AN IN SILICO INVESTIGATION INTO NITROGEN FIXATION

NIMA PAKSERESHT

Ph.D.

University of East Anglia and John Innes Centre
Faculty of Science
Computing Sciences

February 2010

ABSTRACT

One of the most important parts of the nitrogen cycle is nitrogen fixation. We present a computational investigation into two aspects of nitrogen fixation: (1) For the mechanism of nitrogenase action in K. *pneumoniae*, we present the full Thorneley and Lowe model developed from the literature with clarifications and corrections in SBML. Additionally, we perform an analysis of the rate-limiting step and protein ratio in this mechanism. This shows The ratio of proteins plays key role in ammonia production and it decrease the sensitivity of the system to the dissociation constant (k_{-3}, rate limiting state). (2) We study the genetic regulation of nitrogen fixation in *A. vinelandii* and K. *pneumoniae*. To take into account the module-like behavior of the $NifL - NifA$ system in response to environmental signals, we broke the system up into a set of subsystems including the $NifL - NifA$ system in oxidized conditions, the $2 - oxoglutarate$ system, and the $GlnK$ system. Computational studies of these three systems reveal some of the characteristics of the $NifL - NifA$ system in various environmental conditions including (a) oxygen increases the affinity of $NifL$ to interact with $NifA$ 10 fold (b) in contrast to biochemical knowledge that stoichiometric amount of $NifA$ and $NifL$ is available in the system the ratio of $NifL$ to $NifA$ had to be approximately 2 for satisfactory simulations of the system (c) $GlnK$ is a master regulator in the $NifL - NifA$ system, forming an on and off binary system. Finally, we present abstract models of nitrogen fixation in *A. vinelandii* and K. *pneumoniae* based on current understanding of these systems. Comparison of these two models showed that *A. vinelandii* has a more sensitive system to the availability of fixed nitrogen and responds faster than K. *pneumoniae* to environmental conditions. Moreover both models were able to oscillate.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

MY PH.D. YEARS

This thesis is the result of almost four years of hard effort at the University of East Anglia and the John Innes Centre. In the course of my MSc dissertation on bioinformatics I became involved for some time in a systems biology project that I initiated based on my desire to learn this subject area and my interest in nitrogen fixation. I made up my mind to undertake a Ph.D. in the systems biology. Although this was new for my supervisors, we initiated a plan to study this complex system *in silico*.

I enjoyed this exciting project, as it demanded the initiation of new directions, motivation, and hard work. In addition, my student life for the last four years has been highly challenging. Although the studentship converted a 'mission impossible' into a mission possible, it covered only tuition fees so I had to find a way to obtain the rest of my living costs. Gratefully, I was offered a part-time job as a support bioinformatician at the John Innes Centre (20 hours a week). Even though being a full-time Ph.D. student as well as working 20 hours a week is demanding, my inspiration and motivation for my Ph.D. helped me to make progress and finish my thesis.

PROJECT EVOLUTION

We started this project with no prior systems biology knowledge about nitrogen fixation. Furthermore, the complexity of the system and the availability of data were also not clear. Hence I started by literature review of nitrogen fixation and discussing the system with experts to extract relevant knowledge and data. The initial aim was set to investigate full nitrogen fixation in *A. vinelandii* in dynamical conditions. We attempted to model the complete system, however, the complexity of nitrogen fixation and lack of experimental data resulted in an uninformative dynamic model. Meanwhile, I was offered the possibility to work on a related side project concerning the

mechanism of nitrogenase action by David J Lowe, head of computational systems biology at the John Innes Centre at that time. He had developed a pre-steady a state kinetic model of nitrogenase action in 1984, although the model was not available due to technical issues. The aim of the side project was to explore systems biology tools for an example system which is relevant to nitrogen fixation, and also to make Lowe's full model of the system available for the scientific community in the system biology markup language (SBML).

As my understanding of the relevant tools and nitrogen fixation developed, closer inspection concerning the availability of data, the complexity, and the behavior of the NifL – NifA system in response to different environmental signals suggested breaking it down into a more manageable set of sub systems and to investigating the sub systems in steady state conditions instead of dynamical modelling. Modelling the sub systems in steady state revealed some of the characteristics of nitrogen fixation. Finally, the extracted knowledge from modelling the sub systems, the complexity of the complete system, and our experience in first attempt in modelling nitrogen fixation in dynamical conditions allowed me to develop a dynamic regulation model of nitrogen fixation at an abstract level which allowed us to uncover properties of the nitrogen fixation and to reduce the complexity of our model.

PRESENTATION

This thesis consists of seven chapters. In Chapter 1 we review the relevant biological literature, which covers the regulation of the NifL – NifA system in *A. vinelandii*, the mechanism of nitrogenase action, and the NifL – NifA regulatory system in *K. pneumoniae*. It is followed in Chapter 2 by an overview of current computational methods in systems biology including kinetics modelling, the dynamics of transcription, simulation methodologies, and finally a brief description of structural analysis of systems. It also includes a brief description of state of the art systems biology tools.

Chapter 3 is a small part of this thesis in which we present an SBML model of nitrogenase action. The remaining part deals with the main topic of the PhD work, the NifL – NifA regulatory system. We start by modelling the smallest sub systems of the NifL – NifA system (Chapter 4). Then we

extend these sub systems to larger systems (Chapter 5). Chapter 6 puts all the bits and pieces of the last two chapters together, resulting in an abstract model of nitrogen fixation in A. *vinelandii* and K. *pneumoniae*. Finally, in Chapter 7 we presents a brief conclusion.

ACKNOWLEDGMENTS

The author would like to give special thank to his supervisors, Prof. Vincent Moulton, Dr. Richard Morris, and Prof. Ray Dixon, for their help, the ideas that they have contributed in this thesis, their advice, and teaching. He is also grateful to them for the care and consideration that they give him during the course of his Ph.D.

He would like to thank Prof. David lowe for his invaluable guidance in preparing the SBML model of mechanism of nitrogenase action. He is also thankful to his bioinformatics job managers, Mr. Sean Walsh and Dr. John Walshaw, for their support and patience during his Ph.D.

The author would like to thank his group members and people in the computation system biology department in the John Innes Centre for the kind and friendly work environment that they provided.

He is very grateful to his parents for many years of their financial and emotional support, care, and consideration in his education. Without their support he would not have been successful in his education. He also wants to thank his brothers and friends who gave him encouragement to finish his Ph.D. The author also would like to thank his wife and his five month old daughter for their patience and support that gave him during tough and stressful times.

Finally, the author thanks the University of East Anglia and the John Innes Centre for generously funding his tuition fees and laboratory expenses, respectively.

Part I

BACKGROUND AND METHODOLOGY

1

INTRODUCTION

The main source of nitrogen is atmospheric N_2, which comprises approximately 78% of the atmosphere. The process representing the transformation of nitrogen in nature is called the nitrogen cycle [5, 20]. Nitrogen cycle processes include assimilation, ammonification, nitrification, denitrification, anaerobic ammonium oxidation, and nitrogen fixation. Figure 1 presents a schematic overview of the nitrogen cycle.

One of the most important parts of the nitrogen cycle is nitrogen fixation, since usable nitrogen is a limiting nutrient for living organisms. Nitrogen fixation is the conversion of dinitrogen, N_2, to ammonia, NH_3, which is usable for living organisms (Figure 1, r1). Plants are able to absorb ammonia, nitrites, and nitrates from the soil and produce protein in the assimilation process (r2). Animals do not have the enzyme required for catalyzing nitrogen fixation and consume plants as a source of nitrogen. Animals and plants can be decomposed in an ammonification process by decomposers, which results in the release of ammonia back into the environment (r3). Soil ammonia converts to nitrites in a process called nitrification by nitrifying bacteria (r4). Nitrites, which are useable for plants, can be converted to nitrates by nitrifying bacteria (r5). Finally, nitrites return to the atmosphere through conversion to N_2 by denitrifying bacteria in the denitrification process (r6) [5, 20]. In anaerobic ammonium oxidation, nitrite and ammonium are converted directly into atmospheric N_2 (r7).

Although atmospheric nitrogen is abundant, it is not easily accessible since it requires the breakage of one of the strongest covalent bonds in nature [94, 77]. One of the most important recent challenges of agriculture is feeding the growing global population that is anticipated to be 8 billion by the year 2020 [78, 4, 93]. The increasing demand for crops will extend the use of nitrogen-based fertilizers dramatically, though this requires a considerable amount of world energy. In contrast to industrial nitrogen fixation, biological nitrogen fixation is catalyzed by the enzyme nitrogenase in some microorganisms with no cost to world economics [6, 34, 28, 73]. As

nitrogen is often the limiting nutrient for the growth of plants, industry has invested in converting the relatively inert dinitrogen to the more accessible form of ammonia, NH_3. The Haber-Bosch process produces in the order of 100 million tons of industrial fertilizer each year, which contributes to sustaining around 40% of the world population [25, 83, 81]. This process requires temperatures of about 400 °C and pressures of around 400 bar, consuming approximately 1% of the world's annual energy supply. In the light of huge energy resources, the far-reaching environmental issues relating to fertilizers [64], and especially for biofuels and crop productivity [84], the process of biological nitrogen fixation is of prime importance and will remain a focus of research for a long time [47, 71].

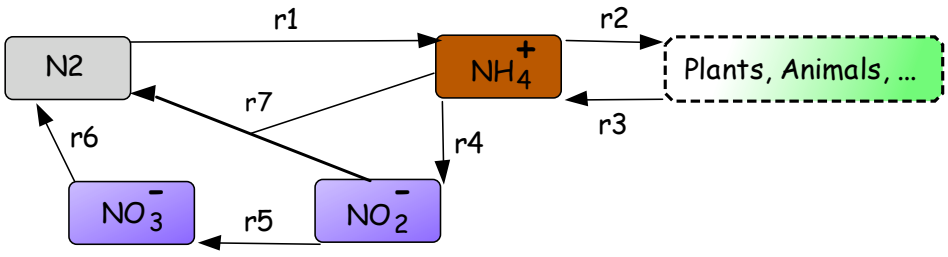


Figure 1: An overview of the nitrogen cycle. See main text for more details.

1.1 NITROGEN FIXATION

Biological nitrogen fixation [98, 28, 35] is performed by a number of different prokaryotes, including bacteria and actinobacteria, collectively known as diazotrophs and archaea. It does not occur in eukaryotes. The main characteristic of diazotrophic prokaryotes is the ability to utilize N_2 as a source of nitrogen for growth. Diazotrophic bacteria have very similar enzyme systems to each other for fixing nitrogen. While most of diazotrophs are free-living bacteria, some diazotrophs form symbiotic associations with plants, supplying fixed nitrogen for plants. The availability of carbon and energy sources, levels of oxygen, and presence of fixed nitrogen are environmental factors that affect nitrogen fixation in free-living diazotrophs. Consequently, the contribution of free-living diazotrophs to the nitrogen input for agriculture depends on the availability of an energy source in their ecosystems [13, 73].

Klebsiella pneumoniae is a common gram-negative bacterium, which has been well studied among the diazotrophs [14]. K. *pneumoniae* is clinically important as it is usually found in the intestine, on the skin, and in the normal flora of the mouth, and it can cause infections. It has a close relation to some well-studied bacteria such as *Escherichia coli* and *Salmonella typhimurium*.

Azotobacter vinelandii is a free-living nitrogen-fixing bacterium, which is an aerobic soil bacterium [14]. *A. vinelandii* is of interest to scientists because it can fix nitrogen under aerobic conditions, conferring considerable energetic benefits, and is capable of synthesizing three different nitrogenase enzymes. *A.vinelandii* has a very high respiratory rate which generates extra ATP, required to support the high-energy demands of nitrogenase [15, 12, 60]. In addition to these capabilities, *A. vinelandii* is able to grow in a variety of organic acids, alcohols, or carbohydrates. Although other bacteria are able to fix nitrogen, there are some major differences between them and *A. vinelandii*. It can fix nitrogen, and unlike others the fixation can occur under high ambient oxygen tension. In this research work, nitrogen fixation by *A. vinelandii* is the main topic because of these unique characteristics [15].

There are two important aspects of nitrogen fixation; the mechanism of the nitrogenase enzyme, catalyzing nitrogen fixation, and the genetic regulation of nitrogen fixation, which is modulated by two key regulatory proteins, *NifA* and *NifL*. In the following sections we review these aspects of nitrogen fixation. K. *pneumoniae* is the focus of our work for the nitrogenase aspect and *A. vinelandii* for genetic regulation.

1.2 NITROGENASE

A family of enzymes known as nitrogenase (EC 1.18.6.1) catalyze biological nitrogen fixation [16, 32]. Nitrogenase contains two components, which are named based on their constituent metals. The nitrogenase cofactors may contain vanadium, V; molybdenum, Mo; or iron, Fe. There are three types of nitrogenase systems available based on their cofactors; the molybdenum- nitrogenase, the vanadium-nitrogenase, and the iron only system. K. *pneumoniae* has a Mo-nitrogenase system. It consists of two oxy-

gen sensitive metalloproteins, an iron (Fe) protein and a molybdenum-iron (MoFe) protein [72, 11]. The vanadium-nitrogenase system has two components. It has a Fe protein which is the same as other nitogenase systems and the second component is a vanadium-iron (VFe) containing protein which is different than two other systems. This type of nitrogenase has been detected in *A. vinelandii* and *A.chroococcum* [13]. Vanadium-nitrogenase has a very similar amino acid sequence to the molybdenum-nitrogenase. The third type of nitrogenase, iron only, contains an iron (Fe) protein and another protein, which is very similar to MoFe protein and VFe protein, while it has only Fe as its cofactor [13]. This type of protein has also been detected in *A. vinelandii* nitrogenase.

Nitrogenase has some unique characteristics, which distinguish it from some other enzymes. It is highly sensitive to oxygen, and is degraded by the reaction of oxygen within the iron component. The high sensitivity of the enzyme to oxygen means the nitrogen-fixing bacteria need some protection mechanism to prevent the degradation. Furthermore, the turnover time of the enzyme is relatively slow, therefore, the organism needs to synthesize large quantities of nitrogenase [10]. Nitrogenase also requires considerable amounts of energy for catalyzing nitrogen fixation. In the following sections we present the properties of the Fe protein, the MoFe protein, and finally the mechanism of nitrogenase action. With regards to K. *pneumoniae* nitrogenase, in the following sections we review the molybdenum- nitrogenase system.

1.2.1 *The Fe protein*

The smaller component of nitrogenase is the Fe protein, which acts as a redox-active agent and transfers electrons to the MoFe protein for reduction of substrates from available electron donor in the system. Although this transfer of electrons is the main function of the Fe protein, it has some other functions. The Fe protein is needed for initial biosynthesis of the MoFe cofactor. Following the biosynthesis of MoFe cofactor, the insertion of the preformed MoFe cofactor into the MoFe protein requires the Fe protein [3]. The Fe protein is ATP-dependant and hydrolyses MgATP using the acquired energy to transfer electrons to the MoFe protein. It has two identical subunits.

The Fe protein of nitrogenase contains one iron sulfur cluster [4Fe-4S], which bridges the two subunits. The oxidation state of the Fe protein can vary through the iron sulfur cluster. The Fe protein has one MgATP binding site in each subunit that binds to two MgATP molecules. Binding of MgATP to the Fe protein induces conformational changes, which facilitate the electron transfer from the Fe protein to the MoFe protein.

1.2.2 *The MoFe protein*

The larger component of nitrogenase is the MoFe protein, which is a $\alpha_2\beta_2$ tetramer, containing two $\alpha\beta$ dimer subunits. Each dimer contains one MoFe cofactor and one P-cluster, [8Fe-7S]. The MoFe cofactor is located in the active site of the protein where the reduction of substrates occurs. The main role of the P-cluster is electron transfer by accepting an electron from the Fe protein and donating it to the MoFe cofactor. The $\alpha\beta$ dimeric units communicate and contact each other through their β subunits [3]. The P cluster bridges between each α and β subunit while the MoFe cofactor is placed on the α subunits.

1.2.3 *The mechanism of nitrogenase action*

As described above, nitrogenase comprises two proteins, the Fe protein and the MoFe protein. The MoFe protein contains a site where electrons accumulate for substrate reduction. In order to pull the nitrogen atoms away from each other in a dinitrogen molecule, a triple bond must be broken, i.e. sufficient energy is required to break all three-paired orbitals. Nitrogenase accomplishes this by accumulating electrons and breaking each bond individually. The Fe protein and MoFe proteins form the Fe protein-MoFe complex to catalyze the substrates. The nitrogenase complex contains one MoFe protein in the middle and two Fe proteins at each side. The overall behavior of nitrogenase is to transfer an electron from the Fe protein to the MoFe protein and reduction of the substrate. The process starts from the nucleotide binding sites in the Fe proteins and goes through the [4Fe-4S] cluster which transfers an electron to the P cluster and finally from the

P cluster to the MoFe cofactor side which is the active side of protein for substrate reduction.

In 1984 Thorneley and Lowe presented a scheme based on kinetic stopped-flow measurements, describing the nitrogenase N_2 turnover cycle [56, 91, 57, 92]. The reaction mechanism of nitrogenase can be separated into two nested cycles; the full cycle of N_2 being reduced, which requires eight electrons and the Fe-protein cycle which provides these electrons one at a time. Each Fe protein cycle delivers one electron and the MoFe protein cycle is the full cycle of the dinitrogen reduction. The delivery of electron is coupled to the hydrolysis of two MgATP molecules which provides energy for electron delivery. The Fe cycle occurs eight times to provide eight electrons which catalyze the reduction of N_2 to NH_3. *In vitro*, nitrogenase activity requires the interaction of the two protein components, a low-potential source of electrons, MgATP and a substrate [90, 89, 85, 36, 17, 29]. The Fe protein binds two MgATPs and, with dithionite as reductant, carries out functions using the redox couple to transfer one electron to the MoFe protein and the hydrolysis of two MgATPs [34]. The Fe protein cycle comprises (1) the Fe protein-MoFe protein complex formation, (2) MgATP hydrolysis and electron transfer which cause the Fe protein to be in an oxidised form with MgADP and the MoFe protein to be in reduction state, (3) complex dissociation, and finally, (4) nucleotide replacement and the Fe protein reduction.

Thorneley and Lowe introduced the MoFe protein cycle to understand the kinetics of the nitrogenase action. In the MoFe protein cycle the electrons and protons are transferred to one-half of MoFe protein. Binding of substrate to the MoFe protein occurs after transfer of 3-4 electrons to the MoFe protein. The MoFe protein cycle presents the full mechanism of nitrogenase actions and the main conclusions of Thorneley and Lowe were that 1) the cycle is controlled by the electron donation of the Fe-protein cycle, 2) protons and electrons are added one-by-one to the MoFe-protein, 3) N_2 binds after the protein has been reduced by 3-4 electrons below the dithionite-reduced level [3, 56, 91, 57, 92].

1.3 THE NIFL-NIFA REGULATORY SYSTEM IN A.VINELANDII

Biological nitrogen fixation is regulated at the transcriptional level in response to multiple environmental signals. The availability of fixed nitrogen, free oxygen, and provision of sufficient energy are three important factors in the regulation of nitrogen fixation [31, 54, 8, 74]. The nitrogen, redox, and energy status regulate the activation of nitrogen fixation genes, *nif*. Nitrogen fixation is energetically very expensive for microorganisms and the source of energy has to be provided for this high-energy consuming process. In addition, the availability of fixed nitrogen controls the synthesis of nitrogenase at the transcriptional level [10]. The other regulatory component is oxygen. The presence of oxygen regulates nitrogen fixation at the transcriptional level. The regulation of nitrogen fixation varies between microorganisms. In the following section we review the regulation of nitrogen fixation in *A. vinelandii*.

1.3.1 *Physiology and genetic regulation of nitrogen fixation*

As mentioned before nitrogen fixation imposes a considerable energy requirement for bacteria. Two ATP molecules are hydrolyzed per electron transfer in each Fe protein cycle. Moreover, the turnover time of nitrogenase is relatively slow forcing the nitrogen-fixing bacteria to synthesize large amounts of nitrogenase, up to 20% of total cell proteins [10]. Therefore, concentration of nitrogenase is regulated at the transcriptional level in response to the availability of fixed nitrogen. The oxygen sensitivity of nitrogenase causes diazotrophs to use different physiological strategies for protection of nitrogenase from oxygen. The consumption of excess oxygen by respiration, compartmentation of nitrogenase, the presence of an oxygen diffusion barrier, conformational protection of the enzyme, and avoidance of oxygen through anaerobic growth are some of the physiological protection mechanisms in diazotrophs. From those strategies, *A. vinelandii* uses a high rate of respiration and production of the oxygen-stable enzyme by binding the enzyme to a protective protein [10, 15, 12, 60].

There are a number of proteins involved in nitrogen fixation which are encoded by *nif* genes. NifL and NifA are two regulatory proteins of the

nitrogen fixation system. NifA is a transcription activator protein, which activates nif genes and NifL is an anti-activator protein, which inactivates NifA by direct protein-protein interaction. In the following sections we discuss the regulation of nitrogen fixation by NifA and NifL. Figure 2 presents a very simplified cartoon of the NifL – NifA system. It shows the regulators and key components of the system. The elements which promote the formation of the NifL – NifA complex have a negative effect on nitrogen fixation, and the components, which enable NifA to escape from a complex, have a positive effect on nitrogen fixation. In this regard, oxygen, GlnK, and ADP act as negative regulators of nitrogen fixation and 2 – oxoglutarate acts as the positive regulator.

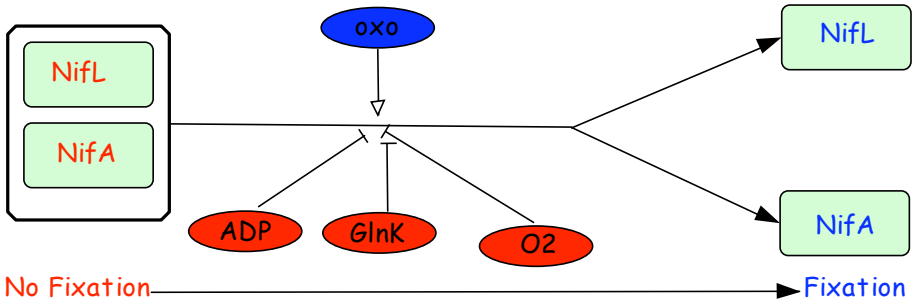


Figure 2: Basic cartoon of the NifL – NifA system. It shows how the system moves from "No Fixation" to "Fixation" in response to regulators. NifA and NifL are indicated in blue when active and in red when passive. The regulators, ADP, GlnK, and O_2 (in red) represent inhibition and 2 – oxoglutarate (in blue) represents activation.

1.3.2 *Regulation of nitrogen fixation by NifA*

NifA protein belongs to the family of proteins called enhancer binding proteins (EBP) which bind to upstream activator sequences and activate transcription of nif genes by interacting with the RNA polymerase sigma factor, σ^{54}. EBPs comprise three domains; a regulatory N-terminal domain, a catalytic domain which interacts with σ^{54} RNA polymerase, and a C-terminal DNA binding domain. In the NifA protein, these domains are called; the GAF domain (37-178), Sigma-54 factor interaction domain (211-439), and C-terminal DNA-binding domain (480-522). GAF domains are small-molecule binding domains and very widespread among regulatory proteins in all kingdoms of life. The N-terminal GAF domain of NifA is

required for interacting with another regulatory protein, NifL [60]. The central domain of NifA belongs to the AAA^+ super-family of ATPase (ATPase associated with different cellular activities). The forming of a ring-shape oligomer is an important characteristic of AAA^+ proteins for their ATPase activities [65]. The GAF domain of NifA interacts with the central domain of NifL in order to regulate its ATPase activity. The C-terminal DNA-binding domain encodes a helix-turn-helix motif structure that binds to the upstream activator sequences.

1.3.3 *The NifL-NifA complex*

NifL is an anti activator of nitrogen fixation which works by prohibiting the activation of transcription of nif genes through a direct protein-protein interaction with NifA. It is a regulatory flavoprotein, which has FAD as the prosthetic group, implying that NifL is redox-sensitive [31]. It inhibits NifA in response to environmental oxygen and fixed nitrogen. NifL in *A. vinelandii* has three main domains: the N-terminal sensory domain, glutamine rich sequence providing a Q linker to the C-terminal kinase like domain. The N-terminal of NifL has two PAS domains, PAS1 and PAS2. The PAS1 domain has FAD and is responsible for the redox-sensing function of NifL. The function of PAS2 is not completely known. The C-terminal region of NifL belongs to the GHKL super family, which contains a HATPase_c domain. ATP-binding proteins such as histidine kinases contain this type of domain. The C-terminal region of the NifL does not have ATP hydrolyse activity although it is needed for $NifL - NifA$ interaction [30]. The C-terminal domain of NifL interacts with the catalytic domain of NifA.

The requirement for stoichiometric amounts of NifA and NifL for inhibition of NifA activity, indicates that the interaction between NifA and NifL is not a catalytic interaction (Figure 3). The central domain of NifA is responsible for transcriptional activation and NifL contacts the central domain of NifA to suppress the transcription activation by inhibiting its catalytic function [63] .

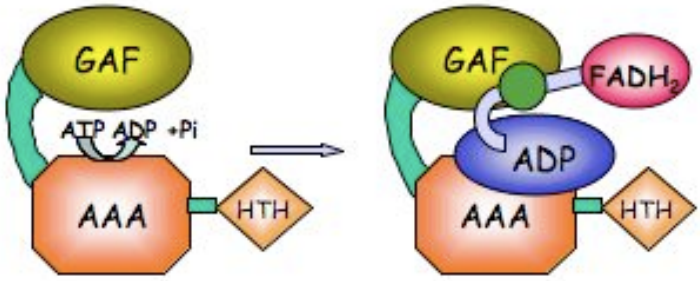


Figure 3: Cartoon of NifL – NifA complex and inhibition mechanism including free NifA and NifA in complex with NifL.

1.3.4 *ADP stimulation of NifL*

NifL in the reduced form does not show any inhibitory activity on NifA. But In the presence of ADP and ATP even the reduced form of NifL shows the inhibitory activity. The C-terminal domain of NifL protein can bind ADP and ATP. Its affinity for ADP is 10 fold higher than ATP [60, 63]. This binding creates a conformational change in the protein, which allows NifL to interact with the central domain of NifA inhibiting the activity of NifA.

1.3.5 *NifA binding to 2-oxoglutarate*

2 – oxoglutarate is a component of the citric acid cycle. It is a key metabolic signal of the carbon and nitrogen status. It is also part of the ammonia assimilation pathway that reflects the nitrogen status. The amino terminal GAF domain of NifA has the capacity to bind to 2 – oxoglutarate in order to regulate nitrogen fixation in response to carbon. 2 – oxoglutarate acts as an activator of NifA, and binds to NifA in the NifL – NifA complex to promote dissociation of the complex. Dissociation of the complex provides free NifA to activate transcription of nif genes.

1.3.6 *Oxygen sensing*

Apart from the physiological protection of nitrogenase under oxidizing conditions, A. vinelandii regulates the system at the transcriptional level

by using the strategy of not wasting energy to transcribe nitrogenase, since the nitrogenase enzyme is very sensitive to oxygen. NifL is a redox-sensitive protein, and its oxidised form, inhibits the activity of NifA. In the oxidized condition the prosthetic group of NifL, FAD, is oxidized and makes the conformational change in NifL that facilitates the interaction of this protein with NifA. This facilitation promotes the formation of the NifL – NifA complex, which inactivates the transcription of *nif* genes [9, 60].

1.3.7 *Nitrogen regulation*

The PII signal transduction proteins are important proteins found in bacteria, which communicate and integrate the status of intracellular fixed nitrogen and carbon to regulate nitrogen fixation. The signal for the status of fixed nitrogen is glutamine, which regulates the uridylylation state of PII protein. Uridylylation of the PII signal transduction protein alters its conformation, thereby changing the interaction with its target. GlnK is a PII signal transduction protein in *A. vinelandii* that regulates nitrogen fixation with respect to nitrogen status. The non-uridylylated form of GlnK interacts with NifL to inhibit the activity of NifA, while the uridylylated form of GlnK cannot interact with NifL. The formation of the GlnK- NifL- NifA complex inhibits the activity of NifA.

In addition to GlnK, *A. vinelandii* contains GlnD, which regulates nitrogen fixation through GlnK modification. The key signal of the nitrogen status is the intracellular concentration of glutamine [10]. Under conditions of fixed nitrogen sufficiency, the concentration of glutamine is high and under conditions of limited fixed nitrogen, the concentration of glutamine is relatively low. Sensing the amount of fixed nitrogen through the concentration of glutamine is integrated by GlnD followed by adjusting the state of GlnK, which can be in two states: the uridylylated form or the non-uridylylated form. The uridylylated form contains uridine monophosphate (UMP) covalently attached to GlnK. The uridylylated form of GlnK, is not able to stimulate the formation of the NifL – NifA complex. GlnD uridlylates GlnK at low concentrations of glutamine when the system is in limited fixed nitrogen conditions. When the glutamine concentration is high, GlnD is not able to uridylylate GlnK. Consequently, in conditions of

fixed nitrogen sufficiency, GlnK prohibits the transcription of the nif genes. In this chapter we analyze the system under conditions of fixed nitrogen sufficiency.

1.3.8 *The NifL-NifA system*

In summary, the three important factors for A. vinelandii which lead the organism to have a highly regulatory system are high sensitivity of nitrogenase to oxygen, high-energy demands of nitrogen fixation, and a low turnover rate of nitrogenase that leads A. vinelandii to produce considerable amounts of nitrogenase.

We divided the NifL – NifA system into four biological meaningful sub-systems to overcome the complexity of the complete NifL – NifA system (See Figure 4). The main sub-systems are NifA and NifL interaction, the 2 – oxoglutarate system, the NifL – NifA system in oxic conditions, and the GlnK system. Furthermore, we investigated the effect of these sub-systems on the dynamics of nif genes transcription. Figure 4 presents a simplified model of the NifL – NifA system. Spiral A in Figure 4 shows the NifA and NifL interaction and effect of 2 – oxoglutarate on the NifL – NifA complex. This part of the system has been named the 2 – oxoglutarate system. Spiral B presents the system in oxic conditions. The system is extended to include GlnK is in Spiral C. Finally Spiral D represents the dynamics of nif transcription in response to interaction of the modules.

Under the fixed nitrogen limitation conditions, A. vinelandii responds to the nitrogen status and stimulates the transcription of nif genes through activating NifA. In this condition the NifL – NifA system responds to 2 – oxoglutarate. Experimental evidence shows the relief of inhibition of NifA in the presence of 2 – oxoglutarate [53]. Interaction of 2 – oxoglutarate with NifA dissociates the complex between NifA and NifL. This relief of inhibition has also been observed when ADP is available in the system [53, 60, 10].

The physiological concentration of 2 – oxoglutarate is about 100 µM under the carbon limiting and nitrogen excess conditions. It increases to 1 mM under conditions of carbon availability and nitrogen limitation. This increase in 2 – oxoglutarate, increases the binding of 2 – oxoglutarate to

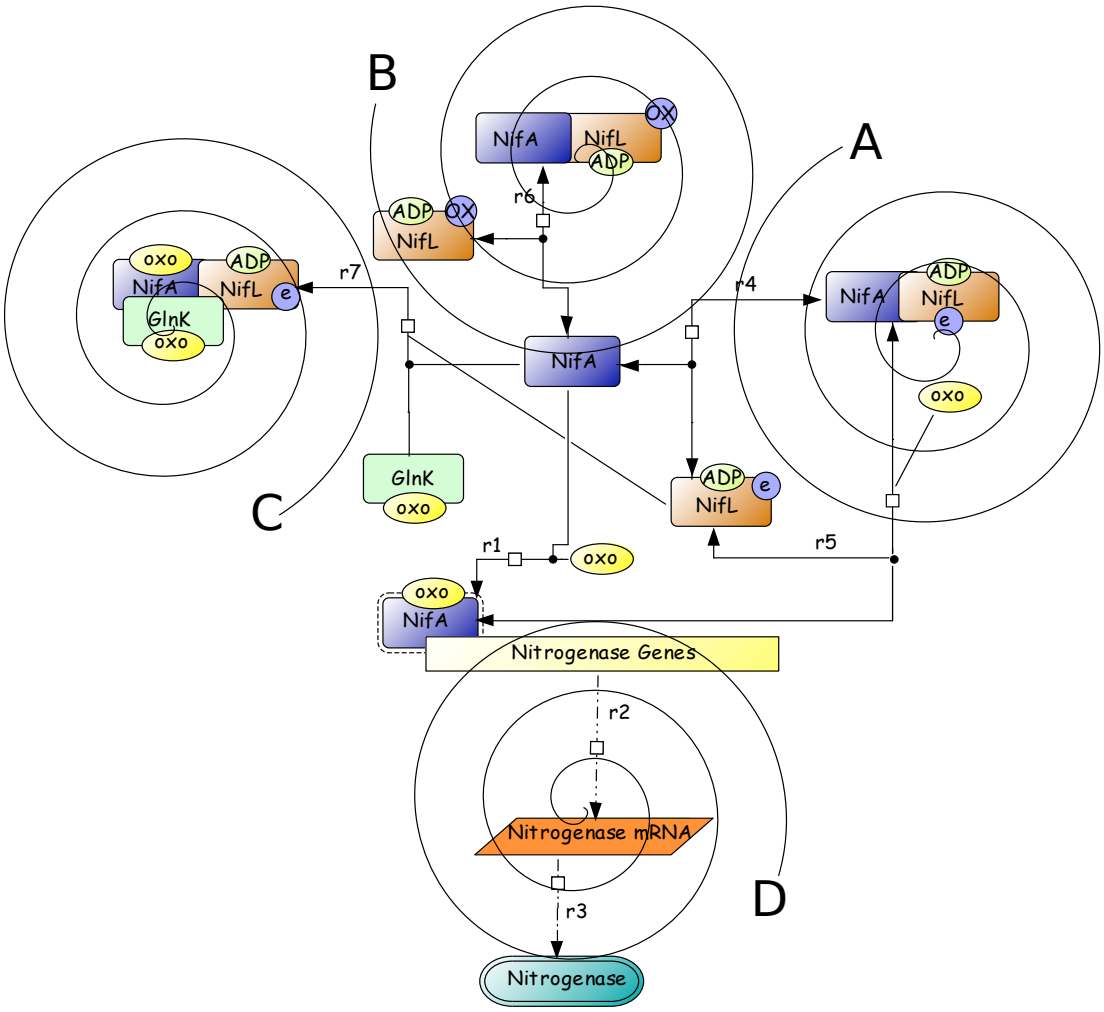


Figure 4: Abstract scheme of the $NifL - NifA$ system. The empty square represents the direction to the product of a reaction. 'OX' represents the oxidized form of the protein and 'e' represents the reduced form. This figure presents how the system reacts to various conditions. The scheme has been categorized into four spirals to modulate the system. $NifL - NifA$ interaction and the $2 - oxoglutarate$ system is presented in Spiral A. Spiral B is the system in oxidized condition and Spiral C summarizes the GlnK system. Spiral D indicates how all these interaction modulate transcription.

NifA to prevent inhibition by NifL and also disassociates the complex of NifL – NifA to provide the active form of NifA. The non-uridylylated form of GlnK can promote the ternary complex of GlnK- NifL- NifA, even in the presence of 2 – oxoglutarate, but under the fixed nitrogen limitation conditions, GlnK is mainly in its uridylylated form and cannot inactivate NifA [60].

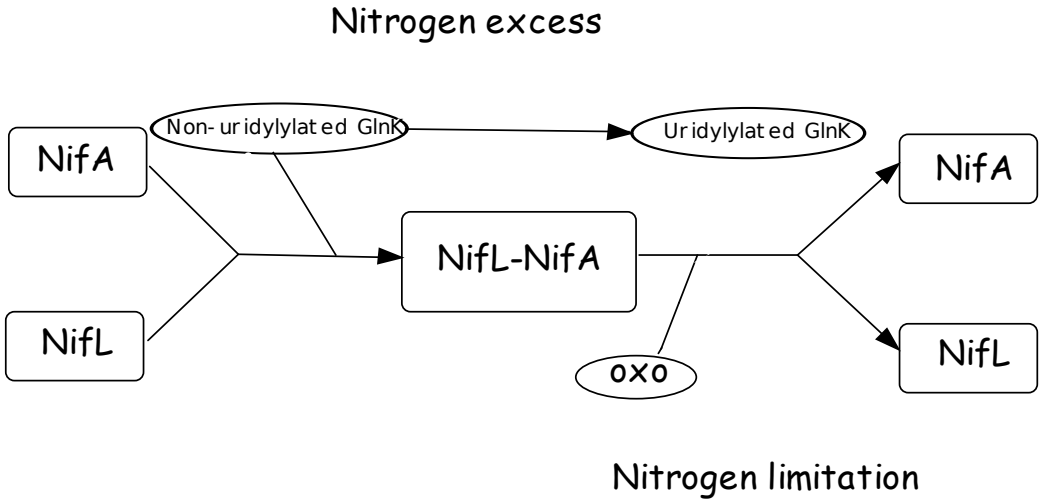


Figure 5: An overview of the NifL – NifA system under the nitrogen excess and nitrogen limitation conditions. Under nitrogen excess conditions the non-uridylylated form of GlnK promotes the formation of NifL – NifA complex. Under nitrogen limited conditions GlnK is converted to uridylylated form and cannot promote the formation of the NifL – NifA complex, while 2 – oxoglutarate (2OG) promotes the dissociation of the complex.

Under nitrogen excess conditions, *A. vinelandii* does not fix nitrogen. Under these conditions GlnK stimulates inhibitory activity of NifL. Under conditions of fixed-nitrogen limitation the GlnK is mainly in the uridylylated form and it cannot interact with NifL while in nitrogen excess conditions it is in the non-uridylylated form and it can interact with NifL [9, 60, 53]. The complex of NifA-NifL-GlnK prevents the activation of transcription by inhibiting NifA, even in the presence of 2 – oxoglutarate and ATP.

NifL is in the reduced form provided that bacterium respires oxygen. Under oxic conditions, NifL is oxidized rapidly to $NifL_{ox}$. The oxidized form of NifL readily interacts with NifA to prevent the activation of *nif* genes. The inhibitory activity of NifL is high in oxic conditions, even in presence of 2 – oxoglutarate, and it has a high affinity to interact with NifA to form the NifL – NifA complex. Unlike reducing conditions, NifL

does not need to interact with ADP in oxic conditions to interact with NifA. This is the strategy that *A. vinelandii* produces to regulate nitrogen fixation at the transcriptional level in oxic conditions.

1.4 THE NIFL-NIFA REGULATORY SYSTEM IN K.PNEUMONIAE

The nitrogen fixation regulatory system in K. *pneumoniae* has a similar set of proteins to *A. vinelandii*. In this section the system in both species will be compared. However, full details of the NifL – NifA system in K. *pneumoniae* are not considered. NifA and NifL are the two main proteins regulating of nitrogen fixation in K. *pneumoniae*. NifA and NifL act in K. *pneumoniae* in a similar manner to *A. vinelandii*.

There are two major differences in the regulatory mechanism of the NifL – NifA system related to nitrogen status between K. *pneumoniae* and *A. vinelandii*. Firstly, transcription of the *glnk* gene is not subject to nitrogen regulation in *A. vinelandii* while it is subject to nitrogen regulation in K. *pneumoniae*. Furthermore, the transcription of *nifa* and *nifl* is also subject to nitrogen regulation. Therefore, the transcription of *glnk*, *nifa*, and *nifl* genes decreases under the conditions of excess fixed nitrogen in K. *pneumoniae*. Secondly, K. *pneumoniae* GlnK acts in an opposite way to *A. vinelandii* GlnK. The non-uridylylated form of GlnK promotes dissociation of the NifL – NifA complex in K. *pneumoniae*.

In summary, in K. *pneumoniae* under excess fixed nitrogen conditions, the transcription of *glnk*, *nifa*, and *nifl* genes are decreased and NifA is incorporated within the NifL – NifA complex when the concentration of GlnK is low and there is not enough GlnK to promote the dissociation of this complex. Under the limited fixed nitrogen conditions, the transcription of *glnk* increases as the transcription of *nifa* and *nifl* genes increase; therefore, GlnK activates nitrogen fixation by promotion of the dissociation of the NifL – NifA complex.

2

COMPUTATIONAL TOOLS AND TECHNIQUES

Systems biology is a new and rapidly expanding field that aims at understanding biological systems at the systems level. Modelling can be used to describe and help to understand a biological system of interest (such as a metabolic pathway, gene regulation, protein-protein interactions, cell signalling, or cell cycle systems) in a precise way. The aim of modelling varies based on the requirement of each system, the availability of data, and the question one wishes to answer. The basic aim of modelling is to describe a specific system in a comprehensive manner, translate a hypothesis into a mathematical model, then make it clearer, confirm or reject it, and extend it to explain new data. Testing and analyzing the system to check for consistent behavior with observed experimental results is crucial. Following the determination of a consistent model, it is desirable to be able to do virtual experiments, which would be time-consuming, very difficult, very expensive or even impossible to do in the lab. Finally, pooling a number of well-defined and related systems together to describe a larger model and investigating the interaction between the models components should be considered as a goal of systems biology. In this chapter, we summarize some of the key modelling techniques that we used in this thesis and also give an overview of some computational techniques that one can use to modelling the biological systems.

2.1 REACTION KINETICS

Kinetics is the study of reaction rates of chemical reactions. The generation of mathematical equations to describe the rates of reactions is the first step, followed by their solution leading to a kinetic model for a set of reactions. The aim of kinetic modelling is to analyze and predict system dynamics based on the kinetic parameters. A kinetic model can be used for hypothesis testing and also compared to experimental observations. Knowledge can

be extracted leading to a better understanding of the system and it can guide the direction of future experiments. Setting the rate equations of the reactions in complex systems is an important task in this type of modelling. In the following sections some of these rate equations are reviewed [82, 19].

2.1.1 *The law of mass action*

Mass action rate laws describe the rate of reactions in elementary steps where one or more reactants react to form products. This law describes the rate of change of any component in the system that is proportional to the number of molecules of those components. The mass action equations can be used for reactions where no intermediates are formed in a single reaction step and with a single transition state[18]. Consider the following reactions for chemical entities x_1 to x_7:

$$x_1 + x_2 \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} x_3, \tag{2.1}$$

$$x_3 + x_4 \underset{k_{-2}}{\overset{k_{+2}}{\rightleftharpoons}} x_5, \tag{2.2}$$

$$x_4 + x_6 \underset{k_{-3}}{\overset{k_{+3}}{\rightleftharpoons}} x_7. \tag{2.3}$$

In these equations, k_i and k_{-i} (for i=1, 2, 3) denote the forward and backward rates of reaction. The rate of each reaction, v, is equal to the forward rate constant multiplied by concentration of the reactants minus the backward rate constant multiplied by the concentration of the products. The following equations give the rates of reactions 2.1, 2.2, and 2.3:

$$\begin{aligned} v_1 &= k_1 x_1 x_2 - k_{-1} x_3, \\ v_2 &= k_2 x_3 x_4 - k_{-2} x_5, \\ v_3 &= k_3 x_4 x_6 + k_{-3} x_7. \end{aligned}$$

From the reaction rates the rate of change of each component of the system can be determined. The reactants have a negative rate on the corresponding reaction as long as they are being used and the products of reactions have a

positive rate on the corresponding reaction since they are being produced. The reactants and products can be involved in more than one reaction and the rates of all corresponding reactions have to be added. The following equations are a system of ordinary differential equations, which model the rate of change of all components of 2.1,2.2, and 2.3 for all components in the system:

$$\begin{aligned}
\dot{x}_1 &= -k_1 x_1 x_2 + k_{-1} x_3, \\
\dot{x}_2 &= -k_1 x_1 x_2 + k_{-1} x_3, \\
\dot{x}_3 &= k_1 x_1 x_2 - k_{-1} x_3 - k_2 x_3 x_4 + k_{-2} x_5, \\
\dot{x}_4 &= -k_2 x_3 x_4 + k_{-2} x_5 - k_3 x_4 x_6 + k_{-3} x_7, \\
\dot{x}_5 &= k_2 x_3 x_4 - k_{-2} x_5, \\
\dot{x}_6 &= -k_3 x_4 x_6 + k_{-3} x_7, \\
\dot{x}_7 &= -k_3 x_4 x_6 + k_{-3} x_7.
\end{aligned}$$

The law of mass action can be used to translate a set of chemical or biochemical reactions in this way into a set of mathematical equations for further analysis. The above methodology was used to form the mathematical equations in most cases in this thesis.

