 PMID: 7968919

98. Mandon K, Kaminski PA, Elmerich C. Functional analysis of the *fixNOQP* region of *Azorhizobium caulinodans*. J Bacteriol. 1994; 176: 2560–2568. https://doi.org/10.1128/jb.176.9.2560-2568.1994 PMID: 8169204

99. Mandon K, Kaminski PA, Mougel C, Desnoues N, Dreyfus B, Elmerich C. Role of the *fixGHI* region of *Azorhizobium caulinodans* in free-living and symbiotic nitrogen fixation. FEMS Microbiol Lett. 1993; 114: 185–189. https://doi.org/10.1111/j.1574-6968.1993.tb06571.x PMID: 8282187

100. Koch HG, Winterstein C, Saribas AS, Alben JO, Daldal F. Roles of the *ccoGHIS* gene products in the biogenesis of the *cbb*3-type cytochrome *c* oxidase. J Mol Biol. 2000; 297: 49–65. https://doi.org/10.1006/jmbi.2000.3555 PMID: 10704306

101. Preisig O, Zufferey R, Hennecke H. The *Bradyrhizobium japonicum fixGHIS* genes are required for the formation of the high-affinity *cbb*3-type cytochrome oxidase. Arch Microbiol. 1996; 165: 297–305. https://doi.org/10.1007/s002030050330 PMID: 8661920

102. Mesa S, Ucurum Z, Hennecke H, Fischer HM. Transcription activation in vitro by the *Bradyrhizobium japonicum* regulatory protein FixK2. J Bacteriol. 2005; 187: 3329–3338. https://doi.org/10.1128/JB.187.10.3329-3338.2005 PMID: 15866917

103. Nellen-Anthamatten D, Rossi P. *Bradyrhizobium japonicum* FixK2, a crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels. J Bacteriol. 1998; 180: 5251–5255. https://doi.org/10.1128/JB.180.19.5251-5255.1998 PMID: 9748464

104. Bauer E, Kaspar T, Fischer H-M, Hennecke H. Expression of the *fixR-nifA* operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR. J Bacteriol. 1998; 180: 3853–3863. https://doi.org/10.1128/JB.180.15.3853-3863.1998 PMID: 9683482

105. Lang C, Smith LS, Long SR. Characterization of novel plant symbiosis mutants using a new multiple gene-expression reporter *Sinorhizobium meliloti* strain. Front Plant Sci. 2018; 9: 76. https://doi.org/10.3389/fpls.2018.00076 PMID: 29467773

106. Mendoza-Suárez MA, Geddes BA, Sánchez-Cañizares C, Ramírez-González RH, Kirchhelle C, Jorrin B, et al. Optimizing *Rhizobium*-legume symbioses by simultaneous measurement of rhizobial competitiveness and N2 fixation in nodules. Proc Natl Acad Sci U S A. 2020; 117: 9822–9831. https://doi.org/10.1073/pnas.1921225117 PMID: 32317381

107. Capela D, Filipe C, Bobik C, Batut J, Bruand C. *Sinorhizobium meliloti* differentiation during symbiosis with Alfalfa: a transcriptomic dissection. Mol Plant-Microbe Interact. 2006; 19: 363–372. https://doi.org/10.1094/MPMI-19-0363 PMID: 16610739

108. Ryu M-H, Zhang J, Toth T, Khokhani D, Geddes BA, Mus F, et al. Control of nitrogen fixation in bacteria that associate with cereals. Nat Microbiol. 2020; 5: 314–330. https://doi.org/10.1038/s41564-019-0631-2 PMID: 31844298

109. Wongdee J, Boonkerd N, Teaumroong N, Tittabutr P, Giraud E. Regulation of nitrogen fixation in *Bradyrhizobium* sp. strain DOA9 involves two distinct NifA regulatory proteins that are functionally redundant during symbiosis but not during free-living growth. Front Microbiol. 2018; 9: 1–11. https://doi.org/10.3389/fmicb.2018.00001 PMID: 29403456

110. Hu Y, Ribbe MW. Nitrogenase assembly. Biochim Biophys Acta—Bioenerg. 2013; 1827: 1112–1122. https://doi.org/10.1016/j.bbabio.2012.12.001 PMID: 23232096

111. Monroe JD, Owens TG, LaRue TA. Measurement of the fractional oxygenation of leghemoglobin in intact detached pea nodules by reflectance spectroscopy. Plant Physiol. 1989; 91: 598–602. https://doi.org/10.1104/pp.91.2.598 PMID: 16667074

112. Gourret J-P, Fernandez-Arias H. Etude ultrastructurale et cytochimique de la différenciation des bactéroïdes de *Rhizobium trifolii* Dangeard dans les nodules de *Trifolium repens* L. Can J Microbiol. 1974; 20: 1169–1181. https://doi.org/10.1139/m74-181 PMID: 4138777

113. Domonkos A, Horvath B, Marsh JF, Halasz G, Ayaydin F, Oldroyd GED, et al. The identification of novel loci required for appropriate nodule development in *Medicago truncatula*. BMC Plant Biol. 2013; 13. https://doi.org/10.1186/1471-2229-13-157 PMID: 24119289

114. Voroshilova VA, Boesten B, Tsyganov VE, Borisov AY, Tikhonovich IA, Priefer UB. Effect of mutations in *Pisum sativum* L. genes blocking different stages of nodule development on the expression of late symbiotic genes in *Rhizobium leguminosarum* bv. *viciae*. Mol Plant-Microbe Interact. 2001; 14: 471–476. https://doi.org/10.1094/MPMI.2001.14.4.471 PMID: 11310734

115. Sousa EHS, Tuckerman JR, Gonzalez G, Gilles-Gonzalez MA. A memory of oxygen binding explains the dose response of the heme-based sensor FixL. Biochemistry. 2007; 46: 6249–6257. https://doi.org/10.1021/bi7003334 PMID: 17487983

116. Moore LJ, Kiley PJ. Characterization of the dimerization domain in the FNR transcription factor. J Biol Chem. 2001; 276: 45744–45750. https://doi.org/10.1074/jbc.M106569200 PMID: 11581261

117. Spiro S, Guest JR. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol Lett. 1990; 75: 399–428. https://doi.org/10.1111/j.1574-6968.1990.tb04109.x PMID: 2248796

118. Green J, Scott C, Guest JR. Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. Advances in Microbial Physiology. Academic Press; 2001. pp. 1–34. https://doi.org/10.1016/s0065-2911(01)44010-0 PMID: 11407111

119. Ingalls BP. Mathematical modelling in systems biology: an introduction. Journal of Chemical Information and Modeling. MIT Press; 2014. https://doi.org/10.1007/s00292-008-1023-1 PMID: 19039618

120. Simon AJ, Vallée-Bélisle A, Ricci F, Watkins HM, Plaxco KW. Using the population-shift mechanism to rationally introduce "hill-type" cooperativity into a normally non-cooperative receptor. Angew Chemie —Int Ed. 2014; 53: 9471–9475. https://doi.org/10.1002/anie.201403777 PMID: 25044647

121. Rey FE, Harwood CS. FixK, a global regulator of microaerobic growth, controls photosynthesis in *Rhodopseudomonas palustris*. Mol Microbiol. 2010; 75: 1007–1020. https://doi.org/10.1111/j.1365-2958.2009.07037.x PMID: 20487293

122. Dispensa M, Thomas CT, Kim MK, Perrotta JA, Gibson J, Harwood CS. Anaerobic growth of *Rhodopseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators. J Bacteriol. 1992; 174: 5803–5813. https://doi.org/10.1128/jb.174.18.5803-5813.1992 PMID: 1522059

123. Egland PG, Harwood CS. BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodopseudomonas palustris* in concert with AadR, an Fnr family member. J Bacteriol. 1999; 181: 2102–2109. https://doi.org/10.1128/JB.181.7.2102-2109.1999 PMID: 10094687

124. Larimer FW, Chain P, Hauser L, Lamerdin J, Malfatti S, Do L, et al. Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. Nat Biotechnol. 2004; 22: 55–61. https://doi.org/10.1038/nbt923 PMID: 14704707

125. Kiers ET, Rousseau RA, West SA, Denlson RF. Host sanctions and the legume-*rhizobium* mutualism. Nature. 2003; 425: 78–81. https://doi.org/10.1038/nature01931 PMID: 12955144

126. Westhoek A, Field E, Rehling F, Mulley G, Webb I, Poole PS, et al. Policing the legume-*Rhizobium* symbiosis: a critical test of partner choice. Sci Rep. 2017; 7: 1419. https://doi.org/10.1038/s41598-017-01634-2 PMID: 28469244

127. Heath KD, Tiffin P. Stabilizing mechanisms in a legume-*rhizobium* mutualism. Evolution (N Y). 2009; 63: 652–662. https://doi.org/10.1111/j.1558-5646.2008.00582.x PMID: 19087187

128. Osorio H, Mettert E, Kiley P, Dopson M, Jedlicki E, Holmes DS. Identification and unusual properties of the master regulator FNR in the extreme acidophile *Acidithiobacillus ferrooxidans*. Front Microbiol. 2019; 10: 1642. https://doi.org/10.3389/fmicb.2019.01642 PMID: 31379789

129. Sambrook J, Russell DW. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring, Harbor Laboratory, New York, 3th edn. New York: Cold Spring Harbor Laboratory; 2001. https://doi.org/10.1128/AEM.71.8.4602

130. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. J Gen Microbiol. 2015; 84: 188–198. https://doi.org/10.1099/00221287-84-1-188 PMID: 4612098

131. Pini F, East AK, Appia-Ayme C, Tomek J, Karunakaran R, Mendoza-Suárez M, et al. Bacterial biosensors for in vivo spatiotemporal mapping of root secretion. Plant Physiol. 2017; 174: 1289–1306. https://doi.org/10.1104/pp.16.01302 PMID: 28495892

132. Poole PS, Schofield NA, Reid CJ, Drew EM, Walshaw DL. Identification of chromosomal genes located downstream of *dctD* that affect the requirement for calcium and the lipopolysaccharide layer of *Rhizobium leguminosarum*. Microbiology. 1994; 140: 2797–2809. https://doi.org/10.1099/00221287-140-10-2797 PMID: 8000544

133. Buchanan-Wollaston V. Generalized transduction in *Rhizobium leguminosarum*. J Gen Microbiol. 1979; 112: 135–142. https://doi.org/10.1099/00221287-112-1-135

134. Choi KH, Schweizer HP. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. Nat Protoc. 2006; 1: 153–161. https://doi.org/10.1038/nprot.2006.24 PMID: 17406227

135. Choi KH, Mima T, Casart Y, Rholl D, Kumar A, Beacham IR, et al. Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. Appl Environ Microbiol. 2008; 74: 1064–1075. https://doi.org/10.1128/AEM.02430-07 PMID: 18156318

136. Quandt J, Hynes MF. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. Gene. 1993; 127: 15–21. https://doi.org/10.1016/0378-1119(93)90611-6 PMID: 8486283

137. Poole PS, Blyth A, Reid CJ, Walters K. Myo-inositol catabolism and catabolite regulation in *Rhizobium leguminosarum* bv. *viciae*. Microbiology. 1994. https://doi.org/10.1099/00221287-140-10-2787

138. Allaway D, Lodwig EM, Crompton LA, Wood M, Parsons R, Wheeler TR, et al. Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids. Mol Microbiol. 2000. https://doi.org/10.1046/j.1365-2958.2000.01884.x PMID: 10792736

139. Tsukada S, Aono T, Akiba N, Lee KB, Liu C Te, Toyazaki H, et al. Comparative genome-wide transcriptional profiling of *Azorhizobium caulinodans* ORS571 grown under free-living and symbiotic conditions. Appl Environ Microbiol. 2009; 75: 5037–5046. https://doi.org/10.1128/AEM.00398-09 PMID: 19542345

B

Supplementary materials for "Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a *Rhizobium*-legume symbiosis"

Contents

B.1 Supplementary 1: Supporting figures

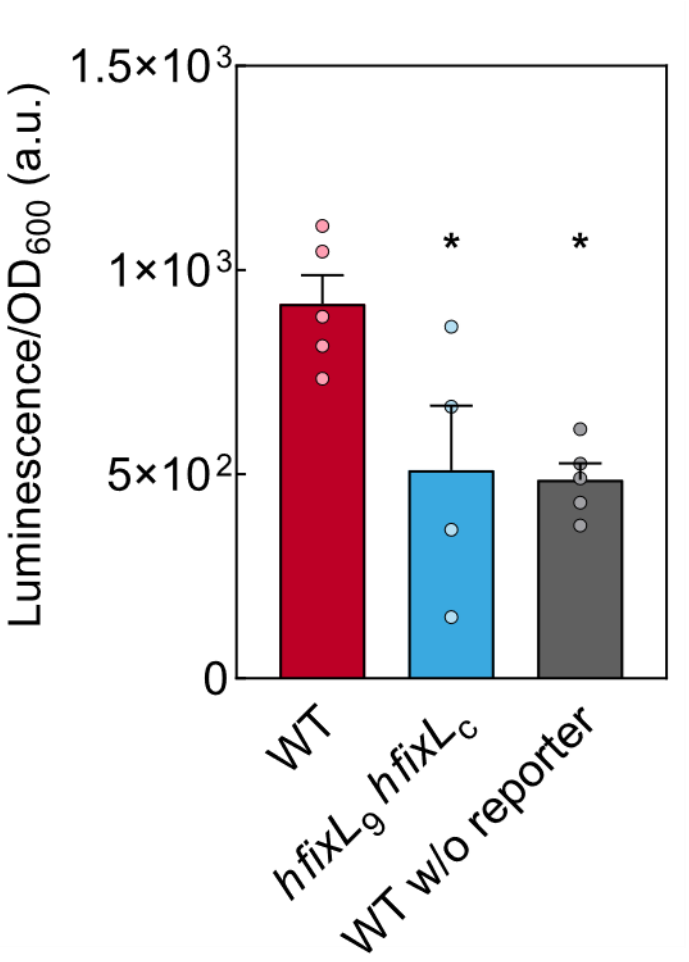


Fig. S1 | hFixL is required for *in planta fixK*$_{9a}$ expression in Rlv3841.

A reporter (pOPS0136) was built with the *luxCDABE* reporter operon fused to the *fixK*$_{9a}$ promoter. The promoter was active in isolated WT Rlv3841 bacteroids (OPS0376), but no luminescence above no-reporter background was recorded in double *hfixL* mutant bacteroids (OPS0528). Data are averages (±SEM) from at least four plants, *P < 0.05; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons.

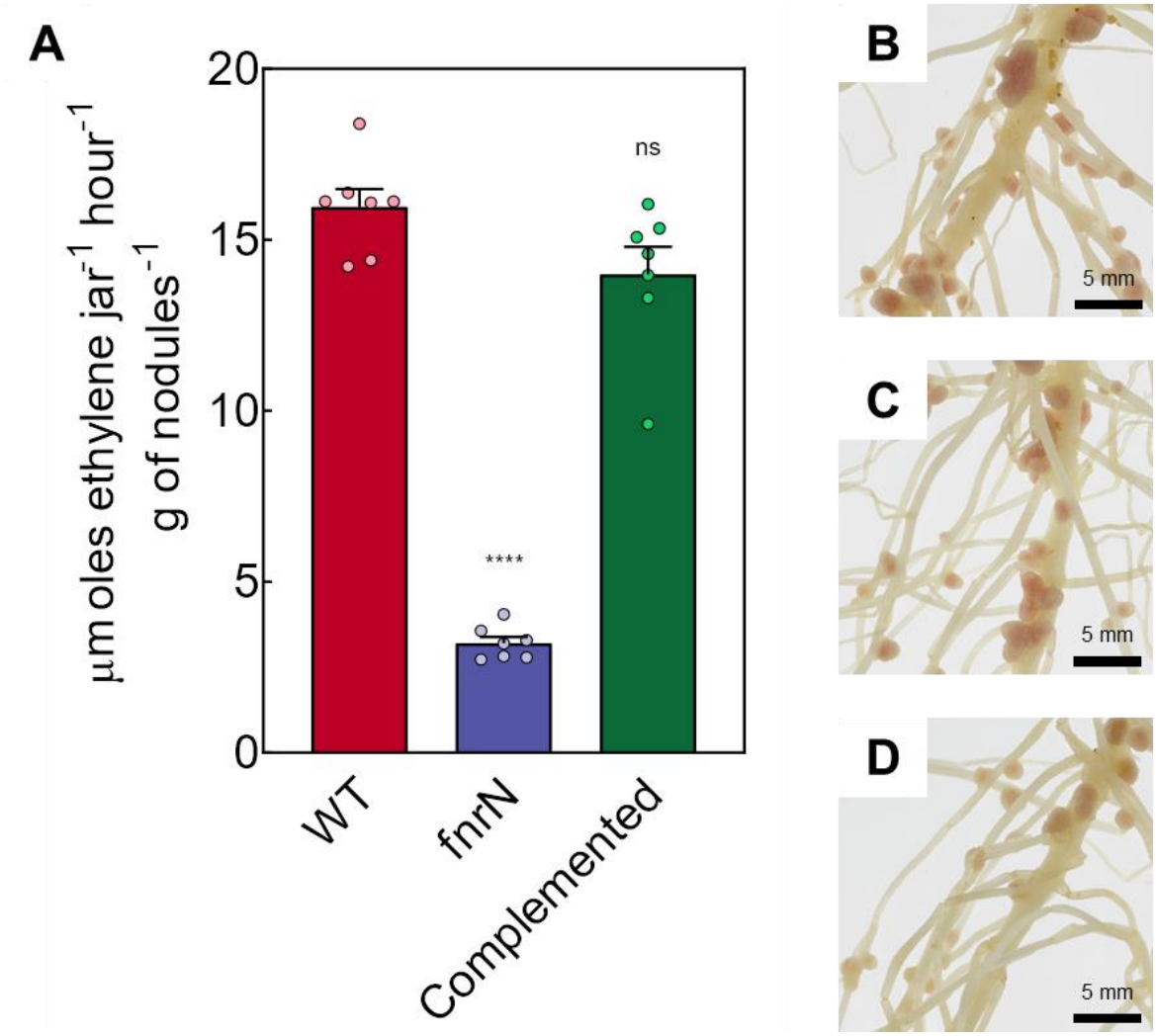


Fig. S2 | Complementation of the Rlv3841 *fnrN* mutant.

(A) Acetylene reduction rates; the activity of the *fnrN* mutant (LMB648) was 20% of WT Rlv3841. The complemented strain (OPS2260) fixed at 88% of WT. Nodules colonized by Rlv3841 WT **(B)**, the complemented *fnrN* mutant **(C)** and the *fnrN* mutant **(D)**. WT and complemented nodules have a red hue characteristic of normal leghaemoglobin content, whilst *fnrN* mutant-infected nodules are pale or brown. The *fnrN* mutant nodules also show minimal elongation, giving them a spherical shape in contrast to the egg-like shape of the WT and complemented strain infected nodules. Acetylene reduction rates are normalised to total weight of nodules per plant. Data are averages (±SEM) from seven plants, ns (not significant) $P \geq 0.05$; ****$P < 0.0001$; by one-way ANOVA with Dunnett's post-hoc test for multiple comparisons. Complementation also restored nodule morphology.

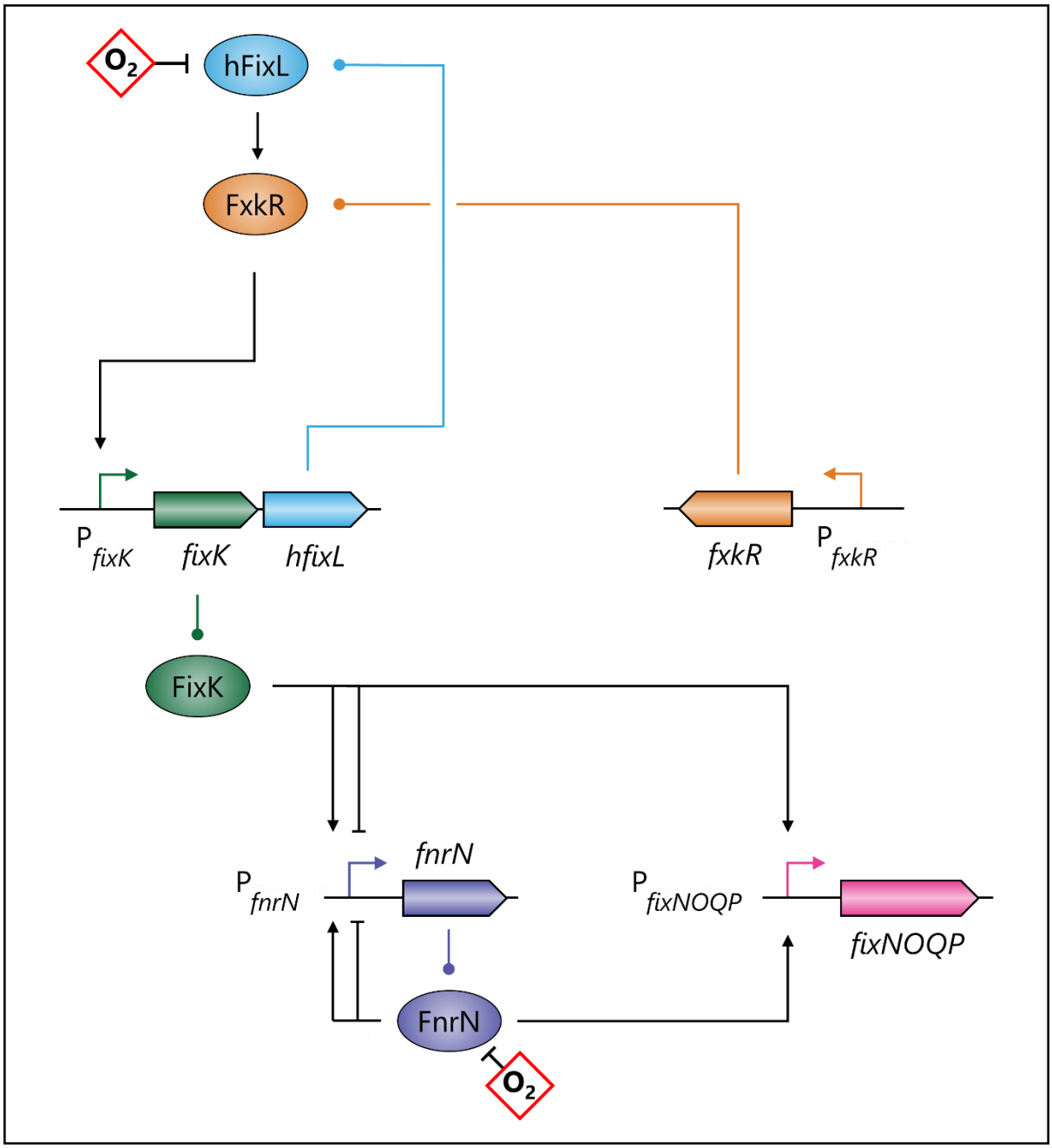


Fig. S3 | Simplified map of the Rlv3841 dual sensor oxygen cascade used for modelling.

Only one copy each is included of *hfixL*, *fixK*, *fxkR* and *fixNOQP*. Both FnrN and FixK can positively and negatively regulate expression of *fnrN*. FxkR negative auto-regulation is not included in the model, nor *fxkR* repression via anaerobox binding. Regulation is indicated with lines ending in arrows (positive regulation) and ending in blunt ends (negative regulation). Translation is shown as lines ending in circles.

B.2 Supplementary 2: Modelling oxygen regulation

1. Overview

In this section we develop a simplified mathematical model of our multi-sensor oxygen (O_2) regulation cascade to further investigate its function and structure. We begin by defining the contribution of each component and input to our system and derive an Ordinary Differential Equation (ODE) model of its dynamics. This model is nondimensionalised to simplify its analysis, and we discuss nominal parameter values with reference to past studies of similar regulatory systems. To derive a tractable model the full complexity of the biochemical systems involved must be greatly simplified, and many secondary, external factors that could influence its behaviour are not included. Therefore, we have aimed to provide a description of our system which captures its key qualitative behaviours, but do not quantitatively fit it to our experimental results (for which many necessary features, particularly *in planta*, may not be measurable).

2. Model Structure

2.1. hFixL/FnrN Oxygen Sensitivity

O_2 can bind to the Per-Arnt-Sim domain of hFixL (L) and the cysteine-rich motif of FnrN (N), in both cases deactivating the protein (reviewed in [1], see [2] and [3] for details respectively). We model the O_2 binding state of each protein using Hill-type saturating functions: the active (i.e. not O_2-bound) concentration of hFixL (L_a) and FnrN (N_a) can therefore be described as [4,5]:

$$N_a(N,X) = N \cdot \frac{K_{X,N}^{n_N}}{K_{X,N}^{n_N} + X^{n_N}}$$

$$L_a(L,X) = L \cdot \frac{K_{X,L}^{n_L}}{K_{X,L}^{n_L} + X^{n_L}}$$

where X is O_2 concentration, $K_{X,N}$ and $K_{L,N}$ are the half-saturating O_2 concentrations for each transcription factor, and n_N and n_L are the Hill coefficients (apparent cooperativity) of O_2 binding in each case.