2.1.2 *Michaelis-Menten kinetics*

Enzymes are proteins which catalyze chemical reactions in biological systems. The Michaelis-Menten equation describes the rate of the enzymatic reactions, and it relates the rate of enzymatic reaction to substrate concentration [19]. We now recall how this equation is derived. The process starts from the binding of a substrate (S) onto an active site of an enzyme (E). This binding produces the enzyme-substrate (E-S) complex. The E-S complex turns into the product (P) and enzyme (E) in an irreversible reaction. The following reactions represent the whole process of S to P catalysis [82]:

$$E + S \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} ES \overset{k_{+2}}{\rightharpoonup} E + P.$$

Under the steady state assumption that the rate of the formation of the ES complex equals the rate of its dissociation we have

$$\begin{aligned} \nu_{formation} &= k_1 \cdot E \cdot S, \\ \nu_{dissociation} &= k_{-1} \cdot ES + k_{+2} \cdot ES. \end{aligned}$$

Assuming $\nu_{formation} = \nu_{dissociation}$, we obtain

$$k_1 \cdot E \cdot S = k_{-1} \cdot ES + k_{+2} \cdot ES.$$

The total concentration of the enzyme is assumed to be constant in the process. E_0 is the total concentration of the enzyme, which is the total available enzyme in the system and so:

$$E_0 = E_{free} + ES.$$

Forming the equation using E_0, we obtain

$$k_1 \cdot (E_0 - ES) \cdot S = k_{-1} \cdot ES + k_{+2} \cdot ES,$$

and, solving for ES, gives

$$ES = \frac{E_0 \cdot S}{S + \frac{k_{+2}+k_{-1}}{k_{+1}}}.$$

Assuming the rate of the whole process is the rate of productivity of P we then obtain:

$$V = k_{+2} \cdot ES = k_{+2} \cdot \frac{E_0 \cdot S}{S + \frac{k_{+2}+k_{-1}}{k_{+1}}}. \tag{1}$$

Here V is the rate of the enzymatic reaction and V_{max} is the maximum speed of the enzymatic reaction in converting the substrate into product. Increasing the concentration of the substrate increases the rate of the enzymatic reaction until it reaches V_{max}, which is k_{+2} times E_0. K_m is the Michaelis-Menten constant, corresponding to the concentration of substrate when the rate reaches the half maximum [19] and it is given by $\frac{k_{+2}+k_{-1}}{k_{+1}}$. Substituting K_m and V_{max} into (1) finally yields the Michaelis-Menten equation:

$$V = \frac{V_{max} \cdot S}{K_m + S}.$$

2.1.3 *Binding*

Binding of an enzyme to a substrate is similar to binding of a protein to a ligand, hormone, antibody, or any other type of protein. The following equations can be applied for simulating a protein-protein or protein-ligand interaction [19]. If P is a protein with only one binding site and L is a ligand, the binding equilibrium equation is:

$$P + L \rightleftharpoons PL,$$

where PL is the protein-ligand complex. Assuming that the concentration of the complex is equal to L_b (ligand in complex), the dissociation constant of this reaction is given by:

$$K_d = \frac{P \cdot L}{L_b}.$$

Assuming that the total concentration of protein (P^t) in the system is equal to the sum of free protein and ligand bounded protein, we obtain:

$$\begin{aligned} P^t &= P + PL, \\ PL &= L_b, \\ P &= P^t - L_b. \end{aligned}$$

Finally, substitutions, the L_b function is given by:

$$L_b = \frac{P^t \cdot L}{K_d + L}.$$

This equation can be applied to model a protein-protein interaction or a protein-ligand interaction.

2.2 DYNAMICS OF TRANSCRIPTION

Cells in different environmental conditions sense signals and respond to them by producing proteins. Interactions between transcription factors, activators, repressors, RNA polymerases, and promoters of genes regulates transcription, leading to the proteins that act in different environmental situations. Regulatory proteins and transcription factors can control the transcription rate of a gene. These controls can be either positive (activation) or negative (repression). Increasing the rate of transcription is positive control and decreasing the rate of transcription is negative control [1].

Assuming Y is a protein produced by regulation of X as a transcription factor or regulator on genes, the productivity of Y is given as a function of active form of concentration of X:

$$\begin{aligned} X &\rightharpoonup Y, \\ \dot{Y} &= f(X). \end{aligned}$$

Note that Y increases with increasing X when it is an activator and decreases with increasing X when it is a repressor. The Hill function describes the rate of production of Y based on the concentration of the regulators. The following equation is the Hill function for an activator [1]:

$$f(X) = \frac{\beta \cdot X^n}{K^n + X^n}.$$

The Hill function for repression is:

$$f(X) = \frac{\beta}{1 + (\frac{X}{K})^n}.$$

The Hill function is an increasing S-shaped (sigmoid) function for an activator starting from zero expression level and reaching the maximal expression level, and for a repressor it is a decreasing S-shaped function starting from the maximal expression level and reaching zero activity. The parameters of the Hill function are K, β, and n. K called the activation coefficient in the activator Hill function and repressor coefficient in the repressor

Hill function. The value of K is half of the concentration X when it reaches the half of its expression level. Therefore, the unit of K is concentration. β is the maximal expression level reached by high concentration of X in an activator system or reached by as low concentration as zero of X in a repressor system. The Hill coefficient n defines the steepness of the function. The Hill function describes the transcription rate where there is only one activator or repressor. It can be used to form a set of ordinary differential equations in combination with other rate formulas [1].

In studying of the dynamics of gene regulation, the rate of change of concentration of Y is not only related to production as described but also related to the rate of protein degradation and dilution. The degradation/dilution of the protein needs to be considered in forming the system. Assuming α is degradation/dilution rate in a unit of the time, the rate of production can be arranged using different rate formulas. The rate of change of concentration of Y is given by:

$$\dot{Y} = \mathrm{production} - \alpha Y.$$

The basic regulation of transcription was presented in the last paragraphs, however, in many situations the transcription is regulated by a network of interactions among the regulators, transcription factors, and other reactants in the cell. The networks of transcription are connected through some components and complex networks are generated to govern the behavior in the cell. Understanding the dynamics of such big networks is very challenging.

One way to understand networks dynamics is to look for common building-block patterns known as network motifs which occur in complex networks higher than randomized networks. There are variety of statistical methods for ranking network motifs such as Bayesian and maximum likelihood-based methods. Out of many possible patterns in networks only few of them tend to be considered as network motifs since these appear in many networks. These network motifs have been detected in diverse systems and have specific functions [1]. Negative auto-regulation, positive auto-regulation, and feed-forward loop are examples of well-studied network motifs (Figure 6).

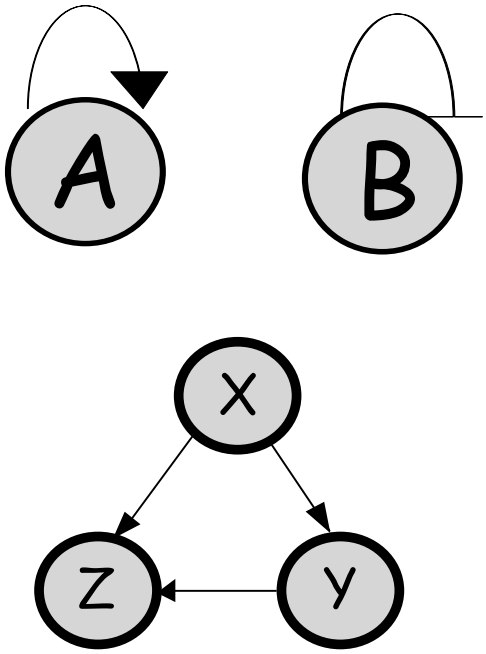


Figure 6: Positive auto-regulation (A), negative auto-regulation (B), and a feed-forward loop (X, Y, and Z).

Negative auto-regulation is a simple example of the network motifs in transcription. We briefly present the benefits of negative auto-regulation in biological systems. Auto-regulation is the regulation of a gene by its own product. The two types of auto-regulation are positive auto-regulation and negative auto-regulation. In this type of motif, the product of a gene represses its own transcription. Negative auto-regulation speeds the response time to the steady state conditions in comparison with simple regulation and positive auto regulation. Fluctuations in production rate of a gene can change the steady state level leading to a change of the steady state concentration of the gene product. Negative auto-regulation motifs promote robustness to fluctuations in production rate by buffering fluctuations through repression [1]. Each biological network motif has its own function. Detecting network motifs can help us to predict the function of biological systems by comparison with known motifs.

2.3 SIMULATION

In the previous sections some reaction rate formulas, which can be used to obtain a set of ordinary differential equations from a set of reactions, were presented. A system of ODEs is a set of differential equations each of which contains an ordinary derivative. A system of ODEs can be solved using numerical integration algorithms. The Euler method and various forms of Runge-Kutta are numerical techniques, which are implemented in many modelling and simulation tools for biological systems.

The Runge-Kutta 4^{th} order method calculates the value of a variable at the next step based on its value at present. In each step of the Runge-Kutta algorithm, the derivatives of four points are evaluated; The initial point, two trial midpoints, and a trial endpoint [68]. A step size of h is the difference between each initial point and the endpoint. Midpoints are the points between initial point and endpoint. The algorithm continues by taking a weighted average of the derivatives. Note that the derivatives at the midpoints receive the greater weight for calculation of the average. The average of the derivatives is added to the variable to calculate the next variable.

For example, consider the following equation

$$\dot{x} = f(t, x).$$

Given step size h and initial conditions, four intermediate values are calculated:

$$\begin{aligned} k_1 &= hf(t_n, x_n), \\ k_2 &= hf(t_n + 1/2h, x_n + 1/2k_1), \\ k_3 &= hf(t_n + 1/2h, x_n + 1/2k_2), \\ k_4 &= hf(t_n + h, x_n + k_3). \end{aligned}$$

In the next step, the next value of the variable x is calculated using the present value of x using the weighted mean of the derivatives:

$$x_{n+1} = x_n + 1/6(k_1 + 2k_2 + 2k_3 + k_4).$$

This process is repeated a fixed number of times. Runge-Kutta is a popular method among the numerical methods for solving differential equations for biological systems. In this thesis, the author used the Runge-Kutta 4^{th} order method for solving ODEs in the computer program that he developed.

2.4 STRUCTURAL ANALYSIS

Kinetic modelling demands a considerable amount of kinetic information, such as rate constants and the total concentrations of reactants. This information is rarely available for complex biological systems. There are some computational techniques which can be used to investigate the biological system with minimal dependency on biochemical data. Structural analysis of a set of ODEs gives properties of the system, which only depend on the structure of the underlying network of interactions and not on the reaction kinetics.

The structure of the model is defined in terms of the stoichiometric reaction scheme, which describes how the metabolites of the system combine together. The structural properties of the system, in contrast to kinetics, do not depend on the environment of the system and on its internal state [70]. Flux analysis, conservation relationships, minimal cut sets, elementary modes, and extreme pathways are all approaches that are based on the stoichiometric matrix [45, 39, 44, 24, 41, 42]. The following section describes briefly an example of these approaches.

The first step in structural analysis is to compute the stoichiometric matrix (N). N has 'm' rows representing the metabolites in the system and 'n' columns representing the corresponding reactions. The following example illustrates how to derive N [70].

$$
\begin{aligned}
* &\rightarrow x_1 \\
x_1 + x_6 &\rightarrow x_2, \\
x_2 + 2x_5 &\rightarrow x_3, \\
x_3 &\rightarrow x_4 + x_5, \\
x_4 &\rightarrow x_5 + x_6 + *.
\end{aligned}
$$

the matrix N is given by:

$$N = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 1 & 1 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix}.$$

The values in N indicate the number of the molecules of the metabolites that are assigned positive for products and negative for reactants. The corresponding system of ODEs is given by

$$\dot{x} = Nv,$$

where x is the of the metabolite concentrations, and v is the vector of rate of reactions, i.e.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 1 & 1 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix}.$$

The aim of many modelling approaches is to describe and understand the behavior of a set of ODEs within a period of time. In contrast, steady state analysis investigates the system in the steady state condition, i.e.

$$\dot{x} = 0.$$

Consequently the steady state equation is given by

$$Nv = 0. \tag{2.4}$$

This system can be solved and the rate of the reactions (fluxes) calculated. The Gauss elimination can be used to solve the system. These type

of systems can be overdetermined, the number equations is more than unknowns, determined, the same number of equations and unknown, and underdetermined, the number equations is less than unknowns. In general, there is no solution for a overdetermined system, there is a unique solution for a determined system, and underdetermined systems can have infinite solutions.

2.5 MODELLING TOOLS AND DATABASES

In this section we shall briefly review some of the tools that may be used to simulate the behavior of biological systems. Many of these tools use SBML, Systems Biology Markup Language, a machine-readable language utilizing XML primary encoding [50, 80, 21, 38]. This was developed about 5 years ago and since then has been extended to represent cell-signalling pathways, metabolic networks, regulatory networks, and other types of biochemical reaction networks. The aim of SBML is to enable users to exchange models between simulation/analysis tools and to publish in electronic format models for the scientific community [38, 21]. SBML is under constant development for both fixing of problems and adding new features. This development has resulted in a series of levels and the latest version, which is not yet supported by all packages, is level 3.

There are many software packages which have been developed with the aim of modelling biochemical systems. We now present an overview of some of the most well known systems biology software packages; Copasi, SBW, CADLive, Celldesigner, CellNetAnalyzer, and SBML-SAT. Supporting SBML, platform availability, free availability, installation, usability and manual, parameter estimation methods, and applicability for different kinds of modelling are all important factors for making a good systems biology program. A more detailed overview of SBML tools is presented in [2, 26, 88, 87].

2.5.1 *Copasi and Gepasi*

Gepasi was one of the first programs developed for systems biology. Pedro Mendes developed it for modelling the dynamics, steady states and control of biochemical systems. It is a Windows based program written in C [61].

Copasi (Complex Pathway Simulator) was developed based on Gepasi by the Mendes and Kummer group. It can be used to simulate and analyse biochemical networks. It can import all levels of SBML and export some versions. Copasi is available for MS Windows, Linux, OSX, and Solaris [33]. It is free for non-commercial use, and its installation is straightforward.

A good manual is available for Copasi. The latest version of Copasi has most modelling methods available for simulating and analysing biological systems. Furthermore, it has a flexible facility for plotting the results. The latest version has the ability to perform steady-state analysis, structural analysis, time-course simulation, metabolic control analysis, parameter scans, optimization, parameter estimation, and sensitivity analysis. For stoichiometric analysis, elementary modes and mass conservation are also available. Genetics algorithms, Hooke & Jeeves, Evolutionary Program, Random search, and Steepest Descent are methods which can be used for parameter estimation and optimization tasks in Copasi. It does not have a graphical interface for representing the biochemical network. Nevertheless, in the author's opinion Copasi is one of the best programs currently available for modelling biochemical systems.

2.5.2 *System Biology Workbench*

System Biology Workbench (SBW) is a package of programs developed by the System Biology Workbench Development Group for simulating and analysing biochemical networks [37, 79]. It is an open source SBML type package, originally Windows based, but now it is available for Linux and OS X operating systems. The Windows version has more tools available from the OS X version. The installation of SBW is easy, and it is well documented. It performs time course simulation, steady state analysis, and stoichiometric analysis. SBW consists of the following programs; JDesigner, Jarnac, SBW Inspector, Structural Analysis tools, and WinSCAMP. JDesigner is the main program, which has a parameter scan facility, however, parameter estimation methods are not included. The author found that SBW was not as capable as Copasi for modelling, but it had better documentation and is probably simpler for novice modellers. In addition, it has a graphical interface for representing the biochemical networks.

2.5.3 *CellDesigner*

CellDesigner is a modelling tool for biochemical networks with a high quality graphical user interface for drawing the biochemical network or gene regulatory system [23]. It is a freely available SBML type program and has been implemented in JAVA so is operating system independent. It has the advantage of SBW simulation modules, and also has time course simulation and parameter scan facilities. The installation of CellDesigner is straight forward, and it is extensively documented. The author found this program simple to use. In particular, the graphical representation of the biochemical network was found to be one of its best aspects. In terms of simulation and modelling methods, it was found to have less capability compared to other software programs.

2.5.4 *CADLIVE*

CADLIVE is a freely available software package for construction and analysis of biochemical systems [49, 48]. This is not SBML compatible. CADLIVE uses genetic algorithms for parameter estimation. The installation of CADLIVE was found to be difficult and time consuming. The authors claim that it is possible to use K_d values in simulations using CADLIVE. A test of the online version showed that it is not possible to use K_d values in simulations as expected from the associated publication.

2.5.5 *CellNetAnalyzer*

CellNetAnalyzer is a further development of FluxAnalyzer [46, 43]. This program was originally developed for the structural analysis of biological networks. It is freely available to academic users but it requires MATLAB version 6.1 or higher. It also needs the optimization toolbox of MATLAB. This program is capable of metabolic flux analysis with flux optimization in flux balance analysis task for an arbitrary linear objective function, analysis of structural properties of biological networks, elementary modes, extreme pathways analysis, minimal cut sets, and analysis of Boolean networks [46, 43].

2.5.6 *SBML-SAT*

SBML-SAT (system biology markup language sensitivity analysis tool) is designed mainly for sensitivity analysis [97]. SBML-SAT is a freely available SBML program, which has been implemented in MATLAB. Although it is free software, it requires MATLAB. It performs time course simulation, local sensitivity analysis, global sensitivity analysis, steady state analysis, and robustness analysis. SBML-SAT was used for many tasks described in this thesis and it was found to be reliable and useful. It runs in Windows, Linux, and OS X platforms, requiring libSBML, SBMLToolbox, and sundialsTB. The installation of SBML-SAT was not as straight forward as the installation of SBMLToolbox and the corresponding libSBML was difficult in some platforms such as Windows Vista and Mac Os X.

2.6 CONCLUSION

In this chapter, we have reviewed some of the standard techniques used to simulate the behavior of biological systems. As we have also seen, there are now many programs available that implement these techniques. We found that most of these programs are very similar. In fact many of them are simply ODE solvers with a graphical user interface. One problem with this type of software is that the programs are not flexible for using with different types of biochemical systems and different types of data. Also they are perfect for a system with large amounts of available data, but such data is seldom available for most biochemical systems. Most of the tools were developed for different purposes and they have their own advantages and disadvantages. Although the programs are generally quite similar, some have more extensive facilities for modelling such as Copasi and SBML-SAT. Based on different modelling tasks and systems, occasionally it was necessary for the author to develop a model specific tool. Copasi and SBML-SAT were the main systems biology programs that were used and CellDesigner was used for drawing biological networks. As we shall see, it was necessary to develop additional code in JAVA, MAPLE, and MATLAB in the modelling of nitrogen fixation.

Part II

NITROGENASE

3

THE THORNELEY AND LOWE MODEL IN SBML

3.1 A REVIEW OF THE THORNELEY AND LOWE MODEL

As we seen in Chapter 1, nitrogenase plays an important role in nitrogen fixation. Diazotrophs regulate nitrogen fixation to respond to the characteristics of nitrogenase such as its oxygen sensitivity and low turnover rate. Thorneley and Lowe revealed the mechanism of nitrogenase action. They introduced two schemes for the mechanism of nitrogenase action, the Fe protein cycle and the MoFe protein cycle. The schemes in Figure 7 & 8 have both been published together with the associated kinetic rate constants, total concentrations, and activity corrections [90, 56, 91, 57, 92, 58, 22, 59, 3, 7].

We first describe the Fe protein cycle which was introduced in Section 1.2.3. *In vitro*, the Fe protein is reduced by dithionite, the usual reducing agent used in such experiments. *In vivo*, the reduction is most probably performed by electron carriers such as ferredoxin or flavodoxin with electrons ultimately coming via NADH, NADPH, or $FADH_2$, from the citric acid cycle or from glycolysis. Dithionite is used as the electron donor in this process. This reduction is accompanied by a replacement of the two MgADP molecules by new MgATP molecules, leading to $Fe_{red}(MgATP)_2$ + MoFe. The next step is the protein complex formation, which requires the Fe protein to be in its reduced state and bound to MgATP. In this complex, MgATP is cleaved to MgADP and phosphate $(P_i)_3$ as shown in Figure 7, where one electron is transferred to the MoFe protein. Next, the complex dissociates, which is believed to be the rate-limiting step not only for the Fe protein cycle but also for the whole substrate conversion when all components are at saturating concentrations. The rate of complex dissociation is approximately 5 s^{-1} under saturating conditions leading to $Fe_{ox}(MgADP)_2$ + $MoFe_{red}$. The electron at the MoFe protein is used for substrate conversion and the Fe protein returns to its starting configuration in the Fe protein cycle.

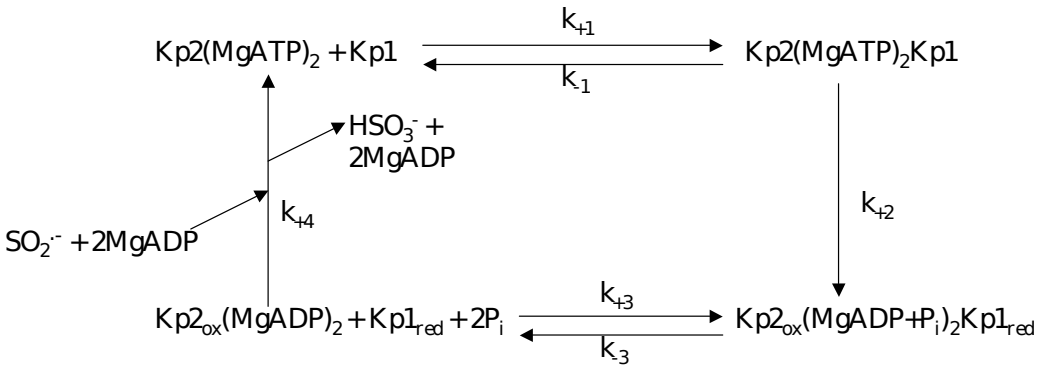


Figure 7: The Fe cycle adapted from Scheme 1 of [57]. Kp1 represents the functional half of the MoFe protein and Kp2 the Fe protein. One electron is transferred from Kp2 to Kp1 with the concomitant hydrolysis of 2MgATP to 2MgADP+2P_i. 'red' means reduced and 'ox' used for oxidized conditions.

The MoFe protein cycle is presented in Figure 8. Starting from the resting state E_0, the MoFe protein is reduced by one electron and protonated once, leading to state E_1H_1. The notation E_iH_j represents the protein with i electrons and j protons added to the initial state of the enzyme. After the next electron-proton transfer step, at E_2H_2, there is a certain probability of hydrogen production taking place which would lead the system back to its resting state. This probability has been shown to depend on the electron flux. Under conditions of low flux a low concentration of the Fe protein compared to the MoFe protein, the E_1H_1 accumulates whereas the E_2H_2 is removed by evolving H_2. At high flux, E_2H_2 will then be reduced and protonated, leading to E_3H_3, which again can produce H_2. It is in this state that dinitrogen can bind. This produces one H_2 and leads to E_3HN_2. H_2 can displace bound N_2 and thereby inhibit its reduction. The state E_3H_3 may also be further reduced to E_4H_4 which then can bind to N_2 with resulting H_2 production. When no substrate is available, H_2 is released reverting back to more oxidised states. The MoFe protein cycle continues to release 8 electrons.

The Fe protein cycle and the MoFe protein cycle are scattered throughout a body of publications spanning over decades [90, 56, 91, 92]. The rapid-quench flow system was used for pre steady state studies of mechanism of nitrogenase action. However, none of the cycles presents the whole reactions of the Thorneley and Lowe model. In addition, the routines to model this system written by Lowe and Thorneley were developed in Fortran and

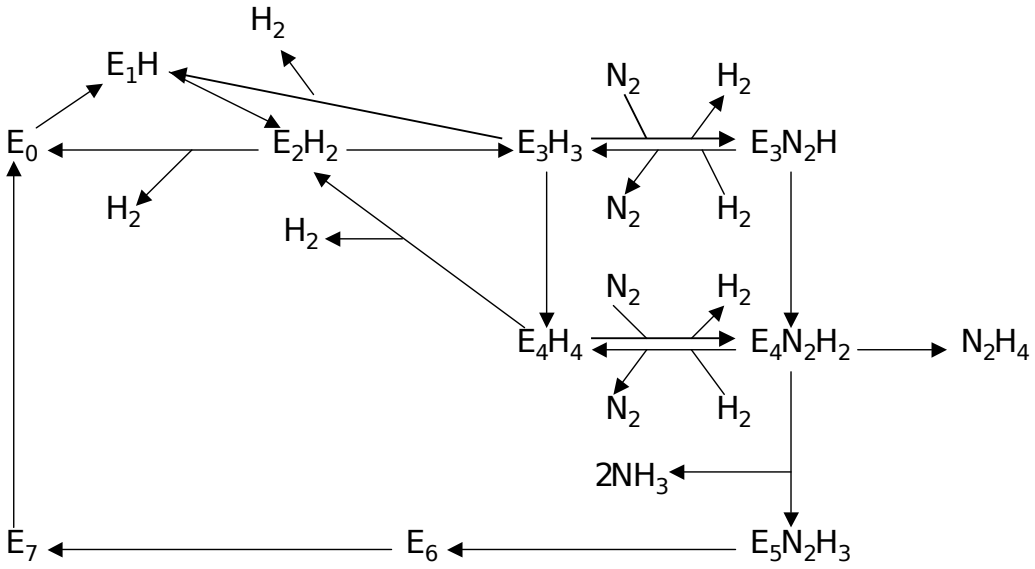


Figure 8: The MoFe cycle adapted from Scheme 2 of [57]. The species E_i represents the electron state of one of two independently functioning halves of Kp1. The process from $E_5N_2H_3$ to E_6, E_7, and E_0 is the same as E_2 to E_3.

require the use of the NAG libraries. For this reason, the exchange and further development of this important model was limited. To remedy this situation, we established a complete SBML model of the Thorneley and Lowe system.

3.2 THE FULL THORNELEY AND LOWE MODEL IN SMBL

In Figure 9 we present all the reactions in Thorneley and Lowe model, which were communicated to the author of this thesis by David Lowe. Thorneley and Lowe published the MoFe cycle but they never published the scheme in Figure 9 that has all the reactions involved in the process of nitrogen fixation. The MoFe protein cycle is a summarized version of this. The scheme in Figure 9 contains very valuable information about the Thorneley and Lowe model and represents distillation of years of research by Thorneley, Lowe, and their colleagues.

To build an SBML model, we extracted data from the original publications and used the scheme in Figure 9. All assumptions made by the original Thorneley and Lowe model have been considered and incorporated into this model. The inactive MoFe protein does not interact with the Fe protein,

although the inactive Fe protein competes with the active one to interact to MoFe protein. The initial reductant source is dithionite. The kinetic parameters for K. *pneumoniae* nitrogenase at 23°C in pH 7.4, 25 mM HEPES buffer with 9 mM ATP, 10 mM $MgCl_2$ and zero ADP, are listed in Table 43 and the total concentrations can be found in Table 2.

We designed the SMBL model, given in Figures 10 to 17, using the Systems Biology Workbench [37, 79] (See Section 2.5.2). O, I, and R represent the oxidized Fe protein, inactivated Fe protein, and reduced Fe protein, respectively. Reduction, oxidization, and inactivation of E_iH_j occurs by R, I, and O in each step to produce E_iH_jr, E_iH_jo, and E_iH_ji. These processes were not presented in the original MoFe protein scheme. There were some species in the scheme that were involved in different parts of the scheme. To monitor these species , x and p were added to E_iH_j making them distinguishable from the same species in other parts of the scheme such as $E_4N_2H_2$ and $E_4N_2H_2x$ which are the same species and were involved in different part of the scheme. We tried a number of simulation tools to explore this scheme, most of which handled the resulting system of ordinary differential equations remarkably well giving consistently close results despite using different integration schemes. An analysis of these differences will not be presented here. For production work and our analysis, Copasi was employed [33] (See Section 2.5.1).

We computed time course data for all the species involved in the Thorneley and Lowe pre-steady state kinetic model. The distance between experimental data and original model together with SBML model was computed for H_2 and NH_3 production. The distance between experimental data and original simulation for NH_3 is 57 and between experimental data and SBML model is 61. These two simulations are in agreement with experimental data. The distance for H_2 for original model is 39 and for SBML model is 58. These simulations are also in agreement with the experimental data. We show simulations for the production of the two most important products, namely H_2 and NH_3 in Figure 18. Thorneley and Lowe published simulations of H_2 and NH_3 in [91, 56]. The simulations of H_2 and NH_3 in the original Thorneley and Lowe model and SBML model of Thorneley and Lowe are in good agreement, although there is further room for fitting computational

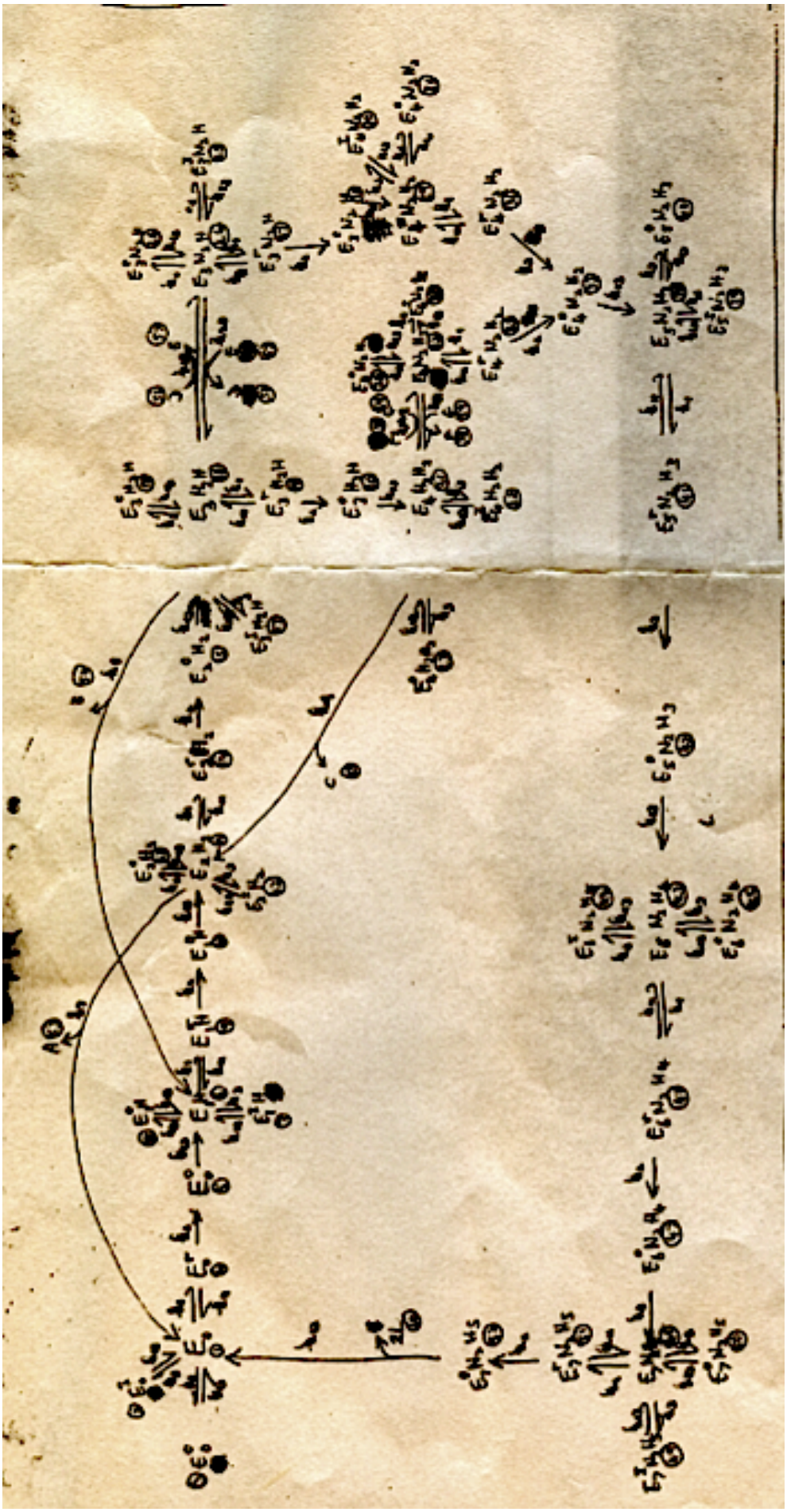


Figure 9: Lowe's notes for the full Thorneley and Lowe scheme.

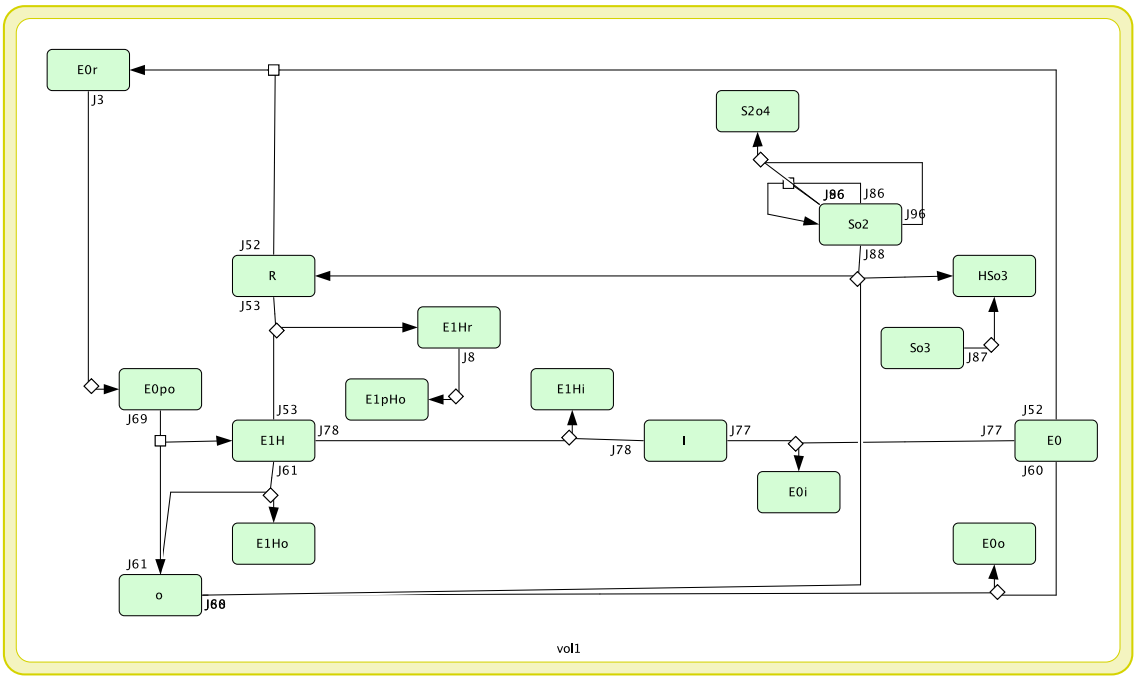


Figure 10: E_0 to E_1 of the Thorneley and Lowe scheme in SBML.

data to experimental data in both models to optimize the parameters of the system.

In addition to the numerical validation of the results, the model has been passed through all available SMBL checking programs which validate the standard SBML format [80, 87, 88, 75, 40]. The author submitted the model to the BioModels database [51]. The identifier of the SBML model of Thorneley and Lowe is MODEL1315592601.

3.3 INVESTIGATION OF THE RATE-LIMITING STEP

As an application of our full SMBL Model, we investigated the rate-limiting step in the mechanism of nitrogenase action. As discussed above, under most conditions dissociation of the Fe protein from the MoFe protein is the rate-limiting step in the system (rate constant k_{-3}). To investigate the importance of k_{-3} we conducted a series of simulations in which we sampled a range of different values for this parameter. For each value, we computed time-series data and analysed ammonia production.

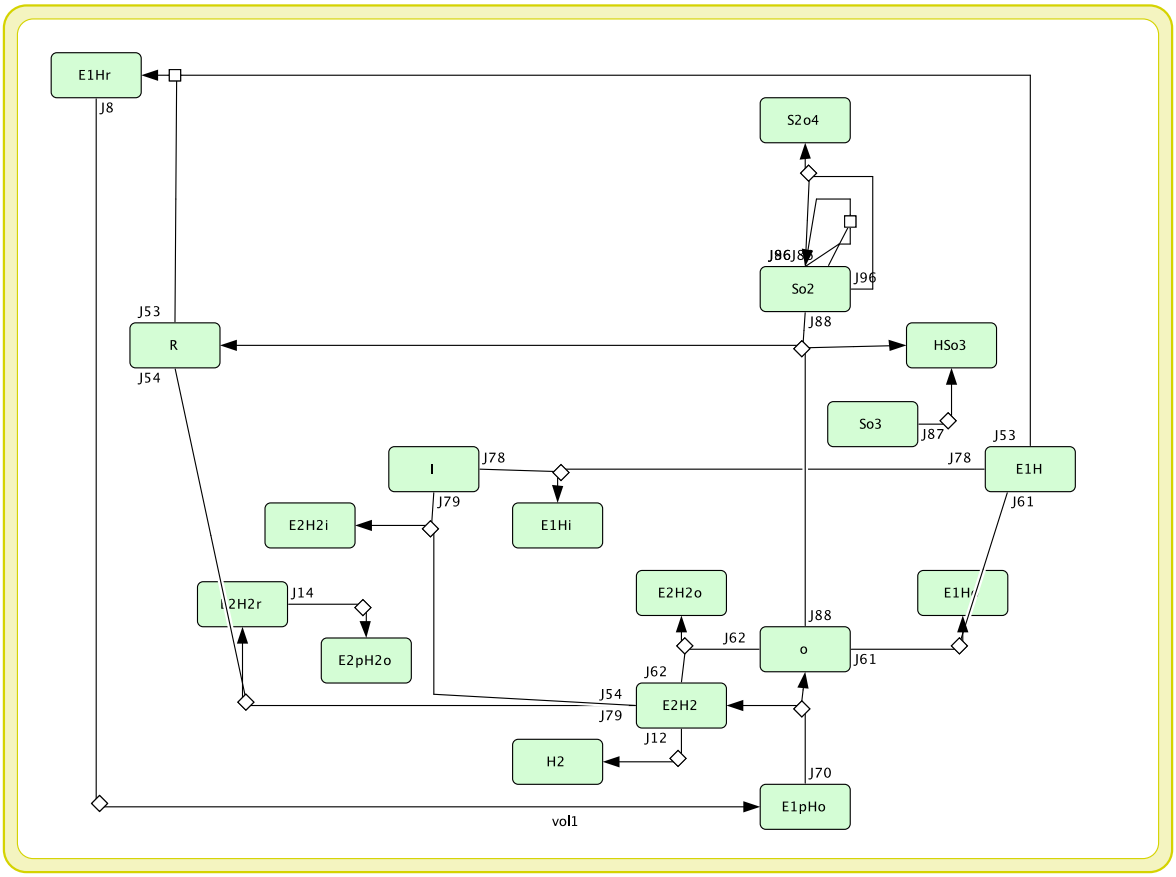


Figure 11: E_1 to E_2 of the Thorneley and Lowe scheme in SBML.

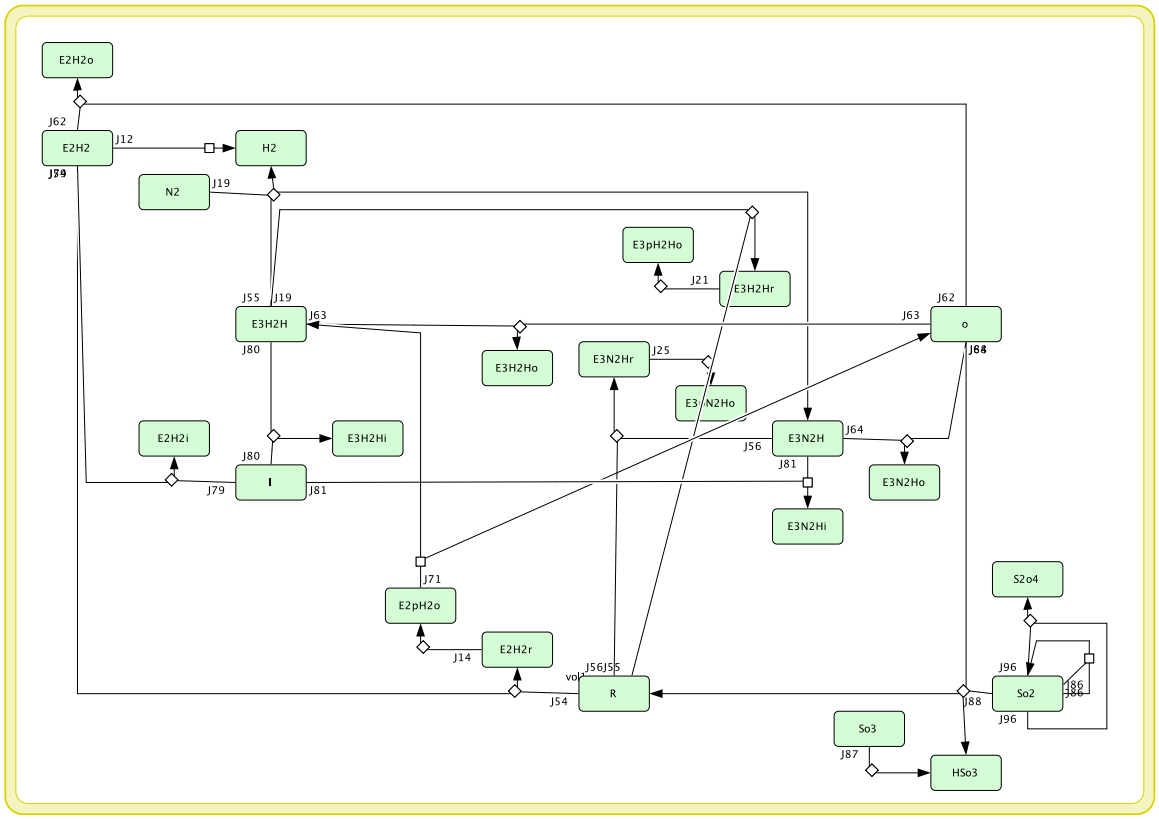


Figure 12: E_2 to E_3 of the Thorneley and Lowe scheme in SBML.

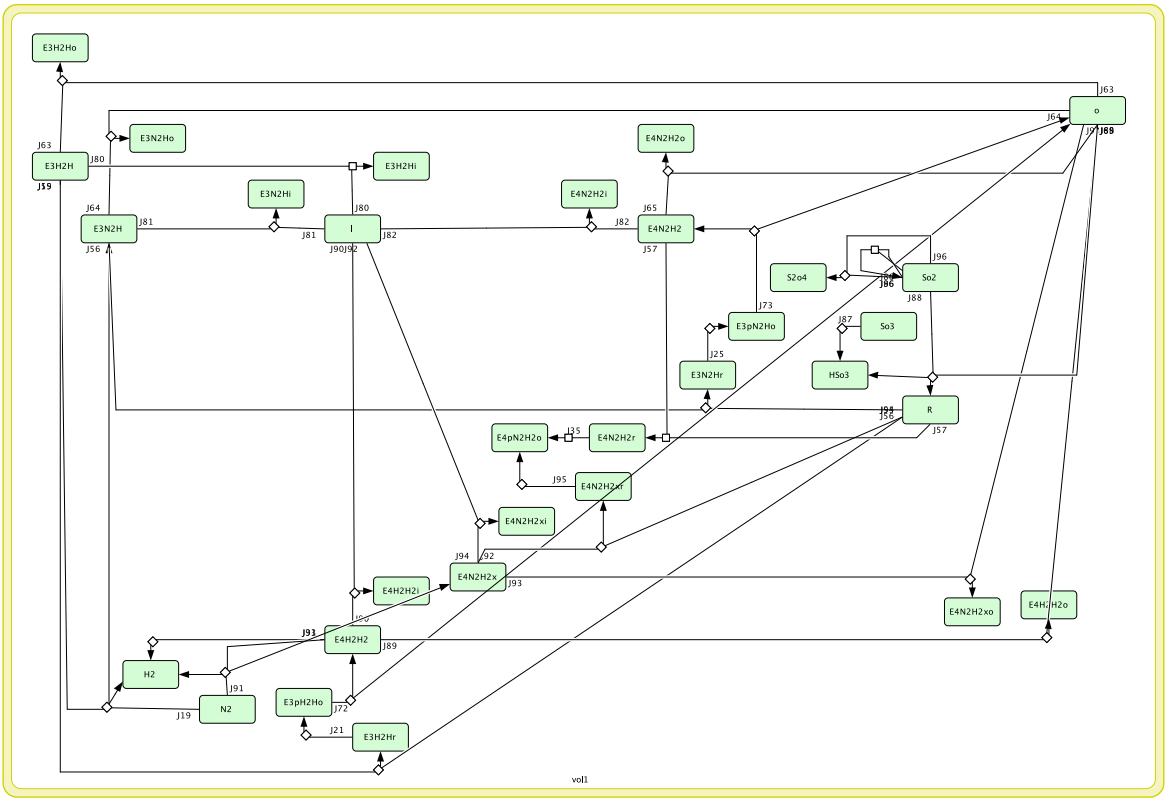


Figure 13: E_3 to E_4 of the Thorneley and Lowe scheme in SBML.

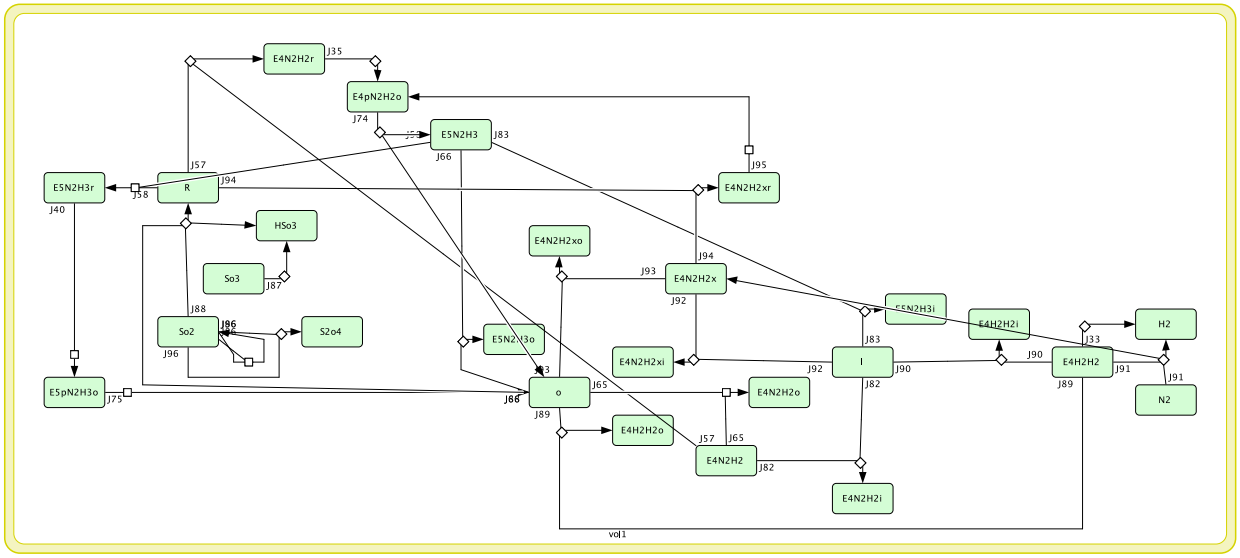


Figure 14: E_4 to E_5 of the Thorneley and Lowe scheme in SBML.

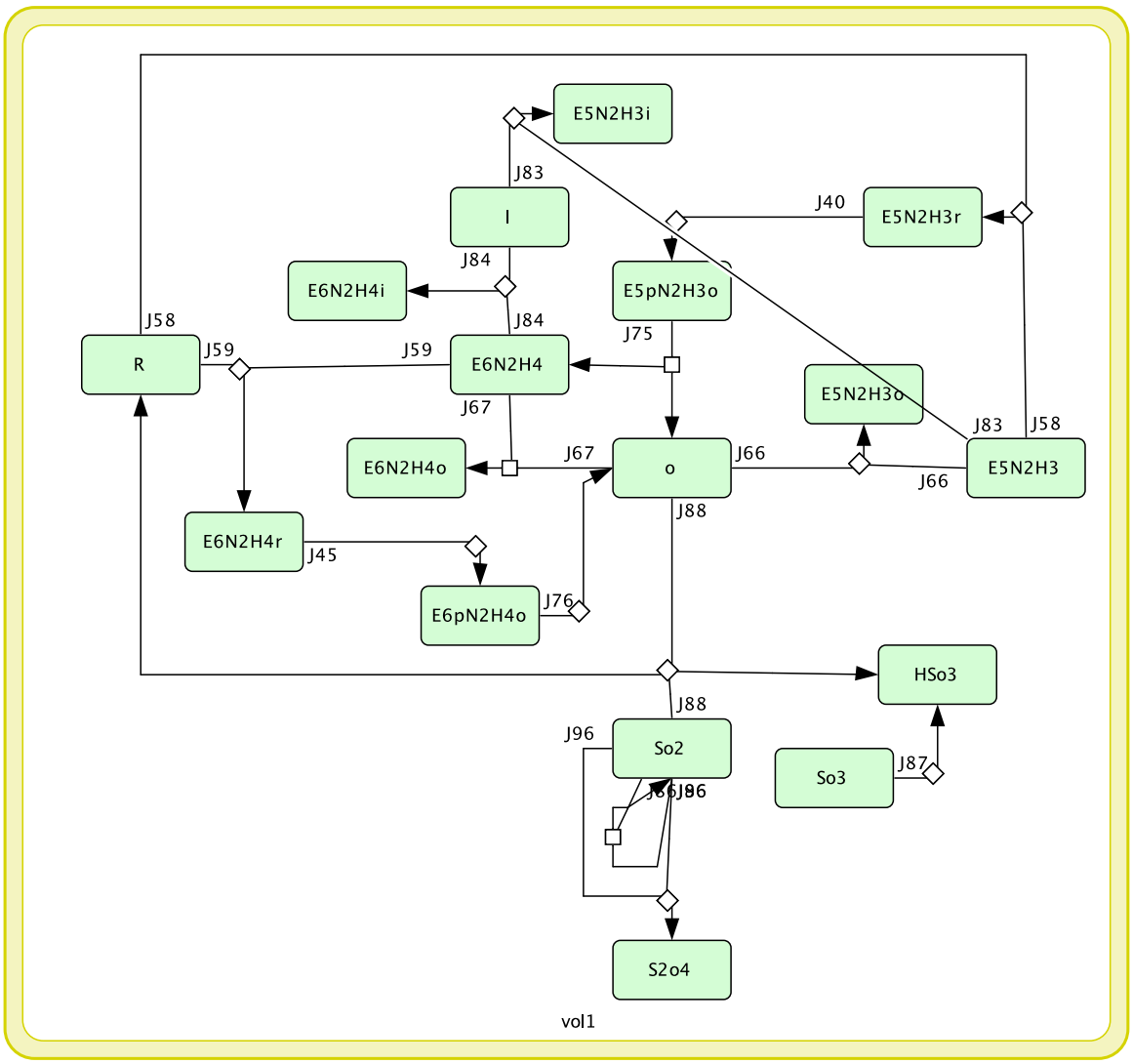


Figure 15: E_5 to E_6 of the Thorneley and Lowe scheme in SBML.

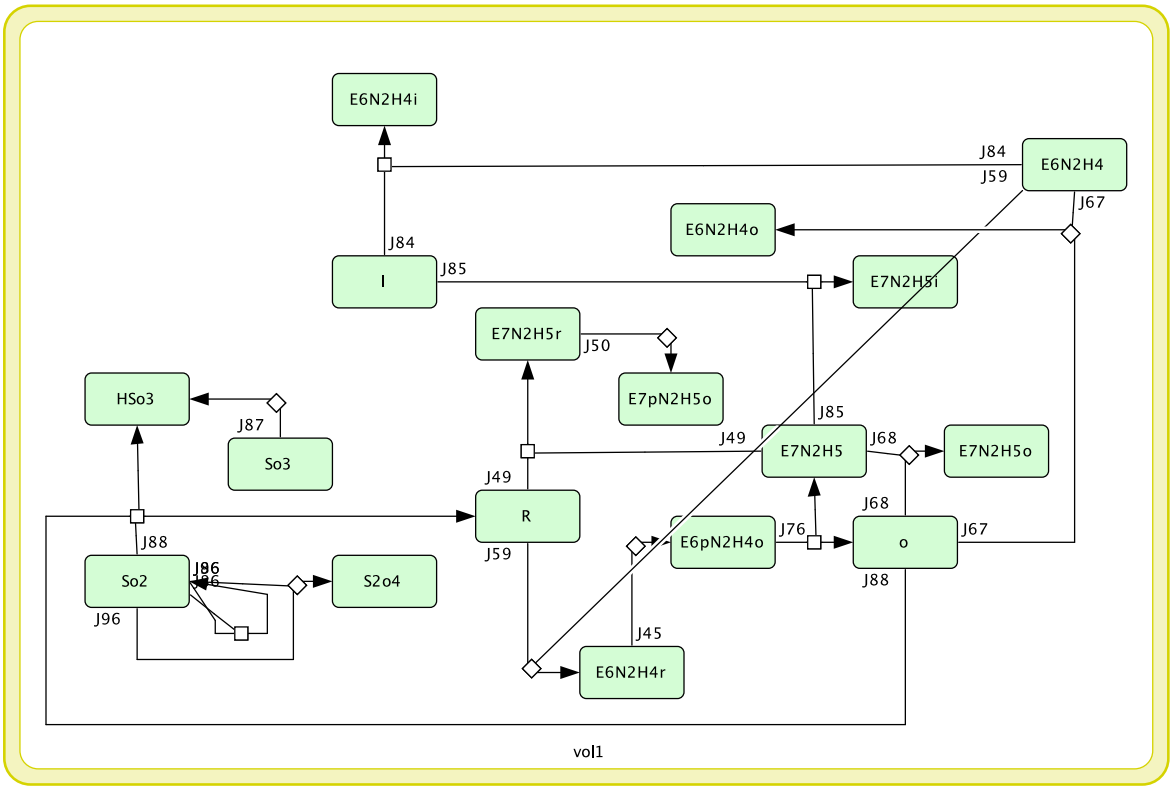


Figure 16: E_6 to E_7 of the Thorneley and Lowe scheme in SBML.

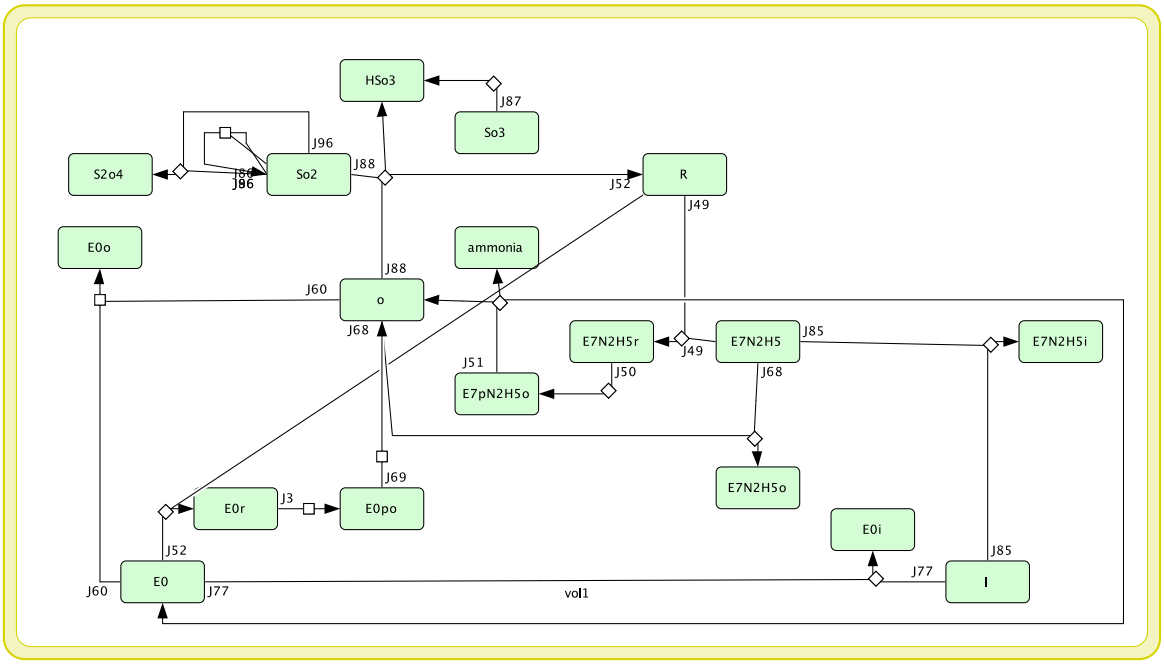


Figure 17: E_7 to E_0 of the Thorneley and Lowe scheme in SBML.

Table 1: The kinetic parameters used in the SBML model for K. *pneumoniae* nitrogenase at 23°C in pH 7.4, 25 mM HEPES buffer with 9 mM ATP, 10 mM $MgCl_2$ and zero ADP. The rate constants k_{+6} and k_{-6} (not listed in the reaction schemes) are for the dissociation of dithionite into $2SO_2^-$ [90].

Rate constant	Value
k_{+1}	5×10^7 $M^{-1}s^{-1}$
k_{-1}	15 s^{-1}
k_{+2}	200 s^{-1}
k_{+3}	4.4×10^6 $M^{-1}s^{-1}$
k_{-3}	6.4 s^{-1}
k_{+4}	3×10^6 $M^{-1}s^{-1}$
k_{+5}	4.4×10^6 $M^{-1}s^{-1}$
k_{-5}	6.4 s^{-1}
k_{+6}	1.2×10^9 $M^{-1}s^{-1}$
k_{-6}	1.7 s^{-1}
k_{+7}	250 s^{-1}
k_{+8}	8 s^{-1}
k_{+9}	400 s^{-1}
k_{-10}	8×10^4 $M^{-1}s^{-1}$
k_{+11}	2.2×10^6 $M^{-1}s^{-1}$
k_{-11}	3×10^4 $M^{-1}s^{-1}$

Table 2: The initial species concentrations used in the SMBL model. E_0 represents the Mo-Fe protein in its resting state, I, R, and O represent the inhibited, reduced and oxidised form of the Fe protein, respectively, and N_2 the nitrogen concentration.

Molecular species	Initial concentration
E_0	3.4×10^{-5} mol/l
I	5.985×10^{-5} mol/l
R	7.315×10^{-5} mol/l
$S_2O_4^{-2}$	0.01 mol/l
O	0.0 mol/l
N_2	0.001 mol/l

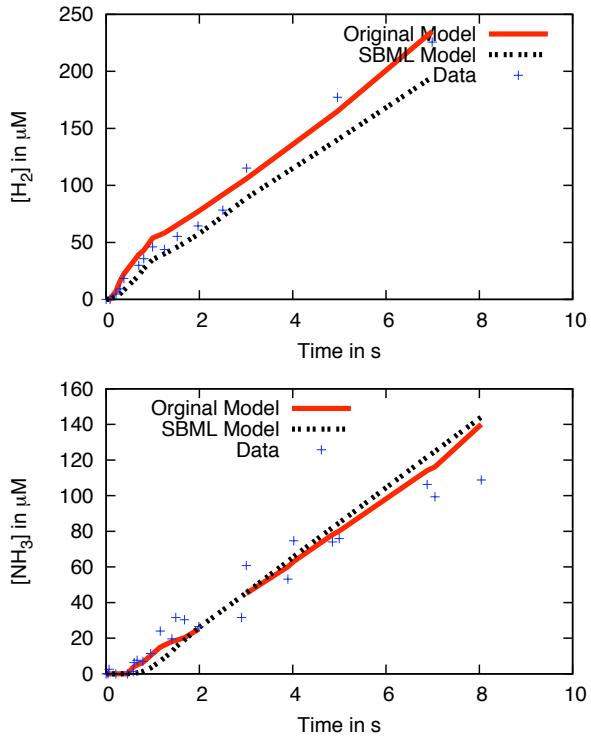
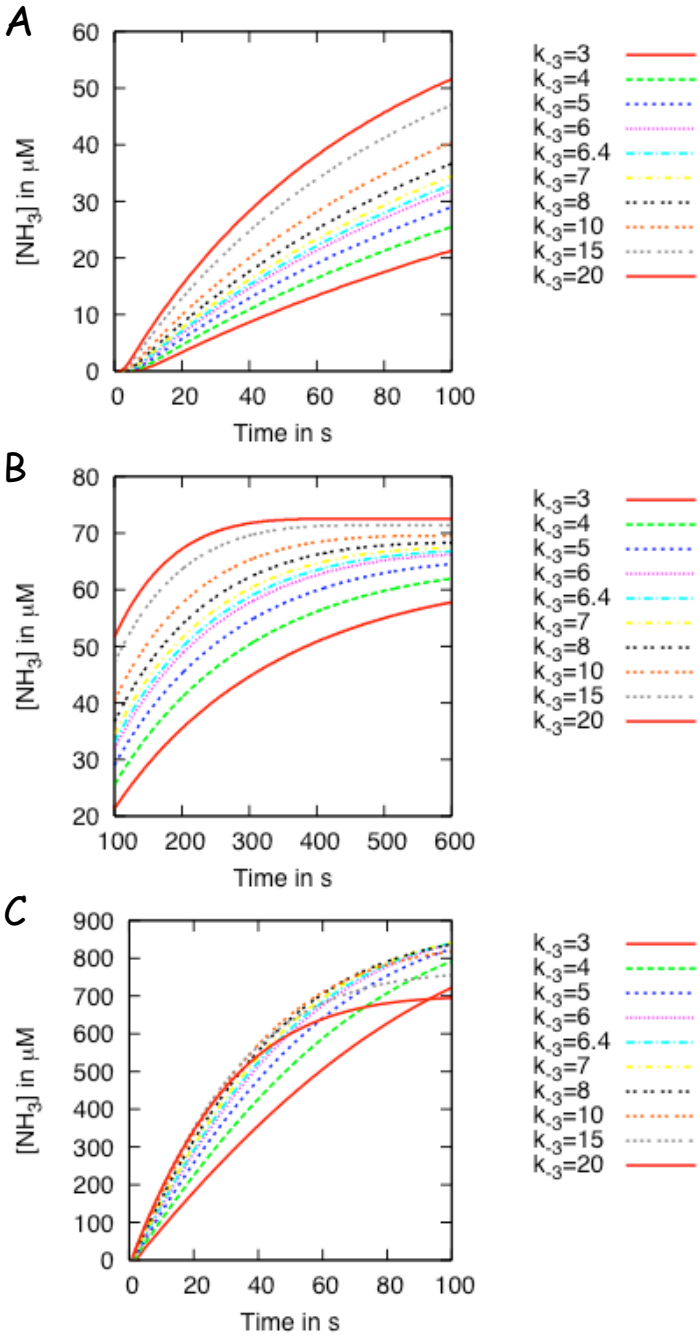


Figure 18: Simulated hydrogen and ammonia concentration together with experimental data points and original Thorneley and Lowe.

As the concentrations of Fe and MoFe protein can vary, these simulations were carried out for different MoFe:Fe ratios. Thorneley and Lowe attempted to investigate NH_3 formation in the nitrogenase system based on N_2 changes in different MoFe:Fe ratios [57]. The ratio of the MoFe protein to Fe protein can change in nitrogenase system and they showed the importance of the MoFe:Fe ratio and N_2 concentration on the formation of NH_3.


A
[NH3] in µM
Time in s
B
[NH3] in µM
Time in s
C
[NH3] in µM
Time in s
k-3=3
k-3=4
k-3=5
k-3=6
k-3=6.4
k-3=7
k-3=8
k-3=10
k-3=15
k-3=20

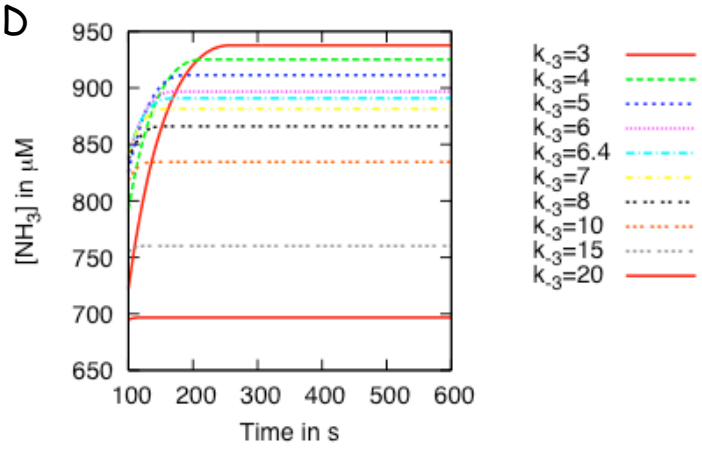


Figure 19: The time evolution of ammonia production for protein ratios of MoFe:Fe 1:1 and 1:2. (A) 1:1 in 100 s. (B) 1:1 in 600 s. (C) 1:2 in 100 s. (D) 1:2 in 600 s.

Table 3: Elasticity and control coefficients for Fe cycle

elasticity	r_1	r_2	r_3	r_4
Kp_2	0.9	0	0	0
Kp_1	0	0	6.4	0
Kp_1-Kp_2	-15	200	0	0
Kp_2o-Kp_1r	0	0	-747	509
Kp_2o	4.9×10^{6}	0	0	0
Kp_1r	0	0	-747	509
control coefficients	r_1	r_2	r_3	r_4
Kp_2	-2×10^{-7}	-1.2×10^{-8}	8.6×10^{-8}	1.2×10^{-7}
Kp_1	-2×10^{-7}	-1.2×10^{-8}	8.6×10^{-8}	1.2×10^{-7}
Kp_1-Kp_2	2.5×10^{-9}	-0.004	0.0019	0.0029
Kp_2o-Kp_1r	1.9×10^{-7}	0.004	-0.002	-0.002
Kp_2o	5×10^{-10}	1.2×10^{-5}	0.0004	-0.0004
Kp_1r	5×10^{-10}	1.2×10^{-5}	0.0004	-0.0004

Simulations of the nitrogenase system were performed between ratios of 1:1 to 1:5 of MoFe:Fe. We kept the concentration of the MoFe protein constant and increased the concentration of the Fe protein. The range of k_{-3} starts from 3 s^{-1} to 20 s^{-1} with a stepwise increase of 1. For a ratio of 1:1, increasing the rate of dissociation, decreases the time required to achieve a steady state (Figure 19). The highest k_{-3} in the simulations is 20 s^{-1}, which leads the system to achieve a steady state in about 300 s while the lowest one is 3 s^{-1}, which does not lead the system into steady state within 600 s. The time required to get to the steady state decreases by increasing the rate of dissociation in ratio of 1:2. The minimum time required to get to the steady state in simulations with a ratio of 1:2 is 100 s while it was 300 s for a ratio of 1:1. A dramatic increase in production of ammonia is the other difference between ratios of 1:1 to 1:2. The maximum ammonia production in a given time for ratio of 1:1 is 70 μM, but it is close to 950 μM in ratio of 1:2. The simulations using a ratio of 1:2 also show that increasing the rate of dissociation decrease the time to get to the steady state, at the same time, it decreases the production of ammonia (Figure 19).

The behavior of the simulations using a ratio of 1:3, 1:4, or 1:5 is similar to the behavior using a ratio of 1:2. The maximum ammonia productions are close to 1000 μM in all three simulations. The minimum time required to get to a steady state decreases to 40 s, 30 s, and 20 s in ratios of 1:3, 1:4, and 1:5, respectively (Table 4). The interval between maximum and minimum ammonia production decreases by increasing the ratio of proteins demonstrating that increasing the ratio of protein, decreases the sensitivity of ammonia production on k_{-3}.

For optimal functioning, this complex enzyme must balance a number of different signals. For example, if one considers the total ammonia production after a time of, say 2 minutes, one can observe the emergence of a maximum at around the experimentally determined value of k_{-3} which is 6.4 s^{-1} (Figure 18). This suggests that this slow enzyme has evolved to maximise the production of NH_3 under the range of protein concentrations that are typical within the cell. Figure 18 also shows that ammonia production is more sensitive to the rate of dissociation of proteins at lower proteins ratio. In higher ratios the system is less sensitive to perturbation of rate of dissociation of the MoFe-Fe complex.

For further analysis of the system we also performed elasticity and control coefficients from metabolic control analysis (MCA). We performed elasticity analysis to measure the sensitivity of the system parameters in response to changes in substrates and products (Table 3). The analysis confirms that k_{-3} is a rate limiting step as r3 which is the reaction of dissociation is strongly inhibited by complex (-747). Control coefficient in that table present the relative changes in concentration in response to changes in parameters. It shows that relative changes of model components concentration is relatively high for r3 leading us to confirm that r3 is a rate limiting step in this process.

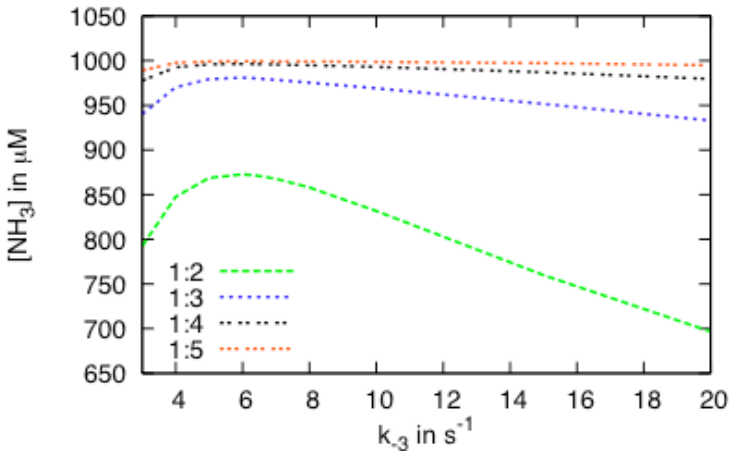


Figure 20: The production of ammonia after 120 s as a function of the rate constant k_3 for different MoFe:Fe protein ratios.

Table 4: Time to steady state based on different ratios of proteins and various values of k_{-3}.

k_{-3}	3	4	5	6	6.4	7	8	10	15	20
1:1	1058	845	749	677	652	606	547	515	432	373
1:2	260	214	192	173	167	162	154	141	128	118
1:3	220	174	147	127	121	115	105	93	77	72
1:4	220	172	141	120	114	106	98	81	63	56
1:5	227	172	143	120	113	104	94	80	60	49

Another recently developed model attempted to compensate for discrepancies between theory and experiment by adjusting the MoFe and Fe protein concentrations to fit the data [95]. Such numerical tasks can also easily be

performed with our model, however, the rationale behind changing the concentrations on a plot-by-plot basis may be questioned. The simulations from our model are overall in better agreement than this other recent attempt. We speculate that this is because we have developed the full model under the guidance from one of the originators, although differences in the numerical simulation scheme may also be a cause.

3.4 CONCLUSION

In this chapter, we have developed and implemented the full Thorneley and Lowe pre-steady state kinetic model of Nitrogenase in K. *pneumoniae* in the Systems Biology Markup Language (SBML) and submitted it to the BioModels database. We validated our model and showed that it is in good agreement with the original experimental data. In addition, we have investigated the rate limiting step by performing a parameter scan of the rate constant k_{-3}. A close inspection of the simulations reveals that at low MoFe:Fe ratios, the production of ammonia is more sensitive to the rate of dissociation constant of the complex of the Fe protein and MoFe protein. However, for increasing MoFe:Fe concentrations, this situation is reversed. The ratio of proteins plays key role in ammonia production and it decrease the sensitivity of the system to the dissociation constant (k_{-3}).

Part III

THE NIFL-NIFA SYSTEM

4

THE 2-OXOGLUTARATE SYSTEM

Modelling nitrogen fixation requires considerable biological knowledge and experimental data. Our starting point was to extract the biological knowledge from different sources so as to identify the system components and master regulators in this complex regulatory system. There are two important factors in modelling the NifL – NifA system. Firstly, most of the experiments for the NifL – NifA system have been performed in steady state conditions [53], and secondly, the low availability of the kinetic data (Appendix C). Rate constants are the most common type of kinetic data used for kinetic modelling, however there have not been measured many rate constants experimentally for the NifL – NifA system. Most of the reactions in the NifL – NifA system are protein-protein interactions or protein-ligand interactions. These interactions are very fast and difficult to measure experimentally. However, data is available for the dissociation constant, which is a measure of the affinity of interaction between two proteins or as protein and a ligand. In this chapter we initiate modelling the NifL – NifA system for *A. vinelandii*.

4.1 METHODOLOGY

We applied two methodologies to model the NifL – NifA sub-systems using dissociation constants, which we name A and B. We now present a detailed description of these two the methodologies.

4.1.1 *Methodology A*

In methodology A, we attempted to form mathematical equations using K_d values, the dissociation constant is equal to the ratio of the backward rate constant to the forward rate constant. As mentioned in the last section, our aim was to model NifL – NifA system in the steady state condition.

Therefore, the first step was arrangement of the steady state equations. Figure 21 presents reaction scheme of an example of the $NifL - NifA$ subsystem. In this section the methodology is presented using the example system below:

$$\begin{aligned} NifL_{red} + ADP &\rightleftharpoons NifL_{red}ADP, \\ NifL_{red}ADP + NifA &\rightleftharpoons NifL_{red}ADPNifA, \\ NifA + 2OG &\rightleftharpoons NifA(2OG). \end{aligned}$$

For ease of presentation we use the following substitutions:

$$\begin{aligned} x_1 &= [NifL_{red}], \\ x_2 &= [ADP], \\ x_3 &= [NifL_{red}ADP], \\ x_4 &= [NifA], \\ x_5 &= [NifL_{red}ADPNifA], \\ x_6 &= [2OG], \\ x_7 &= [NifA(2OG)]. \end{aligned}$$

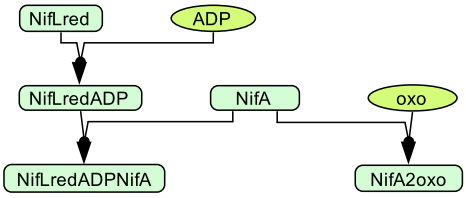


Figure 21: The $2 - oxoglutarate$ system.

Using the methodology described in Chapter 2, the law of mass action was applied to produce a system of ODEs. In the following equations k_i and k_{-i} denote the forward and backward rates.

$$\begin{aligned}
\dot{x}_1 &= -k_1x_1x_2 + k_{-1}x_3,\\
\dot{x}_2 &= -k_1x_1x_2 + k_{-1}x_3,\\
\dot{x}_3 &= k_1x_1x_2 - k_{-1}x_3 - k_2x_3x_4 + k_{-2}x_5,\\
\dot{x}_4 &= -k_2x_3x_4 + k_{-2}x_5 - k_3x_4x_6 + k_{-3}x_7,\\
\dot{x}_5 &= k_2x_3x_4 - k_{-2}x_5,\\
\dot{x}_6 &= k_{-3}x_7 - k_3x_4x_6,\\
\dot{x}_7 &= -k_{-3}x_7 + k_3x_4x_6.
\end{aligned}$$

By assuming $\dot{x}_i$=0 for all i=1,..,7, we obtain:

$$\begin{aligned}
-k_1x_1x_2 + k_{-1}x_3 &= 0,\\
k_1x_1x_2 - k_{-1}x_3 - k_2x_3x_4 + k_{-2}x_5 &= 0,\\
-k_2x_3x_4 + k_{-2}x_5 - k_3x_4x_6 + k_{-3}x_7 &= 0,\\
k_2x_3x_4 - k_{-2}x_5 &= 0,\\
k_{-3}x_7 - k_3x_4x_6 &= 0.
\end{aligned}$$

Solving these equations for the products of the system (x_3, x_5, x_7) gave the following equations:

$$\begin{aligned}
x_3 &= k_1x_1x_2/k_{-1},\\
x_5 &= k_2k_1x_1x_2x_4/k_{-2}k_{-1},\\
x_7 &= k_3x_4x_6/k_{-3}.
\end{aligned}$$

Applying the dissociation constants formulas,

$$\begin{aligned}
K_{d_1} &= k_{-1}/k_1,\\
K_{d_2} &= k_{-2}/k_2,\\
K_{d_3} &= k_{-3}/k_3.
\end{aligned}$$

resulted in:

$$x_3 = x_1 x_2 / K_{d_1}, \tag{1}$$
$$x_5 = x_1 x_2 x_4 / K_{d_2} K_{d_1}, \tag{2}$$
$$x_7 = x_4 x_6 / K_{d_3}. \tag{3}$$

This process shows how a system can be converted to a system of algebraic equations involving dissociation constants instead of forward and backward rate constants. x_1 and x_4 are two main components of the system and these are part of its different products. As the system is in steady state and the total concentration of x_1 and x_4 in the system is constant, the following assumptions were also made (t indicates total):

$$\begin{aligned} x_1 &= x_1^t - x_3 - x_5, \\ x_4 &= x_4^t - x_7 - x_5. \end{aligned}$$

The assumptions can be applied to (1), (2) and 3, giving:

$$x_3 = (x_1^t - x_5) x_2 / (K_{d_1} + x_2), \tag{4}$$
$$\begin{aligned} (x_2) x_5^2 + (x_2 (x_3 - x_1^t - x_4^t + x_7) - K_{d_2} K_{d_1}) x_5 + \\ x_2 (x_1^t x_4^t - x_1^t x_7 - x_3 x_4^t + x_3 x_7) = 0, \end{aligned} \tag{5}$$
$$x_7 = (x_4^t - x_5) x_6 / (K_{d_3} + x_6). \tag{6}$$

In the experiments presented in [53] one or two of the components are variables of the system and the experiments were performed a number of times with different concentrations of the key components. In this particular example x_6, varied in each of the experiments in [53]. In section 3.3.1 we present the results of this model. A simple Maple code was written by author for the purpose of applying methodology A to model the system.

4.1.2 *Methodology B*

In the previous section, a methodology for modelling a biological system in a steady state condition using the dissociation constants instead of forward and backward rate constants was presented. In this section the focus is to present a simpler way to use K_d values.

The dissociation constant indicates the affinity of interaction of two or more components of the system. Since the dissociation constant is given by:

$$K_d = k_{backward}/k_{forward},$$

if we assume the value of the forward rate constant ($k_{forward}$) is equal to one, the dissociation constant is equal to the backward rate constant ($k_{backward}$). This assumption is valid in modelling the systems under the steady state conditions because keeping the ratio of backward and forward constants to the constant value of the dissociation constant, while changing the actual values of rate constants, does not change the steady state concentration of model components. It only alters the time to get to the steady state conditions. There are no standard tools available for modelling using only K_d, although some systems biology tools exist for modelling the system in steady state conditions.

Apart from using available tools for modelling $NifL - NifA$ system, the author of this thesis developed a java code to model the $NifL - NifA$ system in steady state conditions using K_d (The Equilibrium program). The motivation behind the development of the code was to provide specific tools and perform simulation jobs in an automatic way to generate a considerable amount of data for sampling and sensitivity analysis of the $NifL - NifA$ system. Methodology B was used more than Methodology A mainly due to the availability of tools. Methodology A was used for sections 4.2.1 and 4.3.1. Methodology B was used in other sections.

4.1.3 *Parameter estimation*

The availability of biochemical parameters is a limiting factor in modelling biochemical systems. We used an estimation method not only to estimate the missing parameters but also to validate the models by estimating the known parameters. An important part of model development is validation. Comparing the output of the model with experimental data can help to validate a model. Parameter estimation methods can also be used to estimate the known parameters using the available data to compare the estimated value to the original one for validation of the model. We used the genetic

algorithm tool from Copasi to solve parameter estimation problems in this thesis.

The genetic algorithm is a computational technique that is inspired by biological evolution [62]. The genetic algorithm starts from a population of individuals. Each individual has a set of model parameters. Reproduction and selection of the population are the key elements of genetic algorithms. The initial population is generated randomly. The population is updated in each generation by reproduction and selection process. In the reproduction process new individuals are generated by the processes called mutation and cross-over. Mutation randomly alters one or more parameter values in an individual. Genetic algorithms help prevent the algorithm getting trapped in local minima by mutation processes. A cross-over process is used to produce new individuals by combining two individuals. In the cross-over process two individuals swap some of the parameters to generate the new ones. In the selection process some of the individuals, which have a better fitness stay in the population and the others, are deleted and replaced by the individuals that were generated by the mutation and cross-over process. The generation process is repeated until either a good fit is found or the algorithm reaches the defined generation number.

4.1.4 *Stability analysis*

Most of the experimental data that we analysed was generated in steady state conditions. Knowing the steady state points of a system helps to analyse the system and also compare the generated data in that condition to the experimental data. Apart from finding the steady state points, the stability of these points can be determined to characterize the systems. In modelling the $\mathtt{NifL} - \mathtt{NifA}$ system, there were a number of potential models for each sub-system. A stability analysis was used to provide more information about these potential models. Potential models are categorized in two categories of stable and unstable models. This helps narrow down the set of potential models. The author developed a program written in java to calculate the steady state points. Copasi was used to perform stability analysis.

4.1.5 *Sensitivity analysis*

Sensitivity analysis plays an important role in the investigation of biochemical models. It helps to assess the effect of components of the system based on the model output and to determine the key regulators and parameters of the system. We performed a sensitivity analysis to provide information about the $NifL - NifA$ sub-systems. The sensitivity analysis in this thesis was performed in two different ways.

Sampling is the simplest way to carry out a sensitivity analysis. $NifA(2OG)$ was assigned to be the main output of the $NifL - NifA$ system. The $NifL - NifA$ system can be solved for $NifA(2OG)$ many times through sampling components, which reveals some of the characteristics of the system. The range of model components and the number of samples were defined for each component of the system. Uniformly distributed samples were taken from model components in the defined range. Given the parameters and different values for model components as inputs, as many as number of samples the simulations of the $NifL - NifA$ system were run to produce output values. We explored the output values since the variations in outputs are the result of variations in the input. The Equilibrium program was designed and used to perform this task.

Multi-parametric sensitivity analysis (MPSA) can also be used to study the sensitivity of model output to the model components. MPSA introduces uncertainty to the model by injecting randomly generated model component values into the system and statistically evaluating the sensitivity of the system to these values. The procedure of MPSA starts by selecting the parameters as the variables of the system. A set of randomly generated parameters is sampled in a given range by the Latin Hypercube Sampling (LHS) method [96]. This is followed by simulation of the model for each set of parameter values and calculation of the objective function. A set of parameters is then classified into an acceptable or unacceptable sets by comparing the objective function values with a predefined threshold. Finally, the sensitivity of the parameters is evaluated statistically by computing the cumulative frequency of both acceptable and unacceptable cases for each parameter and measuring the separation of the two cumulative frequency distributions using Kolmogorov-Simirnov statistics [96]. SBML-SAT is a

freely available SBML type sensitivity analysis tool, which was used to perform MPSA in this thesis [97]. It can be run to evaluate the sensitivity of both parameters and initial conditions of model components. SBML-SAT has tools to evaluate the sensitivity of the parameters and model components to both the steady state and dynamic response of the system in a given time (integrated response).

4.2 NIFL AND NIFA INTERACTION

Direct interaction of NifA and NifL inactivates the expression of nitrogen fixation genes. This interaction occurs in different redox status and also the presence of ADP and ATP can change the influence of NifL on NifA activity. In the following sections we analyze the simplest sub-systems of the NifL − NifA system in steady state conditions including the NifL and NifA interaction in reduced conditions with ADP, in reduced conditions with ATP, and finally in oxic conditions with ADP. The activity of NifL in transcription of nif genes was determined by measurement of the formation of open promoter complex [74].

4.2.1 *NifL and NifA interaction in reduced conditions with ADP*

The NifL − NifA system switches between oxidized and reduced conditions based on the environmental situation and input signals. We investigated the NifL − NifA system in different redox conditions separately to get a better understanding of their behaviors. The following equations are the reactions in the system in reduced conditions:

$$\mathrm{NifL} + e^{-} \rightarrow \mathrm{NifL_{red}}, \tag{4.1}$$

$$\mathrm{NifL_{red}} + \mathrm{ADP} \underset{k_{d_2}}{\rightleftharpoons} \mathrm{NifL_{red}ADP}, \tag{4.2}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{NifA} \underset{k_{d_3}}{\rightleftharpoons} \mathrm{NifL_{red}ADPNifA}. \tag{4.3}$$

In the presence of reductants and when the system is not in the oxidized condition, NifL interacts with the available reductant in the cell to produce the reduced form of NifL, $\mathrm{NifL_{red}}$ (reaction 4.1). $\mathrm{NifL_{red}}$ is not able to inhibit NifA activity unless it binds to ADP, which causes conformational

changes in $NifL_{red}$ (reaction 4.2) [74, 9, 10]. $NifL_{red}$ can then finally interact with NifA (reaction 4.3).

Table 5 presents the system of ODEs for the above reaction scheme. The law of mass action was used to produce a system of ODEs. The rate of change of all the components of the system is equal to zero in steady state conditions. All forward rate constants have been set to 1 and the backwards rate constants are equal to K_d (See Section 4.1.2).

Table 5: The ODEs for the $NifL - NifA$ system in reduced conditions with ADP. For ease of presentation we used the following substitutions x_1=[NifL], x_2=[e^-], x_3=[$NifL_{red}$], x_4=[ADP], x_5=[$NifL_{red}$ADP], x_6=[NifA], and x_7=[$NifL_{red}$ADPNifA]. k_i and k_{-i} denote the forward and backward rates (i=1, 2, 3).

$$\dot{x}_1 = -k_1x_1x_2$$
$$\dot{x}_2 = -k_1x_1x_2$$
$$\dot{x}_3 = k_1x_1x_2 - k_2x_3x_4 + k_{-2}x_5$$
$$\dot{x}_4 = -k_2x_3x_4 + k_{-2}x_5$$
$$\dot{x}_5 = k_2x_3x_4 - k_{-2}x_5 - k_3x_5x_6 + k_{-3}x_7$$
$$\dot{x}_6 = -k_3x_5x_6 + k_{-3}x_7$$
$$\dot{x}_7 = k_3x_5x_6 - k_{-3}x_7$$

In many *in vitro* experiments for the $NifL - NifA$ system, the truncated form of NifL, NifL(147-519), has been used [74]. The reason behind this is to prohibit NifL from getting into the oxidized form. It loses the ability to sense oxygen, consequently, in this condition NifL is considered to be in the $NifL_{red}$ form. For ease of calculation reaction 4.1 can be ignored and we can assume that the total concentration of NifL is in $NifL_{red}$ form.

The activity of NifL in transcription of *nif* genes is increased by adding NifA to the system (Figure 22 A). The reaction was carried out in the presence of NifL(147-519) while there is no ADP in the system. As expected in the absence of ADP, NifL fails to inhibit NifA. The computational model confirms the availability of free NifA in this condition (Figure 22 B). Reactions 4.2 and 4.3 were used to calculate the steady state concentration of free NifA in the system, where the K_{d_2} is 13 μM and K_{d_3} were set to 0.03 μM. The k_2, k_{-2}, k_3, and k_{-3} are 1 $M^{-1}s^{-1}$, 13 s^{-1}, 1 $M^{-1}s^{-1}$, and 0.03 s^{-1}, respectively (Relevant to reactions 4.2 and 4.3). The assumption was made

that there is either no interaction between NifL and NifA in the absence of ADP or a very low level of interaction. In this condition, since the interaction of NifL and NifA is not significant, almost all NifA is in free form, thus the transcription of nif genes is activated.

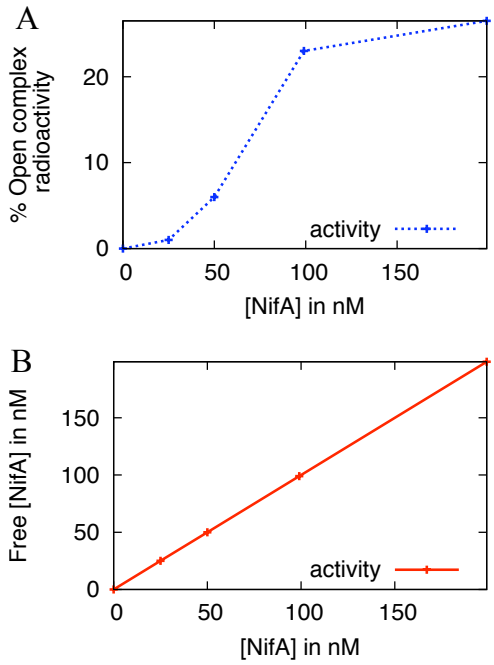


Figure 22: (A) Experimental evidence for activity of NifA on activation of transcription of Nitrogenase. The reaction was performed with 4 mM GTP, 0.2 µM of NifL, and NifA as indicated in graph A [74]. The experiment was performed by experimentalist. (B) Computational evidence for activity of NifA. B presents the change of concentration of NifA which is correspond to activity of NifA. It has the same concentration as A.

ADP plays an important role in the NifL – NifA system in that NifL binds to ADP and causes conformational changes in NifL. This binding activates NifL and the interaction with NifA. An experiment was performed to investigate the effect of ADP on the NifL – NifA system (Figure 23) [53]. The experiment had 0.1 µM of NifA, 4 mM GTP, 0.2 µM of NifL(147-519), and ADP as indicated in Figure 23. Activity of NifA was determined by monitoring the formation of open promoter complexes. NifA can utilise GTP for open complex formation but neither GTP nor GDP are also competent to bind to NifL in the system. Increasing ADP dramatically decreases the activity of NifA in presence of NifL (Figure 23 A).