2.2. Regulation of *fixK/hfixL* Expression

In our model the expression of *fixK* and *hfixL* is controlled by an upstream autoregulatory network; the two-component system (TCS) involving FxkR and hFixL [6,7]. To simplify analysis we proceed by modelling each interaction in this architecture using saturating first-order Hill-type functions (thereby assuming each interaction is non-cooperative), and assume that interactions between O_2/hFixL and hFixL/FxkR occur on a faster timescale than expression of hFixL (i.e. the timescale of transcription & translation) . With these assumptions we can express the rate of change of hFixL concentration (L) as:

$$\frac{dL}{dt} = \beta_0 \frac{R_a}{R_a + k_1} - \delta L + \lambda_1$$

where β_0 is the combined rate of transcription/translation, R_a is the quantity of total FxkR (R, assumed to be expressed at constant concentration) that is active following interaction with hFixL in the TCS, k_1 is the half-saturating constant of the activating promoter P_{fixK}, δ is the rate at which hFixL is degraded/diluted out of the system, and λ is an expression leakage term.

We express R_a in turn as a function of active hFixL as:

$$R_a = R \frac{L_a(L, X)}{L_a(L, X) + k_2}$$

which again has a half saturating constant k_2, and $L_a(L, X)$ is the amount of active (O_2-dependent) hFixL as defined above.

Combining the above expressions for $\frac{dL}{dt}$, L_a and R_a allows us to eliminate R_a and L_a to give:

$$\frac{dL}{dt} = \beta_0 \frac{RL}{(R + k_1)L + k_1 k_2 + k_1 k_2 \frac{X^{n_L}}{K_{X,L}^{n_L}}} - \delta L + \lambda_1$$

By combining parameters (including R, as we have assumed *fxkR* expression is constant), this expression can then be simplified to:

$$\frac{dL}{dt} = \beta_1 \frac{\frac{L}{K_1}}{1 + \frac{L}{K_1} + \left(\frac{X}{K_{X,L}}\right)^{n_L}} - \delta L + \lambda_1$$

where $\beta_1 = \frac{\beta_0 R}{R + k_1}, K_1 = \frac{k_1 k_2}{R + k_1}$

Since *fixK* (F) is co-expressed with *hfixL*, and we assume that downstream processes are not consuming *fixK*, its expression can be expressed using the same equation:

$$\frac{dF}{dt} = \beta_1 \frac{\frac{F}{K_1}}{1 + \frac{F}{K_1} + \left(\frac{X}{K_{X,L}}\right)^{n_L}} - \delta F + \lambda_1$$

2.3. Regulation of *fixNOQP* Expression

The promoter upstream of *fixNOQP* includes an anaerobox, to which active FnrN and FixK can bind to activate transcription [8,9]. Since both transcription factors bind the same motif [10], we model the promoter's response as a competitive binding process with cooperativity of order (n), though we assign different maximal expression rates (β_i's) and binding constants (K_i's) to the two regulators.

$$\Gamma(F, N_a, X) = \beta_2 \frac{\left(\frac{N_a}{K_2}\right)^n}{1 + \left(\frac{N_a}{K_2}\right)^n + \left(\frac{F}{K_3}\right)^n} + \beta_3 \frac{\left(\frac{F}{K_3}\right)^n}{1 + \left(\frac{N_a}{K_2}\right)^n + \left(\frac{F}{K_3}\right)^n}$$

2.4. Regulation of *fnrN* Expression

The promoter upstream of *fnrN* contains a similar distal anaerobox to that regulating *fixNOQP* [8]. It also contains a proximal anaerobox; past studies have shown FnrN can bind this sequence and (by blocking transcription initiation) repress its own expression [5,11]. We therefore model *fnrN* expression as the product of the contribution of the distal anaerobox (section 2.3) and a repression function contributed by the proximal binding with cooperativity as before:

$$\Gamma(F, N, X) \cdot \frac{1}{1 + \left(\frac{N_a}{K_4}\right)^n + \left(\frac{F}{K_5}\right)^n}$$

Here K_4, K_5 are the half saturation binding constants for FnrN and FixK respectively to the proximal anaerobox.

2.5. ODE Model

Combining the above we can describe our system with three linked ODEs of the form:

$$\frac{dF}{dt} = \beta_1 \frac{\frac{F}{K_1}}{1 + \frac{F}{K_1} + \left(\frac{X}{K_{X,L}}\right)^{n_L}} - \delta F + \lambda_1$$

$$\frac{dN}{dt} = \Gamma(F, N, X) \cdot \frac{1}{1 + \left(\frac{N_a}{K_4}\right)^n + \left(\frac{F}{K_5}\right)^n} - \delta N + \lambda_2$$

$$\frac{dY}{dt} = \Gamma(F, N, X) - \delta Y + \lambda_3$$

where we have introduced individual transcriptional leakage parameters $\lambda_{1,2,3}$ for each species, and assume each species is degraded and diluted at an equal rate δ. N_a and Γ are given by:

$$N_a(N, X) = N \cdot \frac{K_{X,N}^{n_N}}{K_{X,N}^{n_N} + X^{n_N}}$$

$$\Gamma(F, N, X) = \frac{\beta_2 \left(\frac{N_a(N,X)}{K_2}\right)^n + \beta_3 \left(\frac{F}{K_3}\right)^n}{1 + \left(\frac{N_a(N,X)}{K_2}\right)^n + \left(\frac{F}{K_3}\right)^n}$$

2.6. ODE Model Nondimensionalisation

To simplify the ODE model derived in section 2.5 we can nondimensionalise several state variables and parameters. This is done in Table S1, where we eliminate δ by re-defining the time parameter (now τ) as a multiple of the degradation timescale. We similarly normalise $K_{1,3,5}$ and $K_{2,4}$ by expression rates β_1 and β_2 respectively, which introduces a new parameter $\bar{\beta}$ that reflects the relative activator effect of FnrN and FixK. In this process we do not nondimensionalise $K_{X,L}$ and $K_{X,N}$, so that their units remain the same as X (O_2 concentration).

Table S1 – Parameter nondimensionalisation

Dimensionless Parameter	Dimensioned Substitution
$\boldsymbol{\tau}$	$t\delta$
$\overline{\boldsymbol{F}}$	$\frac{F\delta}{\beta_1}$
$\overline{\boldsymbol{N}}$	$\frac{N\delta}{\beta_2}$
$\overline{\boldsymbol{Y}}$	$\frac{T\delta}{\beta_2}$
$\overline{\boldsymbol{\beta}}$	$\frac{\beta_3}{\beta_2}$
$\overline{\boldsymbol{K_{1,3,5}}}$	$\frac{K_{1,3,5}\delta}{\beta_1}$
$\overline{\boldsymbol{K_{2,4}}}$	$\frac{K_{2,4}\delta}{\beta_2}$
$\overline{\boldsymbol{\lambda_1}}$	$\frac{\lambda_1}{\beta_1}$
$\overline{\boldsymbol{\lambda_{2,3}}}$	$\frac{\lambda_{2,3}}{\beta_2}$

Completing this nondimensionalisation gives the following simplified system of ODEs:

$$\frac{d\bar{F}}{d\tau} = \frac{\frac{\bar{F}}{\overline{K_1}}}{1 + \frac{\bar{F}}{\overline{K_1}} + \left(\frac{X}{K_{X,L}}\right)^{n_L}} - \bar{F} + \overline{\lambda_1}$$

$$\frac{d\bar{N}}{d\tau} = \Gamma(\bar{F}, \bar{N}, X) \cdot \frac{1}{1 + \left(\frac{\overline{N_a}}{\overline{K_4}}\right)^{n} + \left(\frac{\bar{F}}{\overline{K_5}}\right)^{n}} - \bar{N} + \overline{\lambda_2}$$

$$\frac{d\bar{Y}}{d\tau} = \Gamma(\bar{F}, \bar{N}, X) - \bar{Y} + \overline{\lambda_3}$$

where:

$$N_a(\bar{N}, X) = \bar{N} \cdot \frac{K_{X,N}^{n_N}}{K_{X,N}^{n_N} + X^{n_N}}$$

$$\Gamma(\bar{F}, \bar{N}, X) = \frac{\left(\frac{N_a(\bar{N}, X)}{\overline{K_2}}\right)^n + \bar{\beta}\left(\frac{\bar{F}}{\overline{K_3}}\right)^n}{1 + \left(\frac{N_a(\bar{N}, X)}{\overline{K_2}}\right)^n + \left(\frac{\bar{F}}{\overline{K_3}}\right)^n}$$

With this nondimensionalisation our system's response is largely determined by the five parameters, $\overline{K}_{1,2,3,4,5}$ and $\bar{\beta}$, with the leak terms λ_i playing a smaller role. $K_{X,L}$, $K_{X,N}$, and n determine the location and sensitivity of the oxygen response.

3. Parameter Values

To qualitatively compare our simplified model to the experimental results we must first estimate values for its parameters. Table S2 contains the parameter values used in this study. Many can be estimated from published literature results, or by considering qualitative observations of our system:

$\bar{\beta}$ – For simplicity we will set $\bar{\beta} = 1$, which implies that the activatory effect of FnrN and FixK is equivalent (i.e. $\beta_2 = \beta_3$). A corollary of this assumption is that for our model in nondimensionalised form, $\bar{\beta} \leq 1$ then the equilibrium value of each state variable ($\bar{F}, \bar{N}, \bar{Y}$) will lie in the range $[\lambda_i, 1 + \lambda_i]$.

$\overline{K_1}, \overline{K_2}, \overline{K_3}$ – These three binding constants are set via qualitative comparison between our model and experimental results. $\overline{K_1}$ defines (with $K_{X,L}$) the turn-on point of the autoregulatory loop including FxkR and hFixL. We choose a value ($\overline{K_1} = 0.01$) to satisfy $\overline{K_1} \ll 1$ so that this subsystem is fully activated with $\bar{F} \approx 1$ (for small $\overline{\lambda_1}$) when $X \to 0$. We define the relative magnitudes of $\overline{K_{2,3}}$ as $\overline{K_3} = 10 \times \overline{K_2}$ following experimental observation that FnrN more strongly activates expression than FixK. To set their absolute magnitudes we desire $\overline{K_3} > 1$ such that $\bar{F}$ (recalling the maximum value $\bar{F} \approx 1$) does not saturate the transcription function Γ, and $\overline{K_2} \lesssim \frac{1}{\overline{K_3}^n + 1}$ so that N_a can saturate this function (i.e. displace bound FixK) when $\bar{F} \approx 1$. Consequently, we set $\overline{K_2} = 0.15$ and $\overline{K_3} = 1.5$ which satisfies these relations.

$\overline{K_{4,5}}$ – Past studies of a similar FnrN system[5] observed that when only one promoter location was bound, the binding constant for the proximal (repressing) anaerobox was approximately five times that of the distal (activating) anaerobox (which would imply $\overline{K_4} \approx 5 \cdot \overline{K_2}$). However, cooperativity is also observed between binding at these two locations, which reduces the apparently $\overline{K_4}$ when the distal anaerobox is bound approximately twofold. Consequently, we assume an intermediate value of $\overline{K_{4,5}} \approx 2 \cdot \overline{K_{2,3}}$.

$K_{X,L}$ – Our *in vivo* experiments demonstrated that hFixL mediated activation occurs (at least partially) when O_2 concentration drops to 1%. This is in line with past studies [12,13]. We select a value of $K_{X,L} = 0.3$ such that hFixL mediated activation occurs in our model by this point.

$K_{X,L}$ – Likewise, the *in vivo* experiments demonstrate that FnrN mediated activation does not occur significantly at 1% O_2 concentration, but it does occur at the much lower O_2 levels present *in planta*, and hence we set a smaller value of $K_{X,N} = 0.005$.

$\overline{\lambda_{1,2,3}}$ – Each transcriptional leakiness term is set to the same value $\overline{\lambda_{1,2,3}} = 0.005$, which corresponds to ≈ 0.5% of the maximal nondimensionalised value of each state variable (i.e. $\bar{F}, \bar{N}, \bar{Y} \approx 1$) and implies that the expression of each gene is small in the absence of its activating transcription factors.

n – Previous studies have demonstrated that FnrN binds as a dimer to its target anaerobox and has a sharp sigmoidal binding profile for both the proximal and distal anaeroboxes [5]. Hence we model this binding process (which is assumed to also hold for FixK, which binds the same motif) as cooperative with $n = 2$.

n_L – hFixL exhibits a sharp response to increasing O_2 concentration, which can be explained by hysteretic oxygen binding to the sensor's haem binding domain [4]. This response can be approximated by a Hill function with greater than unity exponent (i.e. $n_L > 1$) [4], and consequently we set $n_L = 2$ for O_2 binding to hFixL.

n_N – We assign $n_N = 1$, which assumes non-cooperative binding between monomeric FnrN and O_2 [14,15].

Table S2 – Nondimensionalised parameter values and definitions used in simulations.

Parameter	Value	Unit	Description
$\overline{K_1}$	0.01	none	Equilibrium constant for autoactivation of hFixL.
$\overline{K_2}$	0.15	none	Equilibrium constant for FnrN binding to distal anaerobox.
$\overline{K_3}$	1.5	none	Equilibrium constant for FixK binding to distal anaerobox.
$\overline{K_4}$	$2 \cdot \overline{K_2}$	none	Equilibrium constant for FnrN binding to proximal anaerobox.
$\overline{K_5}$	$2 \cdot \overline{K_3}$	none	Equilibrium constant for FixK binding to proximal anaerobox.
$\bar{\beta}$	1	none	Relative activation effect of FixK and FnrN.
$\overline{\lambda_{1,2,3}}$	0.005	none	Transcriptional leak rate.
$K_{X,L}$	0.3	% O_2	Equilibrium constant for O_2 binding to hFixL.
$K_{X,N}$	0.005	% O_2	Equilibrium constant for O_2 binding to FnrN.
n	2	none	Hill coefficient for FnrN/FixK to promoter sequences.
n_L	2	none	Hill coefficient for O_2 to hFixL.
n_N	1	none	Hill coefficient for O_2 to FnrN.

References

1. Rutten PJ, Poole PS. Oxygen regulatory mechanisms of nitrogen fixation in rhizobia. In: Poole RK, editor. Advances in Microbial Physiology. Academic Press; 2019. pp. 325–389. doi:10.1016/bs.ampbs.2019.08.001

2. Green J, Crack JC, Thomson AJ, LeBrun NE. Bacterial sensors of oxygen. Curr Opin Microbiol. 2009;12: 145–151. doi:10.1016/j.mib.2009.01.008

3. Jervis AJ, Green J. In vivo demonstration of FNR dimers in response to lower O2 availability. J Bacteriol. 2007;189: 2930–2932. doi:10.1128/JB.01921-06

4. Sousa EHS, Tuckerman JR, Gonzalez G, Gilles-Gonzalez MA. A memory of oxygen binding explains the dose response of the heme-based sensor FixL. Biochemistry. 2007;46: 6249–6257. doi:10.1021/bi7003334

5. Colombo MV, Gutiérrez D, Palacios JM, Imperial J, Ruiz-Argüeso T. A novel autoregulation mechanism of *fnrN* expression in *Rhizobium leguminosarum* bv *viciae*. Mol Microbiol. 2000;36: 477–486. doi:10.1046/j.1365-2958.2000.01867.x

6. Zamorano-Sánchez D, Reyes-González A, Gómez-Hernández N, Rivera P, Georgellis D, Girard L. FxkR provides the missing link in the *fixL-fixK* signal transduction cascade in *Rhizobium etli* CFN42. Mol Plant-Microbe Interact. 2012;25: 1506–1517. doi:10.1094/MPMI-05-12-0136-R

7. Boesten B, Priefer UB. The C-terminal receiver domain of the *Rhizobium leguminosarum* bv. *viciae* FixL protein is required for free-living microaerobic induction of the *fnrN* promoter. Microbiology. 2004;150: 3703–3713. doi:10.1099/mic.0.27323-0

8. Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull KH, et al. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol. 2006;7: R34. doi:10.1186/gb-2006-7-4-r34

9. Lopez O, Morera C, Miranda-Rios J, Girard L, Romero D, Soberon M. Regulation of gene expression in response to oxygen in *Rhizobium etli*: role of FnrN in *fixNOQP* expression and in symbiotic nitrogen fixation. J Bacteriol. 2001;183: 6999–7006. doi:10.1128/JB.183.24.6999-7006.2001

10. Zamorano-Sánchez D, Girard L. FNR-like proteins in rhizobia: past and future. Biological Nitrogen Fixation. Hoboken: John Wiley & Sons, Inc; 2015. pp. 155–166. doi:10.1002/9781119053095.ch15

11. Moris M, Dombrecht B, Xi C, Vanderleyden J, Michiels J. Regulatory role of *Rhizobium etli* CNPAF512 *fnrN* during symbiosis. Appl Environ Microbiol. 2004;70: 1287–1296. doi:10.1128/AEM.70.3.1287-1296.2004

12. Ferrières L, Francez-Charlot A, Gouzy J, Rouillé S, Kahn D. FixJ-regulated genes evolved through promoter duplication in *Sinorhizobium meliloti*. Microbiology. 2004;150: 2335–2345. doi:10.1099/mic.0.27081-0

13. Sciotti MA, Chanfon A, Hennecke H, Fischer HM. Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in *Bradyrhizobium*

japonicum. J Bacteriol. 2003;185: 5639–5642. doi:10.1128/JB.185.18.5639-5642.2003

14. Kiley PJ, Beinert H. Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. FEMS Microbiol Rev. 1998;22: 341–352. doi:10.1111/j.1574-6976.1998.tb00375.x

15. Moore LJ, Kiley PJ. Characterization of the dimerization domain in the FNR transcription factor. J Biol Chem. 2001;276: 45744–45750. doi:10.1074/jbc.M106569200

B.3 Supplementary 3: Strains, plasmids and primers

1. Strains

1.1. *Escherichia coli*

Name	Relevant characteristics	Source
DH5α	F supE44 lacU169 hsdR17 recA1 endA1 gyrA96 thi-1 relA1 (80lacZM15)	Hanahan 1983 [1]
EC100D pir+	F^- mcrA Δ(mrr-hsdRMS-mcrBC) φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL (Str^R) nupG pir^+(DHFR)	Lucigen (Epicentre)
ST18	S17-1 ΔhemA thi pro hsdR-M- chromosomal integrated [RP4-2 Tc::Mu:Kmr::Tn7, Tra+ Trir Str^R]	Thoma and Schobert 2009 [2]
OneShot PIR1	F- Δlac169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA(ΔMluI)::pir-116	ThermoFisher (Invitrogen)

1.2. *Rhizobium leguminosarum* bv. 3841

Name	Relevant characteristics	Source
D5250	WT + pIJ11282 (P*neo*:*luxCDABE* in pIJ11268 backbone, lux positive control)	Frederix et al. 2014 [3]
LMB403	*hfixL*$_c$:pK19 single crossover	This study
LMB495	*hfixL*$_9$::ΩSpec	This study
LMB496	*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19	This study
LMB542	WT + pIJ11268 (promoterless *luxCDABE*, lux negative control)	Frederix et al. 2014 [3]
LMB648	*fnrN*::ΩTet	This study
LMB673	*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19 *fnrN*::ΩTet	This study
OPS0376	WT (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0136 (P*fixK*$_{9a}$:*luxCDABE* in pIJ11268 backbone)	This study
OPS0528	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS0136 (P*fixK*$_{9a}$:*luxCDABE* in pIJ11268 backbone)	This study
OPS1267	WT + pOPS0978 (P*fixNOQP*$_9$:*syfp2* in pOPS0786 backbone)	This study
OPS1268	WT + pOPS0979 (P*nifH*:*syfp2* in pOPS0786 backbone)	This study

OPS1269	WT + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study
OPS1274	LMB648 (*fnrN*::ΩTet) + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1275	LMB648 (*fnrN*::ΩTet) + pOPS0978 (P*fixNOQP*$_{9}$:*syfp2* in pOPS0786 backbone)	This study
OPS1277	LMB648 (*fnrN*::ΩTet) + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study
OPS1278	LMB496 (*hfixL*$_{9}$::ΩSpec *hfixL*$_{c}$:pK19) + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1279	LMB496 (*hfixL*$_{9}$::ΩSpec *hfixL*$_{c}$:pK19) + pOPS0978 (P*fixNOQP*$_{9}$:*syfp2* in pOPS0786 backbone)	This study
OPS1281	LMB496 (*hfixL*$_{9}$::ΩSpec *hfixL*$_{c}$:pK19) + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study
OPS1287	WT + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1294	WT + pOPS0785 (J23106:*syfp2* in pOPS0786 backbone, *syfp2* positive control)	This study
OPS1295	WT + pOPS0786 (Promoterless *syfp2* in pOPS0786 backbone, *syfp2* negative control)	This study
OPS1563	LMB403 (*hfixL*$_{c}$:pK19 single crossover) + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study
OPS1565	LMB495 (*hfixL*$_{9}$::ΩSpec) + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study
OPS1573	LMB403 (*hfixL*$_{c}$:pK19 single crossover) + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1574	LMB403 (*hfixL*$_{c}$:pK19 single crossover) + pOPS0978 (P*fixNOQP*$_{9}$:*syfp2* in pOPS0786 backbone)	This study
OPS1575	LMB495 (*hfixL*$_{9}$::ΩSpec) + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1576	LMB495 (*hfixL*$_{9}$::ΩSpec) + pOPS0978 (P*fixNOQP*$_{9}$:*syfp2* in pOPS0786 backbone)	This study
OPS1808	*ΔfxkR*$_{9}$	This study
OPS1811	OPS1808 (*ΔfxkR*$_{9}$) + pOPS0977 (P*fixNOQP*$_{10}$:*syfp2* in pOPS0786 backbone)	This study
OPS1812	OPS1808 (*ΔfxkR*$_{9}$) + pOPS0978 (P*fixNOQP*$_{9}$:*syfp2* in pOPS0786 backbone)	This study
OPS1813	OPS1808 (*ΔfxkR*$_{9}$) + pOPS0980 (P*fnrN* in pOPS0786 backbone)	This study

OPS2260	LMB648 (*fnrN*::ΩTet) + pOPS1510 (P*lac*:*fnrN* in pOGG280 backbone, genomically integrated by Tn7)	This study
OPS2428	WT + pOPS1593 (P*fixNOQP*$_9$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2429	WT + pOPS1594 (P*fnrN*:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2431	LMB648 (*fnrN*::ΩTet) + pOPS1593 (P*fixNOQP*$_9$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2432	LMB648 (*fnrN*::ΩTet) + pOPS1594 (P*fnrN*:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2434	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1593 (P*fixNOQP*$_9$:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2435	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1594 (P*fnrN*:*syfp2* in pOGG276 backbone, genomically integrated by Tn7)	This study
OPS2468	WT + pOPS1644 (P*fixNOQP*$_{10}$:*syfp2* in pJP2 backbone)	This study
OPS2469	LMB496 (*hfixL*$_9$::ΩSpec *hfixL*$_c$:pK19) + pOPS1644 (P*fixNOQP*$_{10}$:*syfp2* in pJP2 backbone)	This study
OPS2470	WT + pOPS1607 (J23106:*syfp2* in pOGG276 backbone, *syfp2* positive control, genomically integrated by Tn7)	This study
Rlv3841	Wild type *R. leguminosarum* bv. *viciae* 3841, Str^R derivative of strain Rlv300	Johnston and Beringer 1975 [4]