We performed a number of simulations to investigate the influence of ADP on the NifL – NifA system. In simulation 1, we perturbed the ADP concentration using reactions 4.2 and 4.3 to monitor the activity of NifA. The activity of NifA in our computational test is a ratio of free NifA to the total amount of NifA. The simulation based on different concentrations of ADP is in good agreement with the corresponding experimental data.

The physiological concentrations of NifL and NifA were experimentally estimated *in vivo* is 0.01 μM. We carried out simulations with different concentrations of NifA and NifL. In simulation 2 we used the ratio of 2:1 NifL/NifA as used in the *vitro* experiment, but with 10 fold lower concentration of NifA and NifL. As it is indicated in Figure 23 (C), ADP can decrease the activity of NifA but the system is not very sensitive to ADP. It cannot decrease the activity of NifA to less than 75%. In this condition, 75% activity of NifA, is considered to be active for the transcription of *nif* genes.

In the final simulation we used the same concentration of NifL and NifA, 0.01 μM (Figure 23 (D)). The activity of NifA in this condition, based on different concentrations of ADP, is the lowest among all the simulations. Simulation 2 shows that the concentrations of NifL and NifA are around 0.1 μM because the activity of NifA does not decrease when there are 0.01 μM of NifA and 0.02 μM of NifL available in the system. Although the expression of NifL and NifA appears to be transcriptionally coupled resulting in stoichiometric amounts of NifL and NifA in the cell, simulation 3 shows that the ratio of NifA and NifL plays an important role in the activity of NifA as the activity does not decrease when the stoichiometric amounts of NifL and NifA is present. The simulation fits to the experimental data when the amount of NifL is 2 fold higher than NifA. Thus these simulations predict that there is more NifL than NifA in the cell and also their concentrations are 0.1 μM.

4.2.2 *NifL and NifA interaction in reduced conditions with ATP and ADP*

ATP is also able to bind to NifL but with much lower affinity in comparison to ADP. In the presence of ATP, NifL binds NifA and form stable complexes of NifL – NifA. Reactions 4.4 and 4.5 below occur in the presence of ATP, consequently, the system in the presence of ADP and ATP yields reactions

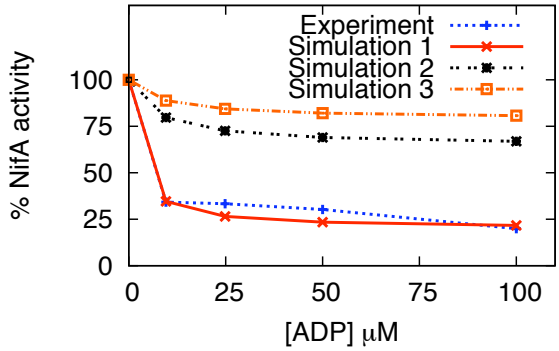


Figure 23: (A) Experimental evidence for the influence of ADP on the NifL – NifA system [53]. (B) Simulation 1: the influence of ADP on NifA activity using the same concentration as in the experiment. (C) Simulation 2: the same computational test as Simulation 1 with a 10 fold less concentration of NifL and NifA. (D) Simulation 3: the same computational test as simulation 1 with 0.01 μM of NifA and NifL. The activity of NifA is division of $NifA_{free}$ to $NifA_{total}$.

4.2, 4.3, 4.4, and 4.5. Considering the ATP hydrolysis activity of NifA, we also added reaction 4.6 to the system.

For a better understanding of the influence of ATP and ADP on the NifL – NifA system and the hydrolysis activity of NifA, sensitivity and also robustness analysis were performed on the system with reactions 4.2, 4.3, 4.4, and 4.5 (model A) and on the system formed from the reactions 4.2, 4.3, 4.4, 4.5, 4.6 (model B). The total concentration of ATP in these simulations was 3500 μM. K_{d_4}, K_{d_5}, and K_{d_6} were set to 130 μM, 0.03 μM, and 800 μM, respectively.

$$NifL_{red} + ATP \underset{k_{d_4}}{\rightleftharpoons} NifL_{red}ATP, \tag{4.4}$$

$$NifL_{red}ATP + NifA \underset{k_{d_5}}{\rightleftharpoons} NifL_{red}ATPNifA, \tag{4.5}$$

$$NifA + ATP \underset{k_{d_6}}{\rightleftharpoons} NifA + ADP + P_i. \tag{4.6}$$

As mentioned above, the multi-parametric sensitivity analysis method was used in the SBML-SAT program. To perform a sensitivity analysis a boundary for the parameters needed to be set. The up/down of 1000 fold of the original parameters were used for parameter space. These boundaries were set to prohibit infinite flexibility for parameter perturbation and also to

give some degree of freedom for perturbation. Figure 24 demonstrates the sensitivity of the integrated response of model components over a certain simulation time. The integrated response of the model output is the total amount of the model output during the time course. As the transcription takes about 60 seconds in bacteria and translation 120 seconds [1] we chose the simulation time to be 1000 seconds.

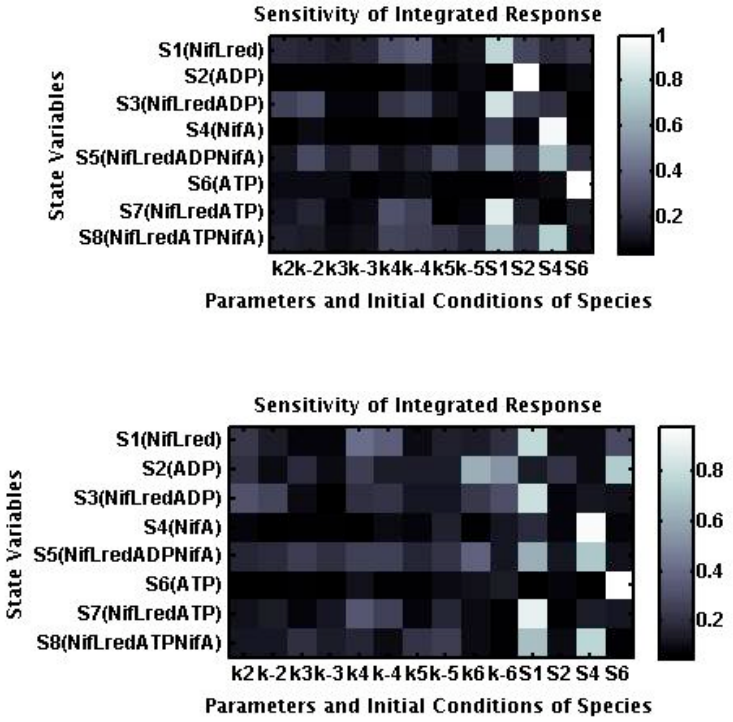


Figure 24: Top plot: The sensitivity of the integrated response of model A against all model parameters and initial conditions. Bottom plot: The sensitivity of the integrated response of model B against all model parameters and initial conditions.

The main output of the $NifL - NifA$ system is free $NifA$ (S4 in Figure 24). Thus, the evaluation of the influence of model parameters and initial conditions on $NifA$ is the main topic of interest. Model A, apart from the high sensitivity of the integrated response of $NifA$ to its total concentration, which is expected, was found to be sensitive to perturbation of $NifL_{red}$. Overall, the total amount of $NifA$ over 1000s does not change by perturbing the model parameters and total concentrations (Figure 24). The values for ATP and ADP are very stable to model perturbations. $NifL_{red}$ and its driven complexes are the most sensitive components of the system, and furthermore,

$NifL_{red}$ is the most critical component among the model parameters and total concentrations in terms of the influence on the model outputs.

For B, the analysis shows that $NifA$ as in model A is sensitive to its total concentration (Figure 24). The main difference between models A and B in terms of sensitivity of integrated response of $NifA$ is that $NifA$ is less sensitive to the total concentration of $NifL_{red}$, and also, it is sensitive to the k_{-5} the backward rate constant for interaction between $NifL_{red}ATP$ and $NifA$. The sensitivity of ATP in model A is the same as that of model B, but ADP is more sensitive in model B against perturbations than model A. $NifL_{red}$ driven complexes have the same behavior in both models A and B. In general, both models A and B react similarly to the perturbation within a timeframe of 1000s.

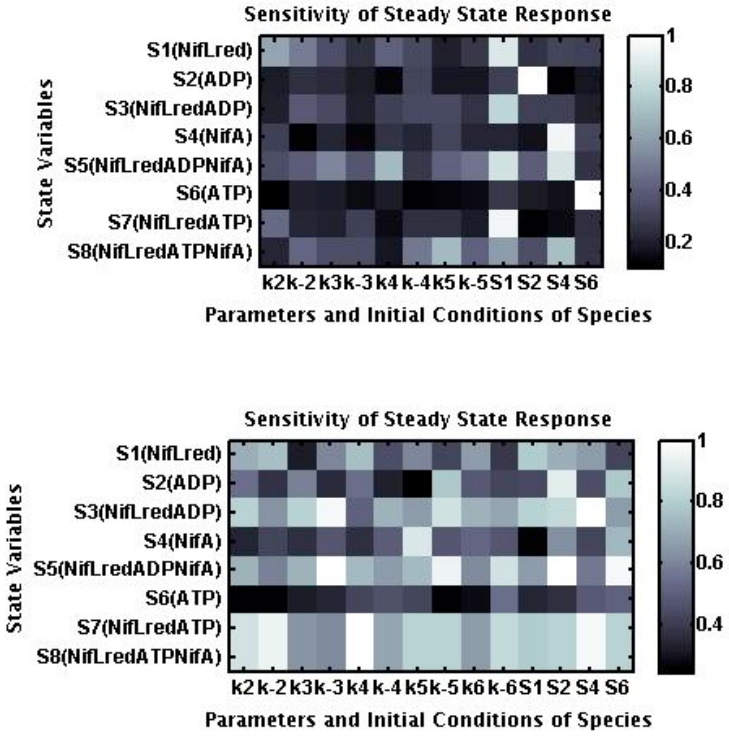


Figure 25: Top graph: The sensitivity of steady state response of model A against all model parameters and initial conditions. Bottom graph: The sensitivity of integrated response of model B for the same conditions.

Figure 25 presents the sensitivity of steady state response of models A and B against the model parameters and initial conditions. This analysis shows the effect of model parameters and initial conditions perturbations on the steady state concentrations of the model outputs. We performed the

sensitivity analysis on the steady state response of models A and B against perturbation of parameters and initial conditions. NifA is not sensitive to these perturbation in model A. Apart from its total concentration it shows some low sensitivity to k_2, the forward rate constant for $NifL_{red}$ and ADP interaction, and k_5, the forward rate constant for the $NifL_{red}ATP$ and NifA interaction. The steady state response of NifA in model B is very different from that of model A. As the analysis shows, apart from the sensitivity of NifA to its total concentration, NifA is more sensitive in model B in comparison to model A. $NifL_{red}$ is the only component of the model B which does not show any effect on NifA. The highest sensitivity of NifA is on the its interaction with $NifL_{red}ATP$, k_5. This implies that ATP plays a more important role in model B than in model A. $NifL_{red}$ driven complexes are dramatically more sensitive in model B than model A, in respect of the model parameters and the initial conditions.

Although performing the sensitivity analysis is informative, it does not show how the system reacts against multiple perturbations. We therefore performed robustness analysis using SBML-SAT. We calculated the robustness of the system in using the following formula [97]:

$$R = \frac{-\sum_{p=1}^{N}(f_p - f_0)^2}{N},$$

where the quantities N, f_p, and f_0 are defined as follows. A parameter space is defined for parameters of the system, which has 1000 fold up/down of the original parameter values in our study. A new set of parameters is generated randomly in the defined space. The next step is the calculation of the model output function under perturbed condition, f_p, and under non-perturbed condition, f_0. The index of robustness, R, is calculated over N perturbations. R is normally a negative number. It is zero when the system is absolutely robust. R can be -∞ indicating that the index is too high and the system is absolutely non-robust.

Table 6 presents the robustness analysis on outputs of both models A and B in steady state and dynamic conditions. Although the main output of both models is NifA, all other model reactants of the system were also assumed as output of the systems and a robustness analysis was performed for these too. In general model A is not robust in steady state conditions but it is robust in dynamic conditions. Whereas model B is robust under both

conditions. In respect to NifA, model B is robust in both conditions while model A is not robust in steady state conditions. The robustness of ATP and ADP are completely opposite in model A and B in steady state conditions.

Table 6: Robustness of model outputs by total parameters and variation of total concentrations. 'A', 'B', 'Steady', 'Integrated' represent model A, model B, steady state conditions, and dynamical conditions, respectively.

Model Output	Steady A	Steady B	Integrated A	Integrated B
$NifL_{red}$	-inf	-5.4171	-1.4936	-1.5679
ADP	-1.0395	-inf	-1.0568	-1.6779
$NifL_{red}$ADP	-inf	-70.69	-2.7987	-2.9403
NifA	-inf	-0.0051	-1.4713	-1.5545
$NifL_{red}$ADPNifA	-inf	-0.0875	-2.8551	-2.8435
ATP	-0.9828	-inf	-1.0022	-0.9471
$NifL_{red}$ATP	-inf	-48.29	-2.5312	-2.3123
$NifL_{red}$ATPNifA	-inf	-49.95	-2.6076	-2.4282

In summary, this study shows that both models A and B are robust against multi-parameter variation and also not sensitive to individual parameter variations in dynamic conditions. Robustness and sensitivity of model A and B are not the same in steady state conditions. NifA is more sensitive to individual variation of the parameters and total concentrations in model B in comparison to model A in steady state conditions, but model B is more robust than model A against multi-parameter variations in steady state conditions. Therefore, the ATP hydrolysis activity of NifA (reaction 4.6) changes a non-robust system into a robust system, although the model components are more sensitive to individual perturbation. It indicates that model B has flexibility to change its parameters and total concentration to some extent, and it can tolerate perturbations by spreading them into the system components.

4.2.3 *NifL and NifA interaction in oxic conditions*

The oxidized form of NifL, $NifL_{ox}$, is competent to inhibit the activity of NifA. Since the experimental data confirms that the oxidized form of NifL

affects the activity of NifA in vitro [67], we attempted to investigate this aspect of the NifL – NifA system. Although the oxidized form of NifL is active as an inhibitor of NifA in the absence of ADP, the presence of ADP increases its inhibitory activity. The system of reactions under oxidising conditions is as follows:

$$\mathrm{NifL} + \mathrm{ox} \underset{k_{d_7}}{\rightarrow} \mathrm{NifL_{ox}}, \tag{4.7}$$

$$\mathrm{NifL_{ox}} + \mathrm{ADP} \underset{k_{d_8}}{\rightleftharpoons} \mathrm{NifL_{ox}ADP}, \tag{4.8}$$

$$\mathrm{NifA} + \mathrm{NifL_{ox}ADP} \underset{k_{d_9}}{\rightleftharpoons} \mathrm{NifL_{ox}ADPNifA}. \tag{4.9}$$

Under oxidising conditions NifL senses the redox state and is converted to $\mathrm{NifL_{ox}}$ (reaction 4.7). Since NifL is converted to $\mathrm{NifL_{ox}}$ in an irreversible reaction, we can leave out reaction 4.7 and assume that the total amount of $\mathrm{NifL_{ox}}$ is NifL to decrease the number of parameters. $\mathrm{NifL_{ox}}$ interacts with ADP to form ADP-bond $\mathrm{NifL_{ox}}$, $\mathrm{NifL_{ox}ADP}$ (reaction 4.8). In the reaction 4.9, $\mathrm{NifL_{ox}ADP}$ interacts with NifA and forms the complex $\mathrm{NifL_{ox}ADPNifA}$.

Our first step to investigate the NifL – NifA interaction under oxidising conditions was to predict the unknown parameters. We used the same parameter values for the interaction of ADP with $\mathrm{NifL_{ox}}$ and $\mathrm{NifL_{red}}$ to reduce the numbers of unknown parameters. Since nitrogenase is highly sensitive to oxygen, reaction 4.9 is crucial to the NifL – NifA system, and it needs to be compared with the corresponding reaction (reaction 4.3) in the reduced state. We then applied the genetic algorithm to estimate the parameter of reaction 4.9 and compare the reduction with the oxidized state. This task was performed in the steady state conditions.

An experiment was performed to investigate the influence of ADP concentration on the inhibition of NifA activity by $\mathrm{NifL_{ox}}$ as described in [67]. Figure 26 presents the experimental data, which was used to estimate K_{d_9} (reaction 4.9). The value of K_{d_3}, which is for the same interaction in the reducted state, is 0.03 µM. One of the steady state points, through which the concentration of ADP leads to complete inhibition of NifA, activity was used to estimate K_{d_9}. The result of our parameter estimation shows that K_{d_9} is around 100 fold less than K_{d_3} (0.00026 µM). The objective function

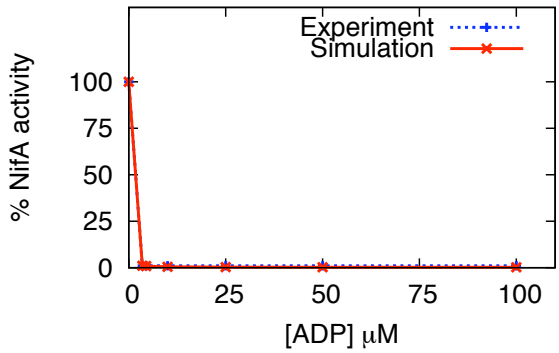


Figure 26: The activity of NifA based on the ADP concentration by oxidized NifL. The Reaction was performed with 125 nM of NifA, 250 nM of $NifL_{ox}$, and ADP as indicated in graph [67]. The activity of NifA is the ratio of $NifA_{free}$ to $NifA_{total}$.

value is 9 and the function evaluation is 4003 which are in accepted region. This investigation indicates that the $NifL_{ox}$ADPNifA complex is 100 fold tighter than that in the reducted state. It also implies that the formation of the complex is faster in oxic conditions than reduced conditions.

The estimated parameter was used to compute the influence of different concentrations of ADP on NifA by $NifL_{ox}$, the result of which is presented in Figure 26. The simulation of the NifA and NifL interaction under oxidising conditions is perfectly in agreement with the biological experiment.

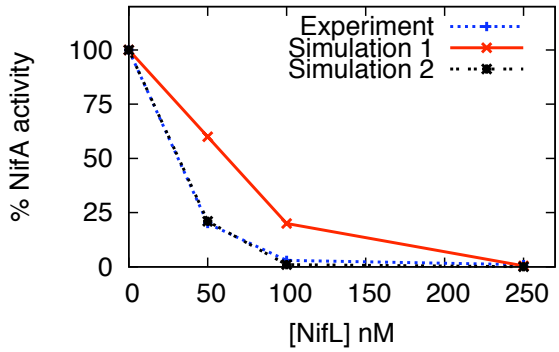


Figure 27: The activity of NifA based on variation of $NifL_{ox}$ concentration. The experiment was performed with 125 nM of NifA, 50 µM of ADP, and the indicated amount of $NifL_{ox}$ [67]. All NifA proteins are active in simulation 1(125 nM of NifA) while only 50% of NifA is active in simulation 2 (65 nM of NifA). The activity of NifA is division of $NifA_{free}$ to $NifA_{total}$.

To further investigate the system under oxidising conditions, we investigated the influence of $NifL_{ox}$ on NifA activity for a constant concentration of ADP. The analogous biological experiment was performed in [67]. The experimental data indicates that the activity of NifA is about 20% when the system contained 50 nM of oxidised NifL (Figure 27). Assuming that NifA protein is 100% active, this is logically impossible, since 50 nM of $NifL_{ox}$ should be able to interact with a maximum of 50 nM of NifA. In contrast the simulation gave dramatically different result when compare with the experiment. As explained above, it is impossible to get to that low activity of NifA with a low concentration of $NifL_{ox}$ simply because there is not enough $NifL_{ox}$ in the system to interact and consequently inhibit NifA.

The simulation suggests that the NifA protein may have not been 100% active. To examine this suggestion, we performed the same computation using 50% active NifA protein. Under these conditions the computation is in agreement with the experiment. This work shows that the activity of NifA could be as low as 50%, and furthermore, it confirms that K_{d_9} is about 100 fold less than K_{d_3}, which makes the complex of $NifL_{ox}$ADPNifA 100 times tighter than $NifL_{red}$ADPNifA.

4.3 THE 2-OXOGLUTARATE SYSTEM

$2-oxoglutarate$ releases NifA from inhibition by NifL under favorable conditions for nitrogen fixation. Since the main regulator in these conditions is $2-oxoglutarate$, we call it the $2-oxoglutarate$ system and in the following section we analyze this system in steady state conditions.

4.3.1 *A model of the 2-oxoglutarate system in steady state conditions*

The response of the $NifL-NifA$ system to $2-oxoglutarate$ has been determined by biochemical experiments [53, 60, 9]. Figure 29 presents the experimental results. The reaction scheme of the $2-oxoglutarate$ system is depicted in Figure 28.

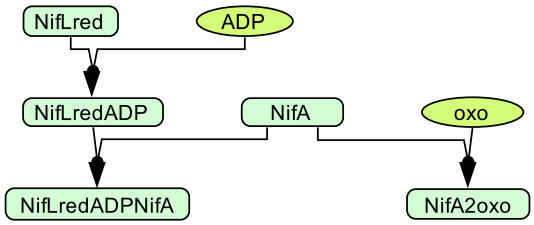


Figure 28: The reaction scheme of the $2-oxoglutarate$ system.

The reactions in the $2-oxoglutarate$ system are as follows:

$$NifL_{red} + ADP \underset{k_{d_1}}{\rightleftharpoons} NifL_{red}ADP, \tag{4.10}$$

$$NifL_{red}ADP + NifA \underset{k_{d_2}}{\rightleftharpoons} NifL_{red}ADPNifA, \tag{4.11}$$

$$NifA + 2OG \underset{k_{d_3}}{\rightleftharpoons} NifA(2OG). \tag{4.12}$$

The experiment was carried out with varying concentrations of $2-oxoglutarate$ to observe the influence of this ligand on the activity of $NifA$. We simulated the $2-oxoglutarate$ system using the same concentrations of $2-oxoglutarate$ as used in the experiment. Table 7 shows the $2-oxoglutarate$ system parameters and concentrations that were used in the experiment and simulations. The simulations are in agreement with the experimental data (Figure 29) and the influence of $2-oxoglutarate$ on the activity of the $NifL-NifA$ system is confirmed. Increasing $2-oxoglutarate$ raises the amount of free released from $NifL$ leading to increasing $NifA$ activity.

In summary, adding $2-oxoglutarate$ to the system increases the rate of reaction 4.12 towards the production of $NifA(2OG)$. There is less free $NifA$ available in the system to interact with $NifL_{red}ADP$, consequently, the rate of reaction 4.11 decreases by adding $2-oxoglutarate$, resulting in more $NifL_{red}ADP$ and less $NifL_{red}ADPNifA$ available in the system. Having more $NifL_{red}ADP$ in the system increases the rate of backward reaction of reaction 4.10. Therefore, adding $2-oxoglutarate$ to the system indirectly increases the availability of $NifL_{red}$ and ADP in the system (Figure 29).

Table 7: Concentrations (C/μM) and Parameters (K_d/μM) of the $2-oxoglutarate$ system related to reactions 4.10, 4.11, and 4.12.

System Component	Value
NifA	0.01
$NifL_{red}$	0.02
ADP	50
$2-oxoglutarate$	0-2000
K_{d_1}	13
K_{d_2}	0.03
K_{d_3}	57

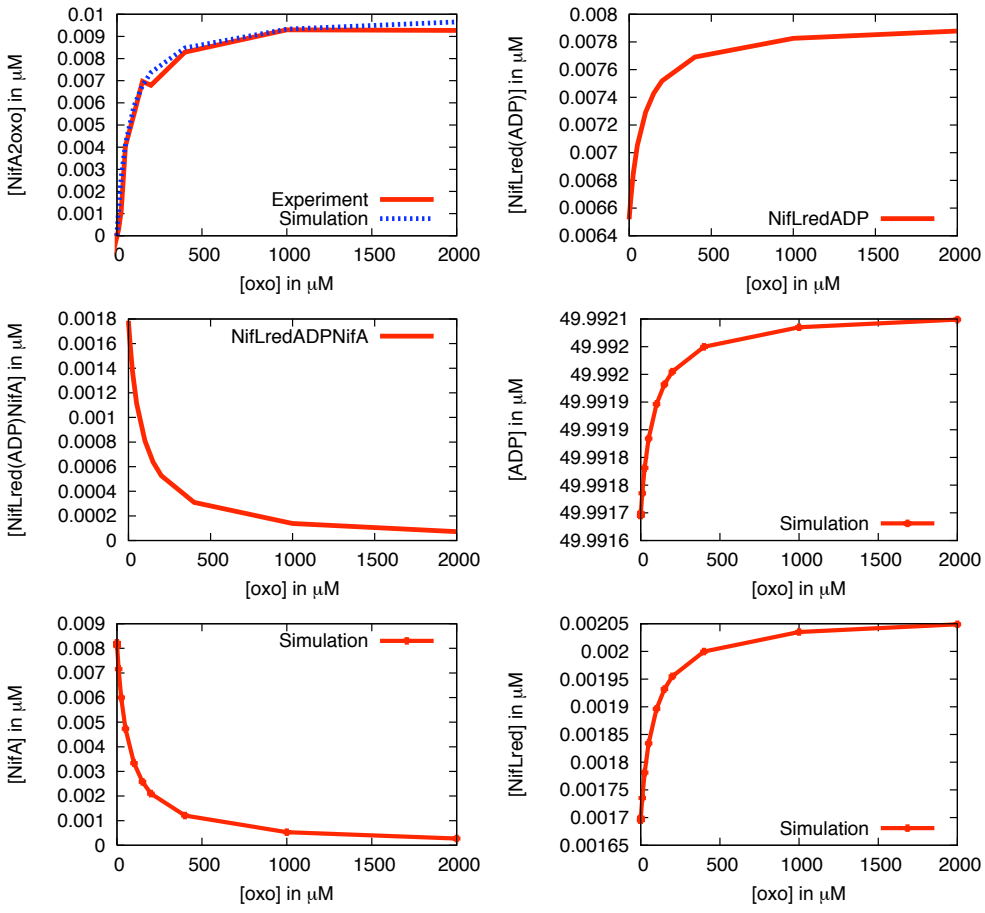


Figure 29: A simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, ADP, NifA(2OG), NifA, and $NifL_{red}$ based on changing the $2-oxoglutarate$ concentration ([2OG]).

4.3.2 *Investigation of the steady state points*

Although the simulation of the $2-oxoglutarate$ system is in agreement with the experimental data, we performed parameter estimation, stability analysis, and sensitivity analysis in order to understand the characteristics of the system. The $2-oxoglutarate$ system was modelled in the steady state condition using dissociation constants. As we have 14 different concentrations of $2-oxoglutarate$ and 14 corresponding steady state concentrations of $NifA(2OG)$ measured experimentally, we performed parameter estimation, forward modelling, stability analysis, and sensitivity analysis for all 14 steady state points using the Copasi and the Equilibrium program as described above.

Chosen Steady State Points

As a representative of the 14 simulations, we present the analysis of three of the steady state points. These three points were selected to monitor the $2-oxoglutarate$ system in the high, low, and medium concentrations of $2-oxoglutarate$. The full analysis of all steady state points is presented in Appendix B. Steady state points 4, 8, and 14 are those points that indicate low, medium, and high concentrations of $2-oxoglutarate$, respectively. We call them points A, B, and C, respectively. In steady state point A, B, and C the concentrations of $2-oxoglutarate$ are 0.5, 50, and 2000 μM, respectively.

Parameter Estimation

In each steady state point, as $NifA(2OG)$ is the main output of the $2-oxoglutarate$ system, the experimental concentration of $NifA(2OG)$ for that point was used to estimate the parameters. In each estimation task in this section one data point was used to estimate three parameters. The concentrations of $NifA(2OG)$ in points A, B, and C are 0.000118, 0.004061, and 0.0092 μM, respectively. Table 8 presents the results of parameter estimation in steady state points. The estimated parameters are close to the original ones for all three steady state points.

Table 8: Estimated Parameters of the $2-oxoglutarate$ system for all three steady state points in comparison to the original parameter values ($K_d/\mu M$).

K_d	Point A	Point B	Point C	Experiment
K_{d_1}	35	19	18	13
K_{d_2}	0.025	0.018	0.02	0.03
K_{d_3}	35	54	115	57

Forward Modelling

In the last section the parameters of the system were estimated in three different conditions. In this section we model the $2-oxoglutarate$ system based on 14 different concentrations of $2-oxoglutarate$ using the estimated parameters. Modelling was performed to investigate the influence of the estimated parameters, which were derived from the behavior of the system within broad range of $2-oxoglutarate$ concentrations. In this regard the concentration of all products of the system, $NifA(2OG)$, $NifL_{red}ADP$, and $NifL_{red}ADPNifA$, were monitored. The $NifA(2OG)$ concentration was compared to experimental data [53]. As the experimental data is not available for $NifL_{red}ADP$ and $NifL_{red}ADPNifA$, they were compared with the simulations with the original parameters.

Figure 30 indicates the result of modelling the $2-oxoglutarate$ system using the estimated parameters in steady state point A. Although the estimated parameters were derived from the conditions of low availability of $2-oxoglutarate$, the simulation of $NifA(2OG)$ is in agreement with the experimental data within the whole range of $2-oxoglutarate$. The simulations of the $NifL_{red}ADP$ using the estimated parameter and the original parameters are not exactly the same. This could be the result of a difference between the estimated parameter and original parameter of the corresponding reaction (K_{d_1}). The two simulations of $NifL_{red}ADPNifA$ are in agreement.

In the Figure 31, the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ are presented for steady state point B. All the simulations in steady state point B are in agreement with the original simulations. The results of the simulations in steady state point C are also close to one another (Figure 32). This investigation shows that the estimated param-

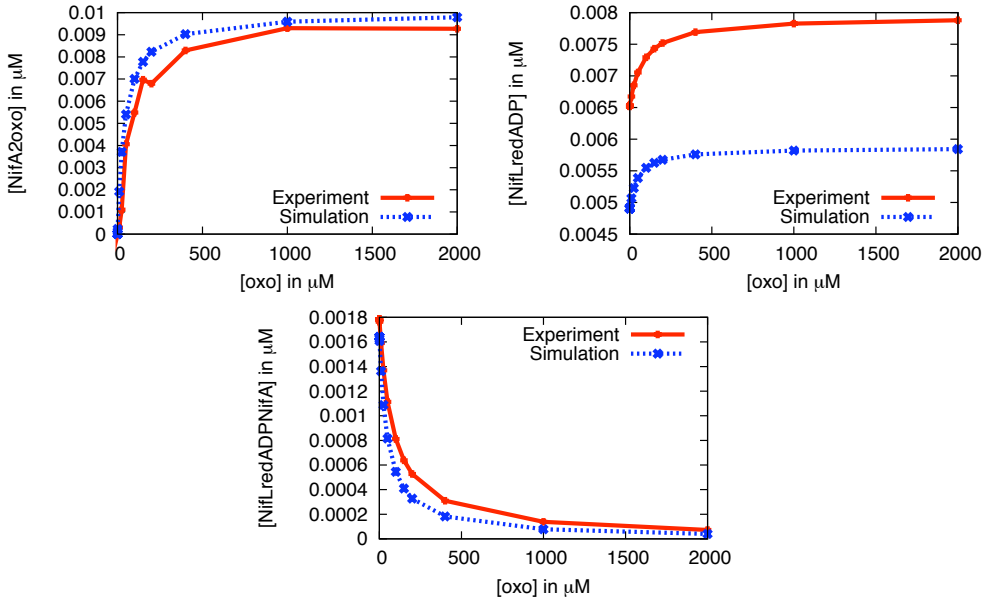


Figure 30: A simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ with the estimated parameters from steady state point A. $NifA(2OG)$ compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ compared to the simulation using the original parameters.

eters from all three conditions of low, high, and medium concentrations of $2-oxoglutarate$ are able to reproduce the expected behavior of the $2-oxoglutarate$ system in a broad range of $2-oxoglutarate$ concentrations.

Steady-state analysis

Stability analysis for steady state points A, B, and C was performed to monitor the effect of $2-oxoglutarate$ on stability and also to investigate if the system shows different characteristics of different concentration of $2-oxoglutarate$. The steady state concentrations of the $2-oxoglutarate$ system components were computed using the estimated parameters in each state. We performed this task to investigate the influence of $2-oxoglutarate$ on the concentration of the system components when the system reaches the steady state.

The results show that the $2-oxoglutarate$ system is generally a stable system and $2-oxoglutarate$ cannot change the stability. Table 9 presents the concentration of metabolites of the $2-oxoglutarate$ system in steady

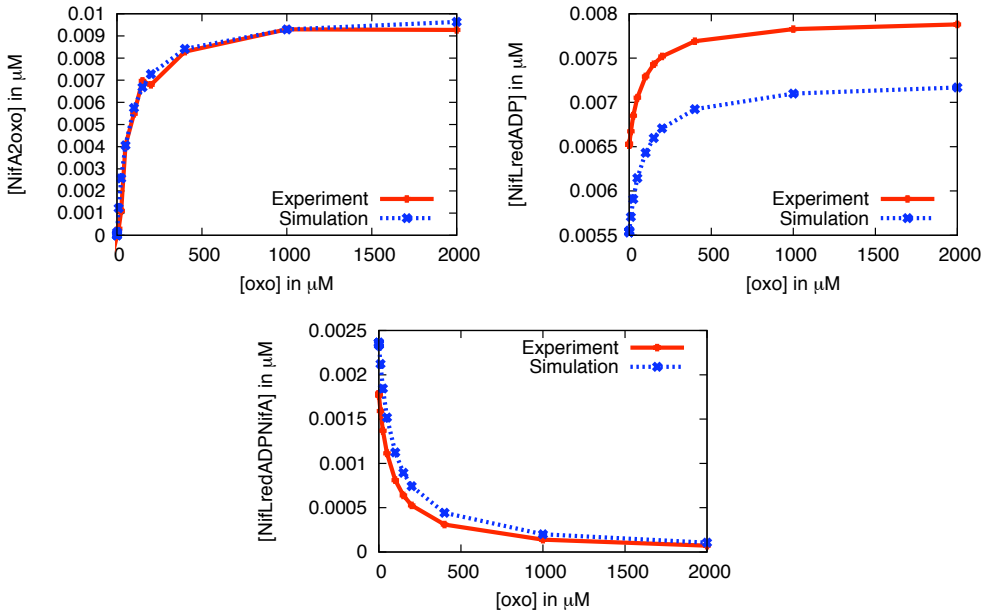


Figure 31: A simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and NifA(2OG) with the estimated parameters from steady state point B. NifA(2OG) is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation using the original parameters.

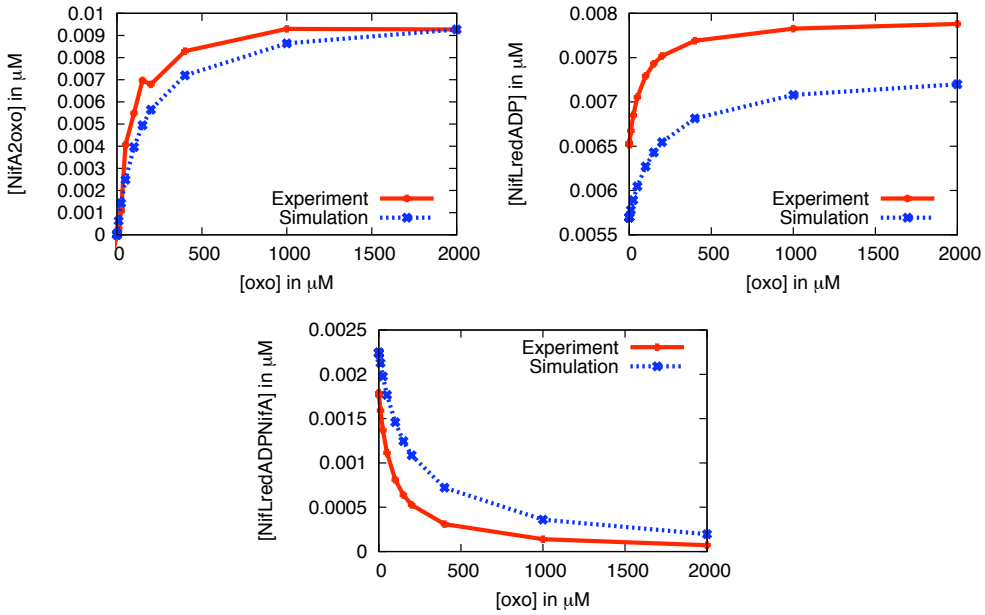


Figure 32: A simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and NifA(2OG) with the estimated parameters from steady state point C. NifA(2OG) is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

state points A, B, and C. In general, this result shows that $NifA(2OG)$, the main component of the $NifL - NifA$ system, continuously increases with added $2-oxoglutarate$. As expected adding $2-oxoglutarate$ dissociates $NifA$ from the $NifL - NifA$ complex, and so there is less $NifL - NifA$ complex available.

Table 9: Concentration values $(C/\mu M)$ of all the components of the $2-oxoglutarate$ system in steady state points A, B, and C.

metabolite	Point A	Point B	Point C
$NifL_{red}$	0.0035	0.0023	0.0026
ADP	49.99	49.99	49.99
$NifL_{red}ADP$	0.0049	0.0061	0.0072
$NifA$	0.0083	0.0044	0.00054
$NifL_{red}ADPNifA$	0.0016	0.0015	0.0002
$2-oxoglutarate$	0.49	49.99	1999.99
$NifA(2OG)$	0.00012	0.004	0.0092

Sensitivity analysis

We performed sensitivity analysis to explore the effect of parameter perturbation on the $2-oxoglutarate$ system and also to better understand the importance of the parameters in the $2-oxoglutarate$ system in conditions of high, low, and medium concentrations of $2-oxoglutarate$. The analysis was performed by sampling the parameters in a defined range, simulating the system with new parameters, and finally comparing the model output with new parameters to model output with the original parameters.

Figure 33 presents the result of the sensitivity analysis of K_{d_1}, K_{d_2}, and K_{d_3} for steady state points A, B, and C. In steady state point A, both of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are sensitive to K_{d_1} in a very similar manner. The highest perturbation of K_{d_1} leads point A to the maximum 20% change in $NifL_{red}ADP$. $NifL_{red}ADPNifA$ is sensitive to perturbation of K_{d_2} and it is more sensitive in decreasing K_{d_2} in comparison to its increasing. K_{d_2} perturbations increase the concentration of $NifL_{red}ADPNifA$ up to 250%. $NifA(2OG)$ is not so sensitive to perturbation of K_{d_1}. The interesting result is that the sensitivity of K_{d_2} on both $NifL_{red}ADP$ and $NifA(2OG)$ is the

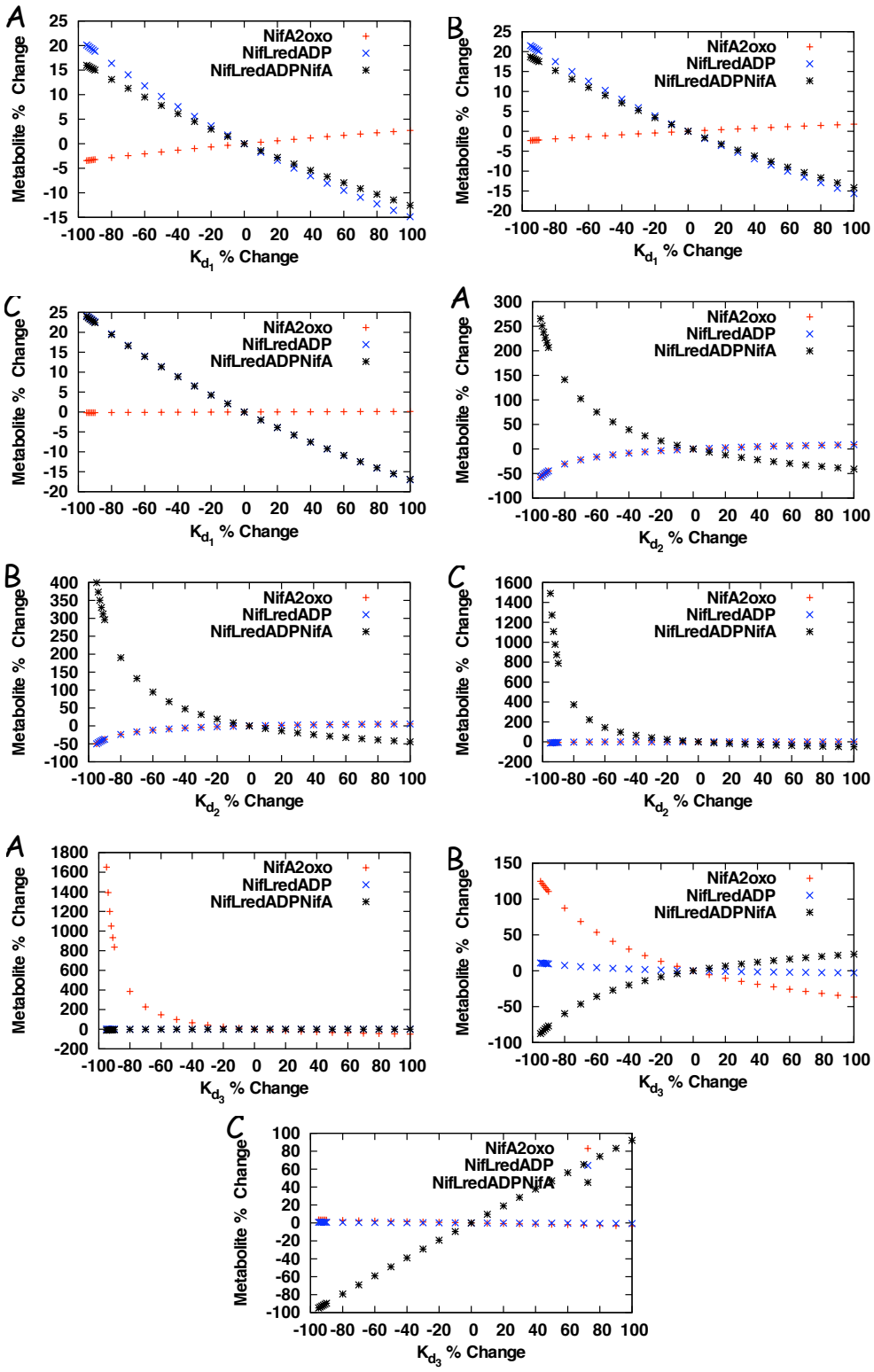


Figure 33: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point A, B, and C.

same in this case. K_{d_3} shows considerably higher sensitivity on productivity of $NifA(2OG)$, though perturbation of K_{d_3} does not induce any effect on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ in this state. 95% decreasing of K_{d_3} results in around 1700 % more $NifA(2OG)$ productivity.

In steady state point B, the sensitivity of K_{d_1} remains the same as steady state point A. The sensitivity of K_{d_2} to $NifA(2OG)$ and $NifL_{red}ADP$ as before indicates the same behavior. The sensitivity of K_{d_2} to $NifL_{red}ADPNifA$ is raised in this steady state point in comparison to steady state point A to a maximum of 400% changes. The sensitivity of K_{d_3} is very different in comparison to where the concentration of $2-oxoglutarate$ is low (steady state point A). The sensitivity of K_{d_3} decreases dramatically on $NifA(2OG)$ from a maximum 1700% to 140% changes in comparison to sensitivity of steady state point A. In contrast, the sensitivity of K_{d_3} increases depending on $NifL_{red}ADPNifA$ in comparison to the sensitivity of steady state point A. The sensitivity of K_{d_3} on $NifL_{red}ADP$ remains the same as steady state point A.

In steady state point C, K_{d_2} is highly sensitive to $NifL_{red}ADPNifA$ and K_{d_2} perturbation can change the $NifL_{red}ADPNifA$ up to maximum of 1500%. K_{d_2} has not shown any sensitivity on $NifA(2OG)$ and $NifL_{red}ADP$. In this state, there is no evidence for sensitivity of K_{d_3} to $NifA(2OG)$ and $NifL_{red}ADP$ while there is a high sensitivity of K_{d_3} to $NifL_{red}ADPNifA$. In this respect the system also presents very different behavior in comparison to steady state points A and B.

4.3.2.1 *Results and summary of the investigation of equilibrium points*

Stability Analysis

We performed stability analysis for all the steady state points with the estimated parameters. The result of this analysis showed that the $2-oxoglutarate$ system is a stable system for all of the states concerned.

Sensitivity Analysis

We performed the sensitivity analysis for all three parameters of the $2-oxoglutarate$ system in all steady state points. As the final step, we present the sensitivity of parameters based on the change of $2-oxoglutarate$ on the maximum increase and decrease of parameters. The maximum increase used

is 100% perturbation and maximum decrease used is about 95% perturbation in (Figure 34).

We found that adding $2-oxoglutarate$ changed the sensitivity of K_{d_1} approximately up to 400 μM of $2-oxoglutarate$. Above that concentration of $2-oxoglutarate$, the sensitivity of K_{d_1} remains unchanged for $NifA(2OG)$, $NifL_{red}ADP$, and $NifL_{red}ADPNifA$. The sensitivity of K_{d_2} shows the same behavior as for K_{d_1} up to the maximum decrease of the parameters, though the sensitivity of K_{d_2} on $NifL_{red}ADPNifA$ increases dramatically by adding $2-oxoglutarate$ in the maximum decrease of the parameters. However, it remains unchanged and insensitive to $NifA(2OG)$ and $NifL_{red}ADP$. Increasing the amount of $2-oxoglutarate$ in the system increases the sensitivity of K_{d_3} to $NifL_{red}ADPNifA$ and decreases the sensitivity of K_{d_3} to $NifA(2OG)$ for the maximum increasing perturbation. Also, in this condition the sensitivity of K_{d_3} on $NifA(2OG)$ was unchanged after 1000 μM of $2-oxoglutarate$. The sensitivity of K_{d_3} remains unchanged and insensitive to $NifL_{red}ADP$ for the maximum increasing perturbation. In the maximum decreasing perturbation, the sensitivity of K_{d_3} to $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ remains unchanged to adding $2-oxoglutarate$ and it is insensitive to $NifA(2OG)$ when the concentration of $2-oxoglutarate$ is around 10-50 μM.

In summary, K_{d_1} is very insensitive to $NifA(2OG)$ in all of the steady state points, insensitive to increasing the concentration of $2-oxoglutarate$, and also insensitive to $NifL_{red}ADP$ and $NifL_{red}ADPNifA$. The sensitivity of K_{d_1} to $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ is exactly the same for high concentrations of $2-oxoglutarate$(1000 μM). K_{d_2} does not show sensitivity to $NifA(2OG)$ and $NifL_{red}ADP$ causing less than 10% changes in low concentration of $2-oxoglutarate$, and it gets less sensitive to them by adding $2-oxoglutarate$. K_{d_2} is sensitive to $NifL_{red}ADPNifA$ and it gets more sensitive to it by adding $2-oxoglutarate$. It is the most sensitive parameter in high concentrations of $2-oxoglutarate$. K_{d_3} is very sensitive to $NifA(2OG)$ in low concentration of $2-oxoglutarate$ causing up to 1800% changes. It is insensitive to $NifA(2OG)$ in high concentration of $2-oxoglutarate$ and adding $2-oxoglutarate$ decrease the sensitivity of K_{d_3} to $NifA(2OG)$. K_{d_3} is insensitive to $NifL_{red}ADP$ in increasing $2-oxoglutarate$. K_{d_3} is insensitive to $NifL_{red}ADPNifA$ in low concentration of $2-oxoglutarate$ while it becomes more sensitive to $NifL_{red}ADPNifA$ by adding $2-oxoglutarate$.

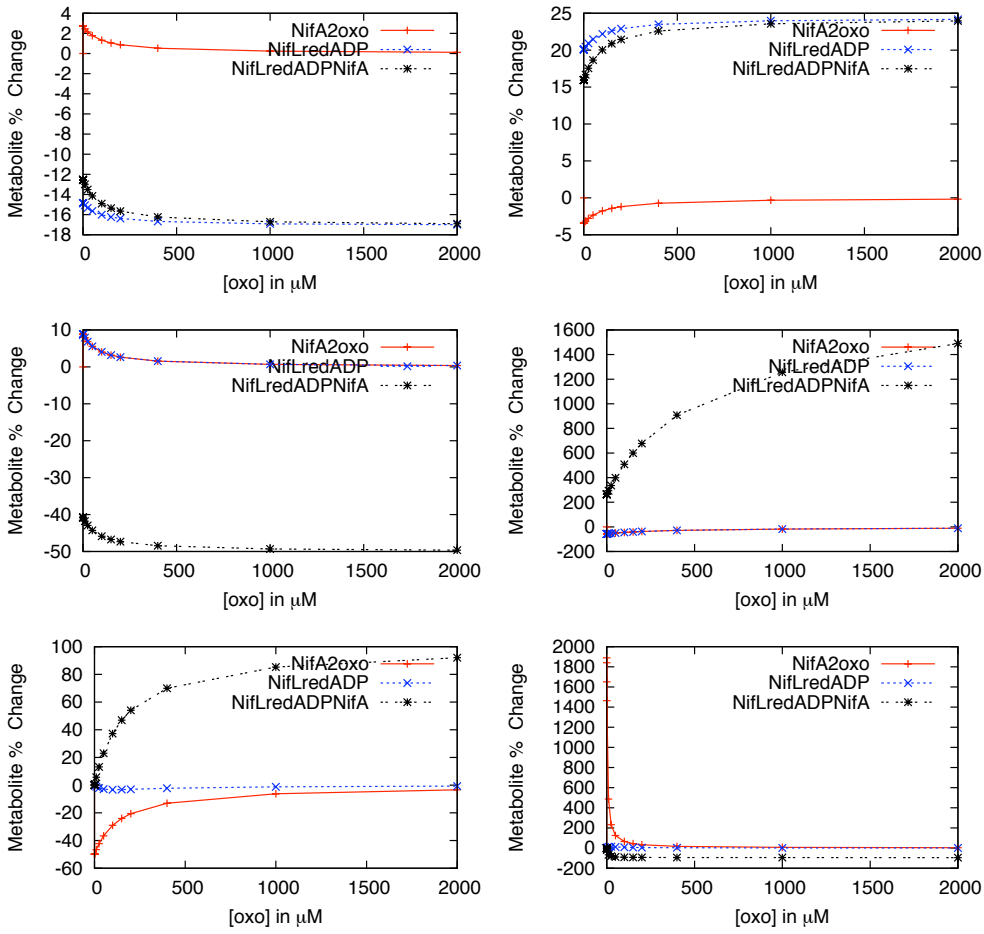


Figure 34: The sensitivity of parameters based on a change of $2-oxoglutarate$ for the maximum increase and decrease of parameters. The plots on the left are for 100% increase and on the right are for 95% decrease. The plots from top to bottom are the sensitivity of K_{d_1},K_{d_2}, and K_{d_3}, respectively.

$NifL_{red}ADPNifA$ is the most sensitive component of the $2-oxoglutarate$ system to perturbation of K_{d_3} in high concentration of $2-oxoglutarate$. This investigation shows that $2-oxoglutarate$ system is sensitive to its parameters in some conditions and reacts differently in different concentrations and parameters.

4.4 CONCLUSION

The complexity of the $NifL-NifA$ system in responding to different environmental signals demands breakage of the system. We first investigated the $NifL-NifA$ system by modelling $NifA$ and $NifL$ interaction in reduced and oxic conditions in the absence of $2-oxoglutarate$ and $GlnK$. The investigation of the $NifL-NifA$ system was then extended by including $2-oxoglutarate$. As $2-oxoglutarate$ plays an important role in the $NifL-NifA$ system we not only modelled that part of the system but we also performed a series of tests to provide more information. We now summarise our investigation of $NifA$ and $NifL$ interaction and the $2-oxoglutarate$ system by characterizing them.

$NifL$ *and* $NifA$ *interaction in reduced conditions with ADP*

- Modelling $NifL$ and $NifA$ interaction in reduced conditions confirms the influence of ADP in increasing inhibitory activity of $NifL$ on $NifA$.
- The ratio of $NifL$ to $NifA$ plays an important role in the inhibitory activity of $NifL$ in the presence of ADP. Although the expression of $NifL$ and $NifA$ appears to be transcriptionally coupled, the simulation shows that $NifL$ needs to be available in twice the amount of $NifA$ to perform its inhibitory activity.
- Not only is the ratio of these two proteins important but also the range of their concentrations. Although the biologists believe they may be $0.01\ \mu M$ in the cell, the simulation reveals that they must be around $0.1\ \mu M$ to explain the experimental data.

NifL *and* NifA *interaction in reduced conditions with ATP and ADP*

- The ATP hydrolyzing activity of NifA increases the robustness of the system against the multi-perturbation of parameters and total concentrations.
- In general NifL and its related products are among the most sensitive components of the system.
- NifA is among the least sensitive components of the system. As the main output of this system is NifA, it shows that the system is robust against perturbations.

NifL *and* NifA *interaction in oxic conditions*

- Modelling NifL and NifA interaction in oxic conditions confirms the inhibitory activity of NifL in these conditions.
- The result of this investigation showed that the $NifL_{ox}ADPNifA$ complex is 100 fold tighter than the same complex in the reduced state leading to faster formation of the complex in the oxic conditions in comparison to the reduced conditions.

The $2-oxoglutarate$ *system*

- Modelling the $2-oxoglutarate$ system confirms the role of $2-oxoglutarate$ in dissociation of the $NifL-NifA$ complex.
- Stability analysis shows that the $2-oxoglutarate$ system is stable in all three states of low, high, and medium concentrations of $2-oxoglutarate$.
- Within a low concentration of $2-oxoglutarate$, $NifA(2OG)$ is very sensitive to decreasing K_{d_3}. 95% decrease in the value of K_{d_3} leads to 1800% increase in the $NifA(2OG)$ concentration.
- A high concentrations of $2-oxoglutarate$, $NifA(2OG)$ is not sensitive to any parameters of the system.

In the next chapter we will continue our investigation by including $\mathrm{Gl}_n\mathrm{K}$.

5

THE GLNK SYSTEM

As mentioned in Chapter 1, GlnK is important in controlling the transcription of nif genes in response to the availability of fixed nitrogen. It acts as a key regulator in A. vinelandii to promote the formation of the NifL – NifA complex and to stop transcription of nif genes by sensing fixed nitrogen. Figure 35 presents the scheme for the GlnK system under conditions of fixed nitrogen sufficiency. In these conditions GlnK is mainly in the non-uridylylated form. Therefore, we modelled the system under these conditions assuming that there is no GlnD in the system to uridylylate GlnK. $NifL_{red}$ interacts with ADP to produce $NifL_{red}ADP$ (Figure 35 r1). $NifL_{red}ADP$ binds NifA to produce $NifL_{red}ADPNifA$, an inhibitory complex that prevents transcription of nif genes (r2). $2-oxoglutarate$ prohibits the formation of the NifL – NifA complex by interacting with NifA (r3). GlnK binds $2-oxoglutarate$ and this reaction produces GlnK2OG (r6), which binds $NifL_{red}ADP$ leading to the formation of $NifL_{red}ADPGlnK2OG$ (r4). This complex is a key component in the system as it can interact with NifA(2OG) (r5). This interaction keeps NifA in the $NifL_{red}ADPGlnK2OGNifA(2OG)$ complex and thus prohibits the transcription of nif genes.

The following are the proposed reactions for the GlnK system; the corresponding ODE model is presented in Table 10:

$$
\begin{aligned}
NifL_{red} + ADP &\underset{k_{d_1}}{\rightleftharpoons} NifL_{red}ADP,\\
NifL_{red}ADP + NifA &\underset{k_{d_2}}{\rightleftharpoons} NifL_{red}ADPNifA,\\
NifA + 2OG &\underset{k_{d_3}}{\rightleftharpoons} NifA(2OG),\\
NifL_{red}ADP + GlnK(2OG) &\underset{k_{d_4}}{\rightleftharpoons} NifL_{red}ADPGlnK(2OG),\\
NifL_{red}ADPGlnK(2OG) + NifA(2OG) &\underset{k_{d_5}}{\rightleftharpoons} NifL_{red}ADPGlnK(2OG)NifA(2OG),\\
GlnK + 2OG &\underset{k_{d_6}}{\rightleftharpoons} GlnK(2OG).
\end{aligned}
$$

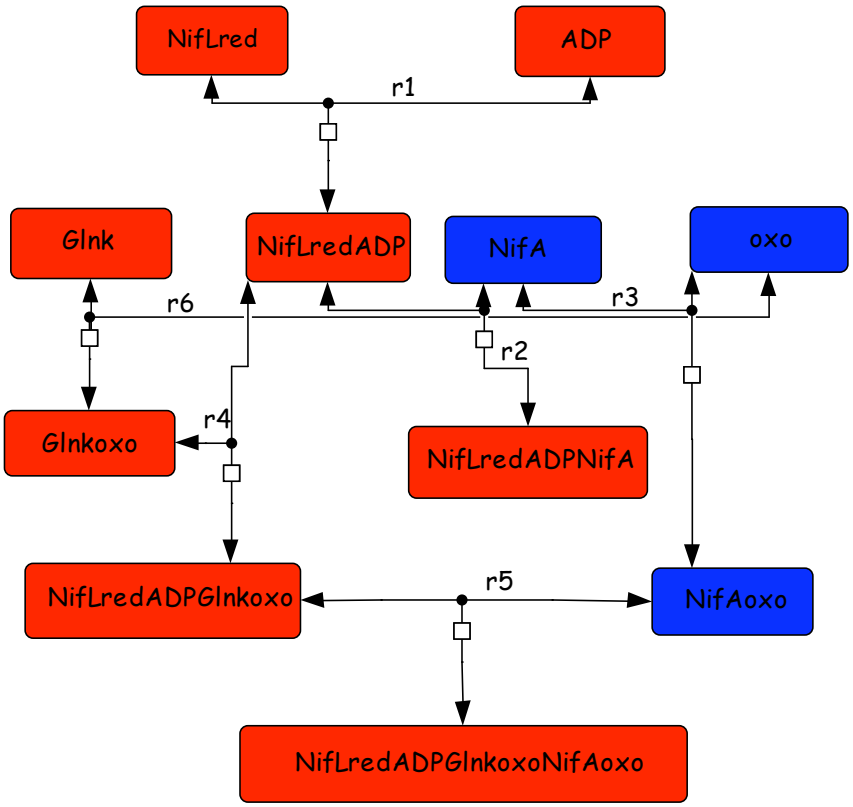


Figure 35: Scheme of the GlnK system. Dots in this figure represent the type of reactions (association and dissociation). The empty square represents the direction to the product of a reaction. Blue components are activators of the GlnK system and red ones are inhibitors. In general, $2-oxoglutarate$ acts as an activator of nitrogen fixation, while GlnK and ADP are inhibitors. $2-oxoglutarate$ acts through NifA, while, GlnK and ADP act through NifL.

Table 10: The ODEs for the $NifL - NifA$ system derived from the set of reactions in the text. The ODE system was generated using mass action equations. Substitutions: x_1=[$NifL_{red}$], x_2=[ADP], x_3=[$NifL_{red}$ADP], x_4=[NifA], x_5=[$NifL_{red}$ADPNifA], x_6=[2OG], x_7=[NifA(2OG)], x_8=[Glnk], x_9=[GlnK(2OG)], x_{11}=[$NifL_{red}$ADPGlnk(2OG)NifA(2OG)], and x_{10}=[$NifL_{red}$ADPGlnk(2OG)]; k_i and k_{-i} denote the forward and backward rates (i=1, 2, 3).

$$\dot{x}_1 = -k_1x_2x_1 + k_{-1}x_3$$
$$\dot{x}_2 = -k_1x_2x_1 + k_{-1}x_3$$
$$\dot{x}_3 = k_1x_2x_1 - k_{-1}x_3 - k_2x_3x_4 + k_{-2}x_5 - k_4x_9x_3 + k_{-4}x_{10}$$
$$\dot{x}_4 = -k_2x_3x_4 + k_{-2}x_5 - k_3x_4x_6 + k_{-3}x_7$$
$$\dot{x}_5 = k_2x_3x_4 - k_{-2}x_5$$
$$\dot{x}_6 = -k_3x_4x_6 + k_{-3}x_7 - k_6x_8x_6 + k_{-6}x_9$$
$$\dot{x}_7 = k_3x_4x_6 - k_{-3}x_7 - k_5x_{10}x_7 + k_{-5}x_{11}$$
$$\dot{x}_8 = -k_6x_8x_6 + k_{-6}x_9$$
$$\dot{x}_9 = -k_4x_9x_3 + k_{-4}x_{10} + k_6x_8x_6 - k_{-6}x_9$$
$$\dot{x}_{10} = k_4x_9x_3 - k_{-4}x_{10} - k_5x_{10}x_7 + k_{-5}x_{11}$$
$$\dot{x}_{11} = k_5x_{10}x_7 - k_{-5}x_{11}$$

Previous experimental data characterised the influence of the inhibitory activity of GlnK on $NifL - NifA$ in the presence of $2-oxoglutarate$ [53]. These experiments demonstrated inhibition of $NifA$ by $NifL_{red}$ in response to different concentrations of GlnK. The concentration of $2-oxoglutarate$ was relatively high in most of these experiments and the ratio of NifL to $NifA$ was 2:1 with concentrations of 0.2 µM and 0.1 µM, respectively. Experiments were performed in steady state conditions and confirmed that adding GlnK dramatically reduces the activity of $NifA(2OG)$ even in presence of $2-oxoglutarate$. Figure 36 presents the effect of GlnK in one of these experiments (Top graph "Experiment") [53].

Using the model in Table 10, the behavior of the GlnK system was simulated to measure its response to six different concentrations of GlnK in steady state conditions. We used the same approach and methodology as we used for the modelling the $2-oxoglutarate$ system in Chapter 4, Section 5.1.2. Table 11 presents the component concentrations and the dissociation constants used. Six simulations corresponding to 3 different concentrations of $NifL_{red}$ and 2 different concentrations of $2-oxoglutarate$ were carried

Table 11: Concentration (C/μM) and parameter (K_d/μM) relevant to the reactions in text and the model in Table 10.

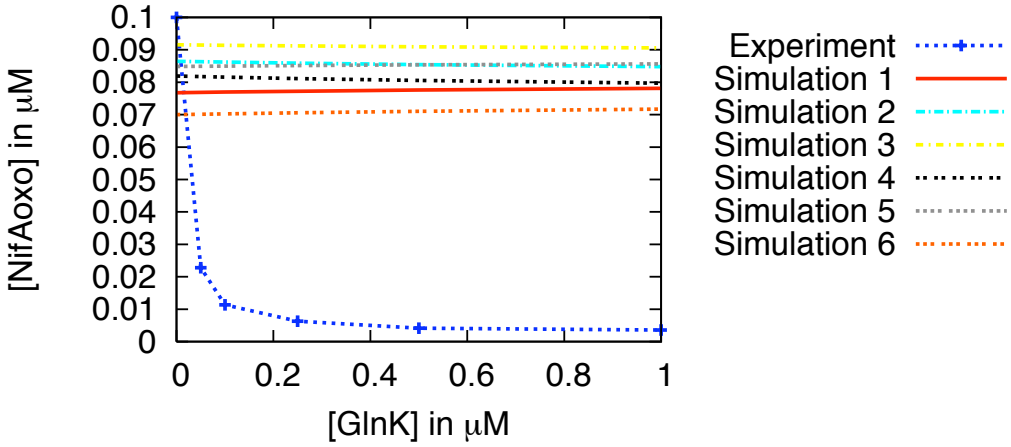
System Component	Value	Source
NifA	0.1	[53]
$NifL_{red}$	0.1,0.2,0.3	[53]
ADP	50	[53]
$2-oxoglutarate$	1000,2000	[53]
GlnK	0.0,0.05,0.1,0.25,0.5,1	[53]
K_{d_1}	20	Provided by biologist
K_{d_2}	0.03	Provided by biologist
K_{d_3}	57	Provided by biologist
K_{d_4}	1	Provided by biologist
K_{d_5}	1	Provided by biologist
K_{d_6}	60	Provided by biologist

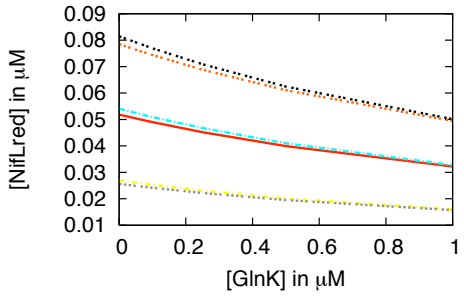
out to investigate the response to GlnK with different ratios of NifL and NifA in addition to two different concentrations of $2-oxoglutarate$.

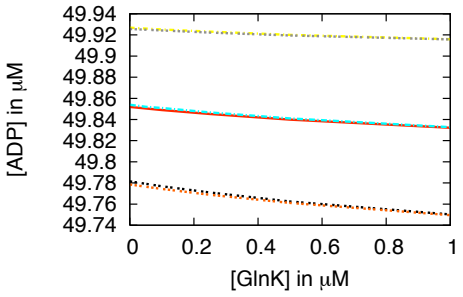
In the first simulation, the same total concentrations of NifA and NifL were used as in the experiment. The ratio of NifL to NifA was 1:2 and the concentration of $2-oxoglutarate$ was 1000 μM. The main output of the system is NifA(2OG), the positive activator of *nif* genes, which we used to monitor the activity of the system. A high concentration of NifA(2OG) represents strong activities of *nif* genes. The total concentration of NifA was 0.1 μM. Thus, in the 100% active system all the available NifA is converted to NifA(2OG) with a concentration of 0.1 μM.

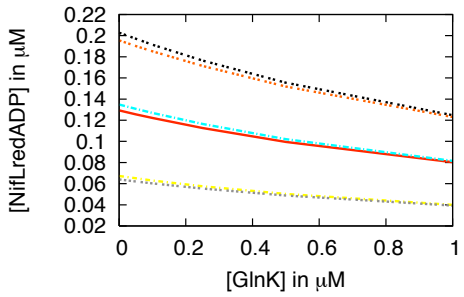
The computational simulation does not show any changes in the concentration of NifA(2OG) when raising the amount of GlnK. Whereas the experiment showed that a very low amount of GlnK reduces the activity of the system dramatically, the computational simulation does not behave like this even with high concentrations of GlnK (Figure 36 top graph "Experiment" and "Simulation 1"). The computational system has the advantage that the concentrations of all the components can be simulated (Figure 36).

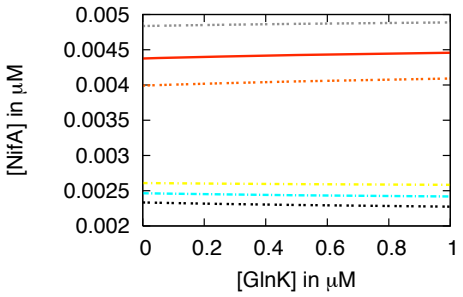
Raising the GlnK concentration does not change the behavior of free NifA, which remains very low at an amount of 4%. The total free NifA and NifA(2OG) concentration constitutes about 83% of NifA in the sys-

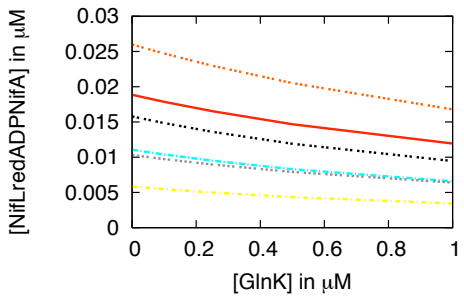

[NifAoxo] in μM
[GlnK] in μM
Experiment
Simulation 1
Simulation 2
Simulation 3
Simulation 4
Simulation 5
Simulation 6

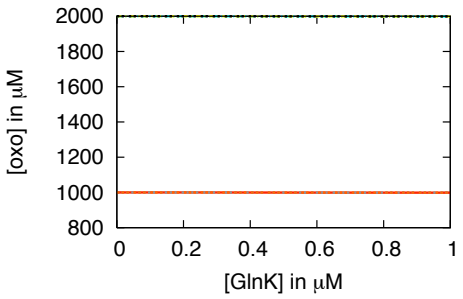

[NifLred] in μM
[GlnK] in μM


[ADP] in μM
[GlnK] in μM


[NifLredADP] in μM
[GlnK] in μM


[NifA] in μM
[GlnK] in μM


[NifLredADPNifA] in μM
[GlnK] in μM


[oxo] in μM
[GlnK] in μM

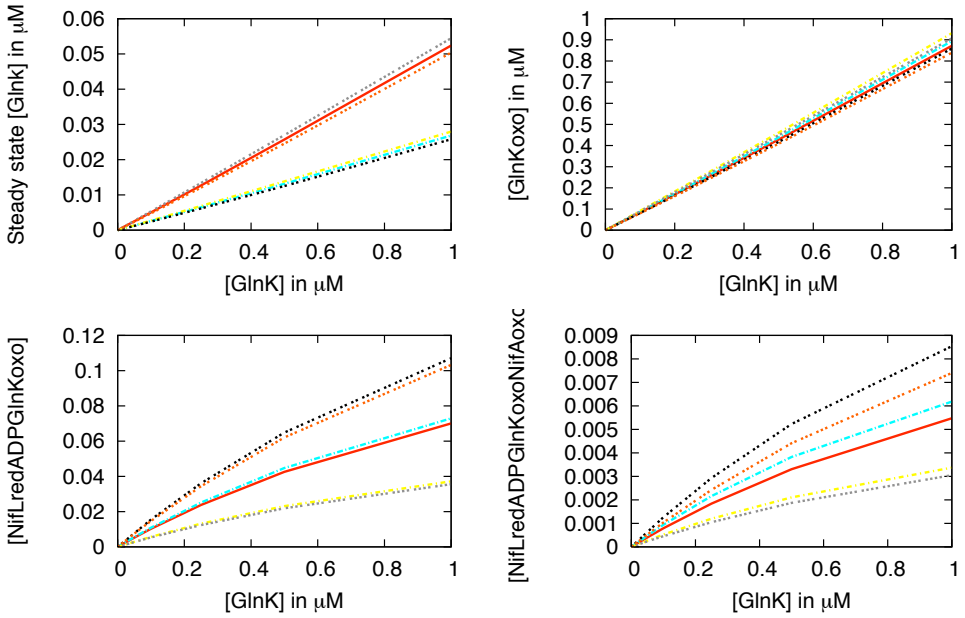


Figure 36: Computational simulation of the GlnK system based on increasing amounts of GlnK. The dissociation constants for all the simulations are as indicated in Table 11. The total concentrations in both experimental data and simulations are the same as shown in Table 11 except for $NifL_{red}$ and $2-oxoglutarate$. Concentrations of these were as follow; Experiment: 0.2 μM $NifL_{red}$ and 1000 μM $2-oxoglutarate$; Simulation 1: 0.2 μM $NifL_{red}$ and 1000 μM $2-oxoglutarate$; Simulation 2: 0.2 μM $NifL_{red}$ and 2000 μM $2-oxoglutarate$; Simulation 3: 2000 μM $2-oxoglutarate$ and 0.1 μM $NifL_{red}$; Simulation 4: 2000 μM $2-oxoglutarate$ and 0.3 μM $NifL_{red}$; Simulation 5: 0.1 μM $NifL_{red}$ and 1000 μM $2-oxoglutarate$; Simulation 6: 0.3 μM $NifL_{red}$ and 1000 μM $2-oxoglutarate$. The first graph presents the activity of the system and shows the influence of GlnK on $NifA(2OG)$ in comparison to the experimental data. Other graphs present the generated data for all other components of the system in the simulation.

tem with the remaining NifA is incorporated in $NifL_{red}ADPNifA$ and $NifL_{red}ADPGlnK2OGNifA(2OG)$.

As was observed in the 2 − oxoglutarate system, different amounts of NifL and NifA and their ratio can change the simulation. We simulated the system with different ratios of NifA and NifL and when two concentrations of 2 − oxoglutarate to observe whether these changes could deliver the experimental data. Simulation 6, in which NifL and 2 − oxoglutarate are 0.3 µM and 1000 µM, respectively, shows the lowest concentration of NifA(2OG) among all of the simulations. Simulation 3, in which NifL and 2 − oxoglutarate are 0.1 µM and 2000 µM, respectively, gave the highest activity. Both computational simulations confirm that a higher ratio of NifL to NifA combined with concentration a lower of 2 − oxoglutarate produce lower activity of the system although none of the simulations show any response to GlnK.

Adding GlnK to the system decreases the amount of $NifL_{red}$. This decrease in $NifL_{red}$ is expected because GlnK2OG interacts with $NifL_{red}ADP$ and produces the complex of $NifL_{red}ADPGlnK2OG$. Consequently, the concentration of the components $NifL_{red}$, $NifL_{red}ADP$, $NifL_{red}ADPNifA$, and ADP decreases when they are sequestered to GlnK in the system. Although GlnK does not show any effect on free NifA, as expected, the simulations with 2000 µM 2 − oxoglutarate have a lower level of free NifA in comparison to the simulations with 1000 µM 2 − oxoglutarate. As anticipated, adding GlnK to the system increases both $NifL_{red}ADPGlnK2OG$ and $NifL_{red}ADPGlnK2OGNifA(2OG)$. The highest increase is in simulation 4 in which NifL is 0.3 µM and the lowest increase is in simulation 5 in which NifL is 0.1 µM, as expected. Considering the effect of GlnK on $NifL_{red}ADPGlnK2OGNifA(2OG)$, the main concern is that increasing GlnK does not change the concentration of the complex as much as expected. Consequently a low amount of NifA remains in $NifL_{red}ADPGlnK2OGNifA(2OG)$ which causes the system to remain in the active mode.

Since the simulation does not fit the experimental data, we perturbed the total concentrations of the key system components in an attempt to obtain closer fit, but this strategy also failed. There are two possible reasons for this. The first might be because of experimental errors in measuring the parameters of the system. The second reason might be the lack of knowledge

about the reactions in the experimental system itself. In this chapter we investigate the effect of GlnK on NifL – NifA system. We will analyse the GlnK system to improve our model and show the possibilities which lead the model to agree with experimental data. For example, the current set of reactions and consequently the ODEs system might be incomplete. There could be a different set of reactions, which can produce quantitatively the same behavior as the experiment with the given parameters. In the next section we explore the GlnK system based on these possibilities. We first focus on estimating system parameters to fit the output with the experimental data. Then we explore various sets of reactions with given parameters to fit to the experimental data.

5.0.1 *Analysis of possible GlnK models*

The process of model exploration was started by forming new GlnK models using additional interactions, which are biologically plausible. 41 different sets were chosen from a larger set of models on the basis of biological knowledge and the literature reviewed in Chapter 1. The generation of potential models was started from the original model. The reactions categorized based on the match to original models and the reactions were switched by the alternative ones to find the minimal adequate reaction which fulfill the experimental observations. Having formed the new system our next step was parameter estimation. In Figure 36, the experimental data shows that, even at low concentrations of GlnK, the system can be inactivated by decreasing the concentration of NifA(2OG) to less than 0.01 μM. We used a genetic algorithm for parameter estimation as described in Chapter 4, section 5.1.3. We then performed parameter estimation for all generated models to check which one is capable of producing the expected behavior with estimated parameters that are close to the original ones.

Furthermore, we performed stability analysis for all 41 systems. We carried out stability analysis using the original parameters and the estimated parameters for all the models. The combination of parameter estimation and stability analysis helped to narrow down the possibilities. The following pages present all the reactions, which were used for all the models. In the

following reactions, D, NA, NL, G, and r represent ADP, NifA, NifL, GlnK, and red (reduction state), respectively.

$$\mathrm{NL_r} + \mathrm{D} \rightleftharpoons \mathrm{NL_rD}, \tag{5.1}$$
$$\mathrm{NL_rD} + \mathrm{NA} \rightleftharpoons \mathrm{NL_rDNA}, \tag{5.2}$$
$$\mathrm{NA} + 2\mathrm{OG} \rightleftharpoons \mathrm{NA}(2\mathrm{OG}), \tag{5.3}$$
$$\mathrm{NL_rD} + \mathrm{G} \rightleftharpoons \mathrm{NL_rDG}, \tag{5.4}$$
$$\mathrm{NL_rDG} + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDGNA}(2\mathrm{OG}), \tag{5.5}$$
$$\mathrm{NL_rDG} + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDGNA} + (2\mathrm{OG}), \tag{5.6}$$
$$\mathrm{NL_rD} + \mathrm{G}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG}), \tag{5.7}$$
$$\mathrm{NL_rDG}(2\mathrm{OG}) + \mathrm{NA}(2\mathrm{OG}) + 2\mathrm{OG} \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})\mathrm{NA}(2\mathrm{OG})_2, \tag{5.8}$$
$$\mathrm{G} + 2\mathrm{OG} \rightleftharpoons \mathrm{G}(2\mathrm{OG}), \tag{5.9}$$
$$\mathrm{NL_rDG}(2\mathrm{OG}) + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})\mathrm{NA}(2\mathrm{OG}), \tag{5.10}$$
$$\mathrm{NL_rD} + \mathrm{NL_rDG}(2\mathrm{OG})\mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDNL_rDG}(2\mathrm{OG})\mathrm{NA}(2\mathrm{OG}), \tag{5.11}$$
$$\mathrm{NL_rD} + \mathrm{G}(2\mathrm{OG})_2 \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})_2, \tag{5.12}$$
$$\mathrm{NL_rDG}(2\mathrm{OG})_2 + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})_2\mathrm{NA}(2\mathrm{OG}), \tag{5.13}$$
$$\mathrm{G} + 2\ (2\mathrm{OG}) \rightleftharpoons \mathrm{G}(2\mathrm{OG})_2, \tag{5.14}$$
$$\mathrm{NL_rD} + \mathrm{NL_rDG}(2\mathrm{OG})_2\mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{NL_rDNL_rDG}(2\mathrm{OG})_2\mathrm{NA}(2\mathrm{OG}), \tag{5.15}$$
$$\mathrm{NL_rDG}(2\mathrm{OG})_2 + \mathrm{NA}(2\mathrm{OG}) + 2\mathrm{OG} \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})_2\mathrm{NA}(2\mathrm{OG})_2, \tag{5.16}$$
$$\mathrm{NL_rDG} + \mathrm{NA} \rightleftharpoons \mathrm{NL_rDGNA}, \tag{5.17}$$
$$\mathrm{NL_rDG}(2\mathrm{OG}) + \mathrm{NA} \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})\mathrm{NA}, \tag{5.18}$$
$$\mathrm{NL_rDG}(2\mathrm{OG})_2 + \mathrm{NA} \rightleftharpoons \mathrm{NL_rDG}(2\mathrm{OG})_2\mathrm{NA}, \tag{5.19}$$
$$\mathrm{G} + \mathrm{NA} \rightleftharpoons \mathrm{GNA}, \tag{5.20}$$
$$\mathrm{G}(2\mathrm{OG}) + \mathrm{NA} \rightleftharpoons \mathrm{NAG}(2\mathrm{OG}), \tag{5.21}$$
$$\mathrm{NAG}(2\mathrm{OG}) + \mathrm{NLrDG} \rightleftharpoons \mathrm{NAG}(2\mathrm{OG})\mathrm{NLrDG}, \tag{5.22}$$
$$\mathrm{G}(2\mathrm{OG})_2 + \mathrm{NA} \rightleftharpoons \mathrm{NAG}(2\mathrm{OG})_2, \tag{5.23}$$
$$\mathrm{NAG}(2\mathrm{OG})_2 + \mathrm{NLrDG} \rightleftharpoons \mathrm{NAG}(2\mathrm{OG})_2\mathrm{NLrDG}, \tag{5.24}$$
$$\mathrm{G} + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{GNA}(2\mathrm{OG}), \tag{5.25}$$
$$\mathrm{G}(2\mathrm{OG}) + \mathrm{NA}(2\mathrm{OG}) \rightleftharpoons \mathrm{G}(2\mathrm{OG})\mathrm{NA}(2\mathrm{OG}), \tag{5.26}$$

$$G(2OG)NA(2OG) + NL_rDG(2OG) \rightleftharpoons NL_rDG(2OG)G(2OG)NA(2OG), \quad (5.27)$$

$$GNA(2OG) + NL_rDG \rightleftharpoons GNA(2OG)NL_rDG, \quad (5.28)$$

$$G(2OG)_2 + NA(2OG) \rightleftharpoons G(2OG)_2NA(2OG), \quad (5.29)$$

$$NL_rDG(2OG)_2 + G(2OG)_2NA(2OG) \rightleftharpoons NL_rDG(2OG)_2G(2OG)_2NA(2OG), \quad (5.30)$$

$$G(2OG)_2NA(2OG) + NLrD \rightleftharpoons G(2OG)_2NA(2OG)NLrD, \quad (5.31)$$

$$G(2OG)NA(2OG) + NLrD \rightleftharpoons G(2OG)NA(2OG)NLrD, \quad (5.32)$$

$$NA + 4\ 2OG \rightleftharpoons NA(2OG)_4, \quad (5.33)$$

$$NL_rDG(2OG)_2 + NA(2OG)_4 \rightleftharpoons NL_rDG(2OG)_2NA(2OG)_4, \quad (5.34)$$

$$NL_rD + 2\ G \rightleftharpoons NL_rD(G)_2, \quad (5.35)$$

$$NL_rD(G)_2 + NA(2OG) \rightleftharpoons NL_rD(G)_2NA(2OG), \quad (5.36)$$

$$NL_rD + 2\ G(2OG) \rightleftharpoons NL_rD(G(2OG))_2, \quad (5.37)$$

$$NL_rD(G(2OG))_2 + NA(2OG) \rightleftharpoons NL_rD(G(2OG))_2NA(2OG), \quad (5.38)$$

$$NL_rD + 2\ G(2OG)_2 \rightleftharpoons NL_rD(G(2OG)_2)_2, \quad (5.39)$$

$$NL_rD(G(2OG)2)_2 + NA(2OG) \rightleftharpoons NL_rD(G(2OG)_2)_2NA(2OG), \quad (5.40)$$

$$NL_r + D + 2OG \rightleftharpoons NL_rD(2OG), \quad (5.41)$$

$$NL_rD + NA + 2OG \rightleftharpoons NL_rDNA(2OG), \quad (5.42)$$

$$NA + 2\ (2OG) \rightleftharpoons NA(2OG)_2, \quad (5.43)$$

$$NL_rD + G + 2OG \rightleftharpoons NL_rDG(2OG), \quad (5.44)$$

$$NL_rD + G \rightleftharpoons NL_rG + D, \quad (5.45)$$

$$NL_rD + G(2OG) \rightleftharpoons NL_rG(2OG) + D, \quad (5.46)$$

$$NL_rD(2OG) + NA \rightleftharpoons NL_rD(2OG)NA, \quad (5.47)$$

$$NL_rD(2OG) + G \rightleftharpoons NL_rD(2OG)G, \quad (5.48)$$

$$NL_rD(2OG)G + NA(2OG) \rightleftharpoons NL_rD(2OG)GNA(2OG), \quad (5.49)$$

$$NL_rDG + NA(2OG)_2 \rightleftharpoons NL_rDGNA(2OG)_2, \quad (5.50)$$

$$NL_rG + NA(2OG) \rightleftharpoons NL_rGNA(2OG), \quad (5.51)$$

$$NL_rG(2OG) + NA(2OG) \rightleftharpoons NL_rG(2OG)NA(2OG). \quad (5.52)$$

Based on the results of the stability analysis and the parameter estimations, we gave the models fail or pass labels. The models with one estimated parameter outside the boundary by 10 fold less or more than the original value failed in the parameter estimation test, otherwise it passed. We carried

out the stability analysis for all the models with estimated parameters and original parameters. The models, which were either unstable or had no steady state, were categorized as a fail in the stability test.

Table 12 presents a summary of the result of the model exploration with parameter estimation and stability analysis. We have categorized the models in three groups A, B, and C. Group A contains the models 3, 25, 28, and 37, which passed both parameter estimation and stability analysis tests. Group B contains the models that failed in the parameter estimation and passed in the stability analysis. This group includes the models 1, 8, 10, 13, 16, 19, 22, 23, 24, 25, 31, 33, 34, 36, 39, 40, and 41. Group C contains the models that failed in both the parameter estimation and stability analysis tests. It includes the models 2, 4, 5, 6, 7, 9, 11, 12, 14, 15, 17, 18, 20, 21, 26, 27, 29, 30, 32, 35, and 38.

Table 12: Summary of model exploration for the GlnK system. Columns M, R, P_o, P_e, and S_o-S_e indicate the model number, reaction numbers of system, original parameters, estimated parameters, and finally the result of stability analysis with original parameters and the result of the stability analysis with the estimated parameters, respectively. The reaction numbers represent the second digit of reaction numbers in text. We have not included Reactions 1, 2, and 3 in the table as they are common to all models, except models 32,36, 37, and 38. Model 32 does not have reaction 3. Model 36 does not have reaction 1 and 2. Model 37 does not have reaction 2, and model 38 does not have reaction 3. Reaction numbers, original parameters, and estimated parameters are in a chronological manner. In the stability analysis column, S, U_s, and U_d denote stable, unstable, and undetermined. The result of the stability analysis is undetermined in case no steady state was found.

M	R	P_o	P_e	S_o-S_e
1	4,5	20,0.03,57,1,1	8,0.0005,1037,0.0006,1.3	S,S
2	4,6	20,0.03,57,1,1	6,0.001,1744,1.5,1.3	S,U_s
3	7,8,9	20,0.03,57,1,1,60	27,0.03,56,2.1,2.3,41	S,S
4	7,9,10	20,0.03,57,1,60,1	2,0.008,51,0.01,1.5e+006,0.3	S,U_d
5	7,9,10,11	20,0.03,57,1,60,1,1	3,0.02,15,0.002,1.7e+006,0.02,1	S,U_d
6	12,13,14,15	20,0.03,57,1,1,60,1	12,0.03,43,0.7,1,50,0.5	U_s,U_s
7	12,16,14	20,0.03,57,1,1,60	8,0.04,229,1,1,84	U_d,U_d
8	7,9,10	20,0.03,57,1,60,1	9,0.001,2354,1.8,45,0.03	S,S

9	12,13,14	20,0.03,57,1,1,60	14,0.005,162,0.5,0.2,51	U_d,U_d
10	4,17	20,0.03,57,1,1	12,0.0006,1257,2.3,1.5	S,S
11	7,9,18	20,0.03,57,1,60,1	14,0.0007,1353,2.5,57,3.5	S,S
12	12,14,19	20,0.03,57,1,60,1	U_d	U_d,U_d
13	4,5,17	20,0.03,57,1,1,1	6,0.0004,970,1,1,0.2	S,S
14	7,9,10,18	20,0.03,57,1,60,1,1	6,0.0009,1440,4.3,27,0.2,2.6	S,U_d
15	12,13,14,19	20,0.03,57,1,1,60,1	U_d	U_d,U_d
16	4,5,20	20,0.03,57,1,1,1	9,0.001,1777,5,1,0.6	S,S
17	7,9,10,20	20,0.03,57,1,60,1,1	7,0.0006,1107,2.4,58,0.1,1.5	S,U_d
18	12,13,14,20	20,0.03,57,1,1,60,1	11,0.001,3069,1.5,0.2,39,0.4	S,U_s
19	4,5,9,21	20,0.03,57,1,1,60,1	1.6,0.001,1351,0.5,0.5,66,0.2	S,S
20	4,5,9,21,22	20,0.03,57,1,1,60,1,1	8,0.0007,1026,2.4,1.2,21,0.8,0.8	S,U_s
21	4,5,14,23,24	20,0.03,57,1,1,60,1,1	6,0.0005,1480,0.3,1.7,64,0.5,2.7	U_d,U_d
22	4,5,25	20,0.03,57,1,1,1	90,0.01,32,0.9,0.5,0.009	S,S
23	7,9,10,26	20,0.03,57,1,60,1,1	18,0.02,262,1.2,10,0.4,0.009	S,S
24	7,9,10,26,27	20,0.03,57,1,60,1,1,1	15,0.02,18,1,52,0.4,0.02,0.07	S,S
25	4,5,25,28	20,0.03,57,1,1,1,1	16,0.03,29,2.4,0.6,0.01,0.4	S,S
26	12,13,14,29	20,0.03,57,1,1,60,1	27,0.04,75,0.6,0.7,55,0.4	U_s,U_s
27	12,13,14,29,30	20,0.03,57,1,1,60,1,1	17,0.02,56,0.6,0.7,60,0.01,0.4	U_s,U_s
28	4,5,9,26	20,0.03,57,1,1,60,1	17,0.01,121,1.4,1.3,45,0.01	S,S
29	4,5,14,29,31	20,0.03,57,1,1,60,1,1	9,0.03,73,0.8,0.6,98,0.01,0.4	U_d,U_d
30	4,5,14,29	20,0.03,57,1,1,60,1	20,0.003,68,1.2,0.9,127,0.009	S,U_d
31	4,5,9,26,32	20,0.03,57,1,1,60,1,1	20,0.02,30,0.9,0.7,28,0.04,0.05	S,S
32	12,14,33,34	20,0.03,1,60,57,1	7,0.02,0.2,37,92,0.8	U_d,U_d
33	35,36	20,0.03,57,1,1	8,0.0008,1457,2.3,1	S,S
34	9,37,38	20,0.03,57,60,1,1	8,0.0005,982,49,2,0.1	S,S
35	14,39,40	20,0.03,57,60,1,1	9,0.005,166,40,1.5,1.3	U_d,U_d
36	41,47,48,49	57,20,0.03,1,1	2127,14,0.001,2.8,0.01	S,S
37	4,5,42	20,57,1,1,0.03	21,51,0.8,1.1,0.03	S,S
38	4,43,50	20,0.03,1,57,1	13,0.02,0.05,46,0.001	S,U_d
39	10,44	20,0.03,57,1,1	10,0.03,184,0.001,0.9	S,S
40	45,51	20,0.03,57,1,1	8,0.001,1357,0.5,0.3	S,S
41	9,46,52	20,0.03,57,60,1,1	4,0.001,1256,23,1.7,1	S,S

The models in group C were not considered further because they did not pass any test. Although all the models of group B passed the stability analysis test, they did not fit the experimental data with the estimated parameters close to the original ones. Considering that there is no strong biological evidence for reaction schemes for the models of group B except model 8, and that the estimated parameters were outside the stimulated boundary of the original parameters, no further investigation was carried out with them except for model 8, which is the original model. Model 8 was selected from this group for further analysis because it is biologically plausible, even though it failed the parameter estimation test.