2. Plasmids

Name	Description	Source
pHP45ΩSpc	pBR322 derivative vector carrying Ω interposon spectinomycin resistance cassette, pHP45 replicon; Amp^R, Spc^R	Fellay et al. 1987 [5]
pHP45ΩTet	pBR322 derivative vector carrying Ω interposon tetracycline resistance cassette, pHP45 replicon; Amp^R, Tet^R	Fellay et al. 1987 [5]
pIJ11268	Broad host range vector based on pJP2 containing promoterless *luxCDABE* operon, used as negative control for Lux assay; Tet^R	Frederix et al. 2014 [3]
pIJ11282	pIJ11268 with neomycin promoter cloned in front of *luxCDABE*, used as positive control for Lux assay; Tet^R	Frederix et al. 2014 [3]
pJET1.2/blunt	*E. coli* vector for cloning PCR products; Amp^R	Thermo Scientific
pJP2	Broad-host-range *gusA* transcriptional promoter probe vector; Tet^R Amp^R	Prell et al. 2012 [6]
pJQ200SK	Suicide vector, pACYC derivative, p15A origin of replication, *lacZ sacB traJ*; $Gent^R$	Quandt and Hynes 1993 [7]
pK19mob	Mobilizable *E. coli* vector for integration mutagenesis (*oriV*), pMB1 replication, RP4 mob; Kan^R	Schafer 1994 [8]
pK19mobSacB	Mobilizable *E. coli* vector for integration mutagenesis (*oriV*, *sacB*), pMB1 replication, RP4 mob; Kan^R	Kirchner and Tauch 2003 [9]
pLMB441	Internal fragment of *hfixL*$_c$ amplified from Rlv3841 with primers pr0988/0989, cloned into pK19mob digested with XbaI.	This work
pLMB581	*hfixL*$_9$ amplified from Rlv3841 with pr1270/1271 cloned into pJET1/2/blunt	This work
pLMB585	*hfixL*$_9$ digested out of pLMB581 with XbaI/XhoI cloned into pJQ200SK, digested with XbaI/XhoI.	This work
pLMB590	ΩSpc from SmaI digested pHP45ΩSpc cloned into pLMB585 digested with StuI (blunted); $Gent^R$ Spc^R	This work
pLMB732	Rlv3841 *fnrN* amplified with primers pr1381/1382 cloned into pJQ200SK at XbaI/XhoI site.	This work
pLMB733	ΩTet from EcoRI digested pHP45ΩTet cloned into pLMB732 digested with MfeI; $Gent^R$ Tet^R	This work
pOGG276	Mobilizable vector for hosting sequences to be genomically inserted via mini-Tn7. R6Kγ; $Gent^R$ (genomic insert) Amp^R (backbone)	This work
pOGG280	Mobilizable vector for hosting sequences to be genomically inserted via mini-Tn7. R6Kγ; Kan^R (genomic insert) Amp^R (backbone)	This work
pOPS0136	P*fixK*$_{9a}$ amplified from Rlv3841 with oxp0287/0288 cloned into pIJ11268 digested with BamHI/KpnI	This work

pOPS0785	Reporter plasmid backbone with an MCS for fusing promoters to *mruby3*. Constitutive *syfp2* expression, used as positive control for *in-vitro* fluorescence assays. Contains parABCDE stability system. $Gent^R$.	This work
pOPS0786	Reporter plasmid backbone with an MCS for fusing promoters to *syfp2*. Constitutive *mruby3* expression. Contains *parABCDE* stability system. Used as negative control for *in-vitro* fluorescence assays. $Gent^R$.	This work
pOPS0977	P*fixNOQP*$_{10}$ amplified from Rlv3841 with primers oxp3039/3040 cloned into pOPS0786 digested with KpnI.	This work
pOPS0978	P*fixNOQP*$_9$ amplified from Rlv3841 with primers oxp3041/3042 cloned into pOPS0786 digested with KpnI.	This work
pOPS0979	P*nifH* amplified from Rlv3841 with primers oxp3043/3044 cloned into pOPS0786 digested with KpnI.	This work
pOPS0980	P*fnrN* amplified from Rlv3841 with primers oxp3045/3046 cloned into pOPS0786 digested with KpnI.	This work
pOPS1199	Fragments upstream (oxp2874/2875) and downstream (oxp2876/oxp2877) of *fxkR*$_9$ amplified from Rlv3841 and cloned into pK19mobSacB digested with HindIII and EcoRI, used to make the markerless mutant.	This work
pOPS1510	Rlv3841 *fnrN* gene amplified with oxp4115/4116 cloned into pOGG280 digested with BsaI.	This work
pOPS1593	P*fixNOQP*$_9$ fused to *syfp2* amplified from pOPS0978 with primers oxp4354/4355 cloned into pOGG276 digested with XbaI.	This work
pOPS1594	P*fnrN* fused to *syfp2* amplified from pOPS0978 with primers oxp4354/4355 cloned into pOGG276 digested with XbaI.	This work
pOPS1607	Reporter plasmid assembled into the pOGG276 backbone by Golden Gate assembly: J23106 promoter, RBStd, *syfp2*, DT16 terminator.	This work
pOPS1644	pJP2 digested with HindIII and XbaI to remove *gfp* reporter, replaced with Rlv3841 P*fixNOQP*$_{10}$ fused to *syfp2*, amplified from pOPS0977 with oxp4550/4551; Tet^R Amp^R	This work
pTNS3	*E. coli* vector expressing *tnsABCD* from PI and lac promoters. Enables Tn7 insertions at the *glmS* site; Amp^R	Choi et al. 2008 [10]
pRK2013	Helper plasmid; mob^+, Kan^R	Ditta et al. 1980 [11]

3. Primers

Name	Description	Sequence (5′-3′)
oxp0283	Forward mapping primer for pOPS0786-based reporter plasmids	AGCGTTCTGAACAAATCC
oxp1331	Reverse mapping primer for pOPS0786-based reporter plasmids	TTTTGAAGACAAAAGCTTATTATTTATACAGCTCATCCATACCCAG
oxp0287	Forward primer for amplification of Rlv3841 P*fixK*$_{9a}$ for cloning into pIJ11268	TTTTGGTACCGATGTCGTCCCCAGTG
oxp0288	Reverse primer for amplification of Rlv3841 P*fixK*$_{9a}$ for cloning into pIJ11268	AAAAGGATCCTGGAACGCCTCTGC
oxp2327	Forward mapping primer for Tn7 integrations into Rlv3841	GATGATCTTCTCGCTGCCGA
oxp2328	Reverse mapping primer for Tn7 integrations into Rlv3841	GCTCTGGCCAATGAGGTTCT
oxp2874	Forward for amplicon upstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	GTCGACTCTAGAGGATCCCCTTCGGGATCATTGGCGCTG
oxp2875	Reverse for amplicon upstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	CGGTGAAGACGTAGCAGTACTCGTCCTCGAAATAGCGCGTCAG
oxp2876	Forward for amplicon downstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	GCTATTTCGAGGACGAGTACTGCTACGTCTTCACCGCCAG
oxp2877	Reverse for amplicon downstream of *fxkR*$_9$, for cloning into pK19mobSacB and markerless mutant generation	TGAATTCGAGCTCGGTACCCTCTTCGGACAGCACATTGAG
oxp3039	Reverse primer for amplification of P*fixNOQP*$_{10}$ from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGATGTCGTCCCCAGTACGCC
oxp3040	Forward primer for amplification of P*fixNOQP*$_{10}$ from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTTACGGCGGCCGCGACAGC
oxp3041	Forward primer for amplification of P*fixNOQP*$_9$ from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGATGTCGTCCCCAGTGCG
oxp3042	Reverse primer for amplification of P*fixNOQP*$_9$ from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTGGAACGCCTCTGCGTCAC
oxp3043	Forward primer for amplification of P*nifH* from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGTTTGGCGTTCCTTCATGTGTTC
oxp3044	Reverse primer for amplification of P*nifH* from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGTCGATGCTGACCGCCTGATC
oxp3045	Reverse primer for amplification of P*fnrN* from Rlv3841 for cloning into pOPS0786	CTTGCTAACCATTTGGTCCTGATCCCTTTTGAAATCCT
oxp3046	Forward primer for amplification of P*fnrN* from Rlv3841 for cloning into pOPS0786	GTGGAGATCTAGAAGGCGCTGTACCTCATGAAAT
oxp3062	Forward mapping primer for pOGG280	GAGCGCTTTTGAAGCTAATTCGA

oxp3063	Reverse mapping primer for pOGG280	TCACTTATCTGGTTGGCCTGC
oxp3115	Forward mapping primer for *fxkR*$_9$ mutagenesis	GGTCGTTGTCTCCAGGCGCG
oxp3156	Reverse mapping primer for *fxkR*$_9$ mutagenesis	TGCGCAGTGGTTGGCTAGGC
oxp4115	Forward primer for amplification of Rlv3841 *fnrN* gene for cloning into pOGG280	TAATGCCGAATTCGGATCCCG CGCTGTACCTCATGAAATG
oxp4116	Reverse primer for amplification of Rlv3841 *fnrN* gene for cloning into pOGG280	CTATCAACAGGAGTCCAAGTA TGCGCTGATCATCCGCTC
oxp4354	Forward primer for amplification of Rlv3841 promoters fused to *syfp2* in pOPS0786	AATTCGGATCCGGAGTCGGTC ACATGTGCATC
oxp4355	Reverse primer for amplification of Rlv3841 promoters fused to *syfp2* in pOPS0786	AGGAGTCCAAGAGCGGGTCG AAAAAAAAAGCCCG
oxp4550	Forward primer for amplification of P*fixNOQP*$_{10}$ fused to *syfp2* from pOPS0977	GTCCGGGTACCATGGATCCAT TACGGCGGCCGCGACAG
oxp4551	Reverse primer for amplification of PfixNOQP10 fused to syfp2 from pOPS0977	CGGACCATGATTACCTCAGTG GTCGAAAAAAAAAGCCCGCA CTGTC
pK19A	Reverse mapping primer for *hfixL*$_c$ mutagenesis, binds in pK19mob	ATCAGATCTTGATCCCCTGC
pr0482	Forward mapping primer for *hfixL*$_c$ mutagenesis, binds in genome	AGTTCGATGTTCGTATCCGAA C
pr0988	Forward primer for amplification of Rlv3841 *hfixL*$_c$ internal fragment for cloning into pK19mob	GCAGGTCGACTCTAGATGGA AGAGCTTCGGACCGAA
pr0989	Reverse primer for amplification of Rlv3841 *hfixL*$_c$ internal fragment for cloning into pK19mob	CCGGGGATCCTCTAGAATATC TCGATCGTCAGACGG
pr1270	Forward primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pJQ200SK	CTCGAGGCTACATCGACCACT ATCTC
pr1271	Reverse primer for amplification of Rlv3841 *hfixL*$_9$ for cloning into pJQ200SK	TCTAGAACACGGGCGTCATCT TCGAC
pr1272	Forward mapping primer for *hfixL*$_9$ mutagenesis	CGGAAGAGCTTCCACGATGA
pr1273	Reverse mapping primer for *hfixL*$_9$ mutagenesis	GCCGTCCGCACCTGTCGTTC
pr1381	Forward primer for amplification of Rlv3841 *fnrN* gene for cloning into pJQ200SK	GCCTAAAGCGCGTCTGGTTC
pr1382	Reverse primer for amplification of Rlv3841 *fnrN* gene for cloning into pJQ200SK	AATAAGCCTGCGGCGCATCC
pr1432	Forward mapping primer for *fnrN* mutagenesis	CTGGGCCATGGTCTCGATCA
pr1433	Reverse mapping primer for *fnrN* mutagenesis	CATAATCTCGGCACCATGGC

References

1. Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 1983;166: 557–580. doi:10.1016/S0022-2836(83)80284-8

2. Thoma S, Schobert M. An improved *Escherichia coli* donor strain for diparental mating. FEMS Microbiol Lett. 2009;294: 127–132. doi:10.1111/j.1574-6968.2009.01556.x

3. Frederix M, Edwards A, Swiderska A, Stanger A, Karunakaran R, Williams A, et al. Mutation of *praR* in *Rhizobium leguminosarum* enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins. Mol Microbiol. 2014;93: 464–478. doi:10.1111/mmi.12670

4. Johnston AWB, Behringer JE. Identification of the *Rhizobium* strains in pea root nodules using genetic markers. J Gen Microbiol. 1975;87: 343–350. doi:10.1099/00221287-87-2-343

5. Fellay R, Frey J, Krisch H. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of Gram-negative bacteria. Gene. 1987;52: 147–154. doi:10.1016/0378-1119(87)90041-2

6. Prell J, Mulley G, Haufe F, White JP, Williams A, Karunakaran R, et al. The PTS Ntr system globally regulates ATP-dependent transporters in *Rhizobium leguminosarum*. Mol Microbiol. 2012;84: 117–129. doi:10.1111/j.1365-2958.2012.08014.x

7. Quandt J, Hynes MF. Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria. Gene. 1993;127: 15–21. doi:10.1016/0378-1119(93)90611-6

8. Schäfer A, Tauch A, Jäger W, Kalinowski J, Thierbach G, Pühler A. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene. 1994;145: 69–73. doi:10.1016/0378-1119(94)90324-7

9. Kirchner O, Tauch A. Tools for genetic engineering in the amino acid-producing bacterium *Corynebacterium glutamicum*. J Biotechnol. 2003;104: 287–299. doi:10.1016/S0168-1656(03)00148-2

10. Choi KH, Mima T, Casart Y, Rholl D, Kumar A, Beacham IR, et al. Genetic tools for select-agent-compliant manipulation of *Burkholderia pseudomallei*. Appl Environ Microbiol. 2008;74: 1064–1075. doi:10.1128/AEM.02430-07

11. Ditta G, Stanfield S, Corbin D, Helinski DR. Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc Natl Acad Sci. 1980;77: 7347–7351. doi:10.1073/pnas.77.12.7347

B.4 Supplementary 4: Data tables for figures 2-5 and S1, S2

Figure 2

All values given as Fluo/OD_{600}.

21% O_2				
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*	P*nifH*
1	1157	1438	8009	3500
2	1025	1212	10446	1934
3	1001	1282	4173	3447

1% O_2				
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*	P*nifH*
1	5504	6924	105397	4207
2	5906	6997	116921	4362
3	5090	6870	84631	3923

Figure 3

All values given as % of the WT average for each respective promoter (columns).

WT			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	113	95	103
2	87	105	97
3	107	115	112
4	93	85	88
5	82	97	92
6	118	103	108
7	87	84	82
8	113	116	118
9	84	92	-
10	116	108	-
11	98	96	-
12	102	104	-

***hfixL*$_9$ *hfixL*$_c$ (LMB496)**			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	22	4	32
2	20	3	25
3	14	3	26
4	13	3	18

***hfixL*$_9$ (LMB495)**			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	28	14	43
2	24	12	31
3	18	12	34
4	14	12	22

***hfixL*$_c$ (LMB403)**			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	109	137	175
2	87	112	115
3	48	106	154
4	42	110	123

Figure 3 data continues on next page.

Figure 3 (continued)

All values given as % of the WT average for each respective promoter (columns).

***fxkR*$_9$ (OPS1808)**			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	21	18	4
2	13	13	3
3	18	17	5
4	14	15	2
5	37	26	-
6	21	19	-
7	30	29	-
8	23	21	-

***fnrN* (LMB648)**			
Replicate no.	P*fixNOQP*$_9$	P*fixNOQP*$_{10}$	P*fnrN*
1	73	99	114
2	76	90	99
3	80	83	106
4	62	80	96

Figure 4

All values given as Fluo/OD_{600}.

P*fnrN* (A)			
Replicate no.	WT	*fnrN* (LMB648)	*hfixL*$_9$ *hfixL*$_c$ (LMB496)
1	13040	1413	4394
2	7306	2868	2014
3	12329	1653	4897
4	13772	2800	1616
5	-	4324	3538

P*fixNOQP*$_9$ (B)			
Replicate no.	WT	*fnrN* (LMB648)	*hfixL*$_9$ *hfixL*$_c$ (LMB496)
1	45572	1947	28891
2	42560	2248	29434
3	43104	2357	29324
4	45162	-	31417
5	-	-	32920

P*fixNOQP*$_{10}$ (C)			
Replicate no.	WT	*fnrN* (LMB648)	*hfixL*$_9$ *hfixL*$_c$ (LMB496)
1	92192	3200	58844
2	71749	4152	43836
3	84645	4796	50489
4	83645	4738	50308
5	85324	4451	40259

Figure 5

All values given as % of the WT average. Note decimals have been eliminated from all strains except LMB673 for clarity.

Replicate no.	**WT**	***hfixL*$_9$ *hfixL*$_c$ (LMB496)**	***fnrN* (LMB648)**	***fxkR*$_9$ (OPS1808)**	***hfixL*$_9$ (LMB495)**	***hfixL*$_c$ (LMB403)**	***fnrN hfixL*$_9$ *hfixL*$_c$ (LMB673)**
1	113	79	21	76	68	42	0.179
2	99	95	13	89	69	101	0.269
3	105	71	9	66	71	85	0.269
4	116	60	18	80	60	69	0.269
5	112	66	14	90	89	73	0.179
6	92	-	-	68	-	-	-
7	87	-	-	100	-	-	-
8	76	-	-	-	-	-	-
9	75	-	-	-	-	-	-
10	103	-	-	-	-	-	-
11	104	-	-	-	-	-	-
12	117	-	-	-	-	-	-
13	101	-	-	-	-	-	-

Figure S1

All values given as Luminescence/OD_{600}.

Replicate no.	**WT**	***hfixL*$_9$ *hfixL*$_c$ (LMB496)**	**WT w/o reporter**
1	734	150	526
2	814	364	374
3	886	665	610
4	1108	861	430
5	1045	-	490

Figure S2

All values given as µmoles ethylene jar^{-1} $hour^{-1}$ g of $nodules^{-1}$

Replicate no.	**WT**	***fnrN* (LMB648)**	***fnrN* complemented (OPS2260)**
1	14.395	2.783	15.074
2	16.117	4.044	13.961
3	16.371	2.819	16.040
4	14.220	3.287	9.609
5	16.081	3.193	14.590
6	16.116	3.565	15.330
7	18.392	2.725	13.298

B.5 Supplementary 5: Figure 8 coordinates

This supplementary document is an Excel spreadsheet that cannot be included in this thesis but is available online from the publisher at this link.

B.6 Supplementary 6: TSS materials & methods

Origin of samples

Pisum sativum cv. Avola was inoculated and grown with Rlv841 as described in the Materials and Methods section of the main text. After 28 days, roots were washed in sterile distilled water and nodules were picked and flash frozen in liquid nitrogen. Bacteroids were isolated by grinding in sterile isolation buffer and separated from plant material by differential centrifugation as described in the Materials and Methods section of the main text. The pellet containing bacteroids was resuspended in RNAlater solution before total RNA extraction as described previously [1].

Differential RNA sequencing (dRNAseq) library preparation and sequencing

Each total RNA sample was quantified with a Qubit 2.0 using the RNA HS assay (Invitrogen Q32852) and then treated with Turbo RNase-free DNase (Ambion AM1907) to remove any genomic DNA contamination. Genomic DNA elimination was confirmed with a Qubit 2.0 using the DNA BR assay (Invitrogen Q32850), followed by a new RNA quantification using the RNA HS assay. Samples were shipped to Vertis Biotechnologies GmbH for preparation of strand specific dRNA sequencing libraries to be sequenced with the Illumina HiSeq2000 platform, as described in [2]. RNA samples were analysed for quality using a MultiNA microchip electrophoresis system (Shimadzu) and ribosomal RNA (rRNA) was depleted with a custom made rRNA capture solution containing a 5:1 ratio of Ribo-Zero Gram-negative Bacteria (Illumina MRZGN126) : Ribo-Zero Plant/Leaf kit (Illumina MRZPL116). cDNA libraries were constructed by fragmenting the RNA with RNaseIII (NEB), and fragments were then poly(A) tailed using *E. coli* poly(A) polymerase (NEB) and any 5' mono-phosphate RNA species were degraded using terminator exonuclease (Epicentre TER51020). The poly(A)

tailed RNA molecules were recovered using oligo(dT) probes bound to magnetic beads. Each sample was then split, and one half of each sample was treated with RNA 5' polyphosphatase (Epicentre RP8092H) to convert 5' triphosphate to 5' monophosphate. The other half of the sample was left untreated. Both halves of each sample were treated equally from this point onwards. An RNA adapter (5'-UUUCCCUACACGACGCUCUUCCGAUCU-3') was ligated to the 5' monophosphorylated RNA. First strand cDNA was synthesised from RNA using an oligo(dT) adapter specific primer with M-MLV reverse transcriptase (Affinityscript Agilent). The cDNA was PCR amplified using HiFi DNA polymerase (Herculase II Fusion DNA polymerase, Agilent) and primers which added a sample specific DNA barcode:

TrueSeq_Sense_primer

5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNN-3'

TrueSeq_Antisense_primer (barcode)

5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNN-TGACTGGAGTTCAGACGTGTGCTCTTCCGATC(dT25)-3')

The cDNA was then purified using an Agencourt AMPure XP kit (Beckman Coulter 10136224). The quantity and quality of the cDNA was analysed using capillary electrophoresis on the MultiNA system. Equimolar amounts of the cDNA from each sample were pooled and sequenced on an Illumina HiSeq2000 lane, with a 100 bp read length. Further details of the method are described in [3].

Read Mapping and TSS identification:

Approximately 87 million reads were sequenced across all samples. The FastQ file from the Illumina HiSeq2000 sequencing run was quality checked using FASTQC and the adapter and multiplex barcode removed by cutadapt [4]. Reads were then mapped to the Rlv3841 genome using Bowtie2 with default parameters [5]. The mapped alignment file (in SAM format) was then sorted, indexed and split it into the separate replicons (Chromosome, pRL7-12), before unique reads were filtered for each replicon and converted to BAM file format using Samtool [6]. The Minus/Plus BAM files were analysed using TSSAR (http://rna.tbi.univie.ac.at/TSSAR), a transcription start site prediction software for dRNAseq data [7]. The output files were downloaded in BED format, with the following parameters (p-value = $1e^{-15}$, noise threshold = 2, merge range = 5). The BED format file had the nucleotide coordinate and a confidence score (out of 1000) for each predicted transcription start site. These BED files alongside the Rlv3841 genome (in FASTA format) and gene annotations (in GFF3 or BED format) were viewed on Integrated Genome Viewer [8]. The complete raw dataset was submitted to the SRA database, with BioProject accession number PRJNA667846.

References

1. Karunakaran R, Ramachandran VK, Seaman JC, East AK, Mouhsine B, Mauchline TH, et al. Transcriptomic analysis of *Rhizobium leguminosarum* biovar *viciae* in symbiosis with host plants *Pisum sativum* and *Vicia cracca*. J Bacteriol. 2009;191: 4002–4014. doi:10.1128/JB.00165-09

2. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiß S, Sittka A, et al. The primary transcriptome of the major human pathogen *Helicobacter pylori*. Nature. 2010;464: 250–255. doi:10.1038/nature08756

3. Bischler T, Tan HS, Nieselt K, Sharma CM. Differential RNA-seq (dRNA-seq) for annotation of transcriptional start sites and small RNAs in *Helicobacter pylori*. Methods. 2015;86: 89–101. doi:10.1016/j.ymeth.2015.06.012

4. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17: 10. doi:10.14806/ej.17.1.200

5. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359. doi:10.1038/nmeth.1923

6. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–2079. doi:10.1093/bioinformatics/btp352

7. Amman F, Wolfinger MT, Lorenz R, Hofacker IL, Stadler PF, Findeiß S. TSSAR: TSS annotation regime for dRNA-seq data. BMC Bioinformatics. 2014;15: 89. doi:10.1186/1471-2105-15-89

8. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. Brief Bioinform. 2013;14: 178–192. doi:10.1093/bib/bbs017

References

[1] Paul J. Rutten and Philip S. Poole. "Oxygen regulatory mechanisms of nitrogen fixation in rhizobia". In: *Adv. Microb. Physiol.* Ed. by Robert K. Poole. Vol. 75. Advances in Microbial Physiology. Academic Press, 2019, pp. 325–389. URL: http://www.sciencedirect.com/science/article/pii/S0065291119300293.

[2] Paul J. Rutten et al. "Multiple sensors provide spatiotemporal oxygen regulation of gene expression in a Rhizobium-legume symbiosis". In: *PLOS Genet.* 17.2 (Feb. 2021). Ed. by Sean Crosson, e1009099. URL: https://dx.plos.org/10.1371/journal.pgen.1009099.

[3] Philip Poole, Vinoy Ramachandran, and Jason Terpolilli. "Rhizobia: from saprophytes to endosymbionts". In: *Nat. Rev. Microbiol.* 16.5 (Jan. 2018), pp. 291–303. URL: http://dx.doi.org/10.1038/nrmicro.2017.171.

[4] Giles E. D. Oldroyd et al. "The Rules of Engagement in the Legume-Rhizobial Symbiosis". In: *Annu. Rev. Genet.* 45.1 (2011), pp. 119–144. URL: http://www.annualreviews.org/doi/10.1146/annurev-genet-110410-132549.

[5] Gordon Conway. "Food for All in the 21st Century". In: *Environ. Sci. Policy Sustain. Dev.* 42.1 (Jan. 2000), pp. 8–18. URL: https://www.tandfonline.com/doi/abs/10.1080/00139150009604857.

[6] John R. Postgate. "New Advances and Future Potential in Biological Nitrogen Fixation". In: *J. Appl. Bacteriol.* 37.2 (June 1974), pp. 185–202. URL: http://doi.wiley.com/10.1111/j.1365-2672.1974.tb00431.x.

[7] Florence Mus et al. "Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes". In: *Appl. Environ. Microbiol.* 82.13 (July 2016), pp. 3698–3710. URL: http://www.ncbi.nlm.nih.gov/pubmed/27084023.