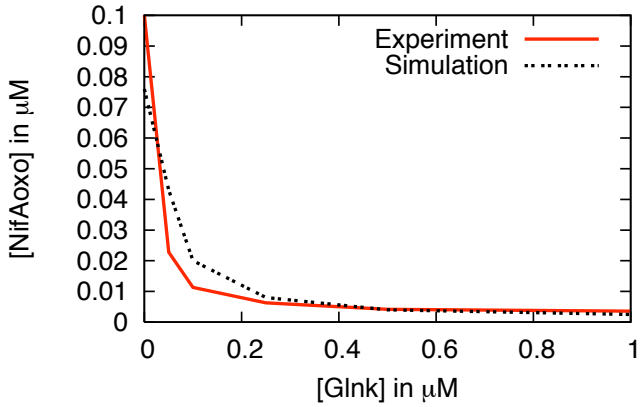


Figure 37: Simulation of model 8 using 0.001 μM as an estimate for $k_{d_{10}}$.

In the first computational test, the parameters for all the models were estimated to check whether there is a set of parameters that is close to the original parameters and capable of fitting the model to the experimental data. There were a set of the parameters, which are capable of leading the simulation of model 8 to be in agreement to the experimental data, however, this set of parameters was not close to the original set.

The dissociation constant between $NifA(2OG)$ and $NifL_{red}ADPGlnK(2OG)$ is the most difficult to estimate under laboratory conditions. Consequently, $K_{d_{10}}$ is the least experimentally reliable parameter of the GlnK system. Setting all other parameters to be the same as the original parameters, the computationally estimated value for $K_{d_{10}}$ (reaction 5.10) is 0.001 μM in comparison to the experimentally estimated value of 1 μM. One data point was used to estimate one parameter in this task. Objective function value is 3.02262 e^{-16} and the function evaluation is 9293 in this estimation task.

As the objective function value is very low, this estimation is very reliable. The simulation of model 8 was compared with the experimental data using a $K_{d_{10}}$ of 0.001 μM. The simulation is in agreement with the experimental data (Figure 37). Consequently, model 8 was kept with the estimated $K_{d_{10}}$ as one of the potential GlnK models.

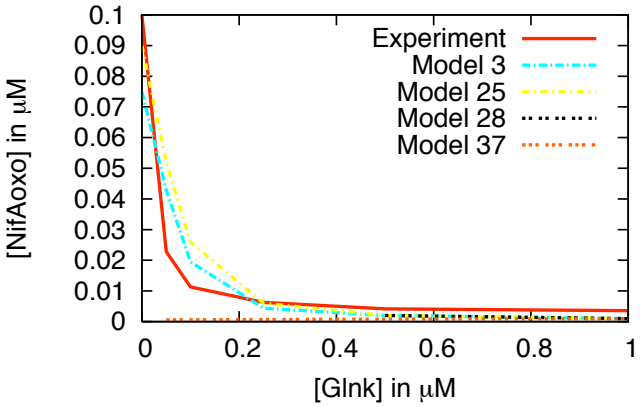


Figure 38: Comparison of the models in Group A with the experimental data for activity of NifA(2OG). Estimated parameters have been used for all the models.

In group A, there are 4 models, which passed both computational tests. Using the estimated data we performed forward modelling for all the models in this group. The task was carried out for all 6 different concentrations of GlnK to compare with the experimental data. Figure 38 shows the result of this analysis.

There is one difference in one of these models, model 3, in comparison to the original model (model 8). Reactions 5.1, 5.2, 5.3, and 5.4 are the same in both models, but model 3 has reaction 5.8 instead of reaction 5.5. In reaction 5.8, we added an additional $2-oxoglutarate$ molecule into the equation as follows:

$$NifL_{red}ADPGlnK(2OG) + NifA(2OG) + 2OG \quad \rightleftharpoons \quad NifL_{red}ADPGlnK(2OG)NifA(2OG)_2.$$

Model 3 passed both computational tests and Figure 38 shows that the simulation fits the experimental data. Therefore, model 3 was kept for

further analysis. Reaction 5.8 in model 3 was divided into two separate reactions for greater clarity:

$$\begin{aligned} \mathrm{NifL_{red}ADPGlnk(2OG)} + \mathrm{NifA(2OG)} &\rightleftharpoons \mathrm{NifL_{red}ADPGlnk(2OG)NifA(2OG)}, \\ \mathrm{NifL_{red}ADPGlnk(2OG)NifA(2OG)} + \mathrm{2OG} &\rightleftharpoons \mathrm{NifL_{red}ADPGlnk(2OG)NifA(2OG)_2}. \end{aligned}$$

Using the new version of model 3, we performed the simulation to compare the result with the experimental data and the old version of model 3 (Figure 39) using the original parameters (Table 12). The simulations of new version and old version of model 3 are very similar and in agreement with experimental data (Figure 39). We therefore use new version of model 3 for further analysis.

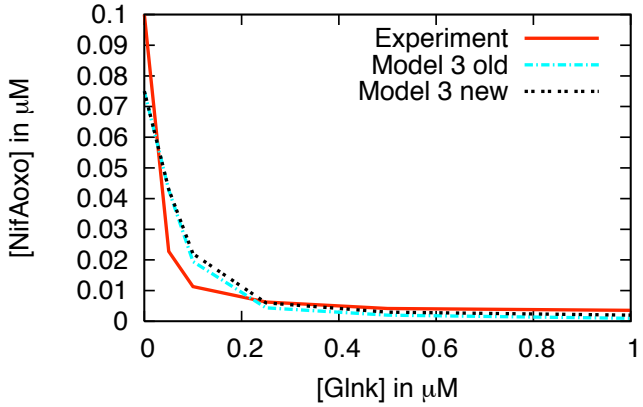


Figure 39: Comparison of the new version of model 3 to the old version and the experimental data.

Model 25 also passed both computational tests. It has the following extra reactions in comparison to our original model, while it does not have reaction 5.9:

$$\begin{aligned} \mathrm{GlnK} + \mathrm{NifA(2OG)} &\rightleftharpoons \mathrm{GlnKNifA(2OG)}, \\ \mathrm{GlnKNifA(2OG)} + \mathrm{NifL_{red}ADPGlnK} &\rightleftharpoons \mathrm{GlnKNifA(2OG)NifL_{red}ADPGlnK}. \end{aligned}$$

Model 25 passed both computational tests and the simulation fitted to the experimental data. However we can argue against this system on the basis of first extra reaction (reaction 5.25) as $\mathrm{NifA(2OG)}$ has been artificially taken away from the system. In this system, disregarding other reactions, only GlnK can titrate away all the $\mathrm{NifA(2OG)}$ as it is present at a higher concentrations than NifA. We did not nominate model 25 for further analysis

because its first extra reaction (reaction 5.25) makes the simulation work merely by taking $NifA(2OG)$ out of the system in an artificial manner.

Model 28 also passed both computational tests and has all the reactions in the original model with the following extra reaction:

$$GlnK(2OG) + NifA(2OG) \rightleftharpoons GlnK(2OG)NifA(2OG).$$

A simulation was performed using estimated parameters. The results show that when the concentration of $GlnK$ is 0.0, 0.05, and 0.25 μM, the steady state cannot be attained. Also model 28 artificially takes $NifA(2OG)$ out of the system in a similar way to model 25. We therefore deleted model 28 from the potential $GlnK$ models.

The last model in Group A, which passed the first two computational tests, is model 37. Model 37 has one difference in comparison to the original model (model 8). In the second reaction of the original model, $NifL_{red}ADP$ interacts with $NifA$ and the interaction produces $NifL_{red}ADPNifA$. In model 37, we added $2-oxoglutarate$ to this equation:

$$NifL_{red}ADP + NifA + 2OG \rightleftharpoons NifL_{red}ADPNifA(2OG).$$

In a similar way to the other models, the simulations of the system were performed with 6 concentrations of $GlnK$. Figure 38 presents the result of this analysis. There are two main points concerning model 37. Firstly, this model has no steady state, when the concentration of $GlnK$ is zero. Secondly, even a very low concentration of $GlnK$ inactivates $NifA(2OG)$ but this result is not supported by the experimental data. The experiment shows that 0.05 μM $GlnK$ decreases the activity of the system to 22 % but in model 37 the same amount of $GlnK$ decreases the activity of the $GlnK$ system dramatically to 0.6 % (Figure 38). Model 37 was therefore deleted from the potential $GlnK$ models.

In summary, computational tests were carried out to narrow down the number of potential $GlnK$ models. Based on these initial tests, we categorized the potential models in Group of A, B, and C. We deleted all the models from Group C. In Group B of the models that failed in the parameter estimation task, model 8 was kept because the model is the original model and is biologically more plausible than the others. In Group A, there were four models, which passed the initial analysis. However, only one of

them, model 3, passed our further analysis and was kept in this group as a potential GlnK model.

5.0.2 *Comparison of potential GlnK models*

In last section, we explored all possible GlnK models and concluded that there were two potential GlnK models: model 8 and model 37. For ease of notation, model 8 will now be called model A and model 37 model B. The following reactions are for model A and the ODE system corresponding to model A is presented in Table 10:

$$\mathrm{NifL_{red}} + \mathrm{ADP} \underset{k_{d_1}}{\rightleftharpoons} \mathrm{NifL_{red}ADP}, \tag{1}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{NifA} \underset{k_{d_2}}{\rightleftharpoons} \mathrm{NifL_{red}ADPNifA}, \tag{2}$$

$$\mathrm{NifA} + 2\mathrm{OG} \underset{k_{d_3}}{\rightleftharpoons} \mathrm{NifA(2OG)}, \tag{3}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{GlnK(2OG)} \underset{k_{d_4}}{\rightleftharpoons} \mathrm{NifL_{red}ADPGlnK(2OG)}, \tag{4}$$

$$\mathrm{NifL_{red}ADPGlnK(2OG)} + \mathrm{NifA(2OG)} \underset{k_{d_5}}{\rightleftharpoons} \mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)}, \tag{5}$$

$$\mathrm{GlnK} + 2\mathrm{OG} \underset{k_{d_6}}{\rightleftharpoons} \mathrm{GlnK(2OG)}. \tag{6}$$

Model B has one extra reaction in comparison to model A. The following reactions are for model B of the GlnK system and the ODE system of model A is presented in Table 13:

$$\mathrm{NifL_{red}} + \mathrm{ADP} \underset{k_{d_1}}{\rightleftharpoons} \mathrm{NifL_{red}ADP}, \tag{1}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{NifA} \underset{k_{d_2}}{\rightleftharpoons} \mathrm{NifL_{red}ADPNifA}, \tag{2}$$

$$\mathrm{NifA} + \mathrm{2OG} \underset{k_{d_3}}{\rightleftharpoons} \mathrm{NifA(2OG)}, \tag{3}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{GlnK(2OG)} \underset{k_{d_4}}{\rightleftharpoons} \mathrm{NifL_{red}ADPGlnK(2OG)}, \tag{4}$$

$$\mathrm{NifL_{red}ADPGlnK(2OG)} + \mathrm{NifA(2OG)} \underset{k_{d_5}}{\rightleftharpoons} \mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)}, \tag{5}$$

$$\mathrm{GlnK} + \mathrm{2OG} \underset{k_{d_6}}{\rightleftharpoons} \mathrm{GlnK(2OG)}, \tag{6}$$

$$\mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)} + \mathrm{2OG} \underset{k_{d_7}}{\rightleftharpoons} \mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)_2}. \tag{7}$$

We first performed simulations using both models and generate data for all components of model A and B so as to compare them. Table 14 presents the total concentrations and parameters, which were used in the simulations for both models. Simulated data was generated for all components of model A and B then the results compared for some of the components (Figure 40). The simulations of NifA(2OG) in both models are almost identical. Furthermore, they are equally good in mimicking the experimental data.

The simulation of free NifA is identical in both models. NifA decreases dramatically by adding GlnK to both models A and B. In high concentrations of GlnK, there a is very low amount of free NifA available to activate the transcription of *nif* genes as expected in models A and B. Adding GlnK to both models decreases NifA and $\mathrm{NifL_{red}ADP}$, consequently the concentration of $\mathrm{NifL_{red}ADPNifA}$ in both of them also decreases.

Adding GlnK to both models A and B increases the concentration of $\mathrm{NifL_{red}ADPGlnK(2OG)}$. Generation of $\mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)}$ from the interaction of $\mathrm{NifL_{red}ADPGlnK(2OG)}$ and NifA(2OG) is the next step in model A and B. The amount of $\mathrm{NifL_{red}ADPGlnK(2OG)NifA(2OG)}$ is increased to the highest level of NifA by adding GlnK to model A leading

Table 13: The ODEs for the GlnK system of model B. The ODE system was generated using mass action equations. Substitutions for ease of presentation are: x_1=[$NifL_{red}$], x_2=[ADP], x_3=[$NifL_{red}$ADP], x_4=[NifA], x_7=[NifA(2OG)], x_9=[GlnK(2OG)], x_8=[Glnk], x_{10}=[$NifL_{red}$ADPGlnK(2OG)], x_5=[$NifL_{red}$ADPNifA], x_{11}=[$NifL_{red}$ADPGlnk(2OG)NifA(2OG)], and x_6=[2OG], x_{12} =[$NifL_{red}$ADPGlnk(2OG)NifA$(2OG)_2$]; k_i and k_{-i} denote the forward and backward rates (i=1, 2, 3).

$$\dot{x}_1 = -k_1x_2x_1 + k_{-1}x_3$$
$$\dot{x}_2 = -k_1x_2x_1 + k_{-1}x_3$$
$$\dot{x}_3 = k_1x_2x_1 - k_{-1}x_3 - k_2x_3x_4 + k_{-2}x_5 - k_4x_9x_3 + k_{-4}x_{10}$$
$$\dot{x}_4 = -k_2x_3x_4 + k_{-2}x_5 - k_3x_4x_6 + k_{-3}x_7$$
$$\dot{x}_5 = k_2x_3x_4 - k_{-2}x_5$$
$$\dot{x}_6 = -k_3x_4x_6 + k_{-3}x_7 - k_6x_8x_6 + k_{-6}x_9 - k_7x_6x_{11} + k_{-7}x_{12}$$
$$\dot{x}_7 = k_3x_4x_6 - k_{-3}x_7 - k_5x_{10}x_7 + k_{-5}x_{11}$$
$$\dot{x}_8 = -k_6x_8x_6 + k_{-6}x_9$$
$$\dot{x}_9 = -k_4x_9x_3 + k_{-4}x_{10} + k_6x_8x_6 - k_{-6}x_9$$
$$\dot{x}_{10} = k_4x_9x_3 - k_{-4}x_{10} - k_5x_{10}x_7 + k_{-5}x_{11}$$
$$\dot{x}_{11} = k_5x_{10}x_7 - k_{-5}x_{11} - k_7x_6x_{11} + k_{-7}x_{12}$$
$$\dot{x}_{12} = k_7x_6x_{11} - k_{-7}x_{12}$$

Table 14: Concentrations (C/μM) and parameters (K_d/μM) of models A and B.

System Component	Value model A	Value model B
NifA	0.1	0.1
$NifL_{red}$	0.2	0.2
ADP	50	50
$2-oxoglutarate$	1000	1000
GlnK	0.0,0.05,0.1,0.25,0.5,1	0.0,0.05,0.1,0.25,0.5,1
K_{d_1}	20	20
K_{d_2}	0.03	0.03
K_{d_3}	57	57
K_{d_4}	1	1
K_{d_5}	0.001	1
K_{d_6}	60	60
K_{d_7}	-	1

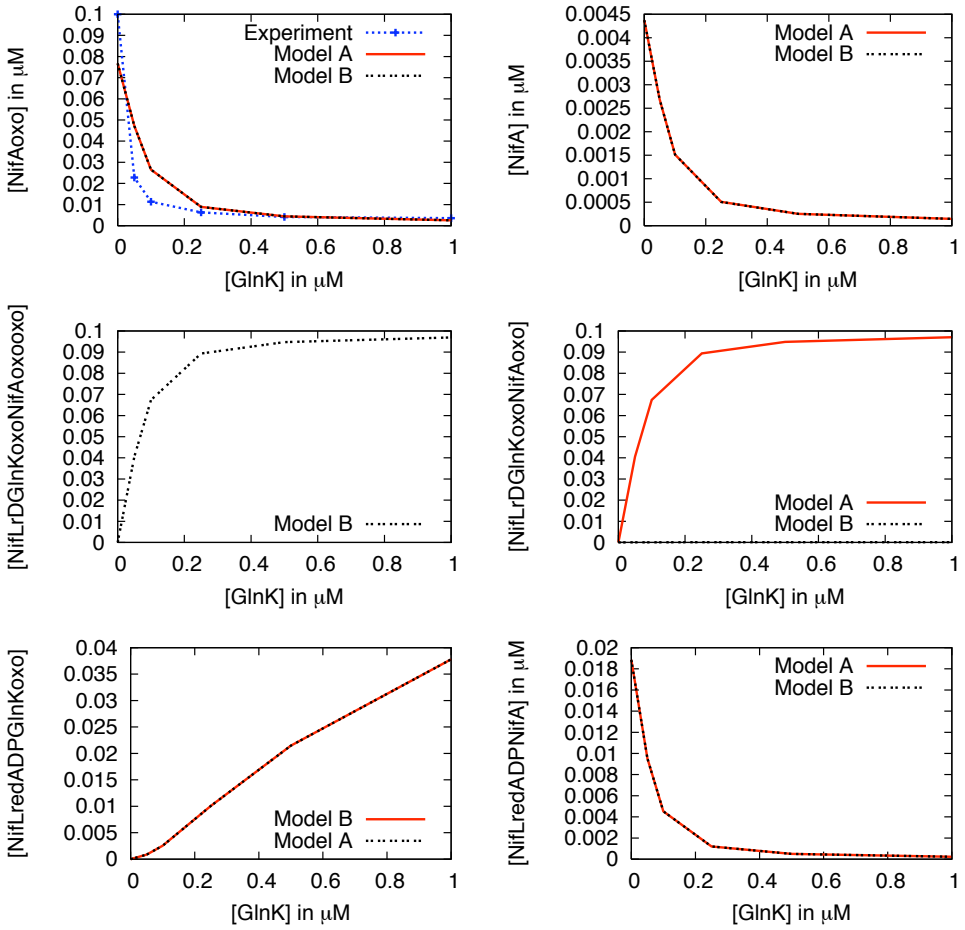


Figure 40: Comparison of simulations of some of the components of model A and B.

to very low availability of free NifA and NifA(2OG). In model A, high affinity of $NifL_{red}ADPGlnK(2OG)$ and NifA(2OG) leads the system to take all the NifA(2OG) to the complex. Model B also takes all the NifA(2OG) to the complex but in a different way. Adding an additional $2-oxoglutarate$ accelerate the rate of reaction 7 as the concentration of $2-oxoglutarate$ is very high (1000 μM) in comparison to NifA (0.1 μM) leading to very low availability of NifA(2OG). This is the main difference between these two models. However, both models drop to the inactive mode by increasing the amount of GlnK to the system and they are equally good in fitting the experimental data. In summary the result of the simulations suggests that there are at least two equally good GlnK models. The first one, model A, is the same as the original model with one different parameter while the second one is different to the original but has the original set of parameters. In the model B next sections we perform a series of tests to try to distinguish between models A and B and also to characterize their properties.

5.0.3 *Test 1*

Four main computational tests were performed with the aim of characterizing the properties of models A and B. These tests are mainly based upon changing the concentrations of GlnK and $2-oxoglutarate$, in order to mimic different environmental conditions. In test 1, $2-oxoglutarate$ was kept in low concentration and GlnK was perturbed. This test was carried out to simulate limited carbon conditions when the $2-oxoglutarate$ concentration is about 100 μM and the concentration of GlnK was varied from 0 to 2.5 μM. Test 1 mimics the conditions when the organism is in carbon limited conditions and transits from low nitrogen to excess nitrogen conditions.

As Figure 41 shows, both models A and B have a low level of NifA(2OG) less than as 25% even when there is no GlnK in the system. By increasing GlnK in this condition the activity of NifA(2OG) decreases. The difference in these two systems in this condition is that the GlnK system drops to a complete switch off condition by increasing GlnK in model A. The complete switch off phenomena in model B is not observed.

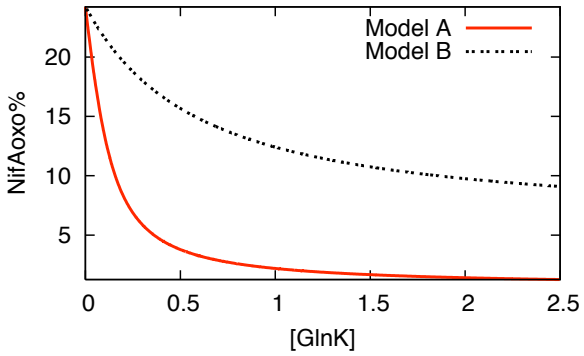


Figure 41: The activity of the GlnK system at the low concentration of $2-$oxoglutarate, whilst varying of GlnK from 0 to 2.5 μM. The activity of the system is measured by calculating the percentage of NifA(2OG) with respect to the total concentration of NifA.

5.0.4 *Test 2*

In test 2 the excess carbon and energy condition were applied to the system and $2-$oxoglutarate raised to 1000 μM. GlnK was varied in the same way as test 1 from 0 to 2.5 μM. Test 2 mimics the conditions when the organism is in carbon excess conditions and transits from low nitrogen to excess nitrogen conditions.

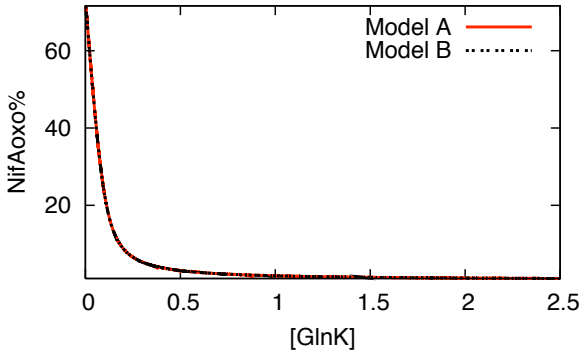


Figure 42: Activity of the GlnK system at high concentrations of $2-$oxoglutarate, whilst varying GlnK from 0 to 2.5 μM. The activity of the system is measured by calculating the percentage of NifA(2OG) with respect to the total concentration of NifA.

As expected, both models A and B have the highest activity when GlnK is not present into the system. Adding GlnK, even within a high concentration of $2-$oxoglutarate , decreases the concentration of NifA(2OG)

in both model A and B dramatically. Comparison of test 1 and 2 shows that, with a low amount of GlnK both model A and B reach zero activity in test 2, while this is not the case in test 1. At high concentrations of $2-oxoglutarate$, model A reaches zero activity at about 0.5 μM GlnK, while at a low concentration model A reaches zero activity at 1.5 μM GlnK.

Considering the reaction schemes in models A and B, there is an interaction between GlnK and $2-oxoglutarate$ in each model, reaction 6. High concentrations of $2-oxoglutarate$ accelerate the forward interaction of $2-oxoglutarate$ and GlnK to push most of the free GlnK to GlnK(2OG). Consequently, at each total concentration of GlnK, the system produces more GlnK(2OG) at high concentrations of $2-oxoglutarate$ in comparison to low concentration of $2-oxoglutarate$(comparison of test 2 to teat 1). To validate this idea, we looked at the concentration of GlnK(2OG) in both conditions in model A. It appears that when the starting concentration of $2-oxoglutarate$ is 0.5 μM, the model reaches a steady state when the concentration of GlnK(2OG) is 0.24 μM, whereas with 1000 μM of $2-oxoglutarate$ leads to 0.36 μM of GlnK(2OG). More GlnK(2OG) concentration at each point leads to greater inhibition of NifA(2OG), which causes the system to stay in an inactive mode at a lower total concentration of GlnK.

5.0.5 *Test 3*

In last two tests the $2-oxoglutarate$ concentration was either low or high, and GlnK in both models was varied. These tests kept the concentration of GlnK low at 0.1 μM while varying the concentration of $2-oxoglutarate$ from 0 to 2500 μM. Test 3 mimics the conditions of the bacteria when under nitrogen deficiency conditions and moving from carbon deficiency to carbon replete conditions, which increases the level of $2-oxoglutarate$.

As $2-oxoglutarate$ is the activator of the $NifL-NifA$ system and the amount of GlnK is low, increasing $2-oxoglutarate$ in test 3 increases the activity of the system (Figure 43). The maximum activity, which can be reached by model A in this condition, is 25% presumably because of the low level of GlnK.

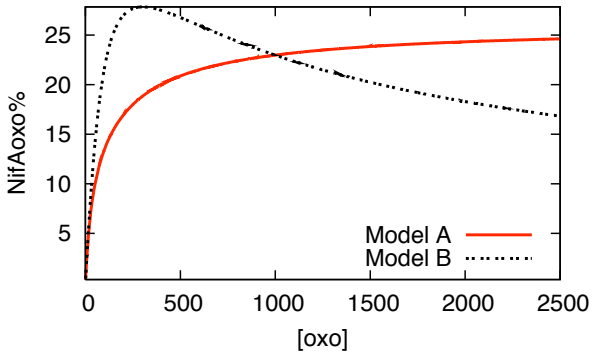


Figure 43: Plots of activity of model A and B at a low concentration of GlnK in response to varying the concentration of $2-oxoglutarate$. The concentration of GlnK is 0.1 μM and $2-oxoglutarate$ varies from 0 to 2500 μM.

The behavior of model B is different to model A in test 3. Model B reaches the maximum activity of 27.5%, but at $2-oxoglutarate$ concentration above 400 μM the NifA(2OG) concentration decreases. This is a very interesting result because it was not apparent from the original experimental observations. This behavior can be because of present of $2-oxoglutarate$ in reaction 7. When the concentration of $NifL_{red}GlnK(2OG)NifA(2OG)$gets higher $2-oxoglutarate$ accelerates the productivity of complex leading to decrease in NifA(2OG).

5.0.6 *Test 4*

In test 3 the GlnK concentration was kept as low as 0.1 μM to examine the behavior of models at relatively low concentrations of GlnK. In test 4 the concentration of GlnK was raised up to 1 μM and $2-oxoglutarate$ was varied from 0 to 2500 μM. Test 4 mimics the conditions for when the bacteria are in excess fixed nitrogen conditions and during the transition from carbon limiting to carbon excess condition, which results in an increase in the level of $2-oxoglutarate$.

The analysis shows that in both models the activity increases transiently in response to the $2-oxoglutarate$ concentration(Fig 38). The activity of the system in model A increases to a maximum of 3% with about 20 μM of $2-oxoglutarate$, while the maximum activity in model B is about 14%

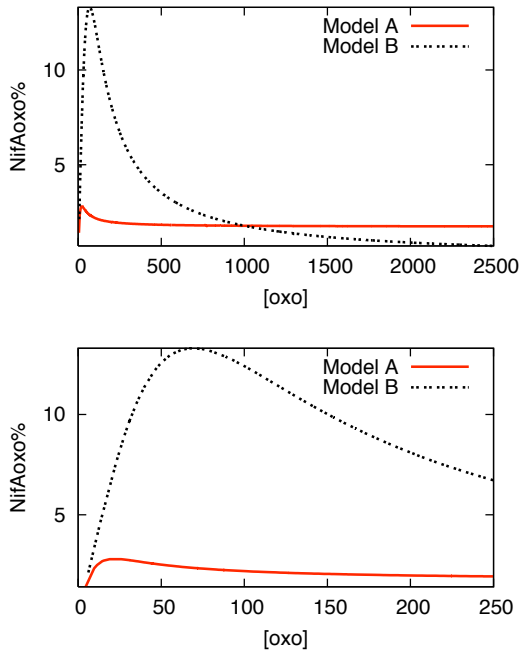


Figure 44: Plots of activity of model A and B at a high concentration of GlnK, 1 μM, with 2 − oxoglutarate concentrations varying from 0 to 2500 μM. The top graph presents the full range of the 2 − oxoglutarate concentrations and the bottom one narrows it down this range to 0 to 250 μM.

reached after the amount of 2 − oxoglutarate has risen to 60 μM. The activity of both models decreases after maximum activity. Assuming 1-2% activity is an inactive mode, both of them finally reach inactivity.

We can argue that the concentration of the GlnK is 1 μM and 10 times the concentration of NifA, 2 − oxoglutarate in certain concentration accelerates the production of GlnK(2OG) in comparison to NifA(2OG) leading to more complex and consequently decrease in NifA(2OG). This is more effective in model B and 2 − oxoglutarate is also one of the reactants of the final product of the system. To summarise, increasing amounts 2 − oxoglutarate leads to more GlnK(2OG), more GlnK(2OG) leads to more $NifL_{red}$ADP GlnK(2OG)NifA(2OG). Finally increasing amounts of 2 − oxoglutarate and $NifL_{red}$ADPGlnK(2OG)NifA(2OG) leads to the formation of more $NifL_{red}$ADPGlnK(2OG)$NifA(2OG)_2$ in model B.

5.1 COMPARISON ON MODEL A AND B

In Chapter 4 we showed that $2-oxoglutarate$ is a key regulator of the $NifL-NifA$ system. The system is in favour of stimulating the synthesis of mRNA transcripts encoding nitrogenase when the fixed nitrogen is a limiting factor. The $2-oxoglutarate$ allows the system to remain in active mode in nitrogen deficiency conditions. On the other hand, there is a mechanism to protect the bacteria from using carbon energy in conditions when the fixation of N_2 is not necessary. In this chapter, this mechanism, i.e. the Glnk system, was investigated in detail.

The $GlnK$ system was modelled initially using a set of reactions and parameters that incorporates current biological knowledge. The results of modelling the $GlnK$ system were far from the experimental observations. We then generated 41 different models for the $GlnK$ system. These systems were investigated through parameter estimation and stability analysis leading to two potential models. In addition to this, four computational tests were performed to monitor the sensitivity of the potential models to $GlnK$ and $2-oxoglutarate$ with the aim of distinguishing the potential models. They showed different behavior in those tests.

It was important to investigate the $NifL-NifA$ system in detail to clarify how the system tunes its activity based on $2-oxoglutarate$ and $GlnK$. $GlnK$ passes the message of nitrogen status to the system and $2-oxoglutarate$ passes the message of carbon energy working together to tune the system in different environmental conditions. In conclusion for characterizing and separating the potential models, the concentrations of $2-oxoglutarate$ and $GlnK$ were perturbed up to 2000 and 2 μM, respectively. The author's program, Equilibrium, was used to simulate the $GlnK$ system based on different $2-oxoglutarate$ and $GlnK$ concentrations. The starting concentration of $2-oxoglutarate$ for simulating of $GlnK$ was 5 μM. Consequently, there were 400 different concentrations of $2-oxoglutarate$ to be investigated by the program. For each concentration of $2-oxoglutarate$ there were 400 different concentrations of $GlnK$. In total, the program therefore simulated the $GlnK$ system 160,000 times with different total concentrations of $2-oxoglutarate$ and $GlnK$.

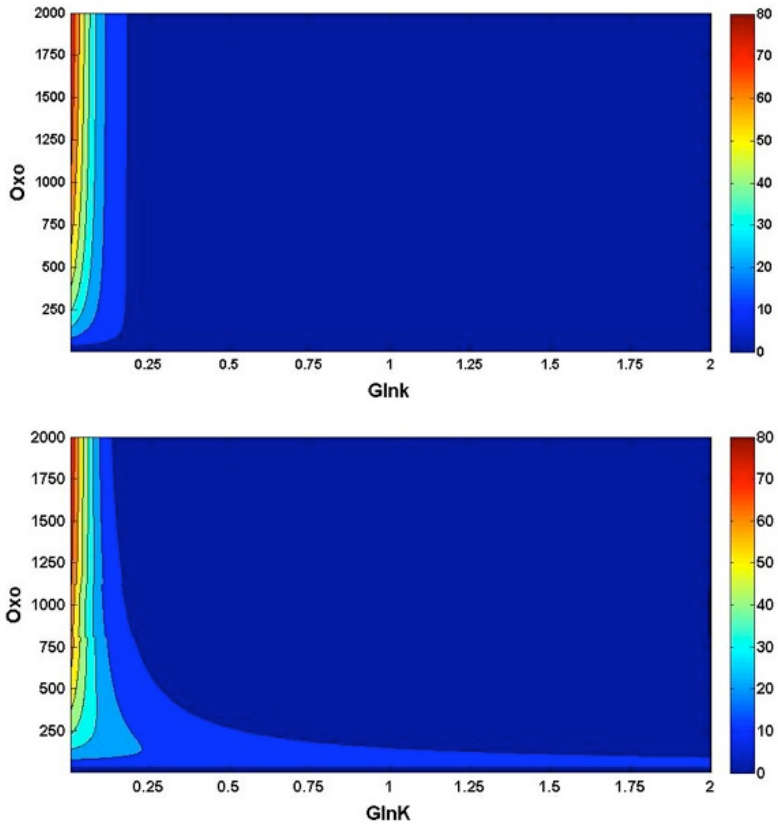


Figure 45: 3D plot presenting the effect of both GlnK and $2-oxoglutarate$ (2OG) at the same time on the activity of the GlnK system. Activity of the system is calculated by taking the ratio of $NifA(2OG)$ to the total amount of $NifA$. The bar shows the activity as a percentage. The upper graph comes from model A and the lower from model B.

The investigation into the activity of the GlnK system using both models A and B indicates that a higher ratio of 2 – oxoglutarate to GlnK increases the activity of the GlnK system. However, the ratio is not the only factor for modulating the activity. The system acts differently in different range of the GlnK and 2 – oxoglutarate concentrations. For low concentrations of GlnK and 2 – oxoglutarate, the concentration of 2 – oxoglutarate raises the activity of the system. At high concentrations of GlnK, even high amounts of 2 – oxoglutarate are not able to change the activity of the system and it stays in a very low range. To investigate the system using the ratio of 2 – oxoglutarate to GlnK, it is important to monitor the activity of the system based on individual concentrations of GlnK and 2 – oxoglutarate and their combinations. Figure 45 presents data for all of 2 – oxoglutarate and GlnK concentrations within the range of 0-2000 µM and 0-2 µM , respectively, for model A and model B (Figure 45).

In summary our findings for GlnK models A and B are as follows:
Model A:

- There is a small range for the 2 – oxoglutarate and GlnK concentrations that leads to high activity of the system.
- Increasing the concentration of GlnK to higher than 0.2 µM switches the system to an inactive mode.
- There is a threshold for the amount of GlnK in the system across which the behavior from active mode to inactive mode switches, disregarding 2 – oxoglutarate concentration.
- In very low concentrations of GlnK the activity of the system is dependent on the 2 – oxoglutarate concentration.
- There is a very narrow range in which the system is fully active. This happens when the concentration of 2 – oxoglutarate is higher than 1000 µM and when the concentration of GlnK is in the range of 0.01 µM.

Model B:

- There is a small range for the $2-oxoglutarate$ and GlnK concentrations that leads the system to high activity in model B, although this range is bigger than the active in model A.
- Increasing the concentration of GlnK decreases the activity of the system. Unlike model A, it does not switch the system to the inactive mode.
- There is no threshold for the amount of GlnK in model B to switch the behavior of the system from active to inactive mode.
- Like model A, when concentrations of GlnK are very low, the activity of model B is the dependent of the $2-oxoglutarate$ concentration.
- There is a very narrow range in model B similar to model A when the system is fully active. This situation occurs, when the concentration of $2-oxoglutarate$ is higher than 1000 μM and concentration of GlnK is around 0.01 μM.
- In model B, even in high concentration of GlnK, the GlnK system shows some degree of activity while the GlnK system in model A does not show any activity in higher than 0.2 μM of GlnK.
- Model B unlike model A is not inactive in the present of 2 μM of GlnK.

5.2 CONCLUSION

The modelling has shown that both model A and B and mimic some of the experimental observations. However, each model has its characteristics leading to two different models responses to the perturbation. We showed that although increasing the concentration of GlnK decreases the activity of the system, model B unlike model A does not switch the system to the inactive mode (Figure 45 and Figure 41). There is strong experimental evidence showing that at a high concentration of GlnK the activity of $NifA(2OG)$ is almost zero [53, 60, 10]. Therefore model B does not mimic the experimental system in this regard.

6

DYNAMIC REGULATION OF NITROGEN FIXATION IN A.VINELANDII AND K.PNEUMONIAE

The complexity of regulation of nitrogen fixation and the low availability of data means that dynamic modelling is a very challenging task. Comparing the dynamics of nitrogen fixation in *A. vinelandii* to *K. pneumoniae*, which regulate nitrogen fixation the same components but in a different way, can help to reveal the characteristics of nitrogen fixation.

We initiated the modelling of the NifL – NifA system in *A. vinelandii* by forming a complete set of reactions involved in nitrogen fixation. Appendix A presents all the reactions involved in system. Three layers of gene, protein, and metabolic were considered in forming the scheme the NifL – NifA system. Most of the reactions are in the protein layer, which contains the protein-protein interactions. Metabolic reactions are in metabolic layer of the scheme and the gene layer contains the transcription and translation reactions. There is currently not much data available for set of reactions. Furthermore, putting all the reactions together increases the complexity of the system. We attempted to model the complete NifL – NifA system dynamically but by doing so the complexity of the system was increased and so the chance of revealing the characteristics of the system was decreased. Consequently modelling the dynamics of nitrogen fixation at an abstract level is more plausible than modelling the complete system.

We have studied the two main sub-systems of the NifL – NifA system in *A. vinelandii* in the last two chapters. We learnt that the system is more sensitive to total concentrations of the system components than to the parameters of the system (Chapter 5). Hence we can set the parameters of the abstract model to arbitrary values, the focus can be shifted to the transcription of model components. Perturbation of the transcription rates changes the concentrations of key components of model. Furthermore we found out that the ratio of NifL to NifA is important in addition to their ranges of concentration (Chapter 5). As NifA is either free or complexed with NifL, our studies showed that the ratio of NifL to NifA is crucial, so

we dismiss all the intermediate steps and only present two forms of active and inactive NifA in this study.

We learnt from Chapter 6 that GlnK is the master regulator in this system. In particular, we showed that, although 2 – oxoglutarate, ADP, and ATP play roles in regulation of nitrogen fixation, the system has two different behaviors based on presence and absence of GlnK [55]. We showed that even low amounts of GlnK reduce the activity of the system dramatically. We can apply this knowledge to model nitrogen fixation abstractly by ignoring some of the interactions of the system and including the active and inactive forms of GlnK to the abstract model.

The global mechanism of nitrogen fixation is deceptively simple. When insufficient fixed nitrogen is available nif genes are transcribed leading to atmospheric nitrogen fixation and once fixed nitrogen is available, the transcription of nif genes slows down. Similarly, if too much oxygen is present, transcription of nif genes slows down. So, in ambient conditions, nitrogen fixation is regulated by the availability of fixed nitrogen (Chapter 1 Section 1.3.8).

The regulation of nitrogen fixation in A. vinelandii is different to that in K. pneumoniae with regards to nitrogen status. Figure 46 presents a very basic model of nitrogen fixation regulation in A. vinelandii and K. pneumoniae. In both species, transcription of the nif genes requires NifA to bind to an upstream promoter site. The transcription of nif genes are then activated leading to more available nitrogenase to catalyze the conversion of atmospheric nitrogen to ammonia [27]. This process is the same in both systems, but the mechanism of responding to production of ammonia in the systems is different. Sensing the nitrogen status and the bacteria's responses to it are the key differences in regulating nitrogen fixation between these species.

Ignoring all the details in sensing the nitrogen in the cell in both systems, we have introduced 'Signal' as the sensor of nitrogen, which includes all the processes involved in sensing nitrogen in the cell. In addition, there was no attempt to model all the reactions in this study. Once NifA, has activated transcription, both systems produce Signal to send the message of nitrogen sufficiency that gets relayed to other components of the system. The next step is from Signal to GlnK which is different for A. vinelandii

and K. pneumoniae (Figure 46). Signal in K. pneumoniae not only acts differently in comparison to Signal in A. vinelandii, but it also does its task in different way. Unlike in A. vinelandii, the Signal in K. pneumoniae acts as a repressor of the glnk gene to decrease the transcription of GlnK [76, 86]. Thus glnk gene is subject to nitrogen regulation in K. pneumoniae, whereas this is not the case for A. vinelandii [10].

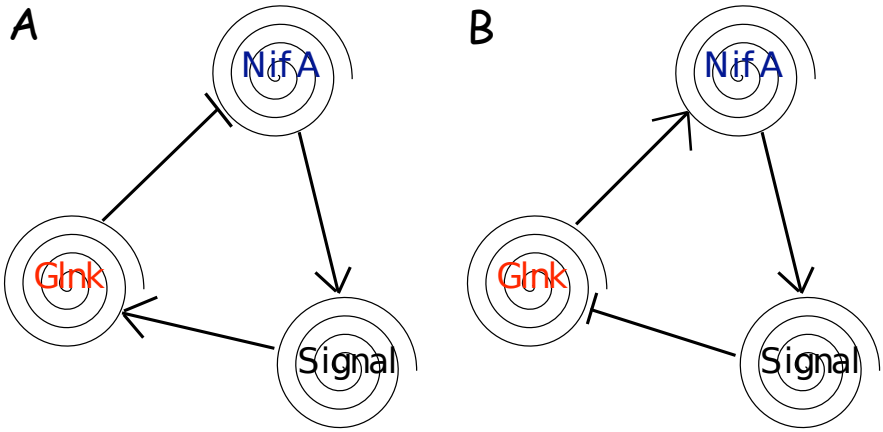


Figure 46: Conceptual model of nitrogen fixation in A. vinelandii (A) and K. pneumoniae (B).

In A. vinelandii Signal only activates GlnK via protein modification, and the glnk gene is not affected by it. GlnK itself also plays different roles in A. vinelandii and K. pneumoniae. GlnK in A. vinelandii inactivates NifA by forming the NifL – NifA complex, which stops activation of nif genes, whereas GlnK in K. pneumoniae promotes the dissociation of NifL – NifA complex to activate the system (Figure 46).

6.1 THE ABSTRACT MODEL OF NITROGEN FIXATION IN A.VINELANDII

Using the conceptual model in Figure 46 A, a set of reactions was formed for A. vinelandii. It consists of degradation reactions, regulatory reactions, and transcriptional reactions:

$$\begin{aligned}
NifA_a + nif &\rightarrow NifA_a + nif + mRNA, && (6.1)\\
Signal + GlnK_i &\rightleftharpoons GlnK_a, && (6.2)\\
GlnK_a + NifA_a &\rightleftharpoons NifA_i, && (6.3)\\
nifa &\rightarrow NifA_a, && (6.4)
\end{aligned}$$

$$\begin{aligned}
glnk &\rightarrow GlnK_a, && (6.5)\\
mRNA + R &\rightarrow Nitrogenase + mRNA + R, && (6.6)\\
Nitrogenase + N_2 &\rightarrow ammonia + Nitrogenase + N_2, && (6.7)\\
ammonia &\rightarrow Signal + ammonia, && (6.8)\\
NifA_a &\rightarrow *, && (6.9)\\
Signal &\rightarrow *, && (6.10)\\
GlnK_i &\rightarrow *, && (6.11)\\
GlnK_a &\rightarrow *, && (6.12)\\
NifA_i &\rightarrow *, && (6.13)\\
mRNA &\rightarrow *, && (6.14)\\
Nitrogenase &\rightarrow *, && (6.15)\\
ammonia &\rightarrow *. && (6.16)
\end{aligned}$$

In this system, $NifA$ is represented in the active form by $NifA_a$ and in the inactive form by $NifA_i$. $NifA_a$ is the free form of $NifA$, which can activate the transcription of nif genes. The $NifA_i$ is interacted form of $NifA$ to NifL. $GlnK$ is also represented as two forms active $Glnk_a$ and inactive, $Glnk_i$. $GlnK$ inactivates $NifA$ in *A. vinelandii*. $GlnK_a$ and $NifA_a$ are products of $glnk$ and $nifa$ genes. R represents the ribosome that interacts with mRNA to translate nitrogenase.

Table 15 presents the rate of all the reactions (7.1)-(7.16) of the system and the corresponding ODE system. The rate of equation 7.1 was calculated using the Hill function for an activator. The rate of all other reactions was calculated using the law of general mass action. The Signal activates $GlnK_i$ (7.2) leading to inactivation of $NifA_a$ by $GlnK_a$ (7.3). Equations 7.4 and 7.5 are the transcriptions reaction of the $nifa$ and $glnk$ genes. Equation

Table 15: The rate of reactions and ODEs for the abstract system of *A. vinelandii*. Substitutions for ease of presentation: y_1=[$NifA_a$], y_2=[*nif*], y_3=[*Signal*], y_4=[$Glnk_i$], y_5=[$Glnk_a$], y_6=[$NifA_i$], y_8=[*nifa*], y_{10}=[*Ribosome*], y_7=[*mRNANitrogenase*], y_9=[*glnk*], y_{13}=[*Ammonia*], y_{12}=[N_2], y_{11}=[*Nitrogenase*].

$v_1 = \frac{B(y_1)^n}{(k_1)^n+(y_1)^n}$	$\dot{y}_1 = -v_3 - v_9 + v_4$
$v_2 = y_3 y_4 k_2 - y_5 k_3$	$\dot{y}_2 = 0$
$v_3 = y_5 y_1 k_4 - y_6 k_5$	$\dot{y}_3 = v_8 - v_2 - v_{10}$
$v_4 = t_1 y_8$	$\dot{y}_4 = -v_2 - v_{11} + v_5$
$v_5 = t_5 y_9$	$\dot{y}_5 = v_2 - v_3 - v_{12}$
$v_6 = y_7 y_{10} k_6$	$\dot{y}_6 = v_3 - v_{13}$
$v_7 = y_{11} y_{12} k_7$	$\dot{y}_7 = v_1 - v_{14}$
$v_8 = y_{13} k_8$	$\dot{y}_8 = 0$
$v_9 = d_1 y_1$	$\dot{y}_9 = 0$
$v_{10} = d_3 y_3$	$\dot{y}_{10} = 0$
$v_{11} = d_4 y_4$	$\dot{y}_{11} = v_6 - v_{15}$
$v_{12} = d_5 y_5$	$\dot{y}_{12} = 0$
$v_{13} = d_6 y_6$	$\dot{y}_{13} = v_7 - v_{16}$
$v_{14} = d_7 y_7$	
$v_{15} = d_{11} y_{11}$	
$v_{16} = d_{13} y_{13}$	

7.6 represents the translation of mRNA to nitrogenase. The fixation of N_2 occurs in equation 7.7. As the source of N_2 is abundant, N_2 was placed deliberately in the products of equation 7.7 to keep the amount of N_2 constant in the system. Signal senses the production of the ammonia in 7.8. Finally, equations 7.9 to 7.16 represent the degradation of the components of the system.

6.2 THE ABSTRACT MODEL OF NITROGEN FIXATION IN K.PNEUMONIAE

There are three differences between the abstract model of K. *pneumoniae* and A. *vinelandii*. Firstly, GlnK is regulated at the transcriptional level in K. *pneumoniae*. Signal acts as a repressor of the corresponding *glnk* gene in equation (7.18). The rate of equation 7.18 is calculated using Hill function for a repressor. Secondly, the transcription of *nifa* gene is controlled by the Signal. It represses the transcription of the *nifa* gene equation (7.19). The rate of equation 7.19 is also calculated using Hill function for a repressor. Thirdly, $GlnK_a$ activates $NifA_i$ to $NifA_a$ equation (7.20). Table 16 presents the rate of all reactions and corresponding ODE system for K. *pneumoniae*.

The conceptual model in Figure 46 B was used to form the abstract model of nitrogen fixation in K. *pneumoniae*, made up of the following reactions:

Table 16: The rate of reactions and ODEs for the abstract system of K. *pneumoniae*. Substitutions for ease of presentation are x_1=[$NifA_a$], x_2=[*nif*], x_3=[*Signal*], x_4=[*glnk*], x_5=[*Glnka*], x_6=[*NifAi*], x_9= [*Ribosome*], x_8=[*nifa*], x_7=[*mRNANitrogenase*], x_{10}=[*Nitrogenase*], x_{11}=[*N2*], x_{12}=[*Ammonia*].

$v_{17} = \frac{Ax_1^n}{(k_1)^n+(x_1)^n}$	$\dot{x}_1 = v_{19} + v_{20} - v_{24}$
$v_{18} = \frac{B}{1+(\frac{x_3}{k_2})^n}$	$\dot{x}_2 = 0$
$v_{19} = \frac{C}{1+(\frac{x_3}{k_3})^n}$	$\dot{x}_3 = v_{23} - v_{d3}$
$v_{20} = x_5x_6k_4 - x_1k_5$	$\dot{x}_4 = 0$
$v_{21} = x_7x_9k_6$	$\dot{x}_5 = v_{18} - v_{20} - v_{26}$
$v_{22} = x_{10}x_{11}k_7$	$\dot{x}_6 = -v_{20} - v_{27}$
$v_{23} = x_{12}k_8$	$\dot{x}_7 = v_{17} - v_{28}$
$v_{24} = d_1x_1$	$\dot{x}_8 = 0$
$v_{25} = d_3x_3$	$\dot{x}_9 = 0$
$v_{26} = d_5x_5$	$\dot{x}_{10} = v_{21} - v_{29}$
$v_{27} = d_6x_6$	$\dot{x}_{11} = 0$
$v_{28} = d_7x_7$	$\dot{x}_{12} = v_{22} - v_{30}$
$v_{29} = d_{10}x_{10}$	
$v_{30} = d_{12}x_{12}$	

$$\begin{aligned}
NifA_a + nif &\rightarrow NifA_a + nif + mRNA, && (6.17)\\
Signal + glnk &\rightarrow Glnk_a + glnk, && (6.18)\\
Signal + nifa &\rightarrow NifA_a + nifa, && (6.19)\\
GlnK_a + NifA_i &\rightleftharpoons NifA_a, && (6.20)\\
mRNA + R &\rightarrow Nitrogenase + mRNA + R, && (6.21)\\
Nitrogenase + N_2 &\rightarrow ammonia + Nitrogenase + N_2, && (6.22)\\
ammonia &\rightarrow signal + ammonia, && (6.23)\\
NifA_a &\rightarrow *, && (6.24)\\
Signal &\rightarrow *, && (6.25)\\
GlnK_a &\rightarrow *, && (6.26)\\
NifA_i &\rightarrow *, && (6.27)\\
mRNA &\rightarrow *, && (6.28)\\
Nitrogenase &\rightarrow *, && (6.29)\\
ammonia &\rightarrow *. && (6.30)
\end{aligned}$$

6.3 COMPARISON OF ABSTRACT MODELS

We simulated nitrogen fixation in both abstract models of *A. vinelandii* and *K. pneumoniae* using the ODEs in Table 15 and Table 16. The same parameter values and total concentrations were used for modelling both systems (Table 17). Since both of these systems regulate nitrogen fixation by regulating $NifA_a$, leading to regulation of *nif* gene transcription and finally productivity of ammonia, we monitored the simulation data for $NifA_a$ and ammonia. These two components of the system were monitored in two opposite environmental conditions. One where there is no ammonia available in the system and the species start fixing N_2, and the other where an excessive amount of ammonia is available at the starting point.

The result of modelling the two species shows that both systems oscillate in prescribed conditions. Figure 47 presents the comparison of these models with respect to $NifA_a$. There is no ammonia available at the starting time of simulations. The amplitude of the first peak in *A. vinelandii* is higher

than the other peaks. The ammonia starvation in *A. vinelandii* leads to this peak, but *K. pneumoniae* does not show any differences between the first peak and others. The amplitude of first peak in concentration of ammonia is higher than the other peaks in both *A. vinelandii* and *K. pneumoniae* (Figure 48). Signal in *K. pneumoniae* controls the transcription of $NifA_a$, therefore production of ammonia increases Signal leading to stop first peak be higher in *K. pneumoniae*. In general, the peaks in *A. vinelandii* are sharper than the peaks in *K. pneumoniae*. This sharper increase and decrease in transcription of $NifA_a$ indicates that *A. vinelandii* has stronger response.

High amounts of ammonia at the starting point of the simulations in both species stop the transcription of $NifA_a$ in *K. pneumoniae* and reduce the transcription of $NifA_a$ in *A. vinelandii* (Figure 49). This is not unexpected as *K. pneumoniae* controls the transcription of $NifA_a$ leading to the halt in transcription for a short period of time, while *A. vinelandii* does not halt transcription completely as there is no $NifA_a$ transcriptional regulation in *A. vinelandii*. Both species reduce the amount ammonia with the same rate (Figure 50).

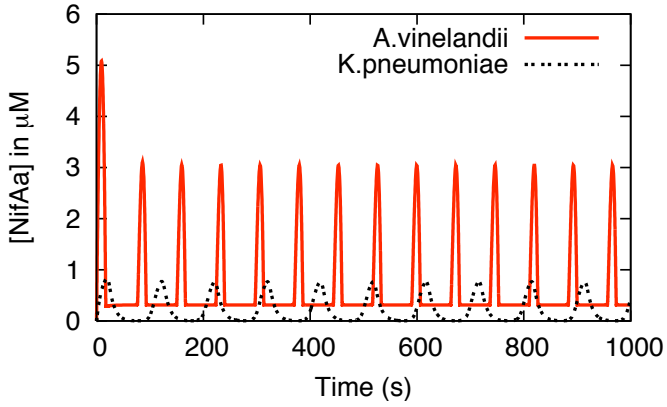


Figure 47: Simulation of $NifA_a$ for both species of *A. vinelandii* and *K. pneumoniae* when no ammonia is available initially.

6.4 PARAMETER PERTURBATION

To investigate the effect of parameter perturbation we performed a number of simulations with different parameters. The effect of perturbation was

Table 17: The parameter values relevant to the ODE model of the abstract system in A. vinelandii and K. pneumoniae.

Model Components	A.vinelandii	K.pneumoniae
k_1	1	1
k_2	1000	1
k_3	10	1
k_4	10000	10000
k_5	100	100
k_6	1	1
k_7	0.1	0.1
k_8	0.0001	0.0001
A	-	0.1
B	0.1	0.1
C	-	0.1
n	4	4
d_1	0.1	0.1
d_3	0.1	0.1
d_4	0.1	-
d_5	-	0.1
d_6	0.1	0.1
d_7	0.1	0.1
d_{10}	-	0.1
d_{11}	0.1	-
d_{12}	-	0.1
d_{13}	0.1	-
t_1	1	-
t_5	1	-

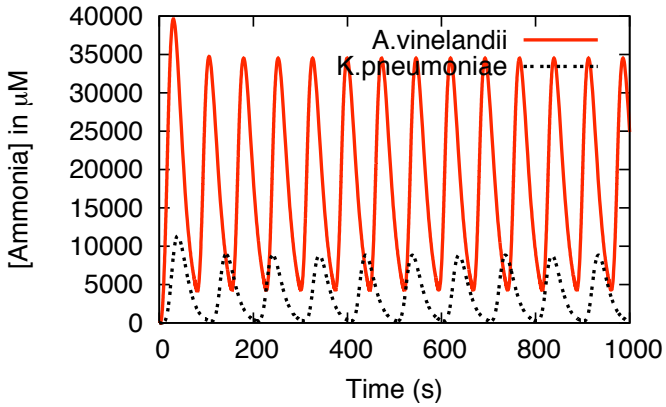


Figure 48: Simulation of ammonia for both species of *A. vinelandii* and K. *pneumoniae* when no ammonia is available initially.

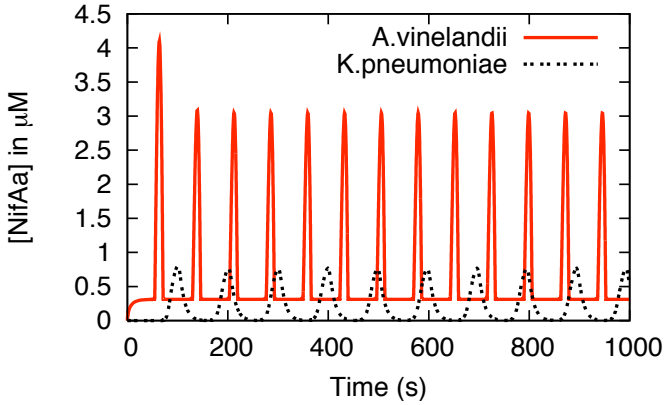


Figure 49: Simulation of $NifA_a$ for both species of *A. vinelandii* and K. *pneumoniae* when 400000 μM ammonia is available initially.

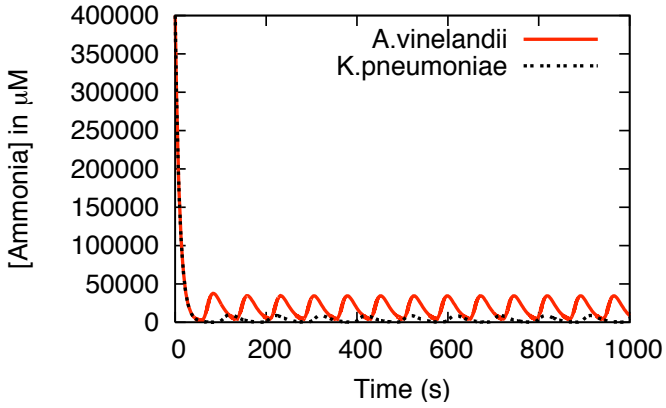


Figure 50: Simulation of ammonia for both species of *A. vinelandii* and K. *pneumoniae* when 400000 μM ammonia is available initially.

investigated in *A. vinelandii*. Here, we present the perturbations of two key rates of the system, rate of transcription of $NifA_a$ and productivity of ammonia. The rate of transcription of *NifA* was perturbed by increasing t_1 from 1 $M^{-1}s^{-1}$ to 1.03 $M^{-1}s^{-1}$ and the productivity of ammonia was perturbed by increasing k_8 from 0.0001 $M^{-1}s^{-1}$ to 0.0005 $M^{-1}s^{-1}$.

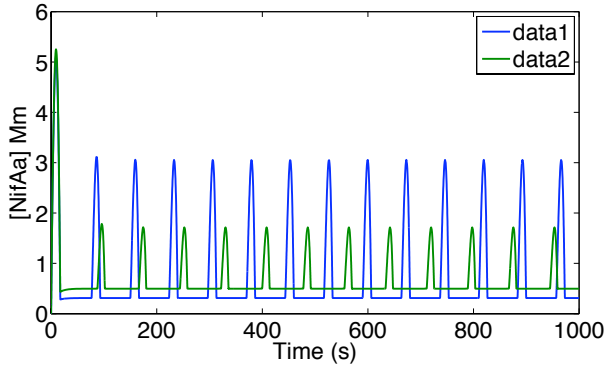


Figure 51: Simulations of $NifA_a$ with the original parameters and with the perturbed parameters for *A. vinelandii*, when no ammonia is available initially. Data1 is the simulation with original parameters. In Data2, the rate of transcription of $NifA_a$ was increased by 3%.

Increasing t_1 decreases the amplitude of spikes in $NifA_a$ transcription (Figure 51) and increases the wavelength, i.e. which is the time between two consecutive spikes. As t_1 increase the amplitude of the spikes tends to zero after the first spike for both $NifA_a$ and ammonia (for t_1 higher than $1.04M^{-1}s^{-1}$ the amplitude is close to zero after first spike). Increasing the level of $NifA_a$ increases the ratio of this protein to $GlnK_a$ leading to lowering the effect of feedback mechanism for the system. This indicates a high sensitivity of the system to t_1. In other words, the system is highly sensitive to the rate of transcription of $NifA_a$.

The system is also sensitive to the productivity of ammonia. Increasing k_8 decreases the amplitude of the spikes in $NifA_a$ transcription and ammonia production ((Figure 52 and Figure 53)). Although the system is sensitive to k_8, the sensitivity to k_8 is lower than the sensitivity to t_1, as 3% perturbation of t_1 can change the concentration of ammonia and $NifA_a$ in almost the same way as a 500% perturbation of k_8. The amplitudes of the spikes tend to zero after a 10 fold increase in k_8.

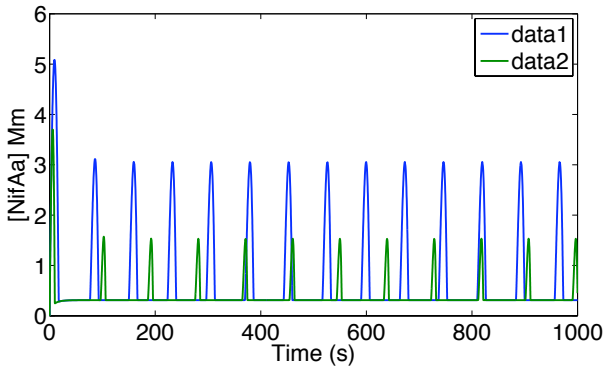


Figure 52: Simulations of $NifA_a$ with the original parameters and with the perturbed parameters for *A. vinelandii*, when no ammonia is available initially. Data1 is the simulation with original parameters. In the Data2, the rate of production of ammonia was increased by 5 fold.

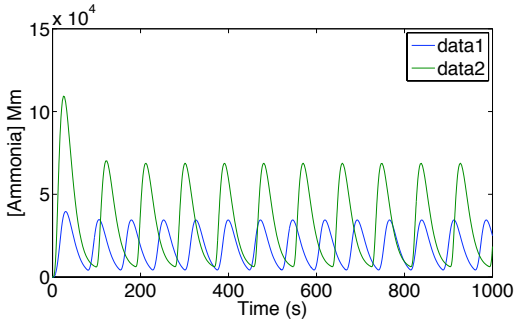


Figure 53: Simulations of ammonia with the original parameters and with the perturbed parameters for *A. vinelandii*, when no ammonia available initially. Data1 is the simulation with original parameters. In the Data2, the productivity of ammonia was increased by 5 fold.

6.5 NIFA$_a$ AGAINST AMMONIA

Availability of ammonia and $NifA_a$ are the main factors of the system. For further analysis, the phase plot of these two components against each other was plotted (Figure 54 and Figure 55) for both species. The result shows that both of them tend to a limit cycle after the first spike. Both systems were simulated with no ammonia available at the start point. The outer curves in both models transit to a limit cycle. Thus, it confirms that both system given the initial conditions stay in oscillation forever. It also shows that the behavior of $NifA_a$ changes for certain concentrations of ammonia. The transcription of $NifA_a$ clearly spikes when ammonia is around 5000 μM and stops after around 20000 μM ammonia in *A. vinelandii*.

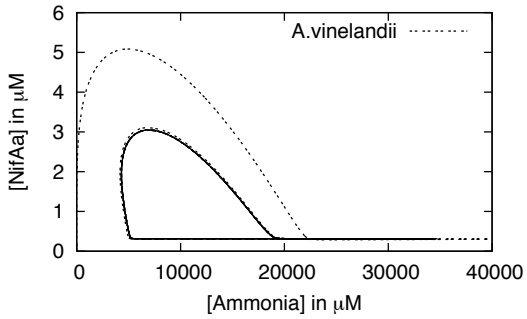


Figure 54: The phase diagram of $NifA_a$ against ammonia when no ammonia is available initially in *A. vinelandii*.

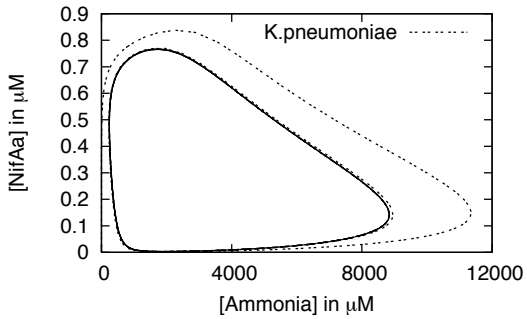


Figure 55: The phase diagram of $NifA_a$ against ammonia when no ammonia is available initially in *K. pneumoniae*.

6.6 CONCLUSION

We developed abstract models that were found to be in agreement with the general properties of *A. vinelandii* and *K. pneumoniae*. From these models we learned that *A. vinelandii* more complete species in comparison to *K. pneumoniae* in respect to nitrogen fixation. The system switches on and off in *A. vinelandii* faster to react to the different environmental conditions. In addition, we learned that given the initial conditions, each system stays in oscillation forever, although, changing the rates of reaction can change the overall behavior. Finally, the system has some critical points of concentration of ammonia leading to change in the behavior of the system. It would be interesting to study this further, but the author had no additional time to undertake a study.

Part IV

CONCLUDING REMARKS

7

CONCLUDING REMARKS

Much effort is currently being made to better understand biological phenomena at the systems level. Biology is evolving from reductionist point of view, which focused on individual cellular components, to systems point of view, driven by the quantitative nature and increasing amounts of biological data [66, 69, 52]. This integrative analysis focuses on system properties of cellular functions. Generally speaking, knowledge about complex biological systems, such as nitrogen fixation, can be separated to 'what we know', 'what we think we know', 'what we know we do not know', and simply 'what we do not know'.

7.1 WHAT WE KNOW → WHAT WE THINK WE KNOW

The importance of understanding of nitrogen fixation has prompted many wet scientists to do research in this field over recent decades, resulting in over six thousand nitrogen fixation related papers plus four hundred hits of nitrogenase kinetics in a recent PubMed search. From what we have presented in this thesis, the following key questions have arisen on the systems biology of nitrogen fixation:

- Is there a limitation to systems biology in regard to the availability of data?
- Is there the possibility of extra interactions in the GlnK system?

7.2 WHAT WE DO NOT KNOW → WHAT WE KNOW

In addition to the questions raised in the last section, we have improved our understanding of nitrogen fixation as follows:

- We have developed a model of nitrogenase action and made it available to the scientific community.

- The availability of stoichiometric amount of NifA and NifL does not fit experimental data, and the ratio of NifL to NifA has to be around 2.
- The concentration of NifA and NifL need to be 10 times more than what we think to fit to experimental data.
- We know that the NifL – NifA system is more sensitive to the total concentrations of its components in comparison to the model parameters.
- A model was presented for the mechanism of the GlnK system.
- The affinity of $NifL_{red}ADPGlnK(2OG)$ to NifA(2OG) can be 1000 times higher than what is currently thought.
- An abstract model of nitrogen fixation in *A. vinelandii* based on current understanding of the system was presented.
- We presented an abstract model of nitrogen fixation in *K. pneumoniae* based on our understanding of that system.
- We showed that *A. vinelandii* has probably evolved a more sensitive system for nitrogen fixation than *K. pneumoniae*, leading to a quicker response to the availability of fixed nitrogen.
- Using the abstract models we have developed, we have shown that given the initial conditions, regulation of nitrogen fixation in both species is probably oscillatory.

7.3 WHAT WE KNOW WE DO NOT KNOW

Our investigation was limited by the complexity of nitrogen fixation and also time. We suggest that more time and effort should be invested in investigating the following:

- More kinetic data needs to be obtained for our system, as we are not able to model the complete system of nitrogen fixation with the current data.

- We do not know how sensitive our abstract model is with regard to changes in parameters.
- We do not know what the boundaries are for the initial conditions of the abstract models so that they stay in oscillation.
- We do not know if adding noise to the abstract returns to former state or goes to another state.

7.4 WHAT WE DO NOT KNOW

- *We do not know many things and many aspects of nitrogen fixation are still a mystery to the author!*

Part V

APPENDIX

A

THE COMPLETE REACTIONS OF NIFL-NIFA SYSTEM

$$\mathrm{NifL_{ox}} + e^{-} \rightleftharpoons \mathrm{NifL_{red}}, \tag{A.1}$$

$$\mathrm{NifL_{ox}} + \mathrm{ADP} \rightleftharpoons \mathrm{NifL_{ox}ADP}, \tag{A.2}$$

$$\mathrm{NifL_{red}} + \mathrm{ADP} \rightleftharpoons \mathrm{NifL_{red}ADP}, \tag{A.3}$$

$$\mathrm{NifL_{red}ADP} + \mathrm{NifA} \rightleftharpoons \mathrm{NifL_{red}ADPNifA}, \tag{A.4}$$

$$\mathrm{NifA} + \mathrm{2OG} \rightleftharpoons \mathrm{NifA(2OG)}, \tag{A.5}$$

$$\mathrm{NifA} + \mathrm{NifL_{ox}ADP} \rightleftharpoons \mathrm{NifL_{ox}ADPNifA}, \tag{A.6}$$

$$\mathrm{NifA(2OG)} + \mathrm{NifL_{ox}ADP} \rightleftharpoons \mathrm{NifL_{ox}ADPNifA(2OG)}, \tag{A.7}$$

$$\mathrm{GlnkATP(2OG)} + \mathrm{NifL_{ox}} \rightleftharpoons \mathrm{NifL_{ox}GlnkATP(2OG)}, \tag{A.8}$$

$$\mathrm{NifL_{ox}ADP} + \mathrm{GlnkATP(2OG)} \rightleftharpoons \mathrm{NifL_{ox}ADPGlnkATP(2OG)}, \tag{A.9}$$

$$\mathrm{NifA(2OG)} + \mathrm{NifL_{ox}GlnKATP(2GO)} \rightleftharpoons \mathrm{NifL_{ox}GlnKATP(2GO)NifA(2GO)}, \tag{A.10}$$

$$\mathrm{NifA(2OG)} + \mathrm{NifL_{ox}ADPGlnKATP(2GO)} \rightleftharpoons \mathrm{NifL_{ox}ADPGlnKATP(2GO)NifA(2OG)}, \tag{A.11}$$

$$\mathrm{2GO} + \mathrm{ATP} + \mathrm{Glnk} \rightarrow \mathrm{GlnKATP(2GO)}, \tag{A.12}$$

$$\mathrm{NifL_{ox}ADPNifA} + e \rightleftharpoons \mathrm{NifL_{red}ADPNifA}, \tag{A.13}$$

$$\mathrm{NifL_{red}} + \mathrm{GlnKATP(2GO)} \rightleftharpoons \mathrm{NifL_{red}GlnKATP(2GO)}, \tag{A.14}$$

$$\mathrm{GlnKATP(2GO)} + \mathrm{NifL_{red}ADP} \rightleftharpoons \mathrm{NifL_{red}ADPGlnKATP(2GO)}, \tag{A.15}$$

$$\mathrm{NifA(2OG)} + \mathrm{NifL_{red}ADPGlnKATP(2GO)} \rightleftharpoons \mathrm{NifL_{red}ADPGlnKATP(2GO)NifA(2OG)}, \tag{A.16}$$

$$\mathrm{NifL_{ox}ADP} + e \rightleftharpoons \mathrm{NifL_{red}ADP}, \tag{A.17}$$

$$\mathrm{NifL_{red}ADPNifA} + \mathrm{2GO} \rightleftharpoons \mathrm{NifL_{red}ADP} + \mathrm{NifA(2OG)}, \tag{A.18}$$

$$\mathrm{NifL_{red}GlnKATP(2GO)} + \mathrm{NifA(2OG)} \rightleftharpoons \mathrm{NifL_{red}GlnKATP(2GO)NifA(2OG)}, \tag{A.19}$$

$$\mathrm{glutamate} + \mathrm{ammonia} \rightarrow \mathrm{glutamine}, \tag{A.20}$$

$$2\mathrm{GO} + \mathrm{glutamine} \rightarrow 2-\mathrm{oxo}-\mathrm{glutamate}, \tag{A.21}$$

$$2-\mathrm{oxo}-\mathrm{glutamate} + \mathrm{ammonia} \rightarrow \mathrm{glutamate}, \tag{A.22}$$

$$\mathrm{NifA(2OG)} + \mathrm{NitrogenaseGenesNifGenes} \rightarrow \mathrm{mRNANitrogenase}, \tag{A.23}$$

$$\mathrm{mRNANitrogenase} \rightarrow \mathrm{Nitrogenase}, \tag{A.24}$$

$$\mathrm{NifA(2OG)} + \mathrm{rnfGenes} \rightarrow \mathrm{mRNArnf}, \tag{A.25}$$

$$\mathrm{mRNArnf} \rightarrow \mathrm{rnfProducts}, \tag{A.26}$$

$$\mathrm{UtaseUR} + \mathrm{glutamine} \rightarrow \mathrm{UtaseUR}-\mathrm{glutamine}, \tag{A.27}$$

$$\mathrm{ATase} + \mathrm{GlnKATP(2GO)} \rightarrow \mathrm{GlnKATP(2GO)ATase}, \tag{A.28}$$

$$\mathrm{ATase} + \mathrm{GlnKATP(2GO)UMP} \rightarrow \mathrm{GlnKATP(2GO)UMPATase}, \tag{A.29}$$

$$\mathrm{GlnKATP(2GO)} + \mathrm{UTP} + \mathrm{UtaseUR} \rightarrow \mathrm{GlnKATP(2GO)UMP}, \tag{A.30}$$

$$\mathrm{UtaseUR}-\mathrm{glutamine} + \mathrm{GlnKATP(2GO)UMP} \rightarrow \mathrm{GlnKATP(2GO)} + \mathrm{UMP}, \tag{A.31}$$

$$\mathrm{glutamineSynthetase} + \mathrm{GlnKATP(2GO)ATase} \rightarrow \mathrm{glutamineSynthetase}-\mathrm{AMP}, \tag{A.32}$$

$$\mathrm{GlnKATP(2GO)UMPATase} + \mathrm{glutamineSynthetaseAMP} \rightarrow \mathrm{glutamineSynthetase} + \mathrm{AMP}, \tag{A.33}$$

$$\mathrm{NifL} + \mathrm{ox} \rightarrow \mathrm{NifL_{ox}}, \tag{A.34}$$

$$\mathrm{NifL} + e \rightarrow \mathrm{NifL_{red}}, \tag{A.35}$$

$$\mathrm{N2} + \mathrm{Nitrogenase} \rightarrow \mathrm{ammonia}. \tag{A.36}$$

B

INVESTIGATION OF THE STEADY STATE POINTS IN THE 2-OXOGLUTARATE SYSTEM

B.1 STEADY STATE POINT 1

In steady state point 1 the concentration of $2-oxoglutarate$ is zero, indicating that there is not any interaction between $NifA$ and $2-oxoglutarate$ and only reaction 6 and 7 happen in $2-oxoglutarate$ system.

Parameter Estimation

Table 18 presents the result of parameter estimation in steady state point. As it is indicated in the Table 18, the estimated parameters are extremely close to the natural.

Table 18: Estimated Parameters ($K_d/\mu M$) for steady state point 1.

K_d	Estimation	Experiment
K_{d_1}	14.34	13
K_{d_2}	0.033	0.03
K_{d_3}	40.05	57

Forward Modeling

Using the estimated parameters in steady state point 1, we performed forward modeling for all steady state points. The Equilibrium program was applied for forward modeling of all points with the estimated parameters. Figure 52 shows the $in\ silico$ production of $NifA(2OG)$ based on different concentrations of $2-oxoglutarate$ in comparison with the experimental

data. The NifA(2OG) simulation is in agreement with the experimental results as it is indicated in Figure 52.

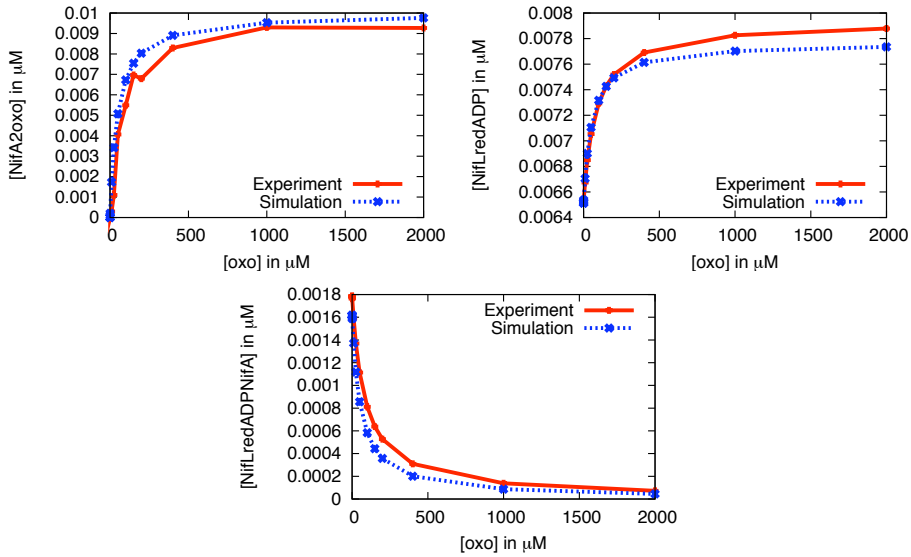


Figure 56: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and NifA(2OG) using the estimated parameters from steady state point 1. NifA(2OG) is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

We extended our in silico experiment to produce data for production of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ with estimated parameters in all 14 different concentrations of $2-oxoglutarate$. The result of this analysis is presented in Figure 52. We compare the simulation of $NifL_{red}ADP$ with the estimated parameters and experimental parameters for all steady state points in $2-oxoglutarate$ system. They are in agreement. We performed the same task for $NifL_{red}ADPNifA$ as it is indicated in Figure 52. There is also agreement between these two simulations.

Steady-state analysis

The stability analysis shows that the system in this state is asymptotically stable using the estimated parameters. Table 19 presents the concentration of metabolites of $2-oxoglutarate$ system in this state, using estimated parameters.

Table 19: Concentration values (C/μM) in steady state conditions for steady state point 1.

metabolite	value
$NifL_{red}$	0.0018
ADP	49.99
$NifL_{red}ADP$	0.0065
NifA	0.0083
$NifL_{red}ADPNifA$	0.0016
$2-oxoglutarate$	0
NifA2oxo	0

Sensitivity analysis

We performed a sensitivity analysis for all three parameters of the $2-oxoglutarate$ system in steady state point. The results of this are presented in Figure 57. Since the concentration of $2-oxoglutarate$ is zero at this point, there is no production of $NifA(2OG)$ and so it is not sensitive to any parameter. Both of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are sensitive to K_{d_1} in a very similar manner. $NifL_{red}ADPNifA$ is quite sensitive to perturbation of K_{d_2} and it is more sensitive to decreasing K_{d_2} in comparn to its increasing. Decreasing K_{d_2} by 95% has the most effect for this steady state point (260% change of $NifL_{red}ADPNifA$). The interesting result for this point is that, although $NifL_{red}ADP$ is production of reaction 1 with dissociation constant K_{d_1}, it is more sensitive to K_{d_2} as it interacts with $NifA$.

B.2 STEADY STATE POINT 2

In steady state point 2 the concentration of $2-oxoglutarate$ is equal to 0.01 μM, and the concentration of $NifA(2OG)$, which was used for parameter estimation is equal to 0. Table 20 presents the result of parameter estimation in this steady state point. As the Table 20 shows the estimated parameters are far from the original values. The concentration of $NifA(2OG)$ which was used in this state is 0 too. As the concentration of $2-oxoglutarate$ is 0.01 μM in this point, there should be some production in $NifA(2OG)$

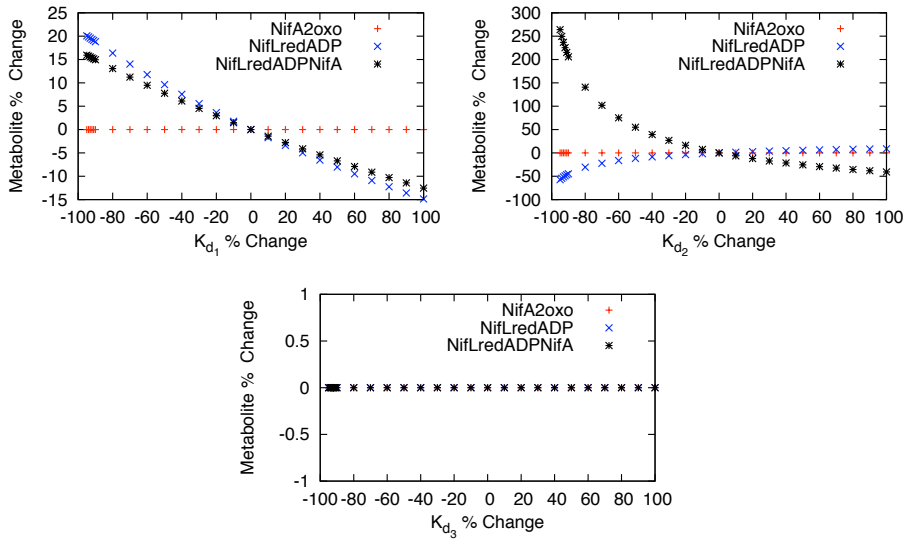


Figure 57: Sensitivity analysis of K_{d_1}, K_{d_2},and K_{d_3} in steady state point 1.

but the experimental value doesn't imply this. The result of this estimation proposes that the activity of NifA(2OG) is more than zero with the 0.01 µM of $2-oxoglutarate$(Experimental error) .

Table 20: Estimated Parameters ($K_d/\mu M$) for steady state point 2.

K_d	Estimation	Experiment
K_{d_1}	0.24	13
K_{d_2}	2.49e-10	0.03
K_{d_3}	8.9e+12	57

Since the estimated parameters are dramatically far from the lab provided parameters in steady state point 2, it makes no sense to perform forward modeling and steady state analysis with this estimated parameters. The concentrations of NifA(2OG), $NifL_{red}ADP$, and $NifL_{red}ADPNifA$ in this steady state point with the original parameters are respectively 1.44e-6 µM , 0.0065 µM , and 0.0017 µM, while in the same state the concentrations of these metabolites using estimated parameters (Table 20) are 4.8e-18 µM, 4.2e-06 µM , and 4.3e-06 µM, respectively, a dramatic difference.

Sensitivity analysis

Figure 58 presents the result of sensitivity analysis for this point. The sensitivity of K_{d_1} on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ is the same as steady state point one and adding $2-oxoglutarate$ does not change the sensitivity of K_{d_1} on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$. $NifA(2OG)$ is not so sensitive to perturbation of K_{d_1}. The effect of perturbation of K_{d_2} on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ is the same as the last steady state point. The interesting result is that the sensitivity of K_{d_2} on both $NifL_{red}ADP$ and $NifA(2OG)$ is the same in this case. K_{d_3} shows a high sensitivity on production of $NifA(2OG)$, though perturbation of K_{d_3} does not produce any effect on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ in this state. Decreasing of K_{d_3} by 95

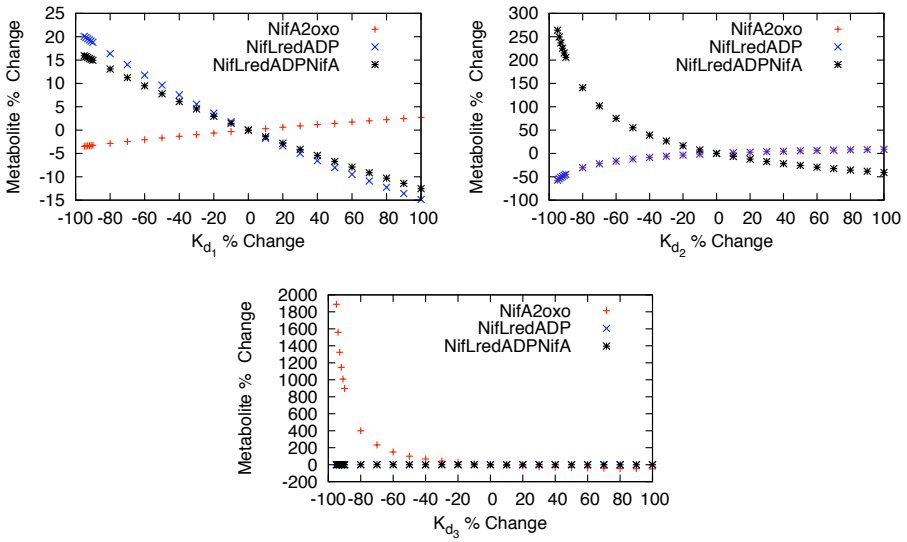


Figure 58: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 2.

B.3 STEADY STATE POINT 3

In steady state point 3 the concentration of $2-oxoglutarate$ is 0.1 µM. Table 21 presents the result of parameter estimation for this steady state point. Although the concentration of $2-oxoglutarate$ is 0.1 µM, the experimental value which was used for $NifA(2OG)$ in this state is 0. This point shows

dramatic difference between estimated parameters and original ones as point 2. The same condition that of the last state does apply through this state.

Table 21: Estimated Parameters ($K_d/\mu M$) for steady state point 3.

K_d	Estimation	Experiment
K_{d_1}	1.79	13
K_{d_2}	5.27e-11	0.03
K_{d_3}	9.52e+13	57

Sensitivity analysis

Figure 59 presents the result of sensitivity analysis for steady state point 3. It is the same same as the last point.

B.4 STEADY STATE POINT 4

In steady state point 4 the concentration of $2-oxoglutarate$ is 0.5 μM. The concentration of $NifA(2OG)$ which was used for parameter estimation was 0.000118 μM.

Parameter Estimation

Table 22 presents the results of parameter estimation for this steady state point. There is a huge difference in estimation for this state in compared to last two. The estimated parameters, however, are very close to original ones.

Table 22: Estimated Parameters ($K_d/\mu M$) for steady state point 4.

K_d	Estimation	Experiment
K_{d_1}	35.23	13
K_{d_2}	0.024	0.03
K_{d_3}	34.98	57

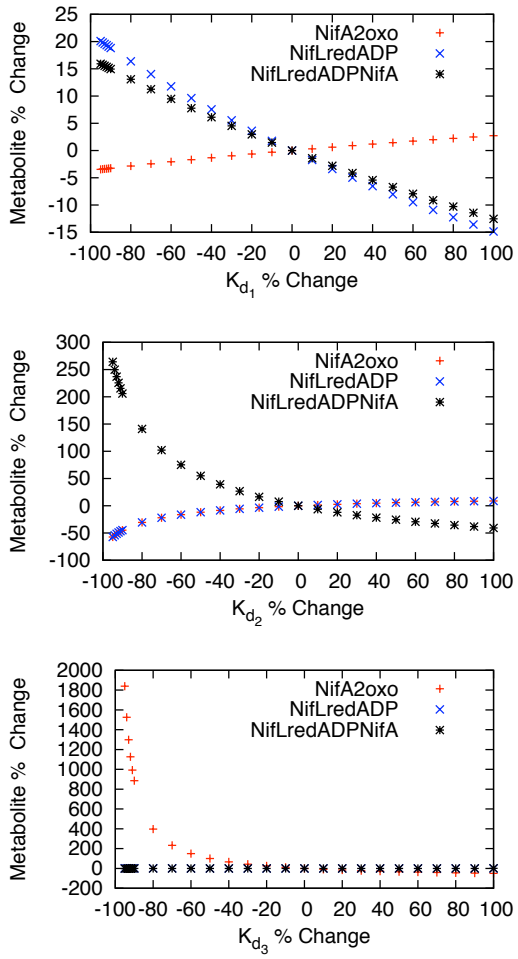


Figure 59: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 3.

Forward Modeling

Using the estimated parameters in steady state point 4 we performed forward modeling for all steady state points. The Equilibrium program was used for forward modeling of all points with the estimated parameters. Figure 60 shows the *in silico* production of NifA(2OG) based on different concentrations of $2-oxoglutarate$ in comparison to the experimental data. NifA(2OG) simulation is in agreement with the experimental results as can be seen in Figure 60.