[8] Stefan Burén, Gema López-Torrejón, and Luis M. Rubio. "Extreme bioengineering to meet the nitrogen challenge". In: *Proc. Natl. Acad. Sci. U. S. A.* 115.36 (Sept. 2018), pp. 8849–8851. URL: http://www.ncbi.nlm.nih.gov/pubmed/30115666.

[9] Achim Dobermann and Kenneth G. Cassman. "Cereal area and nitrogen use efficiency are drivers of future nitrogen fertilizer consumption". In: *Sci. China. Ser. C, Life Sci.* 48.745 (2005), pp. 745–758. URL: https://link.springer.com/content/pdf/10.1007%2FBF03187115.pdf.

[10] Emilio J. Vicente and Dennis R. Dean. "Keeping the nitrogen-fixation dream alive". In: *Proc. Natl. Acad. Sci.* 114.12 (Mar. 2017), pp. 3009–3011. URL: http://www.ncbi.nlm.nih.gov/pubmed/28283657.

[11] Donald E. Canfield, Alexander N. Glazer, and Paul G. Falkowski. "The Evolution and Future of Earth's Nitrogen Cycle". In: *Science (80-.).* 330.6001 (Oct. 2010), pp. 192–196. URL: http://www.sciencemag.org/lookup/doi/10.1126/science.1186120.

[12] Nicolas Gruber and James N. Galloway. "An Earth-system perspective of the global nitrogen cycle." In: *Nature* 451.7176 (Jan. 2008), pp. 293–296. URL: http://www.nature.com/doifinder/10.1038/nature06592.

[13] Vaclav Smil. *Enriching the Earth: Fritz Haber, Carl Bosch, and the Transformation of World Food Production*. Boston: MIT Press, 2001. URL: https://mitpress.mit.edu/books/enriching-earth.

[14] Jan W. Erisman et al. "How a century of ammonia synthesis changed the world". In: *Nat. Geosci.* 1.10 (Oct. 2008), pp. 636–639. URL: http://www.nature.com/articles/ngeo325.

[15] Peter M. Vitousek et al. "Human Alteration of the Global Nitrogen Cycle: Sources and Consequences". In: *Source Ecol. Appl. Ecol. Appl. Ecol. Appl.* 7.3 (Aug. 1997), pp. 737–750. URL: https://doi.org/10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2.

[16] Peter M. Vitousek et al. "Biological nitrogen fixation: Rates, patterns and ecological controls in terrestrial ecosystems". In: *Philos. Trans. R. Soc. B Biol. Sci.* 368.1621 (2013), pp. 1–9. URL: http://dx.doi.org/10.1098/rstb.2013.0119.

[17] Edward C. Cocking, Philip J. Stone, and Michael R. Davey. "Symbiosome-like intracellular colonization of cereals and other crop plants by nitrogen-fixing bacteria for reduced inputs of synthetic nitrogen fertilizers." In: *Sci. China. Ser. C, Life Sci.* 48.4 (Sept. 2005), pp. 888–896. URL: https://pubmed.ncbi.nlm.nih.gov/20549443/.

[18] Christian Rogers and Giles E. D. Oldroyd. "Synthetic biology approaches to engineering the nitrogen symbiosis in cereals". In: *J. Exp. Bot.* 65.8 (2014), pp. 1939–1946. URL: https://pubmed.ncbi.nlm.nih.gov/24687978/.

[19] Gema López-Torrejón et al. "Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast". In: *Nat. Commun.* 7 (Apr. 2016), p. 11426. URL: http://www.ncbi.nlm.nih.gov/pubmed/27126134.

[20] Stefan Burén and Luis M. Rubio. "State of the art in eukaryotic nitrogenase engineering". In: *FEMS Microbiol. Lett.* 365.2 (Jan. 2018), pp. 1–9. URL: https://academic.oup.com/femsle/article/doi/10.1093/femsle/fnx274/4733273.

[21] Robert S. Allen et al. "Expression of 16 Nitrogenase Proteins within the Plant Mitochondrial Matrix". In: *Front. Plant Sci.* 8 (Mar. 2017), p. 287. URL: http://journal.frontiersin.org/article/10.3389/fpls.2017.00287/full.

[22] Barney A. Geddes et al. "Use of plant colonizing bacteria as chassis for transfer of N fixation to cereals". In: *Curr. Opin. Biotechnol.* 32 (Apr. 2015), pp. 216–22. URL: http://dx.doi.org/10.1016/j.copbio.2015.01.004.

[23] Giles E. D. Oldroyd and Ray Dixon. "Biotechnological solutions to the nitrogen problem". In: *Curr. Opin. Biotechnol.* 26 (Apr. 2014), pp. 19–24. URL: http://dx.doi.org/10.1016/j.copbio.2013.08.006.

[24] Nikos Alexandratos and Jelle Bruinsma. "World agriculture towards 2030/2050". In: *Land use policy* 20.4 (2012), p. 375. URL: www.fao.org/3/ap106e/ap106e.pdf.

[25] Allan Downie. "Legume nodulation". In: *Curr. Biol.* 24.5 (Mar. 2014), R184–R190. URL: http://www.ncbi.nlm.nih.gov/pubmed/24602880.

[26] Jeremy D. Murray. "Invasion by Invitation: Rhizobial Infection in Legumes". In: *Mol. Plant-Microbe Interact.* 24.6 (June 2011), pp. 631–639. URL: https://apsjournals.apsnet.org/doi/pdfplus/10.1094/MPMI-08-10-0181.

[27] Mitchell Andrews and Morag E. Andrews. "Specificity in Legume-Rhizobia Symbioses". In: *Int. J. Mol. Sci.* 18.4 (Mar. 2017), p. 705. URL: http://www.mdpi.com/1422-0067/18/4/705.

[28] Rachel M. Wheatley and Philip S. Poole. "Mechanisms of bacterial attachment to roots". In: *FEMS Microbiol. Rev.* 42.4 (Apr. 2018), pp. 448–461. URL: https://academic.oup.com/femsre/article/42/4/448/4975273.

[29] Daniel J. Gage. "Analysis of Infection Thread Development Using Gfp- and DsRed-Expressing Sinorhizobium meliloti". In: *J. Bacteriol.* 184.24 (Dec. 2002), pp. 7042–7046. URL: http://jb.asm.org/cgi/doi/10.1128/JB.184.24.7042-7046.2002.

[30] Giles E. D. Oldroyd and Allan Downie. "Coordinating nodule morphogenesis with rhizobial infection in legumes". In: *Annu. Rev. Plant Biol.* 59.1 (2008), pp. 519–546. URL: https://www.annualreviews.org/doi/pdf/10.1146/annurev.arplant.59.032607.092839.

[31] Carroll P. Vance. "Rhizobium Infection and Nodulation: A Beneficial Plant Disease?" In: *Annu. Rev. Microbiol.* 37.1 (Oct. 1983), pp. 399–424. URL: http://www.annualreviews.org/doi/pdf/10.1146/annurev.mi.37.100183.002151.

[32] Brady Bassett, Robert N. Goodman, and Anton Novacky. "Ultrastructure of soybean nodules. I: release of rhizobia from the infection thread". In: *Can. J. Microbiol.* 23.5 (May 1977), pp. 573–582. URL: http://www.nrcresearchpress.com/doi/10.1139/m77-083.

[33] L. Evans Roth and Gary Stacey. "Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources." In: *Eur. J. Cell Biol.* 49.1 (June 1989), pp. 13–23. URL: http://www.ncbi.nlm.nih.gov/pubmed/2759097.

[34] Francis M. Martin, Stéphane Uroz, and David G. Barker. "Ancestral alliances: plant mutualistic symbioses with fungi and bacteria". In: *Science* 356.6340 (May 2017), pp. 1–9. URL: http://science.sciencemag.org/content/sci/356/6340/eaad4501.full.pdf.

[35] Jürgen Prell et al. "Legumes regulate Rhizobium bacteroid development and persistence by the supply of branched-chain amino acids". In: *Proc. Natl. Acad. Sci.* 106.30 (July 2009), pp. 12477–12482. URL: http://www.ncbi.nlm.nih.gov/pubmed/19597156.

[36] Jacques Batut and Pierre Boistard. "Oxygen control in Rhizobium". In: *Antonie Van Leeuwenhoek* 66.1-3 (1994), pp. 129–150. URL: http://link.springer.com/10.1007/BF00871636.

[37] Cyril A. Appleby. "Leghemoglobin and Rhizobium respiration". In: *Annu. Rev. Plant Physiol.* 35.1 (June 1984), pp. 443–478. URL: http://www.annualreviews.org/doi/pdf/10.1146/annurev.pp.35.060184.002303.

[38] Hans-Martin Fischer. “Genetic regulation of nitrogen fixation in rhizobia”. In: *Microbiol. Rev.* 58.3 (1994), pp. 352–386. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372973/pdf/microrev00022-0070.pdf.

[39] Lance C. Seefeldt, Brian M. Hoffman, and Dennis R. Dean. “Mechanism of Mo-dependent nitrogenase”. In: *Annu. Rev. Biochem.* 78.1 (June 2009), pp. 701–722. URL: https://www.annualreviews.org/doi/pdf/10.1146/annurev.biochem.78.070907.103812.

[40] Luis M. Rubio and Paul W. Ludden. “Maturation of Nitrogenase: a Biochemical Puzzle”. In: *J. Bacteriol.* 187.2 (Jan. 2005), pp. 405–414. URL: http://jb.asm.org/cgi/doi/10.1128/JB.187.2.405-414.2005.

[41] Ray Dixon and Daniel Kahn. “Genetic regulation of biological nitrogen fixation”. In: *Nat. Rev. Microbiol.* 2.8 (2004), pp. 621–631. URL: http://www.ncbi.nlm.nih.gov/pubmed/15263897.

[42] Barbara K. Burgess and David J. Lowe. “Mechanism of molybdenum nitrogenase”. In: *Chem. Rev.* 96.7 (1996), pp. 2983–3012. URL: https://pubs.acs.org/doi/pdfplus/10.1021/cr950055x.

[43] Yilin Hu and Markus W. Ribbe. “Nitrogenase assembly”. In: *Biochim. Biophys. Acta - Bioenerg.* 1827.8-9 (2013), pp. 1112–1122. URL: http://dx.doi.org/10.1016/j.bbabio.2012.12.001.

[44] Vinod K. Shah and Winston J. Brill. “Isolation of an iron-molybdenum cofactor from nitrogenase*”. In: *Proc. Natl. Acad. Sci.* 74.8 (1977), pp. 3249–3253. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC431518/pdf/pnas00030-0151.pdf.

[45] Millie Georgiadis et al. “Crystallographic structure of the nitrogenase iron protein from Azotobacter vinelandii”. In: *Science* 257.5077 (Sept. 1992), pp. 1653–1659. URL: http://www.ncbi.nlm.nih.gov/pubmed/1529353.

[46] James A. Imlay. “Iron-sulphur clusters and the problem with oxygen”. In: *Mol. Microbiol.* 59.4 (Feb. 2006), pp. 1073–1082. URL: http://doi.wiley.com/10.1111/j.1365-2958.2006.05028.x.

[47] Ruud A. De Maagd et al. “Down-Regulation of Expression of the Rhizobium leguminosarum Outer Membrane Protein Gene ropA Occurs Abruptly in Interzone ll-lll of Pea Nodules and Can Be Uncoupled from nif Gene Activation”. In: *Mol. Plant-Microbe Interact.* 7.2 (1994), p. 276. URL: http://www.apsnet.org/publications/mpmi/backissues/Documents/1994Abstracts/Microbe07-276.htm.

[48] Brian D. Shaw. “Oxygen control mechanisms in nitrogen-fixing systems”. In: *Curr. Dev. Biol. Nitrogen Fixat.* Cambridge: Cambridge University Press, 1984. Chap. 5, pp. 111–135. URL: http://agris.fao.org/agris-search/search.do?recordID=US201302645498.

[49] Dario Leister. “Thawing out frozen metabolic accidents”. In: *BMC Biol.* 17.1 (Dec. 2019), p. 8. URL: https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-018-0621-5.

[50] Kathleen Marchal and Jos Vanderleyden. “The "oxygen paradox" of dinitrogen-fixing bacteria”. In: *Biol. Fertil. Soils* 30.5-6 (Mar. 2000), pp. 363–373. URL: https://link.springer.com/content/pdf/10.1007/s003740050017.pdf.

[51] Rhesa N. Ledbetter et al. "The Electron Bifurcating FixABCX Protein Complex from Azotobacter vinelandii: Generation of Low-Potential Reducing Equivalents for Nitrogenase Catalysis". In: *Biochemistry* 56.32 (Aug. 2017), pp. 4177–4190. URL: http://pubs.acs.org/doi/10.1021/acs.biochem.7b00389.

[52] Tomas Edgren and Stefan Nordlund. "The fixABCX Genes in Rhodospirillum rubrum Encode a Putative Membrane Complex Participating in Electron Transfer to Nitrogenase". In: *J. Bacteriol.* 186.7 (Apr. 2004), pp. 2052–2060. URL: http://www.ncbi.nlm.nih.gov/pubmed/15028689.

[53] Michael Udvardi and Philip S. Poole. "Transport and Metabolism in Legume-Rhizobia Symbioses". In: *Annu. Rev. Plant Biol.* 64.1 (Apr. 2013), pp. 781–805. URL: http://www.annualreviews.org/doi/10.1146/annurev-arplant-050312-120235.

[54] David B. Layzell et al. "The regulation of oxygen and its role in regulating nodule metabolism". In: Dordrecht: Springer, 1993, pp. 393–398. URL: https://link.springer.com/chapter/10.1007%2F978-94-017-2416-6_39.

[55] Andrés R. Schwember et al. "Regulation of Symbiotic Nitrogen Fixation in Legume Root Nodules". In: *Plants* 8.9 (Sept. 2019), p. 333. URL: http://www.ncbi.nlm.nih.gov/pubmed/31489914.

[56] Frank R. Minchin. "Regulation of oxygen diffusion in legume nodules". In: *Soil Biol. Biochem.* 29.516 (1997), pp. 88–89. URL: https://www.sciencedirect.com/science/article/abs/pii/S0038071796002040.

[57] John F. Witty and Frank R. Minchin. "Oxygen diffusion in the legume root nodule". In: *Nitrogen Fixat.* Boston: Springer, 1990, pp. 285–292. URL: https://link.springer.com/chapter/10.1007%2F978-1-4684-6432-0_29.

[58] John F. Witty et al. "Nitrogen fixation and oxygen in legume root nodules". English. In: *Oxford Surv. Plant Mol. Cell Biol.* Vol. 3. Oxford: Oxford University Press, 1986, pp. 275–314. URL: http://hdl.handle.net/2160/43314.

[59] Bryan J. King et al. "Regulation of O2 concentration in soybean nodules observed by in situ spectroscopic measurement of leghemoglobin oxygenation". In: *Plant Physiol.* 87.2 (June 1988), pp. 296–9. URL: http://www.ncbi.nlm.nih.gov/pubmed/16666136.

[60] Monica M. Kuzma, Stephen Hunt, and David B. Layzell. "Role of oxygen in the limitation and inhibition of nitrogenase activity and respiration rate in individual soybean nodules". In: *Plant Physiol.* 101.1 (Jan. 1993), pp. 161–169. URL: http://www.ncbi.nlm.nih.gov/pubmed/12231675.

[61] Felix D. Dakora and Craig A. Atkins. "Diffusion of Oxygen in Relation to Structure and Function in Legume Root Nodules". In: *Funct. Plant Biol.* 16.1 (1989), p. 131. URL: https://www.researchgate.net/publication/262994739.

[62] John Draper. "Nitrogen Fixation and oxygen in legume root nodules". In: *Plant, Cell Environ.* 9.4 (June 1986), pp. 353–354. URL: http://agris.fao.org/agris-search/search.do?recordID=US201301759455.

[63] John F. Witty, Leif Skøt, and Niels P. Revsbech. "Direct evidence for changes in the resistance of legume root nodules to O2 diffusion". In: *J. Exp. Bot.* 38.7 (1987), pp. 1129–1140. URL: https://academic.oup.com/jxb/article-abstract/38/7/1129/440716.

[64] Fraser J. Bergersen. "Delivery of O2 to bacteroids in soybean nodule cells: consideration of gradients of concentration of free, dissolved O2 in and near symbiosomes and beneath intercellular spaces". In: *Protoplasma* 191.1-2 (Mar. 1996), pp. 9–20. URL: https://link.springer.com/article/10.1007%2FBF01280821.

[65] Hui Wei and David B. Layzell. "Adenylate-coupled ion movement. A mechanism for the control of nodule permeability to O2 diffusion". In: *Plant Physiol.* 141.1 (May 2006), pp. 280–7. URL: http://www.ncbi.nlm.nih.gov/pubmed/12226357.

[66] Frank R. Minchin, John E. Sheehy, and John F. Witty. "Factors Limiting N2 Fixation by the Legume-Rhizobium Symbiosis". In: *Nitrogen Fixat. Res. Prog.* Dordrecht: Martinus Nijhoff Publishers, 1985, pp. 285–291. URL: https://link.springer.com/chapter/10.1007/978-94-009-5175-4_40.

[67] Peter P. Thumfort, Craig A. Atkins, and David B. Layzell. "A Re-Evaluation of the Role of the Infected Cell in the Control of O2 Diffusion in Legume Nodules". In: *Plant Physiol.* 105.4 (Aug. 1994), pp. 1321–1333. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.105.4.1321.

[68] Ford Denison and Thomas B. Kinraide. "Oxygen-Induced Membrane Depolarizations in Legume Root Nodules". In: *Plant Physiol.* 108.1 (May 1995), pp. 235–240. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.108.1.235.

[69] Frank R. Minchin, Euan K. James, and Manuel Becana. "Oxygen Diffusion, Production Of Reactive Oxygen And Nitrogen Species, And Antioxidants In Legume Nodules". In: *Nitrogen-fixing Legum. Symbioses.* Dordrecht: Springer, 2008, pp. 321–362. URL: https://link.springer.com/chapter/10.1007%2F978-1-4020-3548-7_11.

[70] Allan Downie. "Legume haemoglobins: symbiotic nitrogen fixation needs bloody nodules". In: *Curr. Biol.* 15.6 (Mar. 2005), R196–R198. URL: https://www.sciencedirect.com/science/article/pii/S096098220500268X.

[71] Suman Kundu, James T. Trent, and Mark S. Hargrove. "Plants, humans and hemoglobins". In: *Trends Plant Sci.* 8.8 (2003), pp. 387–393. URL: https://doi.org/10.1016/S1360-1385(03)00163-8.

[72] Thomas Ott et al. "Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development". In: *Curr. Biol.* 15.6 (2005), pp. 531–535. URL: https://pubmed.ncbi.nlm.nih.gov/15797021/.

[73] Kazuya Kawashima et al. "Two types of pea leghemoglobin genes showing different O2-binding affinities and distinct patterns of spatial expression in nodules". In: *Plant Physiol.* 125.2 (Feb. 2001), pp. 641–51. URL: http://www.ncbi.nlm.nih.gov/pubmed/11161022.

[74] Felix D. Dakora. "A functional relationship between leghaemoglobin and nitrogenase based on novel measurements of the two proteins in legume root nodules". In: *Ann. Bot.* 75.1 (1995), pp. 49–54. URL: https://academic.oup.com/aob/article/75/1/49/202710?login=true.

[75] Jonathan B. Wittenberg and Beatrice A. Wittenberg. "Facilitated Oxygen Diffusion by Oxygen Carriers". In: New York: Springer, 2011, pp. 177–199. URL: http://link.springer.com/10.1007/978-1-4612-5890-2_9.

[76] Jonathan B. Wittenberg et al. "Facilitated Oxygen Diffusion: The Role Of Leghemoglobin In Nitrogen Fixation By Bacteroids Isolated From Soybean Root Nodules". In: *J. Biol. Chem.* 249.13 (1974), pp. 4057–4066. URL: http://www.jbc.org/content/249/13/4057.abstract.

[77] John E. Sheehy, Frank R. Minchin, and John F. Witty. "Control of Nitrogen Fixation in a Legume Nodule: an Analysis of the Role of Oxygen Diffusion in Relation to Nodule Structure". In: *Ann. Bot.* 55.4 (Apr. 1985), pp. 549–562. URL: https://academic.oup.com/aob/article-abstract/55/4/549/159746?redirectedFrom=fulltext.

[78] Fraser J. Bergersen. "Distribution of O2 within infected cells of soybean root nodules: a new simulation". In: *Protoplasma* 183.1-4 (Mar. 1994), pp. 49–61. URL: http://link.springer.com/10.1007/BF01276812.

[79] Joachim Schulze. "How are nitrogen fixation rates regulated in legumes?" In: *J. Plant Nutr. Soil Sci.* 167.2 (Apr. 2004), pp. 125–137. URL: http://doi.wiley.com/10.1002/jpln.200320358.

[80] Harvey Millar, David A. Day, and Fraser J. Bergersen. "Microaerobic respiration and oxidative phosphorylation by soybean nodule mitochondria: implications for nitrogen fixation". In: *Plant, Cell Environ.* 18.7 (July 1995), pp. 715–726. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-3040.1995.tb00574.x.

[81] Ivan J. Oresnik, Craig A. Atkins, and David B. Layzell. "The legume symbiosis: C-limited bacteria living within O2 limited plant cells?" In: *Nitrogen Fixat. Fundam. Appl. Proceed. 10th Intl. Cong. Nitrogen Fix., St. Petersbg.* Ed. by Igor A. Tikhonovich et al. Vol. 27. Current Plant Science and Biotechnology in Agriculture. Dordrecht: Springer Netherlands, 1995, p. 601. URL: http://link.springer.com/10.1007/978-94-011-0379-4.

[82] Hwee H. Neo and David B. Layzell. "Phloem glutamine and the regulation of O2 diffusion in legume nodules". In: *Plant Physiol.* 113.1 (Jan. 1997), pp. 259–267. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.113.1.259.

[83] David B. Layzell and Stephen Hunt. "Oxygen and the regulation of nitrogen fixation in legume nodules". In: *Physiol. Plant.* 80.2 (1990), pp. 322–327.

[84] Jesús Montiel et al. "Morphotype of bacteroids in different legumes correlates with the number and type of symbiotic NCR peptides". In: *Proc. Natl. Acad. Sci.* 114.19 (May 2017), pp. 5041–5046. URL: www.pnas.org/cgi/doi/10.1073/pnas.1704217114.

[85] Willem Van de Velde et al. "Plant peptides govern terminal differentiation of bacteria in symbiosis". In: *Science* 327.5969 (Feb. 2010), pp. 1122–1126. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.1184057.

[86] Peter Mergaert et al. "Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis". In: *Proc. Natl. Acad. Sci.* 103.13 (Mar. 2006), pp. 5230–5235. URL: www.pnas.org/cgi/doi/10.1073/pnas.0600912103.

[87] Clemens van de Wiel et al. "The early nodulin transcript ENOD2 is located in the nodule parenchyma (inner cortex) of pea and soybean root nodules". In: *EMBO J.* 9.1 (1990), pp. 1–7. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1460-2075.1990.tb08073.x.

[88] Ryoko Oono, Ford Denison, and Toby Kiers. "Controlling the reproductive fate of rhizobia: How universal are legume sanctions?" In: *New Phytol.* 183.4 (2009), pp. 967–979. URL: https://nph.onlinelibrary.wiley.com/doi/10.1111/j.1469-8137.2009.02941.x.

[89] Cesar Arrese-Igor et al. "Effect of low rhizosphere oxygen on growth, nitrogen fixation and nodule morphology in lucerne". In: *Physiol. Plant.* 89.1 (Sept. 1993), pp. 55–63. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1399-3054.1993.tb01786.x.

[90] Carina Weisbach et al. "Electron microscopic investigation of water occlusions in intercellular spaces in the inner cortex of lucerne nodules". In: *J. Struct. Biol.* 126.1 (1999), pp. 59–71. URL: https://pubmed.ncbi.nlm.nih.gov/10329489/.

[91] Dong Wang et al. "A Nodule-Specific Protein Secretory Pathway Required for Nitrogen-Fixing Symbiosis". In: *Science* 327.5969 (Feb. 2010), pp. 1126–1129. URL: www.sciencemag.org/cgi/content/full/327/5969/1122/DC1.

[92] Keith L. Wycoff et al. "Effects of Oxygen on Nodule Physiology and Expression of Nodulins in Alfalfa". In: *Plant Physiol.* 117.2 (June 1998), pp. 385–395. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.117.2.385.

[93] Jean-Pierre Gourret and Hector Fernandez-Arias. "Etude ultrastructurale et cytochimique de la différenciation des bactéroïdes de Rhizobium trifolii Dangeard dans les nodules de Trifolium repens L." In: *Can. J. Microbiol.* 20.8 (Aug. 1974), pp. 1169–1181. URL: http://www.nrcresearchpress.com/doi/10.1139/m74-181.

[94] Panagiota Mylona, Katharina Pawlowski, and Ton Bisseling. "Symbiotic nitrogen fixation". In: *Plant Cell* 7.7 (July 1995), pp. 869–885. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC160880/pdf/070869.pdf.

[95] Daniel J. Gage and William Margolin. "Hanging by a thread: Invasion of legume plants by rhizobia". In: *Curr. Opin. Microbiol.* 3.6 (Dec. 2000), pp. 613–617. URL: https://www.sciencedirect.com/science/article/pii/S1369527400001491?via%3Dihub.

[96] Jacques Vasse et al. "Correlation between ultrastructural differentiation of bacteriods and nitrogen fixation in alfalfa nodules". In: *J. Bacteriol.* 172.8 (1990), pp. 4295–4306. URL: https://pubmed.ncbi.nlm.nih.gov/2376562/.

[97] Yolanda Hernando et al. "The hypBFCDE operon from Rhizobium leguminosarum biovar viciae is expressed from an Fnr-type promoter that escapes mutagenesis of the fnrN gene". In: *J. Bacteriol.* 177.19 (Oct. 1995), pp. 5661–5669. URL: http://jb.asm.org/lookup/doi/10.1128/jb.177.19.5661-5669.1995.

[98] Eric Soupène et al. "Oxygen as a key developmental regulator of Rhizobium meliloti N2-fixation gene expression within the alfalfa root nodule". In: *Proc. Natl. Acad. Sci.* 92.9 (Apr. 1995), pp. 3759–63. URL: http://www.ncbi.nlm.nih.gov/pubmed/7731979.

[99] Vassily I. Romanov et al. "Anatomy, physiology and biochemistry of root nodules of sprint-2 Fix, a symbiotically defective mutant of pea (Pisum sativum L.)" In: *J. Exp. Bot.* 46.12 (Dec. 1995), pp. 1809–1816. URL: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/46.12.1809.

[100] Antonius C. J. Timmers et al. “Saprophytic Intracellular Rhizobia in Alfalfa Nodules”. In: *Mol. Plant-Microbe Interact.* 13.11 (Nov. 2000), pp. 1204–1213. URL: https://apsjournals.apsnet.org/doi/pdfplus/10.1094/MPMI.2000.13.11.1204.

[101] Oliver Preisig, Denise Anthamatten, and Hauke Hennecke. “Genes for a microaerobically induced oxidase complex in Bradyrhizobium japonicum are essential for a nitrogen-fixing endosymbiosis”. In: *Proc. Natl. Acad. Sci. U. S. A.* 90.8 (Apr. 1993), pp. 3309–3313. URL: https://www.pnas.org/content/pnas/90/8/3309.full.pdf.

[102] Oliver Preisig et al. “A high-affinity cbb3-type cytochrome oxidase terminates the symbiosis-specific respiratory chain of Bradyrhizobium japonicum”. In: *J. Bacteriol.* 178.6 (Mar. 1996), pp. 1532–1538. URL: http://www.ncbi.nlm.nih.gov/pubmed/8626278.

[103] Maria J. Delgado, Eulogio J. Bedmar, and Allan Downie. “Genes involved in the formation and assembly of rhizobial cytochromes and their role in symbiotic nitrogen fixation”. In: *Adv. Microb. Physiol.* 40 (Jan. 1998), pp. 191–231. URL: http://www.sciencedirect.com/science/article/pii/S0065291108601320.

[104] Andreas Schlüter et al. “Functional and regulatory analysis of the two copies of the fixNOQP operon of Rhizobium leguminosarum strain VF39”. In: *Mol. Plant-Microbe Interact.* 10.5 (July 1997), pp. 605–16. URL: http://apsjournals.apsnet.org/doi/10.1094/MPMI.1997.10.5.605.

[105] Vladimir V. Kopat et al. “Evolution of fixNOQP genes encoding cytochrome oxidase with high affinity to oxygen in rhizobia and related bacteria”. In: *Russ. J. Genet.* 53.7 (2017), pp. 766–774. URL: https://link.springer.com/content/pdf/10.1134%2FS1022795417070067.pdf.

[106] Marie-Helene Renalier et al. “A new symbiotic cluster on the pSym megaplasmid of Rhizobium meliloti 2011 carries a functional fix gene repeat and a nod locus”. In: *J. Bacteriol.* 169.5 (May 1987), pp. 2231–2238. URL: http://jb.asm.org/lookup/doi/10.1128/jb.169.5.2231-2238.1987.

[107] Kyung-Bum Lee et al. “The genome of the versatile nitrogen fixer Azorhizobium caulinodans ORS571”. In: *BMC Genomics* 9.1 (June 2008), p. 271. URL: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-9-271.

[108] Daniel Kahn et al. “Rhizobium meliloti fixGHI sequence predicts involvement of a specific cation pump in symbiotic nitrogen fixation”. In: *J. Bacteriol.* 171.2 (Feb. 1989), pp. 929–39. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC209684/pdf/jbacter00168-0319.pdf.

[109] Hans Georg Koch et al. “Roles of the ccoGHIS gene products in the biogenesis of the cbb3-type cytochrome c oxidase”. In: *J. Mol. Biol.* 297.1 (2000), pp. 49–65. URL: https://www.sciencedirect.com/science/article/abs/pii/S0022283600935551.

[110] Karine Mandon et al. “Role of the fixGHI region of Azorhizobium caulinodans in free-living and symbiotic nitrogen fixation”. In: *FEMS Microbiol. Lett.* 114.2 (Dec. 1993), pp. 185–189. URL: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.1993.tb06571.x.

[111] Karine Mandon, Alexandre Kaminski, and Claude Elmerich. “Functional analysis of the fixNOQP region of Azorhizobium caulinodans”. In: *J. Bacteriol.* 176.9 (May 1994), pp. 2560–2568. URL: http://www.ncbi.nlm.nih.gov/pubmed/8169204.

[112] Sarah L. Turner and Peter W. Young. "The glutamine synthetases of rhizobia: Phylogenetics and evolutionary implications". In: *Mol. Biol. Evol.* 17.2 (Feb. 2000), pp. 309–319. URL: https://pdfs.semanticscholar.org/2873/00c06918c661a6536389269ef9f21f9969bd.pdf.

[113] M. Antonio Rogel et al. "Nitrogen-Fixing Nodules with Ensifer adhaerens Harboring Rhizobium tropici Symbiotic Plasmids". In: *Appl. Environ. Microbiol.* 67.7 (2001), pp. 3264–3268. URL: https://aem.asm.org/content/67/7/3264.

[114] Fernando G. Barcellos et al. "Evidence of Horizontal Transfer of Symbiotic Genes from a Bradyrhizobium japonicum Inoculant Strain to Indigenous Diazotrophs Sinorhizobium (Ensifer) fredii and Bradyrhizobium elkanii in a Brazilian Savannah Soil". In: *Appl. Environ. Microbiol.* 73.8 (Apr. 2007), pp. 2635–2643. URL: http://aem.asm.org/cgi/doi/10.1128/AEM.01823-06.

[115] Catherine Masson-Boivin et al. "Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes?" In: *Trends Microbiol.* 17.10 (Oct. 2009), pp. 458–466. URL: http://www.sciencedirect.com/science/article/pii/S0966842X09001644.

[116] Esperanza Martínez-Romero. "Coevolution in Rhizobium-Legume Symbiosis?" In: *DNA Cell Biol.* 28.8 (Aug. 2009), pp. 361–370. arXiv: 1011.1669. URL: https://www.liebertpub.com/doi/10.1089/dna.2009.0863.

[117] Jeffry B. Stock et al. "Two-Component Signal Transduction Systems: Structure-Function Relationships and Mechanisms of Catalysis". In: *Two-Component Signal Transduct.* Washington, D.C.: American Society of Microbiology, Jan. 1995, pp. 25–51. URL: http://www.asmscience.org/content/book/10.1128/9781555818319.chap3.

[118] Ann H. West and Ann M. Stock. "Histidine kinases and response regulator proteins in two-component signaling systems". In: *Trends Biochem. Sci.* 26.6 (June 2001), pp. 369–376. URL: http://linkinghub.elsevier.com/retrieve/pii/S0968000401018527.

[119] Michel A. Sciotti et al. "Disparate oxygen responsiveness of two regulatory cascades that control expression of symbiotic genes in Bradyrhizobium japonicum". In: *J. Bacteriol.* 185.18 (2003), pp. 5639–5642. URL: https://pubmed.ncbi.nlm.nih.gov/12949117/.

[120] Michel David et al. "Cascade regulation of nif gene expression in Rhizobium meliloti". In: *Cell* 54.5 (Aug. 1988), pp. 671–683. URL: http://www.sciencedirect.com/science/article/pii/S0092867488800126.

[121] Alexandre Kaminski and Claudine Elmerich. "Involvement of fixLJ in the regulation of nitrogen fixation in Azorhizobium caulinodans". In: *Mol. Microbiol.* 5.3 (Mar. 1991), pp. 665–673. URL: http://doi.wiley.com/10.1111/j.1365-2958.1991.tb00738.x.

[122] Denise Anthamatten and Hauke Hennecke. "The regulatory status of the fixL- and fixJ-like genes in Bradyrhizobium japonicum may be different from that in Rhizobium meliloti". In: *MGG Mol. Gen. Genet.* 225.1 (1991), pp. 38–48. URL: http://www.ncbi.nlm.nih.gov/pubmed/2000090.

[123] Thomas Patschkowski, Andreas Schlüter, and Ursula B. Priefer. "Rhizobium leguminosarum bv. viciae contains a second fnr/fixK-like gene and an unusual fixL homologue". In: *Mol. Microbiol.* 21.2 (July 1996), pp. 267–280. URL: http://doi.wiley.com/10.1046/j.1365-2958.1996.6321348.x.

[124] Cecilia Hertig et al. "Rhizobium meliloti regulatory gene fixJ activates transcription of R. meliloti nifA and fixK genes in Escherichia coli". In: *J. Bacteriol.* 171.3 (Mar. 1989), pp. 1736–1738. URL: http://www.ncbi.nlm.nih.gov/pubmed/2646295.

[125] Socorro Mesa et al. "Comprehensive assessment of the regulons controlled by the FixLJ-FixK2-FixK1 cascade in Bradyrhizobium japonicum". In: *J. Bacteriol.* 190.20 (Oct. 2008), pp. 6568–6579. URL: http://www.ncbi.nlm.nih.gov/pubmed/18689489.

[126] Christine Bobik, Eliane Meilhoc, and Jacques Batut. "FixJ: a Major Regulator of the Oxygen Limitation Response and Late Symbiotic Functions of Sinorhizobium meliloti". In: *J. Bacteriol.* 188.13 (July 2006), pp. 4890–4902. URL: http://jb.asm.org/cgi/doi/10.1128/JB.00251-06.

[127] Marie Foussard et al. "Regulation of Nitrogen Fixation Gene Expression in Rhizobia: An Overview". In: *Biol. Nitrogen Fixat. 21st Century*. Kluwer Academic Publishers, 1998, pp. 101–106. URL: http://link.springer.com/10.1007/978-94-011-5159-7%7B%5C_%7D33.

[128] Lourdes Girard et al. "Differential regulation of fixN-reiterated genes in Rhizobium etli by a novel fixL—fixK cascade". In: *Mol. Plant-Microbe Interact.* 13.12 (Dec. 2000), pp. 1283–1292. URL: http://apsjournals.apsnet.org/doi/10.1094/MPMI.2000.13.12.1283.

[129] David Zamorano-Sánchez et al. "FxkR provides the missing link in the fixL-fixK signal transduction cascade in Rhizobium etli CFN42". In: *Mol. Plant-Microbe Interact.* 25.11 (Oct. 2012), pp. 1506–1517. URL: https://apsjournals.apsnet.org/doi/abs/10.1094/MPMI-05-12-0136-R.

[130] Inge D'hooghe et al. "Structural and functional analysis of the fixLJ genes of Rhizobium leguminosarum biovar phaseoli CNPAF512". In: *MGG Mol. Gen. Genet.* 249.1 (1995), pp. 117–126. URL: https://link.springer.com/content/pdf/10.1007%2FBF00290243.pdf.

[131] Martine Moris et al. "Regulatory role of Rhizobium etli CNPAF512 fnrN during symbiosis". In: *Appl. Environ. Microbiol.* 70.3 (Mar. 2004), pp. 1287–1296. URL: http://www.ncbi.nlm.nih.gov/pubmed/15006745.

[132] Alma Reyes-González et al. "Expanding the regulatory network that controls nitrogen fixation in Sinorhizobium meliloti: elucidating the role of the two-component system hFixL-FxkR". In: *Microbiology* 162.6 (June 2016), pp. 979–988. URL: http://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.000284.

[133] Jodi R. Trzebiatowski, Daniel M. Ragatz, and Frans J. De Bruijn. "Isolation and Regulation of Sinorhizobium meliloti 1021 Loci Induced by Oxygen Limitation". In: *Appl. Environ. Microbiol.* 67.8 (Aug. 2001), pp. 3728–3731. URL: http://www.ncbi.nlm.nih.gov/pubmed/11472955.

[134] Weimin Gong et al. "Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction". In: *Proc. Natl. Acad. Sci.* 95.26 (Dec. 1998), pp. 15177–15182. URL: http://www.ncbi.nlm.nih.gov/pubmed/9860942.

[135] Jeffrey Green et al. "Bacterial sensors of oxygen". In: *Curr. Opin. Microbiol.* 12.2 (Apr. 2009), pp. 145–151. URL: http://www.sciencedirect.com/science/article/pii/S1369527409000058.

[136] Seiji Yamada et al. "Structure of PAS-Linked Histidine Kinase and the Response Regulator Complex". In: *Structure* 17.10 (2009), pp. 1333–1344.

[137] John Stigter. "Regulation of Azorhizobium caulinodans ORS571 nitrogen fixation (NIF/FIX) genes". PhD thesis. Wageningen: Landbouwuniversiteit te Wageningen, 1994. URL: http://edepot.wur.nl/206538.

[138] Gareth S. A. Wright et al. "Architecture of the complete oxygen-sensing FixL-FixJ two-component signal transduction system". In: *Sci. Signal.* 11.525 (Apr. 2018), pp. 1–12. URL: http://www.ncbi.nlm.nih.gov/pubmed/29636388.

[139] Eduardo Henrique Silva Sousa, Gonzalo Gonzalez, and Marie Alda Gilles-Gonzalez. "Oxygen blocks the reaction of the FixL-FixJ complex with ATP but does not influence binding of FixJ or ATP to FixL". In: *Biochemistry* 44.46 (2005), pp. 15359–15365. URL: https://pubs.acs.org/doi/10.1021/bi051661h.

[140] Eduardo H. S. Sousa et al. "Signal transduction and phosphoryl transfer by a FixL hybrid kinase with low oxygen affinity: Importance of the vicinal PAS domain and receiver aspartate". In: *Biochemistry* 52.3 (Jan. 2013), pp. 456–465. URL: http://pubs.acs.org/doi/10.1021/bi300991r.

[141] Bing Hao et al. "Structure-based mechanism of O2 sensing and ligand discrimination by the FixL heme domain of Bradyrhizobium japonicum". In: *Biochemistry* 41.43 (2002), pp. 12952–12958. URL: https://pubmed.ncbi.nlm.nih.gov/12390021/.

[142] Kenton R. Rodgers and Gudrun S. Lukat-Rodgers. "Insights into heme-based O2 sensing from structure-function relationships in the FixL proteins". In: *J. Inorg. Biochem.* 99.4 (2005), pp. 963–977. URL: https://pubmed.ncbi.nlm.nih.gov/15811514/.

[143] Eduardo H. S. Sousa et al. "A memory of oxygen binding explains the dose response of the heme-based sensor FixL". In: *Biochemistry* 46.21 (2007), pp. 6249–6257. URL: https://pubmed.ncbi.nlm.nih.gov/17487983/.

[144] Augusto F. Lois et al. "Autophosphorylation and phosphatase activities of the oxygen-sensing protein FixL of Rhizobium meliloti are coordinately regulated by oxygen". In: *J. Biol. Chem.* 268.6 (Feb. 1993), pp. 4370–4375. URL: http://www.jbc.org/content/267/30/21864.full.pdf.

[145] Anne Galinier et al. "Phosphorylation of the Rhizobium meliloti FixJ protein induces its binding to a compound regulatory region at the fixK promoter". In: *J Biol Chem* 269.38 (1994), pp. 23784–23789. URL: http://www.jbc.org/content/269/38/23784.full.pdf.

[146] Jorg Schumacher et al. "Structural Basis for Signal Transduction within the FixJ Transcriptional Activator". In: *Nitrogen Fixat. From Mol. to Crop Product.* Vol. 408. November. Kluwer Academic Publishers, 2000, pp. 99–100. URL: http://link.springer.com/10.1007/0-306-47615-0_36.

[147] Patrice Gouet et al. "Structural transitions in the FixJ receiver domain". In: *Structure* 7.12 (1999), pp. 1517–1526. URL: https://pubmed.ncbi.nlm.nih.gov/10647182/.

[148] Sandra Da Re et al. "Intramolecular signal transduction within the FixJ transcriptional activator: In vitro evidence for the inhibitory effect of the phosphorylatable regulatory domain". In: *Nucleic Acids Res.* 22.9 (1994), pp. 1555–1561. URL: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/22.9.1555.

[149] Philippe Roche et al. "Molecular dynamics of the FixJ receiver domain: movement of the beta4-alpha4 loop correlates with the in and out flip of Phe101". In: *Protein Sci.* 11.11 (Dec. 2002), pp. 2622–30. URL: http://doi.wiley.com/10.1110/ps.0218802.

[150] Catherine Birck et al. "Conformational changes induced by phosphorylation of the FixJ receiver domain". In: *Structure* 7.12 (Jan. 1999), pp. 1505–1515. URL: https://www.sciencedirect.com/science/article/pii/S0969212600883410.

[151] Sandra Da Re et al. "Phosphorylation-induced dimerization of the FixJ receiver domain". In: *Mol. Microbiol.* 34.3 (Nov. 1999), pp. 504–511. URL: http://doi.wiley.com/10.1046/j.1365-2958.1999.01614.x.

[152] J. Peter W. Young et al. "The genome of Rhizobium leguminosarum has recognizable core and accessory components". In: *Genome Biol.* 7.4 (2006), R34. URL: http://www.ncbi.nlm.nih.gov/pubmed/16640791.

[153] Wellinson G. Guimarães et al. "Insights into signal transduction by a hybrid FixL: Denaturation study of on and off states of a multi-domain oxygen sensor". In: *J. Inorg. Biochem.* 172 (July 2017), pp. 129–137. URL: https://linkinghub.elsevier.com/retrieve/pii/S0162013416305372.

[154] Bert Boesten and Ursula B. Priefer. "The C-terminal receiver domain of the Rhizobium leguminosarum bv. viciae FixL protein is required for free-living microaerobic induction of the fnrN promoter". In: *Microbiology* 150.11 (2004), pp. 3703–3713. URL: https://pubmed.ncbi.nlm.nih.gov/15528657/.

[155] Ralf Heermann and Kirsten Jung. "Stimulus Perception and Signaling in Histidine Kinases". In: *Bact. Signal.* Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2010, pp. 135–161. URL: http://doi.wiley.com/10.1002/9783527629237.ch8.

[156] David Nellen-Anthamatten and Patrick Rossi. "Bradyrhizobium japonicum FixK2, a crucial distributor in the FixLJ-dependent regulatory cascade for control of genes inducible by low oxygen levels". In: *J. Bacteriol.* 180.19 (Oct. 1998), pp. 5251–5255. URL: http://www.ncbi.nlm.nih.gov/pubmed/9748464.

[157] Lionel Ferrières and Daniel Kahn. "Two distinct classes of FixJ binding sites defined by in vitro selection". In: *FEBS Lett.* 517.1-3 (Apr. 2002), pp. 185–189. URL: http://doi.wiley.com/10.1016/S0014-5793%2802%2902618-2.

[158] Frances Waelkens et al. "Molecular genetic analysis of the Rhizobium meliloti fixK promoter: identification of sequences involved in positive and negative regulation". In: *Mol. Microbiol.* 6.11 (June 1992), pp. 1447–1456. URL: http://doi.wiley.com/10.1111/j.1365-2958.1992.tb00865.x.

[159] Bao Ton-Hoang et al. "Promoter-specific involvement of the FixJ receiver domain in transcriptional activation". In: *J. Mol. Biol.* 312.4 (2001), pp. 583–589. URL: https://pubmed.ncbi.nlm.nih.gov/11575915/.

[160] Stefan Weidner et al. "Genome Sequence of Sinorhizobium meliloti Rm41". In: *Genome Announc.* 1.1 (Feb. 2013), e00013–12. URL: http://www.ncbi.nlm.nih.gov/pubmed/23405285.

[161] Lionel Ferrières et al. "FixJ-regulated genes evolved through promoter duplication in Sinorhizobium meliloti". In: *Microbiology* 150.7 (2004), pp. 2335–2345. URL: https://pubmed.ncbi.nlm.nih.gov/15256575/.

[162] José I. Jiménez-Zurdo, Fernando M. García-Rodríguez, and Nicolás Toro. "The Rhizobium meliloti putA gene: Its role in the establishment of the symbiotic interaction with alfalfa". In: *Mol. Microbiol.* 23.1 (Jan. 1997), pp. 85–93. URL: http://doi.wiley.com/10.1046/j.1365-2958.1997.1861555.x.

[163] José I. Jiménez-Zurdo. "Characterization of a Rhizobium meliloti Proline Dehydrogenase Mutant Altered in Nodulation Efficiency and Competitiveness on Alfalfa Roots". In: *Mol. Plant-Microbe Interact.* 8.4 (1995), p. 492. URL: https://www.researchgate.net/publication/14611955.

[164] Natalie D. King, David Hojnacki, and Mark R. O'Brian. "The Bradyrhizobium japonicum proline biosynthesis gene proC is essential for symbiosis". In: *Appl. Environ. Microbiol.* 66.12 (2000), pp. 5469–5471. URL: https://aem.asm.org/content/66/12/5469.

[165] Huamin Li et al. "Functional analysis of the fixL/fixJ and fixK genes in Azospirillum brasilense Sp7". In: *Ann. Microbiol.* 60.3 (2010), pp. 469–480. URL: https://link.springer.com/article/10.1007/s13213-010-0065-9.

[166] Albert I. Loroch, Bao G. Nguyen, and Robert A. Ludwig. "Interactive regulation of Azorhizobium nifA transcription via overlapping promoters". In: *J. Bacteriol.* 177.24 (Dec. 1995), pp. 7210–7221. URL: http://www.ncbi.nlm.nih.gov/pubmed/8522530.

[167] Alexandre Kaminski et al. "Regulation of nitrogen fixation in Azorhizobium caulinodans: identification of a fixK-like gene, a positive regulator of nifA". In: *Mol. Microbiol.* 5.8 (Aug. 1991), pp. 1983–1991. URL: http://doi.wiley.com/10.1111/j.1365-2958.1991.tb00820.x.

[168] Christopher L. Kitts and Robert A. Ludwig. "Azorhizobium caulinodans respires with at least four terminal oxidases". In: *J. Bacteriol.* 176.3 (1994), pp. 886–895. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC205126/.

[169] Rong Gao and Ann M. Stock. "Biological Insights from Structures of Two-Component Proteins". In: *Annu. Rev. Microbiol.* 63.1 (2009), pp. 133–154. arXiv: NIHMS150003. URL: https://pubmed.ncbi.nlm.nih.gov/19575571/.

[170] Socorro Mesa et al. "Bradyrhizobium japonicum NnrR, a denitrification regulator, expands the FixLJ-FixK2 regulatory cascade". In: *J. Bacteriol.* 185.13 (2003), pp. 3978–3982. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC161565/.

[171] Manuel J. Granados-Baeza et al. "Novel reiterated Fnr-type proteins control the production of the symbiotic terminal oxidase cbb3 in Rhizobium etli CFN42". In: *Mol. Plant-Microbe Interact.* 20.10 (Oct. 2007), pp. 1241–1249. URL: http://apsjournals.apsnet.org/doi/10.1094/MPMI-20-10-1241.

[172] Jason C. Crack et al. "Superoxide-mediated amplification of the oxygen-induced switch from [4Fe-4S] to [2Fe-2S] clusters in the transcriptional regulator FNR". In: *Proc. Natl. Acad. Sci.* 104.7 (2007), pp. 2092–2097. URL: www.pnas.orgcgidoi10.1073pnas.0609514104.

[173] Stephen Spiro and John R. Guest. "FNR and its role in oxygen-regulated gene expression in Escherichia coli". In: *FEMS Microbiol. Lett.* 75.4 (Aug. 1990), pp. 399–428. URL: https://academic.oup.com/femsre/article-lookup/doi/10.1111/j.1574-6968.1990.tb04109.x.

[174] Socorro Mesa, Hauke Hennecke, and Hans-Martin Fischer. "A multitude of CRP/FNR-like transcription proteins in Bradyrhizobium japonicum". In: *Biochem. Soc. Trans.* 34.1 (2006), pp. 156–159. URL: https://pubmed.ncbi.nlm.nih.gov/16417509/.

[175] Jason C. Crack and Nick E. Le Brun. "Redox-Sensing Iron–Sulfur Cluster Regulators". In: *Antioxid. Redox Signal.* 29.18 (Dec. 2018), pp. 1809–1829. URL: http://online.liebertpub.com/doi/10.1089/ars.2017.7361.

[176] John R. Guest et al. "The FNR modulon and FNR-regulated gene expression". In: *Regul. gene Expr. Escherichia coli.* Boston: Springer US, 1996, pp. 317–342. URL: http://link.springer.com/10.1007/978-1-4684-8601-8_16.

[177] David Zamorano-Sánchez and Lourdes Girard. "FNR-like proteins in rhizobia: past and future". In: *Biol. Nitrogen Fixat.* Vol. 1. Hoboken: John Wiley & Sons, Inc, July 2015, pp. 155–166. URL: http://doi.wiley.com/10.1002/9781119053095.ch15.

[178] Socorro Mesa et al. "Posttranslational control of transcription factor FixK2, a key regulator for the Bradyrhizobium japonicum-soybean symbiosis". In: *Proc. Natl. Acad. Sci.* 106.51 (2009), pp. 21860–21865. URL: https://www.pnas.org/content/106/51/21860.

[179] Stephen Spiro. "The FNR family of transcriptional regulators". In: *Antonie Van Leeuwenhoek* 66.1-3 (1994), pp. 23–36. URL: https://link.springer.com/article/10.1007/BF00871630.

[180] Heinz Körner, Heidi J. Sofia, and Walter G. Zumft. "Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: Exploiting the metabolic spectrum by controlling alternative gene programs". In: *FEMS Microbiol. Rev.* 27.5 (2003), pp. 559–592. URL: https://pubmed.ncbi.nlm.nih.gov/14638413/.

[181] Socorro Mesa et al. "Transcription activation in vitro by the Bradyrhizobium japonicum regulatory protein FixK2". In: *J. Bacteriol.* 187.10 (2005), pp. 3329–3338. URL: https://jb.asm.org/content/187/10/3329.

[182] Jacques Batut et al. "fixK, a gene homologous with fnr and crp from Escherichia coli, regulates nitrogen fixation genes both positively and negatively in Rhizobium meliloti". In: *EMBO J.* 8.4 (Apr. 1989), pp. 1279–86. URL: http://onlinelibrary.wiley.com/doi/10.1002/j.1460-2075.1989.tb03502.x/full.

[183] Delia Gutiérrez et al. "FnrN controls symbiotic nitrogen fixation and hydrogenase activities in Rhizobium leguminosarum biovar viciae UPM791". In: *J. Bacteriol.* 179.17 (Sept. 1997), pp. 5264–5270. URL: http://www.ncbi.nlm.nih.gov/pubmed/9286975.

[184] Jacques Batut et al. "Oxygen Regulation of Nitrogen Fixation Gene Expression in Rhizobium Meliloti". In: *Adv. Mol. Genet. Plant-Microbe Interact.* Dordrecht: Springer, 1993, pp. 183–191. URL: http://link.springer.com/10.1007/978-94-017-0651-3_20.

[185] Bernard L. Dreyfus and Yvon R. Dommergues. "Stem nodules of the tropical legume, Sesbania rostrata". In: *Curr. Perspect. Nitrogen Fixat.* Ed. by A. H. Gibson and W. E. Newton. Australian Academy of Science, 1981, p. 471. URL: http://www.documentation.ird.fr/hor/fdi:16530.

[186] Bernard L. Dreyfus, Claudine Elmerich, and Yvon R. Dommergues. *Free-living Rhizobium strain able to grow on N2 as the sole nitrogen source*. Feb. 1983. URL: http://www.ncbi.nlm.nih.gov/pubmed/16346220.

[187] Christiane Gebhardt et al. "Nitrogen-fixing Growth in Continuous Culture of a Strain of Rhizobium sp. Isolated from Stem Nodules on Sesbania rostrata". In: *J. Gen. Microbiol.* 130.4 (Apr. 1984), pp. 843–848. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-130-4-843.

[188] Oswaldo Lopez et al. "Regulation of gene expression in response to oxygen in Rhizobium etli: role of FnrN in fixNOQP expression and in symbiotic nitrogen fixation". In: *J. Bacteriol.* 183.24 (Dec. 2001), pp. 6999–7006. URL: http://www.ncbi.nlm.nih.gov/pubmed/11717256.

[189] Jean-Marc M. Reyrat et al. "Oxygen-regulated in vitro transcription of Rhizobium meliloti nifA and fixK genes". In: *J. Bacteriol.* 175.21 (1993), pp. 6867–6872. URL: http://jb.asm.org/content/175/21/6867.full.pdf.

[190] Hans-Martin Fischer. "Environmental regulation of rhizobial symbiotic nitrogen fixation genes". In: *Trends Microbiol.* 4.8 (Aug. 1996), pp. 317–320. URL: http://linkinghub.elsevier.com/retrieve/pii/0966842X96100494.

[191] Ines Kullik et al. "Bradyrhizobium japonicum has two differentially regulated, functional homologs of the σ54 gene (rpoN)". In: *J. Bacteriol.* 173.3 (1991), pp. 1125–1138. URL: https://pubmed.ncbi.nlm.nih.gov/1991712/.

[192] Jan Michiels et al. "The Rhizobium etli rpoN locus: DNA sequence analysis and phenotypical characterization of rpoN, ptsN, and ptsA mutants." In: *J. Bacteriol.* 180.7 (1998), pp. 1729–40. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC107084/?tool=pmcentrez&report=abstract.

[193] Stephen Spiro and John R. Guest. "Regulation and over-expression of the fnr gene of Escherichia coli". In: *J. Gen. Microbiol.* 133.12 (Dec. 1987), pp. 3279–3288. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-133-12-3279.

[194] Luzia Reutimann, Socorro Mesa, and Hauke Hennecke. "Autoregulation of fixK 2 gene expression in Bradyrhizobium japonicum". In: *Mol. Genet. Genomics* 284.1 (July 2010), pp. 25–32. URL: http://link.springer.com/10.1007/s00438-010-0547-2.

[195] Hélène Bergès et al. "A glutamine-amidotransferase-like protein modulates FixT anti-kinase activity in Sinorhizobium meliloti". In: *BMC Microbiol.* 1.6 (2001), pp. 1–9. URL: http://www.biomedcentral.com/1471-2180/1/6.

[196] Anne-Marie Garnerone et al. "Inhibition of the FixL sensor kinase by the FixT protein in Sinorhizobium meliloti". In: *J. Biol. Chem.* 274.45 (1999), pp. 32500–32506. URL: https://www.sciencedirect.com/science/article/pii/S0021925819515848.

[197] Anne-Marie Garnerone et al. "Mode of Action of the FixT Repressor Protein of Sinorhizobium Meliloti". In: *Highlights Nitrogen Fixat. Res.* New York: Kluwer Academic / Plenum Publishers, 1999, pp. 195–199. URL: http://link.springer.com/10.1007/978-1-4615-4795-2_39.

[198] Tadeo F. Fernandez-Göbel et al. "Redox systemic signaling and induced tolerance responses during soybean–Bradyrhizobium japonicum interaction: Involvement of nod factor receptor and autoregulation of nodulation". In: *Front. Plant Sci.* 10 (Feb. 2019), p. 141. URL: https://www.frontiersin.org/article/10.3389/fpls.2019.00141/full.

[199] Karen M. Page and Mary L. Guerinot. "Oxygen control of the Bradyrhizobium japonicum hemA gene". In: *J. Bacteriol.* 177.14 (1995), pp. 3979–3984. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC177127/.

[200] Scott D. McGinnis and Mark R. O'Brian. "The Rhizobial hemA Gene Is Required for Symbiosis in Species with Deficient delta-Aminolevulinic Acid Uptake Activity". In: *Plant Physiol.* 108.4 (Aug. 1995), pp. 1547–1552. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.108.4.1547.

[201] Nicolás Gómez-Hernández et al. "Regulation and Symbiotic Role of nirK and norC Expression in Rhizobium etli". In: *Mol. Plant-Microbe Interact.* 24.2 (Feb. 2011), pp. 233–245. URL: http://apsjournals.apsnet.org/doi/10.1094/MPMI-07-10-0173.

[202] Andreas Schlüter et al. "The Rhizobium leguminosarum FnrN protein is functionally similar to Escherichia coli Fnr and promotes heterologous oxygen-dependent activation of transcription". In: *Mol. Microbiol.* 6.22 (Nov. 1992), pp. 3395–3404. URL: http://doi.wiley.com/10.1111/j.1365-2958.1992.tb02207.x.

[203] Marianne Gamper, Axel Zimmermann, and Dieter Haas. "Anaerobic regulation of transcription initiation in the arcDABC operon of Pseudomonas aeruginosa". In: *J. Bacteriol.* 173.15 (1991), pp. 4742–4750. URL: https://pubmed.ncbi.nlm.nih.gov/1906871/.

[204] Sergio Colonna-Romano et al. "An Fnr-like protein encoded in Rhizobium leguminosarum biovar viciae shows structural and functional homology to Rhizobium meliloti fixK". In: *MGG Mol. Gen. Genet.* 223.1 (1990), pp. 138–147. URL: https://link.springer.com/article/10.1007/BF00315806.

[205] Mara V. Colombo et al. "A novel autoregulation mechanism of fnrN expression in Rhizobium leguminosarum bv viciae". In: *Mol. Microbiol.* 36.2 (Apr. 2000), pp. 477–486. URL: http://doi.wiley.com/10.1046/j.1365-2958.2000.01867.x.

[206] Carmen Sánchez-Cañizares et al. "Genomic diversity in the endosymbiotic bacterium Rhizobium leguminosarum". In: *Genes (Basel).* 9.2 (Jan. 2018), p. 60. URL: http://www.mdpi.com/2073-4425/9/2/60.

[207] Jeffrey Green, Colin Scott, and John R. Guest. "Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP". In: *Adv. Microb. Physiol.* Vol. 44. February. Academic Press, 2001, pp. 1–34. URL: https://linkinghub.elsevier.com/retrieve/pii/S0065291101440100.

[208] Kai U. Vollack et al. "Multiple transcription factors of the FNR family in denitrifying Pseudomonas stutzeri: Characterization of four fnr-like genes, regulatory responses and cognate metabolic processes". In: *Mol. Microbiol.* 31.6 (1999), pp. 1681–1694. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2958.1999.01302.x.

[209] Anjana Ray and Huw D. Williams. "The effects of mutation of the anr gene on the aerobic respiratory chain of Pseudomonas aeruginosa". In: *FEMS Microbiol. Lett.* 156.2 (Jan. 2006), pp. 227–232. URL: https://academic.oup.com/femsle/article-abstract/156/2/227/523304.

[210] Matthias Ebert et al. "FnrL and Three Dnr regulators are used for the metabolic adaptation to low oxygen tension in Dinoroseobacter shibae". In: *Front. Microbiol.* 8.APR (2017), p. 642. URL: http://www.ncbi.nlm.nih.gov/pubmed/28473807.

[211] Delia Gutiérrez et al. "Symbiotic Expression of Hydrogenase and Nitrogenase Activities of Rhizobium leguminosarum bv. Viciae are Controlled by FnrN". In: *Biol. Nitrogen Fixat. 21st Century*. Dordrecht: Springer, 1998, pp. 286–286. URL: http://link.springer.com/10.1007/978-94-011-5159-7_155.

[212] Jason J. Terpolilli, Graham A. Hood, and Philip S. Poole. "What Determines the Efficiency of N2-Fixing Rhizobium-Legume Symbioses?" In: *Adv. Microb. Physiol.* Vol. 60. Elsevier/Academic Press, 2012, pp. 325–389. URL: http://linkinghub.elsevier.com/retrieve/pii/B978012398264300005X.

[213] Denise Anthamatten, Barbara Scherb, and Hauke Hennecke. "Characterization of a fixLJ-regulated Bradyrhizobium japonicum gene sharing similarity with the Escherichia coli fnr and Rhizobium meliloti fixK genes". In: *J. Bacteriol.* 174.7 (Apr. 1992), pp. 2111–2120. URL: http://www.ncbi.nlm.nih.gov/pubmed/1551834.

[214] Jason Crack, Jeffrey Green, and Andrew J. Thomson. "Mechanism of Oxygen Sensing by the Bacterial Transcription Factor Fumarate-Nitrate Reduction (FNR)". In: *J. Biol. Chem.* 279.10 (Mar. 2004), pp. 9278–9286. URL: http://www.jbc.org/lookup/doi/10.1074/jbc.M309878200.

[215] Gottfried Unden and Martin Trageser. "Oxygen regulated gene expression in Escherichia coli: control of anaerobic respiration by the FNR protein". In: *Antonie Van Leeuwenhoek* 59.2 (Feb. 1991), pp. 65–76. URL: http://link.springer.com/10.1007/BF00445650.

[216] Laura J. Moore, Erin L. Mettert, and Patricia J. Kiley. "Regulation of FNR dimerization by subunit charge repulsion". In: *J. Biol. Chem.* 281.44 (Nov. 2006), pp. 33268–33275. URL: http://www.ncbi.nlm.nih.gov/pubmed/16959764.

[217] Laura J. Moore and Patricia J. Kiley. "Characterization of the dimerization domain in the FNR transcription factor". In: *J. Biol. Chem.* 276.49 (Dec. 2001), pp. 45744–45750. URL: http://www.ncbi.nlm.nih.gov/pubmed/11581261.

[218] Donna M. Bates et al. "Substitution of leucine 28 with histidine in the Escherichia coli transcription factor FNR results in increased stability of the [4Fe-4S]2+ cluster to oxygen". In: *J. Biol. Chem.* 275.9 (Mar. 2000), pp. 6234–6240. URL: http://www.ncbi.nlm.nih.gov/pubmed/10692418.

[219] Victoria R. Sutton et al. "Kinetic analysis of the oxidative conversion of the [4Fe-4S]2+ cluster of FNR to a [2Fe-2S]2+ cluster". In: *J. Bacteriol.* 186.23 (Dec. 2004), pp. 8018–8025. URL: http://www.ncbi.nlm.nih.gov/pubmed/15547274.

[220] Victoria R. Sutton et al. "Superoxide Destroys the [2Fe-2S]2+ Cluster of FNR from Escherichia coli". In: *Biochemistry* 43.3 (2004), pp. 791–798. URL: https://pubs.acs.org/doi/abs/10.1021/bi0357053.

[221] Stephen Spiro and John R. Guest. "Activation of the lac Operon of Escherichia coli by a mutant FNR protein". In: *Mol. Microbiol.* 1.3 (Nov. 1987), pp. 53–58. URL: http://doi.wiley.com/10.1111/j.1365-2958.1987.tb00526.x.

[222] Martin Trageser and Gottfried Unden. "Role of cysteine residues and of metal ions in the regulatory functioning of FNR, the transcriptional regulator of anaerobic respiration in Escherichia coli". In: *Mol. Microbiol.* 3.5 (May 1989), pp. 593–599. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2958.1989.tb00206.x.

[223] Beth A. Lazazzera, Donna M. Bates, and Patricia J. Kiley. "The activity of the Escherichia coli transcription factor FNR is regulated by a change in oligomeric state". In: *Genes Dev.* 7.10 (1993), pp. 1993–2005. URL: http://genesdev.cshlp.org/content/7/10/1993.full.pdf.

[224] Adrian J. Jervis and Jeffrey Green. "In vivo demonstration of FNR dimers in response to lower O2 availability". In: *J. Bacteriol.* 189.7 (Apr. 2007), pp. 2930–2932. URL: http://www.ncbi.nlm.nih.gov/pubmed/17277055.

[225] Natalia Khoroshilova et al. "Iron-sulfur cluster disassembly in the FNR protein of Escherichia coli by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity". In: *Proc. Natl. Acad. Sci.* 94.12 (June 1997), pp. 6087–6092. URL: http://www.ncbi.nlm.nih.gov/pubmed/9177174.

[226] Gary Sawers. "Identification and molecular characterization of a transcriptional regulator from Pseudomonas aeruginosa PAO1 exhibiting structural and functional similarity to the FNR protein of Escherichia coli". In: *Mol. Microbiol.* 5.6 (June 1991), pp. 1469–1481. eprint: 1502.05928. URL: https://pubmed.ncbi.nlm.nih.gov/1787797/.

[227] Padma-Sheela Jayaraman et al. "The nirB promoter of Escherichia coli: location of nucleotide sequences essential for regulation by oxygen, the FNR protein and nitrite". In: *Mol. Microbiol.* 2.4 (July 1988), pp. 527–530. URL: http://doi.wiley.com/10.1111/j.1365-2958.1988.tb00059.x.

[228] Padma-Sheela Jayaraman, Jeffrey A. Cole, and Stephen J. W. Busby. "Mutational analysis of the nucleotide sequence at the FNR-dependent nirB promoter in Escherichia coli". In: *Nucleic Acids Res.* 17.1 (1989), pp. 135–145. URL: https://academic.oup.com/nar/article-abstract/17/1/135/1191680.

[229] Sophie R. D. Clark, Ivan J. Oresnik, and Michael F. Hynes. "RpoN of Rhizobium leguminosarum bv. viciae strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation". In: *MGG Mol. Gen. Genet.* 264.5 (2001), pp. 623–633. URL: https://pubmed.ncbi.nlm.nih.gov/11212917/.

[230] Andreas Schlüter et al. "Function and regulatory characteristics of FnrN, an oxygen-responsive transcriptional activator in Rhizobium leguminosarum bv. viciae". In: *New Horizons Nitrogen Fixat.* (1993), p. 493. URL: https://www.springer.com/gp/book/9780792322078.

[231] Tomas Ruiz-Argüeso, Jose M. Palacios, and Juan Imperial. "Regulation of the hydrogenase system in Rhizobium leguminosarum". In: *Plant Soil* 230.1 (2001), pp. 49–57. URL: http://link.springer.com/10.1023/A:1004578324977.

[232] Ray Dixon. "The Genetic Complexity of Nitrogen Fixation". In: *Microbiology* 130.11 (Nov. 1984), pp. 2745–2755. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-130-11-2745.

[233] Ann M. Hirsch and Carol A. Smith. "Effects of Rhizobium meliloti nif and fix mutants on alfalfa root nodule development". In: *J. Bacteriol.* 169.3 (1987), pp. 1137–1146. URL: https://jb.asm.org/content/169/3/1137.

[234] Marta Martínez et al. "Symbiotic autoregulation of nifA expression in Rhizobium leguminosarum bv. viciae". In: *J. Bacteriol.* 186.19 (Oct. 2004), pp. 6586–6594. URL: http://www.ncbi.nlm.nih.gov/pubmed/15375140.

[235] Christopher D. Earl, Clive W. Ronson, and Frederick M. Ausubel. "Genetic and structural analysis of the Rhizobium meliloti fixA, fixB, fixC, and fixX genes". In: *J. Bacteriol.* 169.3 (Mar. 1987), pp. 1127–1136. URL: http://www.ncbi.nlm.nih.gov/pubmed/3029021.

[236] Resie M. P. Schetgens et al. "Identification and phenotypical characterization of a cluster of fix genes, including a nif regulatory gene, from Rhizobium leguminosarum PRE". In: *MGG Mol. Gen. Genet.* 200.3 (Aug. 1985), pp. 368–374. URL: http://link.springer.com/10.1007/BF00425719.

[237] John R. Postgate. "Biology Nitrogen Fixation: Fundamentals". In: *Philos. Trans. R. Soc. B Biol. Sci.* 296.1082 (Jan. 1982), pp. 375–385. URL: http://www.jstor.org/stable/2395691.

[238] Isabel Martinez-Argudo et al. "Nitrogen fixation: key genetic regulatory mechanisms". In: *Biochem. Soc. Trans.* 33.1 (Feb. 2005), pp. 152–156. URL: https://pdfs.semanticscholar.org/19ea/6719cfc471fedefff4cf499709f88ea4e999.pdf.

[239] Alexandre Kaminski and Claudine Elmerich. "The control of Azorhizobium caulinodans nifA expression by oxygen, ammonia and by the HF-I-like protein, NrfA". In: *Mol. Microbiol.* 28.3 (May 1998), pp. 603–13. URL: http://www.ncbi.nlm.nih.gov/pubmed/9632262.

[240] Isabel Martinez-Argudo et al. "The NifL-NifA System: A Multidomain Transcriptional Regulatory Complex That Integrates Environmental Signals". In: *J. Bacteriol.* 186.3 (Feb. 2004), pp. 601–610. URL: http://www.ncbi.nlm.nih.gov/pubmed/14729684.

[241] Ivo Baldani et al. “Analysis of nitrogen fixation and regulatory genes in the sugar cane endophyte Acetobacter diazotrophicus”. In: *Nitrogen Fixat. with Non-Legumes*. Dordrecht: Springer, 2011, pp. 11–19. URL: http://link.springer.com/10.1007/978-94-011-5232-7_2.

[242] Zhenhua Yao et al. “Complementation analyses of Sinorhizobium meliloti nifA mutant with different originated nifA genes”. In: *Chinese Sci. Bull.* 51.22 (Nov. 2006), pp. 2748–2754. URL: http://link.springer.com/10.1007/s11434-006-2203-0.

[243] Venkatesan Sundaresan et al. “Klebsiella pneumoniae nifA product activates the Rhizobium meliloti nitrogenase promoter”. In: *Nature* 301.5902 (Feb. 1983), pp. 728–732. URL: http://www.nature.com/articles/301728a0.

[244] Jim L. Beynon, Mark K. Williams, and Frank C. Cannon. “Expression and functional analysis of the Rhizobium meliloti nifA gene”. In: *EMBO J.* 7.1 (1988), pp. 7–14. URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16453824.

[245] Ariel Alvarez-Morales and Hauke Hennecke. “Expression of Rhizobium japonicum nifH and nifDK operons can be activated by the Klebsiella pneumoniae nifA protein but not by the product of ntrC”. In: *MGG Mol. Gen. Genet.* 199.2 (1985), pp. 306–314. URL: https://link.springer.com/content/pdf/10.1007%2FBF00330273.pdf.

[246] Hans-Martin Fischer and Hauke Hennecke. “Direct response of Bradyrhizobium japonicum nifA-mediated nif gene regulation to cellular oxygen status”. In: *MGG Mol. Gen. Genet.* 209.3 (Oct. 1987), pp. 621–626. URL: http://link.springer.com/10.1007/BF00331174.

[247] Marta Martínez et al. “Novel arrangement of enhancer sequences for NifA-dependent activation of the hydrogenase gene promoter in Rhizobium leguminosarum bv. viciae”. In: *J. Bacteriol.* 190.9 (May 2008), pp. 3185–3191. URL: http://www.ncbi.nlm.nih.gov/pubmed/18310336.

[248] David J. Studholme and Ray Dixon. “Domain Architectures of 54-Dependent Transcriptional Activators”. In: *J. Bacteriol.* 185.6 (Mar. 2003), pp. 1757–1767. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC150144/pdf/1271.pdf.

[249] Matthew Bush and Ray Dixon. “The Role of Bacterial Enhancer Binding Proteins as Specialized Activators of 54-Dependent Transcription”. In: *Microbiol. Mol. Biol. Rev.* 76.3 (Sept. 2012), pp. 497–529. URL: http://www.ncbi.nlm.nih.gov/pubmed/22933558.

[250] Florence Arsène, P. Alexandre Kaminski, and Claudine Elmerich. “Modulation of NifA activity by P(II) in Azospirillum brasilense: Evidence for a regulatory role of the NifA N-terminal domain”. In: *J. Bacteriol.* 178.16 (1996), pp. 4830–4838. URL: https://jb.asm.org/content/178/16/4830.

[251] David K. Berger et al. “In vitro studies of the domains of the nitrogen fixation regulatory protein NIFA”. In: *J. Bacteriol.* 177.1 (1995), pp. 191–199. URL: http://jb.asm.org/content/177/1/191.full.pdf.

[252] Florence Arsène, P. Alexandre Kaminski, and Claudine Elmerich. “Control of Azospirillum brasilense NifA activity by P(II): Effect of replacing Tyr residues of the NifA N-terminal domain on NifA activity”. In: *FEMS Microbiol. Lett.* 179.2 (1999), pp. 339–343. URL: https://doi.org/10.1111/j.1574-6968.1999.tb08747.x.

[253] Xiaoxiao Zou et al. "Identification and functional characterization of NifA variants that are independent of GlnB activation in the photosynthetic bacterium Rhodospirillum rubrum". In: *Microbiology* 154.9 (Sept. 2008), pp. 2689–2699. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.2008/019406-0.

[254] Enrique Morett and Martin Buck. "NifA-dependent in vivo protection demonstrates that the upstream activator sequence of nif promoters is a protein binding site". In: *Proc. Natl. Acad. Sci.* 85.24 (Dec. 1988), pp. 9401–9405. URL: http://www.ncbi.nlm.nih.gov/pubmed/2849102.

[255] Martin Buck et al. "Upstream activator sequences are present in the promoters of nitrogen fixation genes". In: *Nature* 320.6060 (Mar. 1986), pp. 374–378. URL: http://www.nature.com/articles/320374a0.

[256] Ariel Avarez-Morales et al. "Activation of the Bradyrhizobium japonicum nifH and nifDK operons is dependent on promoter-upstream DNA sequences". In: *Nucleic Acids Res.* 14.10 (1986), pp. 4207–4227. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC339856/pdf/nar00279-0231.pdf.

[257] Marcel Gubler. "Fine-tuning of nif and fix gene expression by upstream activator sequences in Bradyrhizobium japonicum". In: *Mol. Microbiol.* 3.2 (1989), pp. 149–159. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2958.1989.tb01804.x.

[258] Alexander J. Ninfa, Lawrence J. Reitzer, and Boris Magasanik. "Initiation of transcription at the bacterial glnAp2 promoter by purified E. coli components is facilitated by enhancers". In: *Cell* 50.7 (1987), pp. 1039–1046. URL: https://doi.org/10.1016/0092-8674(87)90170-X.

[259] Felix Hauser et al. "Dissection of the Bradyrhizobium japonicum NifA+σ54 regulon, and identification of a ferredoxin gene (fdxN) for symbiotic nitrogen fixation". In: *Mol. Genet. Genomics* 278.3 (Aug. 2007), pp. 255–271. URL: http://link.springer.com/10.1007/s00438-007-0246-9.

[260] Katharina Pflüger-Grau and Boris Görke. "Regulatory roles of the bacterial nitrogen-related phosphotransferase system". In: *Trends Microbiol.* 18.5 (2010), pp. 205–214. URL: http://www.sciencedirect.com/science/article/pii/S0966842X1000020X.

[261] Yew-Seng S. Ho, Lisa M. Burden, and James H. Hurley. "Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor". In: *EMBO J.* 19.20 (Oct. 2000), pp. 5288–5299. URL: http://emboj.embopress.org/cgi/doi/10.1093/emboj/19.20.5288.

[262] Eva Huala and Frederick M. Ausubel. "The central domain of Rhizobium meliloti NifA is sufficient to activate transcription from the R. meliloti nifH promoter". In: *J. Bacteriol.* 171.6 (June 1989), pp. 3354–3365. URL: http://www.ncbi.nlm.nih.gov/pubmed/2722751.

[263] Alexander J. Ninfa and Peng Jiang. "PII signal transduction proteins: Sensors of α-ketoglutarate that regulate nitrogen metabolism". In: *Curr. Opin. Microbiol.* 8.2 (Apr. 2005), pp. 168–173. URL: https://www.sciencedirect.com/science/article/pii/S1369527405000226.

[264] Ruth A. Schmitz, Kai Klopprogge, and Roman Grabbe. "Regulation of nitrogen fixation in Klebsiella pneumoniae and Azotobacter vinelandii: NifL, transducing two environmental signals to the nif transcriptional activator NifA". In: *J. Mol. Microbiol. Biotechnol.* 4.3 (2002), pp. 235–242. URL: https://pubmed.ncbi.nlm.nih.gov/11931553/.

[265] Isabel Martinez-Argudo, Richard Little, and Ray Dixon. "Role of the amino-terminal GAF domain of the NifA activator in controlling the response to the antiactivator protein NifL". In: *Mol. Microbiol.* 52.6 (May 2004), pp. 1731–1744. URL: http://doi.wiley.com/10.1111/j.1365-2958.2004.04089.x.

[266] Ray Dixon et al. "Analysis of regulation of Klebsiella pneumoniae nitrogen fixation (nif) gene cluster with gene fusions". In: *Nature* 286.5769 (July 1980), pp. 128–132. URL: https://www.nature.com/articles/286128a0.

[267] Susan Hill et al. "Azotobacter vinelandii NIFL is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch". In: *Proc. Natl. Acad. Sci.* 93.5 (1996), pp. 2143–2148. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC39924/pdf/pnas01509-0418.pdf.

[268] Sydneiu Moreno et al. "Phenotype of a Rhizobium leguminosarum ntrC mutant". In: *Res. Microbiol.* 143.2 (1992), pp. 161–171. URL: https://www.sciencedirect.com/science/article/abs/pii/0923250892900059.

[269] Martin H. Drummond, Asunción Contreras, and Lesley A. Mitchenall. "The function of isolated domains and chimaeric proteins constructed from the transcriptional activators NifA and NtrC of Klebsiella pneumoniae". In: *Mol. Microbiol.* 4.1 (Jan. 1990), pp. 29–37. URL: http://doi.wiley.com/10.1111/j.1365-2958.1990.tb02012.x.

[270] Maria Arnott et al. "Deletion analysis of the nitrogen fixation regulatory gene nifL of Klebsiella pneumoniae". In: *Arch. Microbiol.* 151 (1989), pp. 180–182. URL: https://link.springer.com/content/pdf/10.1007%2FBF00414436.pdf.

[271] Richard Little and Ray Dixon. "The amino-terminal GAF domain of Azotobacter vinelandii NifA binds 2-oxoglutarate to resist inhibition by NifL under nitrogen-limiting conditions". In: *J. Biol. Chem.* 278.31 (2003), pp. 28711–28718. URL: https://pubmed.ncbi.nlm.nih.gov/12759352/.

[272] Jong-Kay Liu et al. "Alternative Function of the Electron Transport System in Azotobacter vinelandii: Removal of Excess Reductant by the Cytochrome d Pathway". In: *Appl. Environ. Microbiol.* 61.11 (Nov. 1995), pp. 3998–4003. URL: http://www.ncbi.nlm.nih.gov/pubmed/16535163.

[273] James W. Drozd and John R. Postgate. "Effects of Oxygen on Acetylene Reduction, Cytochrome Content and Respiratory Activity of Azotobacter chroococcum". In: *J. Gen. Microbiol.* 63.1 (Sept. 2009), pp. 63–73. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-63-1-63.

[274] Wael Sabra et al. "Effect of Oxygen on Formation and Structure of Azotobacter vinelandii Alginate and Its Role in Protecting Nitrogenase". In: *Appl. Environ. Microbiol.* 66.9 (Sept. 2000), pp. 4037–4044. eprint: 0204013 (cond-mat). URL: http://www.ncbi.nlm.nih.gov/pubmed/16535163.

[275] Vicky Buchanan-Wollaston et al. “Role of the nifA gene product in the regulation of nif expression in Klebsiella pneumoniae”. In: *Nature* 294.5843 (Dec. 1981), pp. 776–778. URL: http://www.ncbi.nlm.nih.gov/pubmed/6119619.

[276] Hans-Martin Fischer, Thomas Bruderer, and Hauke Hennecke. “Essential and non-essential domains in the Bradyrhizobium japonicum nifa protein: Identification of indispensable cysteine residues potentially involved in redox reactivity and/or metal binding”. In: *Nucleic Acids Res.* 16.5 (Mar. 1988), pp. 2207–2224. URL: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/16.5.2207.

[277] Siiri E. Iismaa and John M. Watson. “The nifA gene product from Rhizobium leguminosarum biovar trifolii lacks the N-terminal domain found in other NifA proteins”. In: *Mol. Microbiol.* 3.7 (July 1989), pp. 943–955. URL: http://doi.wiley.com/10.1111/j.1365-2958.1989.tb00244.x.

[278] Nathalie Michel-Reydellet and P. Alexandre Kaminski. “Azorhizobium caulinodans P(II) and GlnK proteins control nitrogen fixation and ammonia assimilation”. In: *J. Bacteriol.* 181.8 (Apr. 1999), pp. 2655–2658. URL: http://www.ncbi.nlm.nih.gov/pubmed/10198037.

[279] Emanuel M. Souza et al. “Sequence and structural organization of a nifA-like gene and part of a nifB-like gene of Herbaspirillum seropedicae strain Z78”. In: *J. Gen. Microbiol.* 137.7 (July 1991), pp. 1511–1522. URL: http://mic.microbiologyresearch.org/content/journal/micro/10.1099/00221287-137-7-1511.

[280] Reiner Krey, Alfred Pühler, and Werner Klipp. “A defined amino acid exchange close to the putative nucleotide binding site is responsible for an oxygen-tolerant variant of the Rhizobium meliloti NifA protein”. In: *MGG Mol. Gen. Genet.* 234.3 (Sept. 1992), pp. 433–41. URL: http://www.ncbi.nlm.nih.gov/pubmed/1406589.

[281] Hans-Martin Fischer et al. “Critical spacing between two essential cysteine residues in the interdomain linker of the Bradyrhizobium japonicum NifA protein”. In: *FEBS Lett.* 255.1 (Sept. 1989), pp. 167–171. URL: http://www.sciencedirect.com/science/article/pii/001457938981083X.

[282] Werner Klipp et al. “The Rhizobium meliloti fdxN gene encoding a ferredoxin-like protein is necessary for nitrogen fixation and is cotranscribed with nifA and nifB.” In: *MGG Mol. Gen. Genet.* 216.2-3 (Apr. 1989), pp. 293–302. URL: http://link.springer.com/10.1007/BF00334368.

[283] Evelyne Bauer et al. “Expression of the fixR-nifA operon in Bradyrhizobium japonicum depends on a new response regulator, RegR”. In: *J. Bacteriol.* 180.15 (Aug. 1998), pp. 3853–3863. URL: https://jb.asm.org/content/180/15/3853.

[284] David W. Nees, Pascal A. Stein, and Robert A. Ludwig. “The Azorhizobium caulinodans nifA gene: Identification of upstream-activating sequences including a new element, the ’anaerobox’”. In: *Nucleic Acids Res.* 16.20 (Oct. 1988), pp. 9839–9853. URL: https://academic.oup.com/nar/article/16/20/9839/2378589.

[285] Katharina Pawlowski, Ulrike Klosse, and Frans J. de Bruijn. "Characterization of a novel Azorhizobium caulinodans ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism". In: *MGG Mol. Gen. Genet.* 231.1 (Dec. 1991), pp. 124–138. URL: http://link.springer.com/10.1007/BF00293830.

[286] Katharina Pawlowski et al. "Cloning and characterization of nifA and ntrC genes of the stem nodulating bacterium ORS571, the nitrogen fixing symbiont of Sesbania rostrata: Regulation of nitrogen fixation (nif) genes in the free living versus symbiotic state". In: *MGG Mol. Gen. Genet.* 206.2 (Feb. 1987), pp. 207–219. URL: http://link.springer.com/10.1007/BF00333576.

[287] Pascal Ratet et al. "The Azorhizobium caulinodans nitrogen-fixation regulatory gene, nifA, is controlled by the cellular nitrogen and oxygen status". In: *Mol. Microbiol.* 3.6 (June 1989), pp. 825–838. URL: http://doi.wiley.com/10.1111/j.1365-2958.1989.tb00231.x.

[288] Beat Thöny, Denise Anthamatten, and Hauke Hennecke. "Dual control of the Bradyrhizobium japonicum symbiotic nitrogen fixation regulatory operon fixR nifA: analysis of cis- and trans-acting elements". In: *J. Bacteriol.* 171.8 (1989), pp. 4162–4169. URL: https://jb.asm.org/content/171/8/4162.

[289] Choong H. Kim, Donald R. Helinski, and Gary Ditta. "Overlapping transcription of the nifA regulatory gene in Rhizobium meliloti". In: *Gene* 50.1-3 (1986), pp. 141–148. URL: http://www.sciencedirect.com/science/article/pii/0378111986903197.

[290] Wynne W. Szeto et al. "Identification and characterization of the Rhizobium meliloti ntrC gene: R. meliloti has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells". In: *J. Bacteriol.* 169.4 (1987), pp. 1423–1432. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC211963/.

[291] Gary Ditta et al. "The nifA gene of Rhizobium meliloti is oxygen regulated". In: *J. Bacteriol.* 169.7 (July 1987), pp. 3217–3223. URL: http://www.ncbi.nlm.nih.gov/pubmed/2439489.

[292] Enrique Morett, Hans-Martin Fischer, and Hauke Hennecke. "Influence of oxygen on DNA binding, positive control, and stability of the Bradyrhizobium japonicum NifA regulatory protein". In: *J. Bacteriol.* 173.11 (June 1991), pp. 3478–3487. URL: http://www.ncbi.nlm.nih.gov/pubmed/2045367.

[293] Belén Brito et al. "Hydrogenase genes from Rhizobium leguminosarum bv. viciae are controlled by the nitrogen fixation regulatory protein NifA". In: *Proc. Natl. Acad. Sci.* 94.June (1997), pp. 6019–6024. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC20993/pdf/pq006019.pdf.

[294] Beat Thöny and Hauke Hennecke. "The -24/-12 promoter comes of age". In: *FEMS Microbiol. Lett.* 63.4 (Dec. 2006), pp. 341–357. URL: https://academic.oup.com/femsre/article-abstract/5/4/341/496924.

[295] Jan Michiels et al. "Characterization of the Rhizobium leguminosarum biovar phaseoli nifA gene, a positive regulator of nif gene expression". In: *Arch. Microbiol.* 161.5 (1994), pp. 404–408. URL: https://link.springer.com/content/pdf/10.1007/BF00288950.pdf.

[296] Beat Thöny et al. "The symbiotic nitrogen fixation regulatory operon (fixRnifA) of Bradyrhizobium japonicum is expressed aerobically and is subject to a novel, nifA -independent type of activation". In: *Nucleic Acids Res.* 15.20 (1987), pp. 8479–8499. URL: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/15.20.8479.

[297] Traki Benhassine et al. "Interaction of an IHF-like protein with the Rhizobium etli nifA promoter". In: *FEMS Microbiol. Lett.* 271.1 (2007), pp. 20–26. URL: https://academic.oup.com/femsle/article-abstract/271/1/20/500525.

[298] Roseli Wassem et al. "Control of autogenous activation of Herbaspirillum seropedicae nifA promoter by the IHF protein". In: *FEMS Microbiol. Lett.* 212.2 (July 2002), pp. 177–192. URL: https://academic.oup.com/femsle/article-abstract/212/2/177/632598.

[299] Roseli Wassem et al. "Two roles for integration host factor at an enhancer-dependent nifA promoter". In: *Mol. Microbiol.* 35.4 (Feb. 2000), pp. 756–764. URL: http://doi.wiley.com/10.1046/j.1365-2958.2000.01746.x.

[300] Humberto Barrios et al. "Overlapping promoters for two different RNA polymerase holoenzymes control Bradyrhizobium japonicum nifA expression". In: *J. Bacteriol.* 177.7 (1995), pp. 1760–1765. URL: https://jb.asm.org/content/177/7/1760.

[301] Sydney Kustu et al. "Expression of sigma 54 (ntrA)-dependent genes is probably united by a common mechanism." In: *Microbiol. Rev.* 53.3 (1989), pp. 367–76. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC372741/.

[302] Gary N. Gussin, Clive W. Ronson, and Frederick M. Ausubel. "Regulation of Nitrogen Fixation Genes". In: *Annu. Rev. Genet.* 20.1 (Dec. 1986), pp. 567–591. URL: http://www.annualreviews.org/doi/10.1146/annurev.ge.20.120186.003031.

[303] Ralf Emmerich, Hauke Hennecke, and Hans-Martin Fischer. "Evidence for a functional similarity between the two-component regulatory systems RegSR, ActSR, and RegBA (PrrBA) in α-proteobacteria". In: *Arch. Microbiol.* 174.5 (2000), pp. 307–313. URL: https://pubmed.ncbi.nlm.nih.gov/11131020/.

[304] Andrea Lindemann et al. "New target genes controlled by the Bradyrhizobium japonicum two-component regulatory system RegSR". In: *J. Bacteriol.* 189.24 (2007), pp. 8928–8943. URL: https://jb.asm.org/content/189/24/8928.

[305] Andrea Lindemann et al. "Host-specific symbiotic requirement of BdeAB, a RegR-controlled RND-type efflux system in Bradyrhizobium japonicum". In: *FEMS Microbiol. Lett.* 312.2 (2010), pp. 184–191. URL: https://core.ac.uk/download/pdf/85213641.pdf.

[306] Maria J. Torres et al. "The Global Response Regulator RegR Controls Expression of Denitrification Genes in Bradyrhizobium japonicum". In: *PLoS One* 9.6 (2014), e99011. URL: https://pubmed.ncbi.nlm.nih.gov/24949739/.

[307] Carl E. Bauer and Jiang Wu. *RegB/RegA, a global redox-responding two-component system*. New York, 2008. URL: https://doi.org/10.1128/MMBR.68.2.263-279.2004.

[308] Ralf Emmerich et al. "Phosphorylation, dephosphorylation and DNA-binding of the Bradyrhizobium japonicum RegSR two-component regulatory proteins". In: *Eur. J. Biochem.* 263.2 (July 1999), pp. 455–463. URL: https://febs.onlinelibrary.wiley.com/doi/full/10.1046/j.1432-1327.1999.00517.x.

[309] Emmanuel Salazar et al. "Characterization of the NifA-RpoN regulon in Rhizobium etli in free life and in symbiosis with Phaseolus vulgaris". In: *Appl. Environ. Microbiol.* 76.13 (2010), pp. 4510–4520. URL: https://aem.asm.org/content/76/13/4510.

[310] Víctor De Lorenzo and José Pérez-Martín. "Regulatory noise in prokaryotic promoters: How bacteria learn to respond to novel environmental signals". In: *Mol. Microbiol.* 19.6 (1996), pp. 1177–1184. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2958.1996.tb02463.x.

[311] Robert L. Robson and John R. Postgate. "Oxygen and hydrogen in biological nitrogen fixation". In: *Ann. Rev. Mar. Sci.* 34 (1980), pp. 183–207. URL: http://www.annualreviews.org/doi/pdf/10.1146/annurev.mi.34.100180.001151.

[312] Karel R. Schubert and Harold J. Evans. "Hydrogen evolution: A major factor affecting the efficiency of nitrogen fixation in nodulated symbionts". In: *Proc. Natl. Acad. Sci.* 73.4 (Apr. 1976), pp. 1207–1211. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC430231/pdf/pnas00034-0233.pdf.

[313] Jenjira Wongdee et al. "Regulation of nitrogen fixation in Bradyrhizobium sp. strain DOA9 involves two distinct NifA regulatory proteins that are functionally redundant during symbiosis but not during free-living growth". In: *Front. Microbiol.* 9.JUL (2018), pp. 1–11. URL: https://www.frontiersin.org/article/10.3389/fmicb.2018.01644/full.

[314] Zhexian Tian et al. "Transcriptome analysis of Sinorhizobium meliloti nodule bacteria in nifA mutant background". In: *Chinese Sci. Bull.* 51.17 (Sept. 2006), pp. 2079–2086. URL: http://link.springer.com/10.1007/s11434-006-2092-2.

[315] Wim De Vries et al. *Assessing planetary and regional nitrogen boundaries related to food security and adverse environmental impacts*. Sept. 2013. URL: https://www.sciencedirect.com/science/article/abs/pii/S1877343513000833.

[316] Will Steffen et al. "Planetary boundaries: guiding human development on a changing planet". In: *Science* 347.6223 (Feb. 2015). eprint: 461472a (10.1038). URL: http://www.ncbi.nlm.nih.gov/pubmed/21764740.

[317] Ana R. Fox et al. "Major cereal crops benefit from biological nitrogen fixation when inoculated with the nitrogen-fixing bacterium Pseudomonas protegens Pf-5 X940". In: *Environ. Microbiol.* 18.10 (2016), pp. 3522–3534. URL: https://pubmed.ncbi.nlm.nih.gov/27198923/.

[318] Kerrie Farrar, David Bryant, and Naomi Cope-Selby. "Understanding and engineering beneficial plant–microbe interactions: plant growth promotion in energy crops". In: *Plant Biotechnol. J.* 12.9 (Dec. 2014), pp. 1193–1206. URL: https://onlinelibrary.wiley.com/doi/10.1111/pbi.12279.

[319] Jianguo Yang et al. "Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity". In: *Proc. Natl. Acad. Sci.* 114.12 (2017), E2460–E2465. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1620058114.

[320] Xiaomeng Liu et al. "Combined Assembly and Targeted Integration of Multigene for Nitrogenase Biosynthetic Pathway in Saccharomyces cerevisiae." In: *ACS Synth. Biol.* (2019). URL: http://www.ncbi.nlm.nih.gov/pubmed/31117360.

[321] Stefan Burén et al. "Biosynthesis of the nitrogenase active-site cofactor precursor NifB-co in Saccharomyces cerevisiae". In: *Proc. Natl. Acad. Sci.* (2019), p. 201904903. URL: http://www.pnas.org/content/early/2019/11/22/1904903116.abstract.

[322] Deng Liu et al. "Engineering nitrogen fixation activity in an oxygenic phototroph". In: *MBio* 9.3 (2018), pp. 1–12. URL: https://mbio.asm.org/content/9/3/e01029-18.

[323] Jianguo Yang et al. "Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology". In: *Proc. Natl. Acad. Sci.* 115.36 (2018). URL: http://www.pnas.org/content/115/36/E8509.

[324] Álvaro Eseverri et al. "Use of synthetic biology tools to optimize the production of active nitrogenase Fe protein in chloroplasts of tobacco leaf cells". In: *Plant Biotechnol. J.* 18.9 (Sept. 2020), pp. 1882–1896. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/pbi.13347.

[325] Perrin H. Beatty and Allen G. Good. "Future prospects for cereals that fix nitrogen". In: *Science (80-.).* 333.6041 (July 2011), pp. 416–417. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.1209467.

[326] Qi Cheng and Wenli Sun. *Expression of Nitrogenase Structural Scaffold Genes in Higher Plant*. Tech. rep. 2018, p. 38. URL: https://www.researchgate.net/publication/333103297_BAOJ_Biotechnology_Expression_of_Nitrogenase_Structural_Scaffold_Genes_in_Higher_Plant.

[327] Nan Xiang et al. "Using synthetic biology to overcome barriers to stable expression of nitrogenase in eukaryotic organelles". In: *Proc. Natl. Acad. Sci.* (June 2020), p. 202002307. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.2002307117.

[328] Karsten Temme, Dehua Zhao, and Christopher A. Voigt. "Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca". In: *Proc. Natl. Acad. Sci.* 109.18 (2012), pp. 7085–7090. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.1120788109.

[329] Xia Wang et al. "Using Synthetic Biology to Distinguish and Overcome Regulatory and Functional Barriers Related to Nitrogen Fixation". In: *PLoS One* 8.7 (July 2013). Ed. by Szabolcs Semsey, e68677. URL: http://dx.plos.org/10.1371/journal.pone.0068677.

[330] Ray A. Dixon and John R. Postgate. "Transfer of nitrogen-fixation genes by conjugation in Klebsiella pneumoniae". In: *Nature* 234.5323 (Nov. 1971), pp. 47–48. URL: http://www.nature.com/articles/234048a0.

[331] Robin Van Velzen et al. "Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses". In: (2018). URL: http://www.pnas.org/content/pnas/early/2018/04/30/1721395115.full.pdf.

[332] George C. DiCenzo et al. "Genomic resources for identification of the minimal N 2 -fixing symbiotic genome". In: *Environ. Microbiol.* 18.8 (Sept. 2016), pp. 2534–2547. URL: http://doi.wiley.com/10.1111/1462-2920.13221.

[333] Jovelyn Unay and Xavier Perret. "Synthetic Plasmids to Challenge Symbiotic Nitrogen Fixation Between Rhizobia and Legumes". In: 2019, pp. 3–18. URL: https://doi.org/10.1007/978-981-13-5767-1_1.

[334] Michael J. Smanski et al. "Functional optimization of gene clusters by combinatorial design and assembly". In: *Nat. Biotechnol.* 32.12 (2014), pp. 1241–1249. URL: http://dx.doi.org/10.1038/nbt.3063.

[335] Barney A. Geddes et al. "Engineering transkingdom signaling in plants to control gene expression in rhizosphere bacteria". In: *Nat. Commun. 18-34342* 10.1 (Dec. 2019), p. 3430. URL: http://www.nature.com/articles/s41467-019-10882-x.

[336] Min-Hyung Ryu et al. "Control of nitrogen fixation in bacteria that associate with cereals". In: *Nat. Microbiol.* 5.2 (Feb. 2020), pp. 314–330. URL: https://doi.org/10.1038/s41564-019-0631-2.

[337] Catherine Masson-Boivin and Joel L. Sachs. "Symbiotic nitrogen fixation by rhizobia — the roots of a success story". In: *Curr. Opin. Plant Biol.* 44 (2018), pp. 7–15. URL: https://www.sciencedirect.com/science/article/abs/pii/S1369526617301243.

[338] Andrew W. B. Johnston and John E. Behringer. "Identification of the Rhizobium strains in pea root nodules using genetic markers". In: *J. Gen. Microbiol.* 87.2 (Apr. 1975), pp. 343–350. URL: https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-87-2-343.

[339] Simon H. Miller et al. "Host-specific regulation of symbiotic nitrogen fixation in Rhizobium leguminosarum biovar trifolii". In: *Microbiology* 153.9 (Sept. 2007), pp. 3184–3195. URL: http://www.ncbi.nlm.nih.gov/pubmed/17768261.

[340] Joseph Sambrook and David W. Russell. *Molecular cloning: a laboratory manual*. 3rd. New York: Cold Spring Harbor Laboratory, 2001. URL: https://books.google.co.uk/books/about/Molecular_Cloning.html?id=Bosc5JVxNpkC&redir_esc=y.

[341] John E. Beringer. "R factor transfer in Rhizobium leguminosarum". In: *J. Gen. Microbiol.* 84 (2015), pp. 188–198. URL: https://doi.org/10.1099/00221287-84-1-188.

[342] Francesco Pini et al. "Bacterial biosensors for in vivo spatiotemporal mapping of root secretion". In: *Plant Physiol.* 174.3 (July 2017), pp. 1289–1306. URL: http://www.ncbi.nlm.nih.gov/pubmed/28495892.

[343] Douglas Hanahan. "Studies on transformation of Escherichia coli with plasmids". In: *J. Mol. Biol.* 166.4 (1983), pp. 557–580. URL: https://pubmed.ncbi.nlm.nih.gov/6345791/.

[344] Sabrina Thoma and Max Schobert. "An improved Escherichia coli donor strain for diparental mating". In: *FEMS Microbiol. Lett.* 294.2 (May 2009), pp. 127–132. URL: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2009.01556.x.

[345] Marijke Frederix et al. "Mutation of praR in Rhizobium leguminosarum enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins". In: *Mol. Microbiol.* 93.3 (Aug. 2014), pp. 464–478. URL: http://doi.wiley.com/10.1111/mmi.12670.

[346] Rémy Fellay, Joachim Frey, and Henry Krisch. "Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for in vitro insertional mutagenesis of Gram-negative bacteria". In: *Gene* 52.2-3 (1987), pp. 147–154. URL: https://pubmed.ncbi.nlm.nih.gov/3038679/.

[347] Jürgen Prell et al. "The PTS Ntr system globally regulates ATP-dependent transporters in Rhizobium leguminosarum". In: *Mol. Microbiol.* 84.1 (Apr. 2012), pp. 117–129. URL: http://doi.wiley.com/10.1111/j.1365-2958.2012.08014.x.

[348] Jürgen Quandt and Michael F. Hynes. "Versatile suicide vectors which allow direct selection for gene replacement in Gram-negative bacteria". In: *Gene* 127.1 (May 1993), pp. 15–21. URL: https://pubmed.ncbi.nlm.nih.gov/8486283/.

[349] Andreas Schäfer et al. "Small mobilizable multi-purpose cloning vectors derived from the Escherichia coli plasmids pK18 and pK19: selection of defined deletions in the chromosome of Corynebacterium glutamicum". In: *Gene* 145.1 (July 1994), pp. 69–73. URL: https://pubmed.ncbi.nlm.nih.gov/8045426/.

[350] Oliver Kirchner and Andreas Tauch. "Tools for genetic engineering in the amino acid-producing bacterium Corynebacterium glutamicum". In: *J. Biotechnol.* 104.1-3 (Sept. 2003), pp. 287–299. URL: https://pubmed.ncbi.nlm.nih.gov/12948646/.

[351] Kyoung H. Choi et al. "Genetic tools for select-agent-compliant manipulation of Burkholderia pseudomallei". In: *Appl. Environ. Microbiol.* 74.4 (Feb. 2008), pp. 1064–1075. URL: https://pubmed.ncbi.nlm.nih.gov/18156318/.

[352] Gary Ditta et al. "Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of Rhizobium meliloti". In: *Proc. Natl. Acad. Sci.* 77.12 (Dec. 1980), pp. 7347–7351. URL: http://www.ncbi.nlm.nih.gov/pubmed/7012838.

[353] Emma Lodwig et al. "Regulation of L-Alanine Dehydrogenase in Rhizobium leguminosarum bv. viciae and Its Role in Pea Nodules". In: *J. Bacteriol.* 186.3 (Feb. 2004), pp. 842–849. URL: https://pubmed.ncbi.nlm.nih.gov/14729712/.

[354] Barney A. Geddes, Marcela A. Mendoza-Suárez, and Philip S. Poole. "A Bacterial Expression Vector Archive (BEVA) for Flexible Modular Assembly of Golden Gate-Compatible Vectors". In: *Front. Microbiol.* 9.JAN (Jan. 2019), p. 3345. URL: https://www.frontiersin.org/article/10.3389/fmicb.2018.03345/full.

[355] Stephan Heeb et al. "Small, Stable Shuttle Vectors Based on the Minimal pVS1 Replicon for Use in Gram-Negative, Plant-Associated Bacteria". In: *Mol. Plant-Microbe Interact.* 13.2 (2000), pp. 232–237. URL: https://apsjournals.apsnet.org/doi/pdf/10.1094/MPMI.2000.13.2.232.

[356] Adrian J. Tett et al. "Regulatable vectors for environmental gene expression in Alphaproteobacteria". In: *Appl. Environ. Microbiol.* 78.19 (2012), pp. 7137–7140. URL: https://aem.asm.org/content/78/19/7137.

[357] Ernst Weber et al. "A modular cloning system for standardized assembly of multigene constructs". In: *PLoS One* 6.2 (Feb. 2011). Ed. by Jean Peccoud. URL: https://dx.plos.org/10.1371/journal.pone.0016765.

[358] Carola Engler et al. "Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes." In: *PLoS One* 4.5 (May 2009). Ed. by Jean Peccoud, e5553. arXiv: 1403.5698. URL: http://dx.plos.org/10.1371/journal.pone.0005553.

[359] Carola Engler et al. "A Golden Gate modular cloning toolbox for plants". In: *ACS Synth. Biol.* 3.11 (Nov. 2014), pp. 839–843. URL: https://pubs.acs.org/doi/10.1021/sb4001504.

[360] Vicky Buchanan-Wollaston. "Generalized transduction in Rhizobium leguminosarum". In: *J. Gen. Microbiol.* 112.1 (May 1979), pp. 135–142. URL: https://doi.org/10.1099/00221287-112-1-135.

[361] Kyoung H. Choi and Herbert P. Schweizer. "mini-Tn7 insertion in bacteria with single attTn7 sites: example Pseudomonas aeruginosa". In: *Nat. Protoc.* 1.1 (2006), pp. 153–161. URL: https://www.nature.com/articles/nprot.2006.24.

[362] Graham A. Hood. "Physiological response of Rhizobium leguminosarum during bacteroid development". PhD thesis. 2013. URL: https://ueaeprints.uea.ac.uk/id/eprint/48693/.

[363] Philip S. Poole et al. "Myo-inositol catabolism and catabolite regulation in Rhizobium leguminosarum bv. viciae". In: *Microbiology* (1994). URL: https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-140-10-2787.

[364] David Allaway et al. "Identification of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea bacteroids". In: *Mol. Microbiol.* (2000). URL: https://pubmed.ncbi.nlm.nih.gov/10792736/.

[365] Shuhei Tsukada et al. "Comparative genome-wide transcriptional profiling of Azorhizobium caulinodans ORS571 grown under free-living and symbiotic conditions". In: *Appl. Environ. Microbiol.* 75.15 (2009), pp. 5037–5046. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2725507/.

[366] Lucie McMurtry and Poole Philip. *Determining the induction and regulation of nif and fix N fixation genes in Rhizobium leguminosarum.* Tech. rep. 2015.

[367] Fábio Madeira et al. "The EMBL-EBI search and sequence analysis tools APIs in 2019". In: *Nucleic Acids Res.* 47.W1 (July 2019), pp. 636–641. URL: https://academic.oup.com/nar/article/47/W1/W636/5446251.

[368] Andrew M. Waterhouse et al. "Jalview Version 2 - A multiple sequence alignment editor and analysis workbench". In: *Bioinformatics* 25.9 (May 2009), pp. 1189–1191. URL: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp033.

[369] Shennan Lu et al. "CDD/SPARCLE: The conserved domain database in 2020". In: *Nucleic Acids Res.* 48.D1 (Jan. 2020), pp. D265–D268. URL: https://pubmed.ncbi.nlm.nih.gov/31777944/.

[370] Aron Marchler-Bauer et al. "CDD: A Conserved Domain Database for the functional annotation of proteins". In: *Nucleic Acids Res.* 39.SUPPL. 1 (Jan. 2011). URL: https://pubmed.ncbi.nlm.nih.gov/21109532/.

[371] Helga Thorvaldsdóttir, James T. Robinson, and Jill P. Mesirov. "Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration". In: *Brief. Bioinform.* 14.2 (Mar. 2013), pp. 178–192. URL: https://academic.oup.com/bib/article/14/2/178/208453.

[372] James T. Robinson et al. "Integrative genomics viewer". In: *Nat. Biotechnol.* 29.1 (Jan. 2011), pp. 24–26. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3346182/.

[373] Áki J. Láruson, Sam Yeaman, and Katie E. Lotterhos. *The importance of genetic redundancy in evolution.* Sept. 2020. URL: http://www.cell.com/article/S0169534720301166/fulltext.

[374] Giulio Tononi, Olaf Sporns, and Gerald M. Edelman. "Measures of degeneracy and redundancy in biological networks". In: *Proc. Natl. Acad. Sci. U. S. A.* 96.6 (1999), pp. 3257–3262. URL: https://www.pnas.org/content/96/6/3257.

[375] Yann S. Dufour, Patricia J. Kiley, and Timothy J. Donohue. "Reconstruction of the core and extended regulons of global transcription factors". In: *PLoS Genet.* 6.7 (July 2010), pp. 1–20. URL: http://www.ncbi.nlm.nih.gov/pubmed/20661434.

[376] Oliver Preisig, Rachel Zufferey, and Hauke Hennecke. "The Bradyrhizobium japonicum fixGHIS genes are required for the formation of the high-affinity cbb3-type cytochrome oxidase". In: *Arch. Microbiol.* 165.5 (May 1996), pp. 297–305. URL: http://link.springer.com/10.1007/s002030050330.

[377] Ramakrishnan Karunakaran et al. "Transcriptomic analysis of Rhizobium leguminosarum biovar viciae in symbiosis with host plants Pisum sativum and Vicia cracca". In: *J. Bacteriol.* 191.12 (June 2009), pp. 4002–4014. URL: http://www.ncbi.nlm.nih.gov/pubmed/19376875.

[378] Robert T. Green et al. "Transcriptomic analysis of rhizobium leguminosarum bacteroids in determinate and indeterminate nodules". In: *Microb. Genomics* 5.2 (2019). URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6421345/.

[379] Claus Lang, Lucinda S. Smith, and Sharon R. Long. "Characterization of novel plant symbiosis mutants using a new multiple gene-expression reporter Sinorhizobium meliloti strain". In: *Front. Plant Sci.* 9.February (2018), p. 76. URL: http://journal.frontiersin.org/article/10.3389/fpls.2018.00076/full.

[380] Aleksandr Gavrin et al. "Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalization, HOPS suppression, and TIP1g retargeting in Medicago". In: *Plant Cell* 26.9 (Sept. 2014), pp. 3809–3822. URL: https://academic.oup.com/plcell/article/26/9/3809/6100361.

[381] Marcela A. Mendoza-Suárez et al. "Optimizing Rhizobium-legume symbioses by simultaneous measurement of rhizobial competitiveness and N2 fixation in nodules." In: *Proc. Natl. Acad. Sci. U. S. A.* 117.18 (May 2020), pp. 9822–9831. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1921225117.

[382] Delphine Capela et al. "Sinorhizobium meliloti differentiation during symbiosis with Alfalfa: a transcriptomic dissection". In: *Mol. Plant-Microbe Interact.* 19.4 (Apr. 2006), pp. 363–372. URL: https://apsjournals.apsnet.org/doi/10.1094/MPMI-19-0363.

[383] Brett J. Ferguson et al. "Legume nodulation: the host controls the party". In: *Plant. Cell Environ.* 42.1 (Jan. 2019), pp. 41–51. URL: http://doi.wiley.com/10.1111/pce.13348.

[384] Sharon R. Long. "Rhizobium-legume nodulation: life together in the underground". In: *Cell* 56.2 (Jan. 1989), pp. 203–214. URL: http://www.sciencedirect.com/science/article/pii/0092867489908933.

[385] Barbara Łotocka, Joanna Kopcińska, and Monika Skalniak. "Review article: the meristem in indeterminate root nodules of Faboideae". In: *Symbiosis* 58.1-3 (2012), pp. 63–72. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3589660/.

[386] Claudia Popp and Thomas Ott. "Regulation of signal transduction and bacterial infection during root nodule symbiosis". In: *Curr. Opin. Plant Biol.* 14.4 (Aug. 2011), pp. 458–467. URL: https://doi.org/10.1016/j.pbi.2011.03.016.

[387] Jonathan D. Monroe, Thomas G. Owens, and Thomas A. LaRue. "Measurement of the fractional oxygenation of leghemoglobin in intact detached pea nodules by reflectance spectroscopy". In: *Plant Physiol.* 91.2 (Oct. 1989), pp. 598–602. URL: http://www.plantphysiol.org/lookup/doi/10.1104/pp.91.2.598.

[388] Agota Domonkos et al. "The identification of novel loci required for appropriate nodule development in Medicago truncatula". In: *BMC Plant Biol.* 13.1 (Oct. 2013). URL: https://bmcplantbiol.biomedcentral.com/articles/10.1186/1471-2229-13-157.

[389] Vera A. Voroshilova et al. "Effect of mutations in Pisum sativum L. genes blocking different stages of nodule development on the expression of late symbiotic genes in Rhizobium leguminosarum bv. viciae". In: *Mol. Plant-Microbe Interact.* 14.4 (2001), pp. 471–476. URL: https://pubmed.ncbi.nlm.nih.gov/11310734/.

[390] Limin Zheng, Robert H. White, and Dennis R. Dean. "Purification of the Azotobacter vinelandii nifV-encoded homocitrate synthase". In: *J. Bacteriol.* 179.18 (1997), pp. 5963–5966. URL: https://jb.asm.org/content/179/18/5963.

[391] Luis M. Rubio and Paul W. Ludden. "Biosynthesis of the Cofactor of Nitrogenase". In: *Annu. Rev. Microbiol.* 62.1 (Oct. 2008), pp. 93–111. URL: http://www.annualreviews.org/doi/10.1146/annurev.micro.62.081307.162737.

[392] Jacqueline Y. Quinn et al. "SBOL Visual: A Graphical Language for Genetic Designs". In: *PLoS Biol.* 13.12 (Dec. 2015), pp. 1–9. URL: http://dx.plos.org/10.1371/journal.pbio.1002310.

[393] Jacob Beal et al. "Communicating Structure and Function in Synthetic Biology Diagrams". In: *ACS Synth. Biol.* (July 2019), acssynbio.9b00139. URL: https://pubs.acs.org/doi/pdf/10.1021/acssynbio.9b00139?rand=9lhm61lq.

[394] Jason R. Kelly et al. "Measuring the activity of BioBrick promoters using an in vivo reference standard". In: *J. Biol. Eng.* 3.34 (2009), p. 4. URL: http://www.jbioleng.org/content/3/1/4.

[395] Isabel Webb. "Characterisation of the fixABCX operon in symbiotic nitrogen fixation". PhD thesis. University of East Anglia, 2016. URL: https://ueaeprints.uea.ac.uk/id/eprint/63286/.

[396] Carmen Sánchez-Cañizares et al. “Global control of bacterial nitrogen and carbon metabolism by a PTS Ntr -regulated switch”. In: *Proc. Natl. Acad. Sci.* (2020). URL: https://www.pnas.org/content/117/19/10234.

[397] Jan Lisec et al. “Gas chromatography mass spectrometry–based metabolite profiling in plants”. In: *Nat. Protoc.* 1.1 (June 2006), pp. 387–396. URL: http://www.nature.com/articles/nprot.2006.59.

[398] Rafael Gómez-Bombarelli, Emilio Calle, and Julio Casado. “Mechanisms of lactone hydrolysis in neutral and alkaline conditions”. In: *J. Org. Chem.* 78.14 (July 2013), pp. 6868–6879. URL: https://pubs.acs.org/doi/full/10.1021/jo400258w.

[399] John T. Sullivan, Steven D. Brown, and Clive W. Ronson. “The NifA-RpoN Regulon of Mesorhizobium loti Strain R7A and Its Symbiotic Activation by a Novel LacI/GalR-Family Regulator”. In: *PLoS One* 8.1 (2013). URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0053762.

[400] Clive W. Ronson, B. Tracy Nixon, and Frederick M. Ausubel. “Conserved domains in bacterial regulatory proteins that respond to environmental stimuli”. In: *Cell* 49.5 (June 1987), pp. 579–581. URL: https://linkinghub.elsevier.com/retrieve/pii/0092867487905307.

[401] Peng Jiang and Alexander J. Ninfa. “Regulation of autophosphorylation of Escherichia coli nitrogen regulator II by the PII signal transduction protein”. In: *J. Bacteriol.* 181.6 (1999), pp. 1906–1911. URL: https://pubmed.ncbi.nlm.nih.gov/10074086/.

[402] Daniel R. Brown et al. “Combinatorial stress responses: Direct coupling of two major stress responses in escherichia coli”. In: 1.9 (Sept. 2014), pp. 315–317. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5349134/.

[403] Sacha De Carlo et al. “The structural basis for regulated assembly and function of the transcriptional activator NtrC”. In: *Genes Dev.* 20.11 (2006), pp. 1485–1495. URL: http://genesdev.cshlp.org/content/20/11/1485.abstract.

[404] Gary N. Gussin, Clive W. Ronson, and Frederick M. Ausubel. “Regulation of nitrogen fixation genes”. In: *Annu. Rev. Genet.* 20 (1986), pp. 567–591. URL: https://www.annualreviews.org/doi/10.1146/annurev.ge.20.120186.003031.

[405] Susan C. Porter et al. “Oligomerization of NTRC at the glnA enhancer is required for transcriptional activation”. In: *Genes Dev.* 7.11 (1993), pp. 2258–2273. URL: http://genesdev.cshlp.org/content/7/11/2258.full.pdf.

[406] Emanuel M. Souza et al. “Expression of the nifA gene of Herbaspirillum seropedicae: Role of the NtrC and NifA binding sites and of the -24/-12 promoter element”. In: *Microbiology* 146.6 (June 2000), pp. 1407–1418. URL: https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-146-6-1407.

[407] David W. Ow and Frederick M. Ausubel. “Regulation of nitrogen metabolism genes by nifA gene product in Klebsiella pneumoniae”. In: *Nature* 301.5898 (1983), pp. 307–313. URL: https://www.nature.com/articles/301307a0.

[408] Rachel M. Wheatley et al. “Role of O2 in the Growth of Rhizobium leguminosarum bv. viciae 3841 on Glucose and Succinate”. In: *J. Bacteriol.* 199.1 (Jan. 2017). Ed. by Anke Becker, e00572–16. URL: http://jb.asm.org/lookup/doi/10.1128/JB.00572-16.

[409] Federico E. Rey and Caroline S. Harwood. "FixK, a global regulator of microaerobic growth, controls photosynthesis in Rhodopseudomonas palustris". In: *Mol. Microbiol.* 75.4 (Feb. 2010), pp. 1007–1020. URL: http://doi.wiley.com/10.1111/j.1365-2958.2009.07037.x.

[410] Marilyn Dispensa et al. "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators". In: *J. Bacteriol.* 174.18 (1992), pp. 5803–5813. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC207109/.

[411] Paul G. Egland and Caroline S. Harwood. "BadR, a new MarR family member, regulates anaerobic benzoate degradation by Rhodopseudomonas palustris in concert with AadR, an Fnr family member". In: *J. Bacteriol.* 181.7 (1999), pp. 2102–2109. URL: https://pubmed.ncbi.nlm.nih.gov/10094687/.

[412] Frank W. Larimer et al. "Complete genome sequence of the metabolically versatile photosynthetic bacterium Rhodopseudomonas palustris". In: *Nat. Biotechnol.* 22.1 (Jan. 2004), pp. 55–61. URL: https://www.nature.com/articles/nbt923.

[413] E. Toby Kiers et al. "Host sanctions and the legume-rhizobium mutualism". In: *Nature* 425.6953 (2003), pp. 78–81. URL: https://pubmed.ncbi.nlm.nih.gov/12955144/.

[414] Annet Westhoek et al. "Policing the legume-Rhizobium symbiosis: a critical test of partner choice". In: *Sci. Rep.* 7.1 (May 2017), p. 1419. URL: http://www.nature.com/articles/s41598-017-01634-2.

[415] Katy D. Heath and Peter Tiffin. "Stabilizing mechanisms in a legume-rhizobium mutualism". In: *Evolution (N. Y).* 63.3 (Mar. 2009), pp. 652–662. URL: http://doi.wiley.com/10.1111/j.1558-5646.2008.00582.x.

[416] Rebecca T. Batstone et al. "The evolution of symbiont preference traits in the model legume Medicago truncatula". In: *New Phytol.* 213.4 (2017), pp. 1850–1861. URL: https://nph.onlinelibrary.wiley.com/doi/full/10.1111/nph.14308.

[417] Héctor Osorio et al. "Identification and unusual properties of the master regulator FNR in the extreme acidophile Acidithiobacillus ferrooxidans". In: *Front. Microbiol.* 10 (July 2019), p. 1642. URL: https://www.frontiersin.org/article/10.3389/fmicb.2019.01642/full.

[418] Elmar L. Kannenberg and Nicholas J. Brewin. "Expression of a cell surface antigen from Rhizobium leguminosarum 3841 is regulated by oxygen and pH". In: *J. Bacteriol.* 171.9 (1989), pp. 4543–4548. URL: https://pubmed.ncbi.nlm.nih.gov/2768181/.

[419] Patricia R. Rosas. "Análisis funcional y fenotípico de una mutante fnr de Rhizobium etli CFN42". PhD thesis. Universidad Autónoma Metropolitana, 2012. URL: http://tesiuami.izt.uam.mx/uam/aspuam/presentatesis.php?recno=17310&docs=UAMI17310.pdf.

[420] Janos Z. Kelemen et al. "Wide-range transcriptional modulating effect of ntrR under microaerobiosis in Sinorhizobium meliloti". In: *Mol. Genet. Genomics* 272.3 (2004), pp. 275–289. URL: https://pubmed.ncbi.nlm.nih.gov/15365818/.

[421] Mario Soberón et al. "A Purine-Related Metabolite Negatively Regulates fixNOQP Expression in Sinorhizobium meliloti by Modulation of fixK Expression". In: *Mol. Plant-Microbe Interact.* 14.4 (Apr. 2001), pp. 572–576. URL: https://apsjournals.apsnet.org/doi/pdfplus/10.1094/MPMI.2001.14.4.572.

[422] Maria Soberón et al. "Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic terminal oxidase cbb3 in Rhizobium etli." In: *J. Bacteriol.* 179.22 (2016), pp. 6887–6893. URL: https://pubmed.ncbi.nlm.nih.gov/9371431/.

[423] Maria Soberón et al. "Genetic evidence for 5-aminoimidazole-4-carboxamide ribonucleotide (AICAR) as a negative effector of cytochrome terminal oxidase cbb 3 production in Rhizobium etli". In: *MGG Mol. Gen. Genet.* 254.6 (May 1997), pp. 665–673. URL: https://link.springer.com/content/pdf/10.1007/s004380050464.pdf.

[424] Philippe Remigi et al. "Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts". In: 24.1 (Jan. 2016), pp. 63–75. URL: https://doi.org/10.1016/j.tim.2015.10.007.

[425] Camille Clerissi et al. "Parallels between experimental and natural evolution of legume symbionts". In: *Nat. Commun.* 9.1 (2018), pp. 1–13. URL: http://dx.doi.org/10.1038/s41467-018-04778-5.

[426] Jian B. Xie et al. "Comparative Genomic Analysis of N 2 -Fixing and Non-N 2 -Fixing Paenibacillus spp.: Organization, Evolution and Expression of the Nitrogen Fixation Genes". In: *PLoS Genet.* 10.3 (Mar. 2014). Ed. by Paul M. Richardson, e1004231. URL: https://dx.plos.org/10.1371/journal.pgen.1004231.