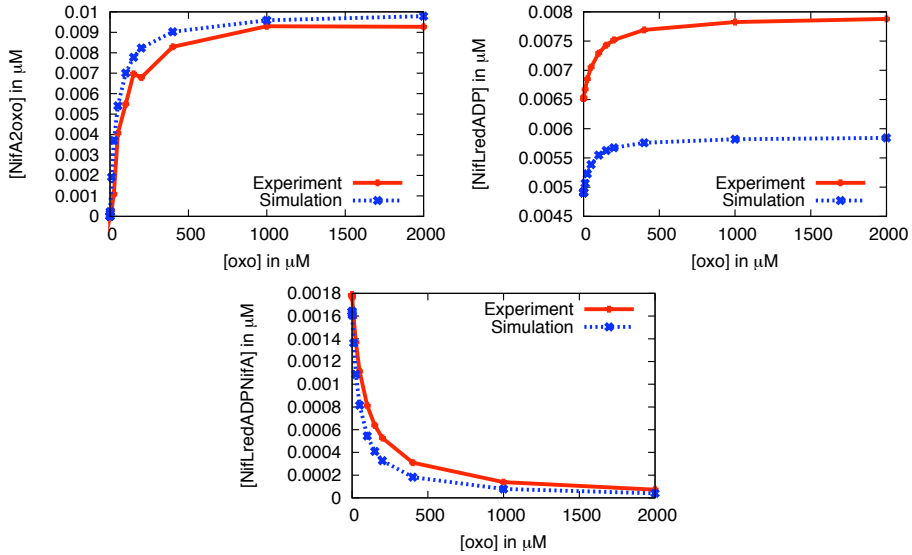


Figure 60: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and NifA(2OG) using estimated parameters from steady state poin 4. NifA(2OG) is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters

We further extended our *in silico* experiment to produce data for the production of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ with estimated parameters. Figure 60 presents the result of this analysis. We then compared the simulation of $NifL_{red}ADPNifA$ with the estimated parameters and experimental parameters for all steady state points in the $2-oxoglutarate$ system. These two simulations are qualitatively in agreement, which implies this that the estimation task was performed based on available data of NifA(2OG) but not $NifL_{red}ADP$. We then carried out the same analysis for $NifL_{red}ADPNifA$. In Figure 60 the comparison of computational simulations with experimen-

tally verified and computationally estimated parameters is presented. This shows agreement between these two simulations.

Steady-state analysis

Stability analysis of $2-oxoglutarate$ system in steady state point 4 shows that this system in this state is asymptotically stable using estimated parameters. Table 23 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 23: Concentration values $(C/\mu M)$ in steady state conditions for steady state point 4.

metabolite	value
$NifL_{red}$	0.00346268
ADP	49.9935
$NifL_{red}ADP$	0.00491276
NifA	0.00825744
$NifL_{red}ADPNifA$	0.00162456
$2-oxoglutarate$	0.499882
NifA2oxo	0.000117999

Sensitivity analysis

Figure 61 presents the result of the sensitivity analysis in steady state point 4. As it is shown, the sensitivity of K_{d_1} and K_{d_2} on model components are the same as last points, though perturbation of K_{d_3} has less effect on $NifA(2OG)$.

B.5 STEADY STATE POINT 5

In steady state point 5 the concentration of $2-oxoglutarate$ is 1 μM and the concentration of $NifA(2OG)$ used for parameter estimation is 0. Table 24 presents the result of parameter estimation for this steady state point. The estimation is similar to steady state points 2 and 3 while the amount

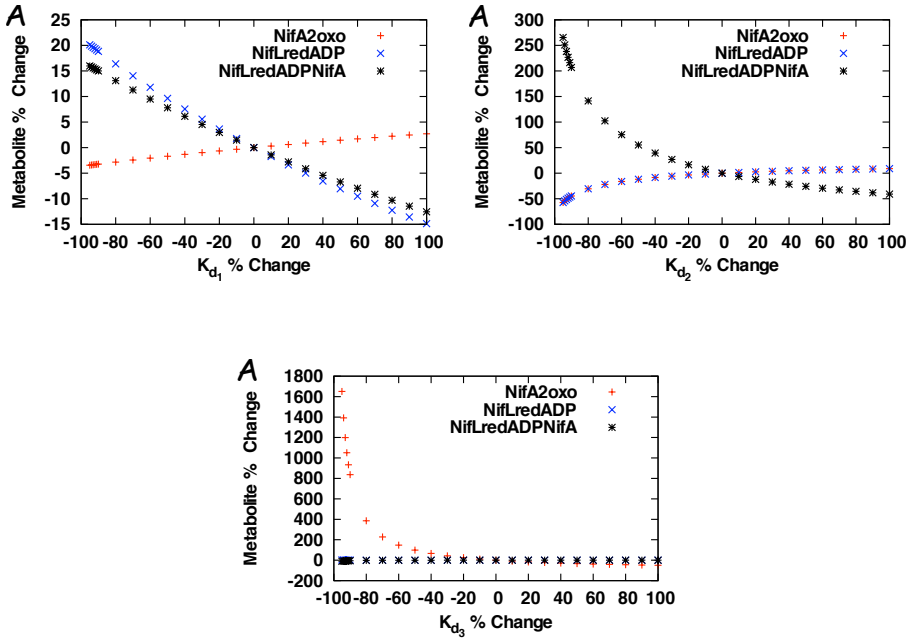


Figure 61: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 4.

of NifA(2OG) is zero. There is a dramatic difference between estimated parameters and original ones. The argument against using this point is similar to points 2 and 3.

Table 24: Estimated Parameters ($K_d/\mu M$) for steady state point 5.

K_d	Estimation	Experiment
K_{d_1}	0.44	13
K_{d_2}	7.095e-11	0.03
K_{d_3}	7.28e+13	57

Sensitivity analysis

Figure 62 presents the result of sensitivity analysis in steady state point 5. As can be seen the sensitivity of K_{d_1} and K_{d_2} on model components are the same as last points, though perturbation of K_{d_3} has less effect on NifA(2OG).

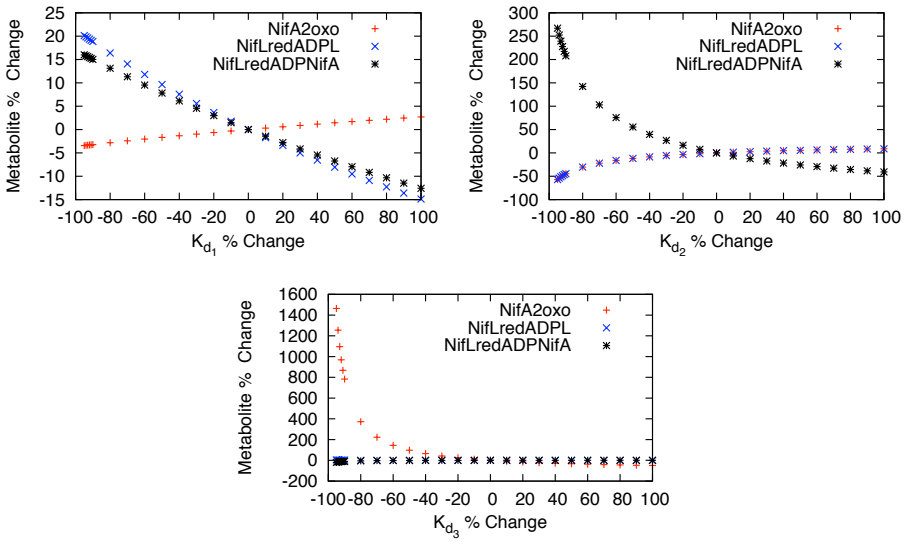


Figure 62: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 5.

B.6 STEADY STATE POINT 6

In steady state point 6 the concentration of $2-oxoglutarate$ is 10 μM while the concentration of $NifA(2OG)$ which was used for parameter estimation is 0.00029 μM .

Parameter Estimation

Table 25 presents the result of parameter estimation for this steady state point. The estimated parameters are fairly close to original ones, though, not as close as steady state points 1 and 4.

Table 25: Estimated Parameters ($K_d/\mu M$) for steady state point 6.

K_d	Estimation	Experiment
K_{d_1}	6.71	13
K_{d_2}	0.017	0.03
K_{d_3}	240.23	57

Forward Modeling

Figure 63 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$. The estimated parameters in steady state point 6 are not as close as estimated parameters in steady state point 1 and 4, consequently, the simulations are not as good as mentioned steady state points.

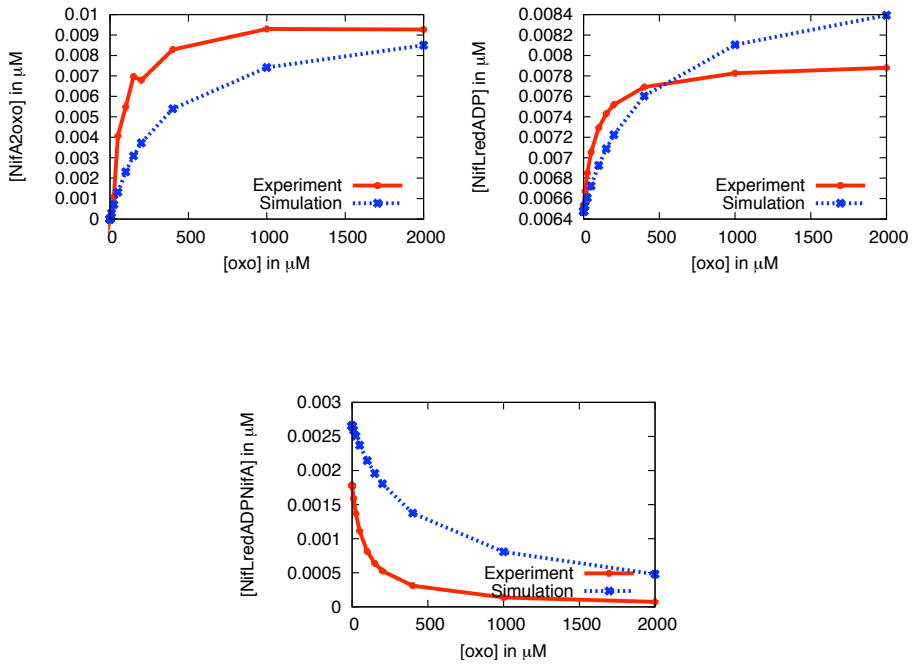


Figure 63: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 6. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

Stability analysis of $2-oxoglutarate$ system in steady state point 6 shows that this system in this state is asymptotically stable using estimated parameters. Table 26 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 26: Concentration values (C/μM) in steady state conditions for steady state point 6.

metabolite	value
$NifL_{red}$	0.00087
ADP	49.99
$NifL_{red}ADP$	0.0065
NifA	0.0071
$NifL_{red}ADPNifA$	0.0025
$2-oxoglutarate$	9.99
NifA2oxo	0.00029

Sensitivity analysis

Figure 64 presents the result of sensitivity analysis for steady state point 6. Sensitivity of K_{d_1} remains the same as the last points. The sensitivity of K_{d_2} to $NifA(2OG)$ and $NifL_{red}ADP$, shows the same behavior as before. But sensitivity of K_{d_2} on $NifL_{red}ADPNifA$ increases in comparison to last points. As can be seen in Figure 64 the sensitivity of K_{d_3} on $NifA(2OG)$ jumps down from 1600 % change to 500 % change for a 95% decrease perturbation. Up to this point there was no evidence for sensitivity of K_{d_3} on $NifL_{red}ADP$ and $NifL_{red}ADPNifA$, but from this steady state point $NifL_{red}ADPNifA$ starts to show sensitivity on perturbation of K_{d_3}.

B.7 STEADY STATE POINT 7

In steady state point 7 the concentration of $2-oxoglutarate$ is 25 μM and the concentration of $NifA(2OG)$ which was used for parameter estimation was 0.0011 μM.

Parameter Estimation

Table 27 presents the result of parameter estimation for this steady state point. The estimated parameters are close to original ones.

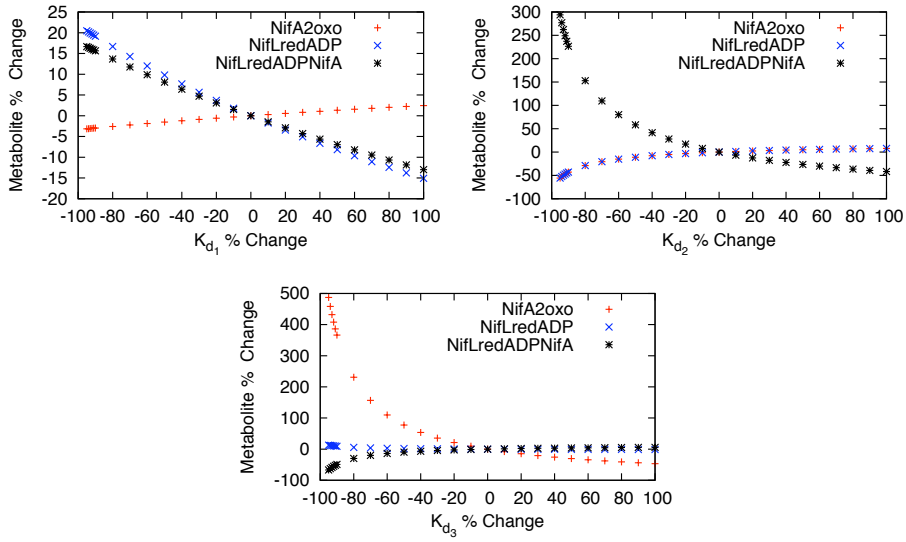


Figure 64: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 6.

Table 27: Estimated Parameters ($K_d/\mu M$) for steady state point 7.

K_d	Estimation	Experiment
K_{d_1}	16.15	13
K_{d_2}	0.010	0.03
K_{d_3}	134.42	57

Forward Modeling

Figure 65 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$. The estimated K_{d_3} for this point is closer than the estimated one in steady state point 6 and the simulation is in better agreement with experimental data for the case of $NifA(2OG)$ and simulations with the original parameters in the case of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$.

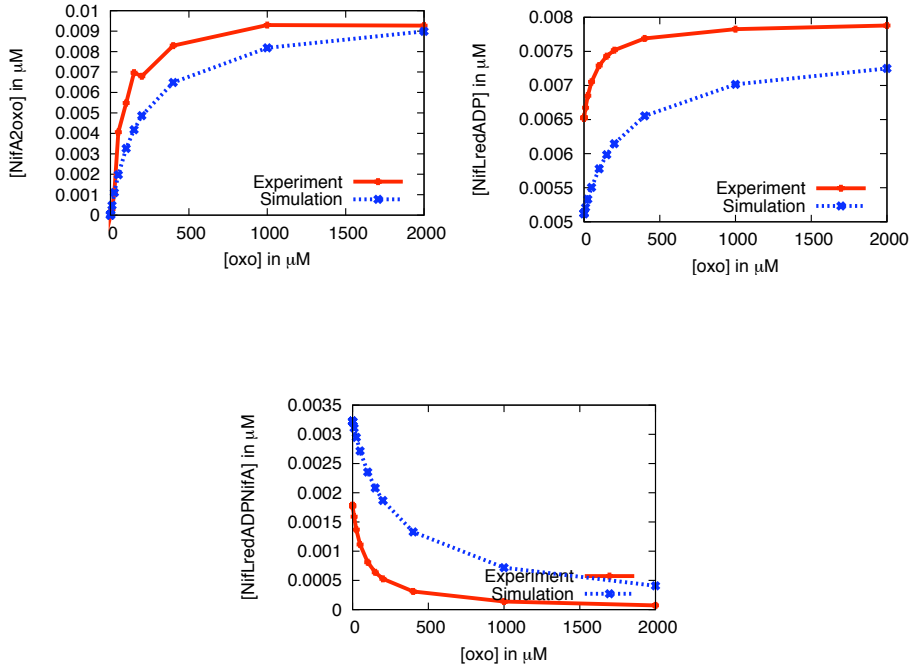


Figure 65: This figure demonstrates the simulation of $NifA(2OG)$, $NifL_{red}ADP$, and $NifL_{red}ADPNifA$with the estimated parameters of steady state point 7. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 7 shows that this system in this state is asymptotically stable using estimated parameters. Table 28 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 28: Concentration values $(C/\mu M)$ in steady state conditions for steady state point 7.

metabolite	value
$NifL_{red}$	0.0017
ADP	49.99
$NifL_{red}ADP$	0.0053
NifA	0.0059
$NifL_{red}ADPNifA$	0.0029
$2-oxoglutarate$	24.99
NifA2oxo	0.0011

Sensitivity analysis

In Figure 66 the result of sensitivity analysis in steady state point 7 is presented. Sensitivity of K_{d_1} remained the same as the last points. The sensitivity of K_{d_2} on $NifA(2OG)$ and $NifL_{red}ADP$, as always, shows the same behavior. Sensitivity of K_{d_2} on $NifL_{red}ADPNifA$ is raised in this steady state point. Sensitivity of K_{d_3} decreases on $NifA(2OG)$ and increases on $NifL_{red}ADPNifA$.

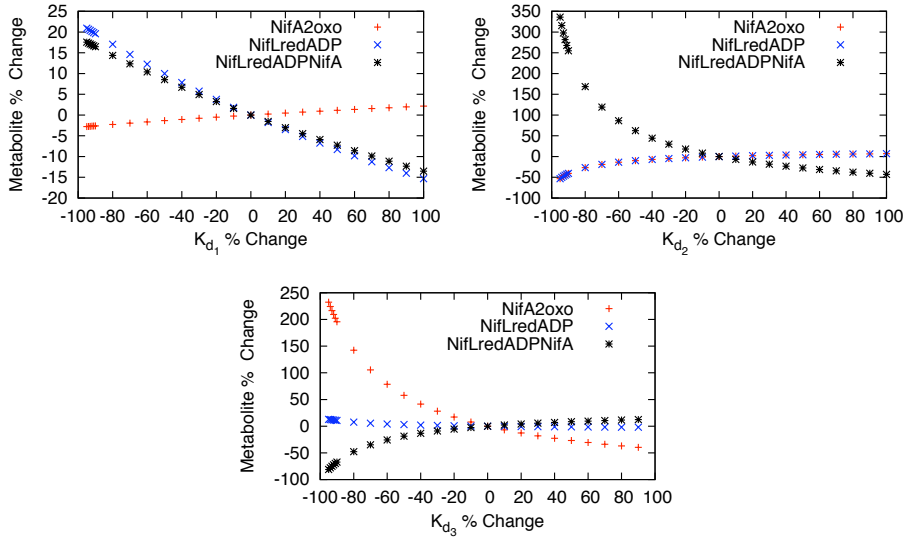


Figure 66: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in the steady state point 7.

B.8 STEADY STATE POINT 8

In steady state point 8, the concentration of $2-oxoglutarate$ is 50 μM and the concentration of $NifA(2OG)$ which was used for parameter estimation is 0.0040 μM .

Parameter Estimation

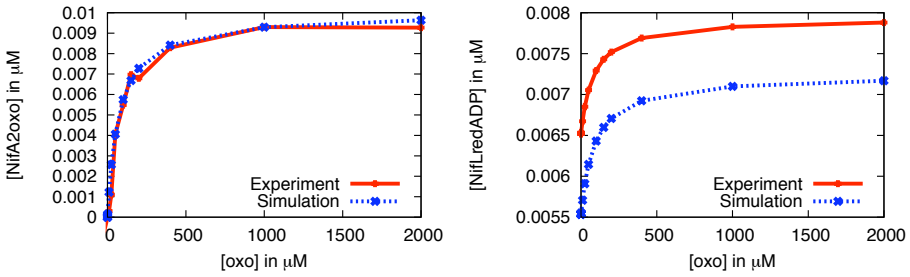
In the Table 29 the result of parameter estimation in this steady state point is presented. The estimated parameters are very close to original ones.

Table 29: Estimated Parameters ($K_d/\mu M$) for steady state point 8.

K_d	Estimation	Experiment
K_{d_1}	19.01	13
K_{d_2}	0.017	0.03
K_{d_3}	54.42	57

Forward Modeling

In the Figure 65, the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ are presented. The set of estimated parameters in this steady state point are in very agreement within the simulation and experimental data for $NifA(2OG)$ and in between simulations in $NifL_{red}ADPNifA$.



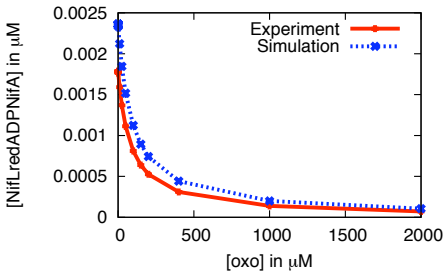


Figure 67: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 8. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 8 indicates that in this state the system is asymptotically stable using the estimated parameters. Table 30 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 30: Concentration values $(C/\mu M)$ in steady state conditions for steady state point 8.

metabolite	value
$NifL_{red}$	0.0023
ADP	49.99
$NifL_{red}ADP$	0.0061
NifA	0.0044
$NifL_{red}ADPNifA$	0.0015
$2-oxoglutarate$	49.99
NifA2oxo	0.0040

Sensitivity analysis

In Figure 68 the result of the sensitivity analysis in steady state point 8. The sensitivity of K_{d_1} remains the same as the last point. The sensitivity of K_{d_2} on $NifA(2OG)$ and $NifL_{red}ADP$ shows the same behavior as always.

The sensitivity of K_{d_2} on $NifL_{red}ADPNifA$ is raised in this steady state point. The sensitivity of K_{d_3} decreases on $NifA(2OG)$ and increases on $NifL_{red}ADPNifA$.

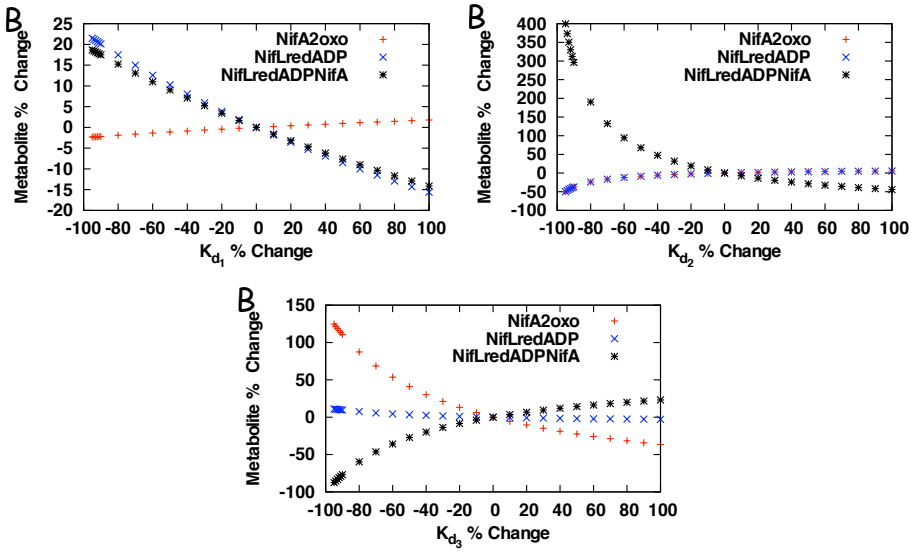


Figure 68: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 8.

B.9 STEADY STATE POINT 9

In steady state point 9 the concentration of $2-oxoglutarate$ is 50 μM and the concentration of $NifA(2OG)$ which was used for parameter estimation is 0.0054 μM .

Parameter Estimation

In the Table 31 the result of the parameter estimation in this steady state point is presented. The estimated parameters are very close to original ones.

Forward Modeling

Figure 69 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$. The simulations are in agreement with the experimental data in $NifA(2OG)$ and also simulation with the original parameters in $NifL_{red}ADPNifA$.

Table 31: Estimated Parameters ($K_d/\mu M$) for steady state point 9.

K_d	Estimation	Experiment
K_{d_1}	19.85	13
K_{d_2}	0.03	0.03
K_{d_3}	67.57	57

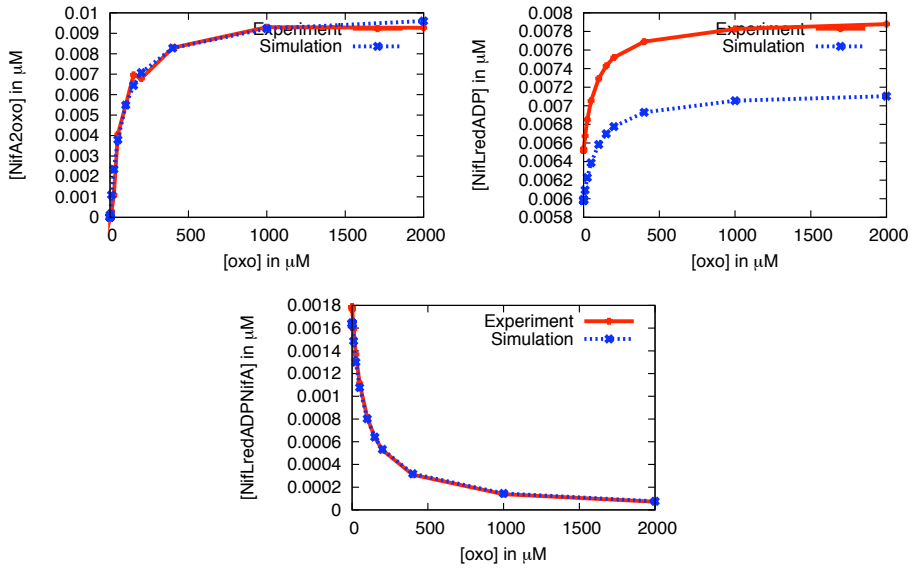


Figure 69: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 9.

Steady-state analysis

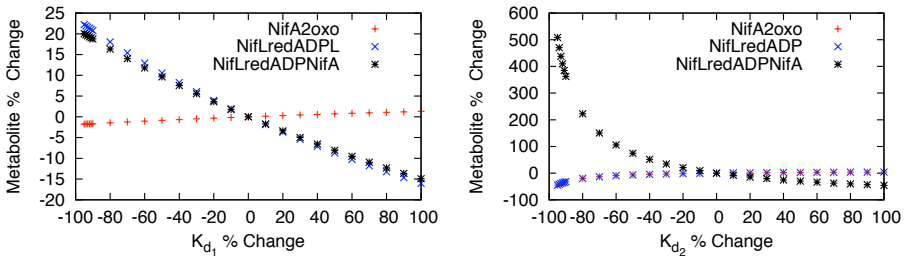
The stability analysis of $2-oxoglutarate$ system in steady state point 9 shows that this system in this state is asymptotically stable using the estimated parameters. Table 32 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 32: Concentration values $(C/\mu M)$ in steady state conditions for steady state point 9.

metabolite	value
$NifL_{red}$	0.0056
ADP	49.99
$NifL_{red}ADP$	0.0041
NifA	0.0016
$NifL_{red}ADPNifA$	0.00022
$2-oxoglutarate$	99.99
NifA2oxo	0.0081

Sensitivity analysis

Figure 70 presents the result of sensitivity analysis in steady state point 9. The behavior of sensitivity is similar to the last point. $NifL_{red}ADPNifA$ is very sensitive to K_{d_2}. In contrast to the last steady state points in which K_{d_3} was more sensitive to $NifA(2OG)$ than $NifL_{red}ADPNifA$, K_{d_3} is more sensitive to $NifL_{red}ADPNifA$ in comparison to $NifA(2OG)$.



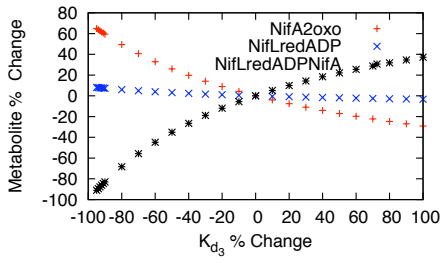


Figure 70: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in the steady state point 9.

B.10 STEADY STATE POINT 10

In steady state point 10 the concentration of $2-oxoglutarate$ is 100 μM and the concentration of $NifA(2OG)$ which was used for parameter estimation is 0.0069 μM.

Parameter Estimation

In Table 33 the result of parameter estimation in this steady state point is presented. The estimated parameters are close to original ones.

Table 33: Estimated Parameters ($K_d/\mu M$) for steady state point 10.

K_d	Estimation	Experiment
K_{d_1}	14.69	13
K_{d_2}	0.027	0.03
K_{d_3}	51.6	57

Forward Modeling

Figure 71 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$. The simulations are in agreement with the experimental data in $NifA(2OG)$ and simulation with the original parameters in $NifL_{red}ADPNifA$. It also shows closer simulations in $NifL_{red}ADP$.

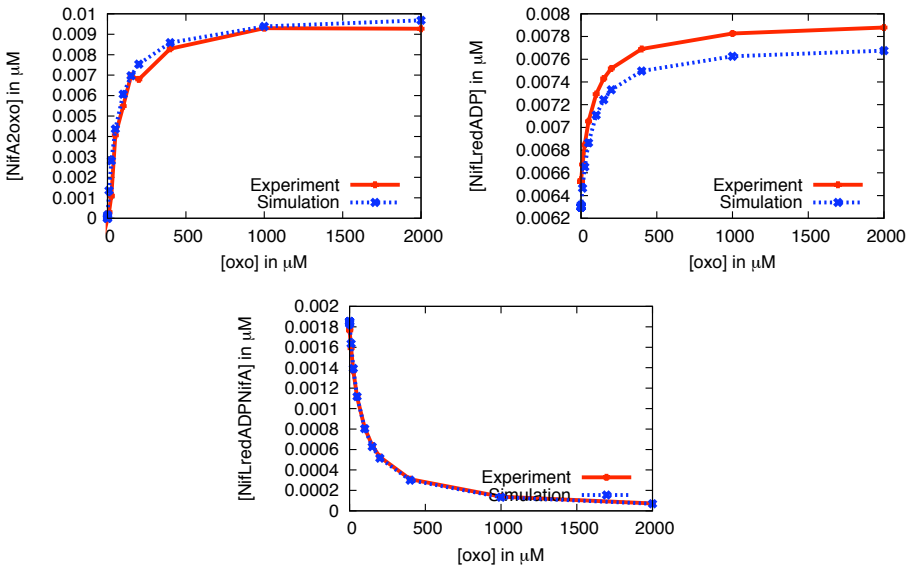


Figure 71: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 10. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Table 34: Concentration values $(C/\mu M)$ in steady state conditions for steady state point 10.

metabolite	value
$NifL_{red}$	0.0021
ADP	49.99
$NifL_{red}ADP$	0.0072
NifA	0.0024
$NifL_{red}ADPNifA$	0.00062
$2-$oxoglutarate	149.99
NifA2oxo	0.0069

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 10 shows that in this state the system is asymptotically stable the using estimated parameters. Table 34 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Sensitivity analysis

Figure 72 presents the result of the sensitivity analysis in steady state point 10. The behavior of sensitivity is similar to that of the last point.

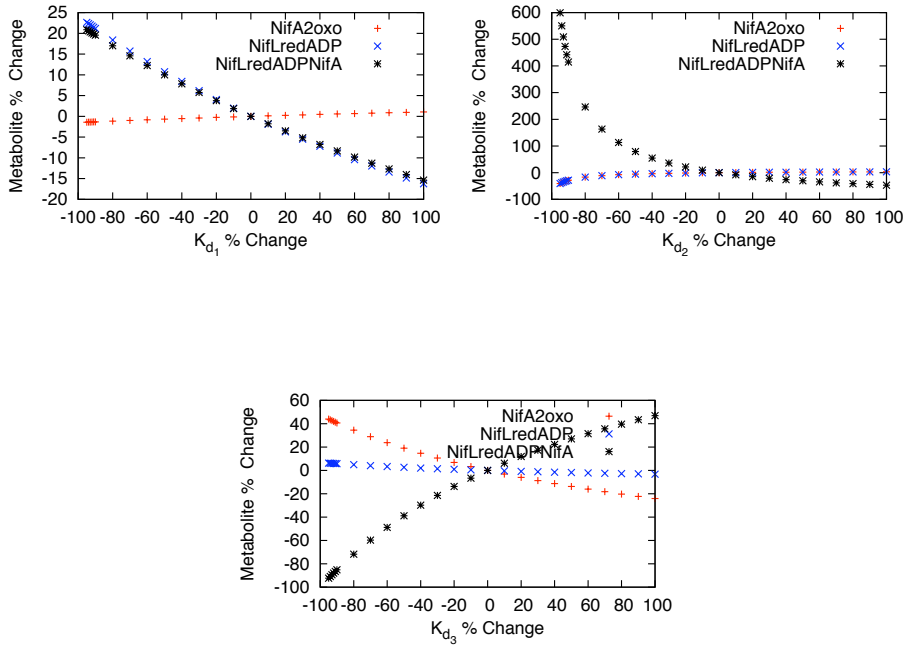


Figure 72: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 10.

B.11 STEADY STATE POINT 11

In steady state point 11 the concentration of $2-oxoglutarate$ is 200 μM and the concentration of $NifA2oxo$ which was used for parameter estimation is 0.0067 μM.

Parameter Estimation

Table 35 presents the result of the parameter estimation in this steady state point. The estimated parameters are close to original ones.

Table 35: Estimated Parameters ($K_d/\mu M$) for steady state point 11.

K_d	Estimation	Experiment
K_{d_1}	14.04	13
K_{d_2}	0.033	0.03
K_{d_3}	77.49	57

Forward Modeling

In Figure 73 the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ are presented. The estimated parameters in this steady state point are the best simulations amongst all the steady state points we checked.

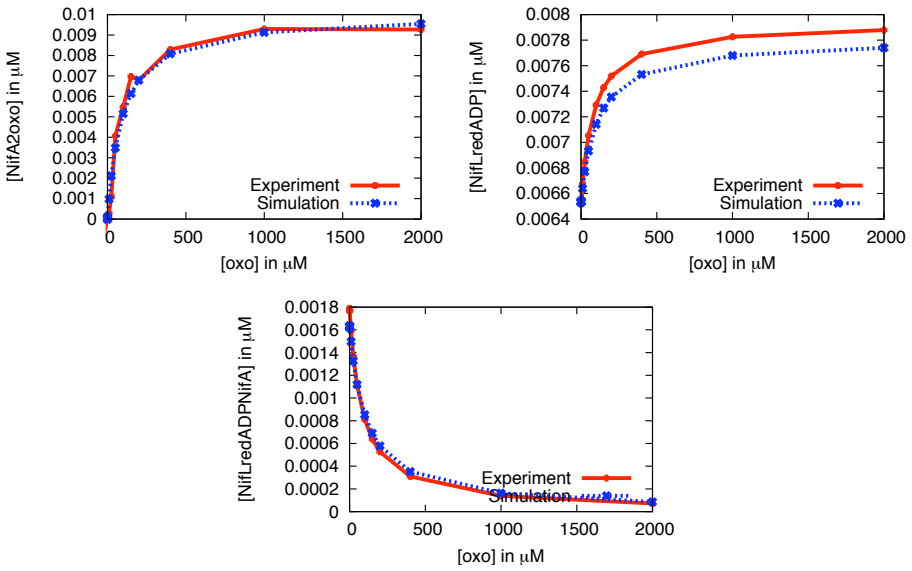


Figure 73: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 11. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 11 shows that at the corresponding state this system is asymptotically stable using estimated parameters. Table 36 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 36: Concentration values $(C/\mu M)$ in the steady state conditions for steady state point 11.

metabolite	value
$NifL_{red}$	0.002
ADP	49.99
$NifL_{red}ADP$	0.007
NifA	0.0026
$NifL_{red}ADPNifA$	0.00057
$2-oxoglutarate$	199.99
NifA2oxo	0.00678

Sensitivity analysis

Figure 74 presents the result of the sensitivity analysis in steady state point 11. The behavior of sensitivity is similar to last point. The sensitivity of $NifA(2OG)$ to perturbation of K_{d_3} is low for this steady state point.

B.12 STEADY STATE POINT 12

In steady state point 12 the concentration of $2-oxoglutarate$ is 400 μM and the concentration of $NifA(2OG)$ which was used for the parameter estimation is 0.0082 μM.

Parameter Estimation

Table 37 presents the result of parameter estimation in this steady state point. The estimated parameters are close to original ones.

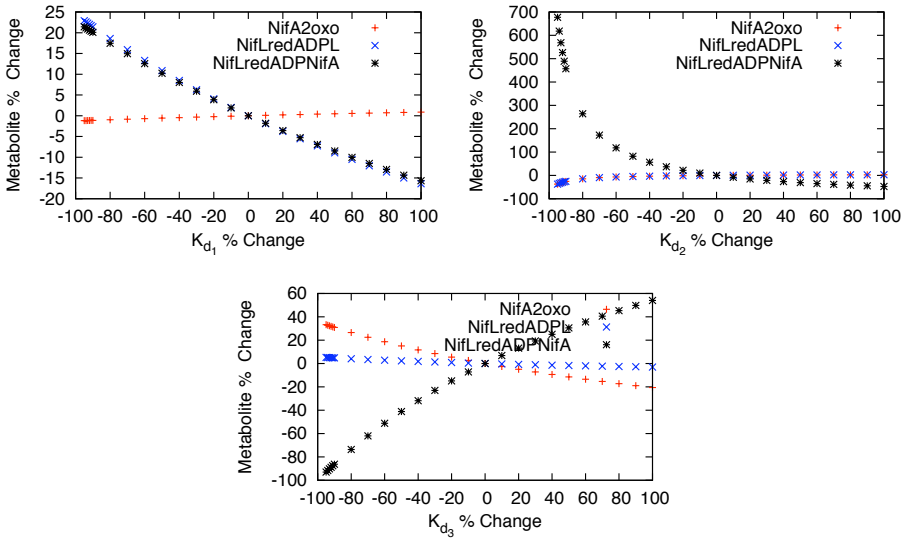


Figure 74: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 11.

Table 37: Estimated Parameters ($K_d/\mu M$) for steady state point 12.

K_d	Estimation	Experiment
K_{d_1}	9.38	13
K_{d_2}	0.045	0.03
K_{d_3}	69.9	57

Forward Modeling

Figure 75 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ for this steady state point. The simulations are in agreement with the data.

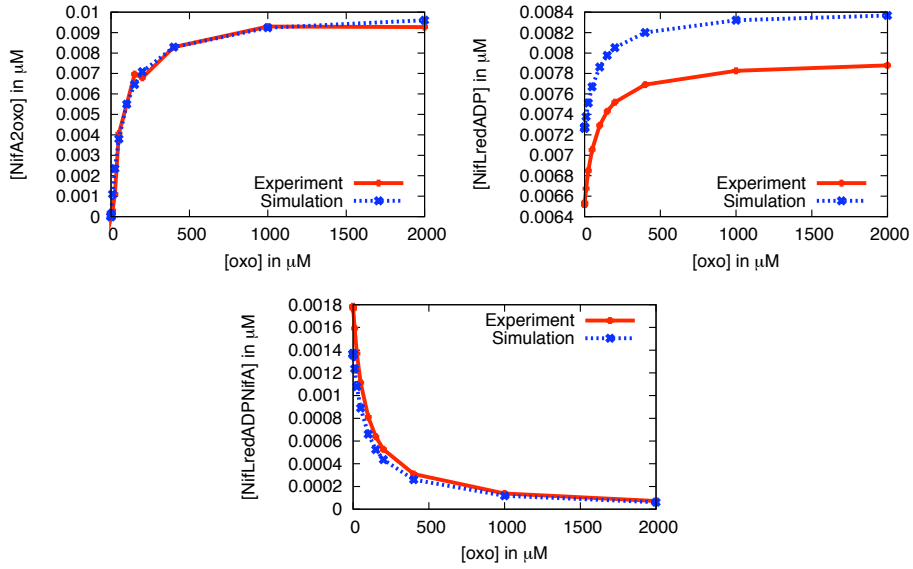


Figure 75: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 12. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 12 shows that at the corresponding state this system is asymptotically stable using estimated parameters. Table 38 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Sensitivity analysis

Figure 76 presents the result of the sensitivity analysis in steady state point 12. The behavior of sensitivity change of parameters is similar to that of the

Table 38: Concentration values (C/μM) in steady state conditions for steady state point 12.

metabolite	value
$NifL_{red}$	0.0015
ADP	49.99
$NifL_{red}ADP$	0.0082
NifA	0.0014
$NifL_{red}ADPNifA$	0.00025
$2-oxoglutarate$	399.99
NifA2oxo	0.0082

last point and the sensitivity of $NifA(2OG)$ to perturbation of K_{d_3} is very low.

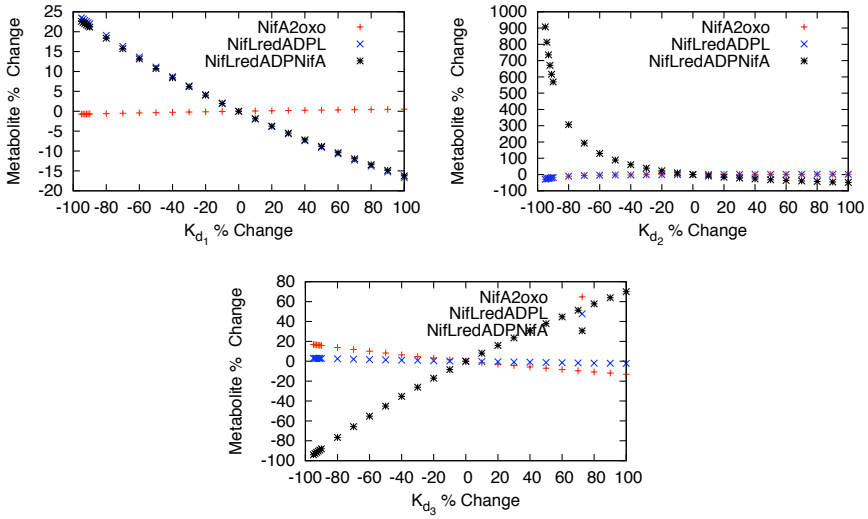


Figure 76: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 12.

B.13 STEADY STATE POINT 13

In steady state point 13 the concentration of $2-oxoglutarate$ is 1000 μM and the concentration of $NifA(2OG)$ which was used for the parameter estimation is 0.0092 μM.

Parameter Estimation

Table 39 presents the result of parameter estimation in this steady state point. The estimated parameters are close to original ones.

Table 39: Estimated Parameters ($K_d/\mu M$) for steady state point 13.

K_d	Estimation	Experiment
K_{d_1}	18.45	13
K_{d_2}	0.034	0.03
K_{d_3}	62.96	57

Forward Modeling

In the Figure 77 the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ for this steady state point are presented. The simulations are in agreement with the data.

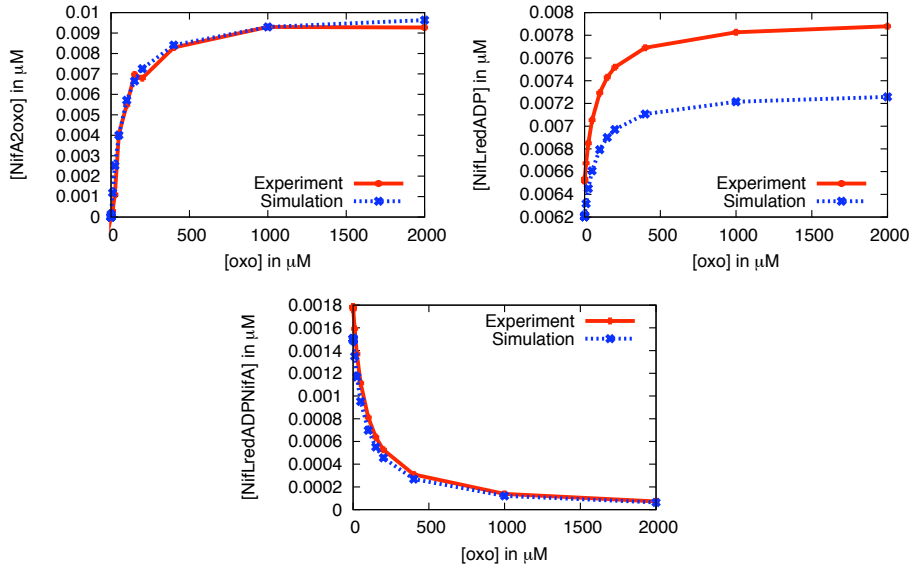


Figure 77: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 13. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 13 shows that at the corresponding state this system is asymptotically stable using estimated parameters. Table 40 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 40: Concentration values $(C/\mu M)$ in the steady state conditions for steady state point 13.

$metabolite$	$value$
$NifL_{red}$	0.0026
ADP	49.99
$NifL_{red}ADP$	0.0072
NifA	0.00058
$NifL_{red}ADPNifA$	0.00012
$2-oxoglutarate$	999.91
NifA2oxo	0.0092

Sensitivity analysis

Figure 78 presents the result of the sensitivity analysis in steady state point 13. The sensitivity behavior is similar to that of the last point. There is a slight change in the sensitivity of K_{d_1} since the sensitivity K_{d_1} on both of $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are exactly the same for this state. The sensitivity of K_{d_2} on $NifL_{red}ADPNifA$ is considerably increased. In this state, there is no evidence for sensitivity of K_{d_3} on $NifA(2OG)$ while there is a high sensitivity of K_{d_3} on $NifL_{red}ADPNifA$.

B.14 STEADY STATE POINT 14

In steady state point 14 the concentration of $2-oxoglutarate$ is 2000 μM and the concentration of $NifA(2OG)$ which was used for the parameter estimation is 0.0092 μM.

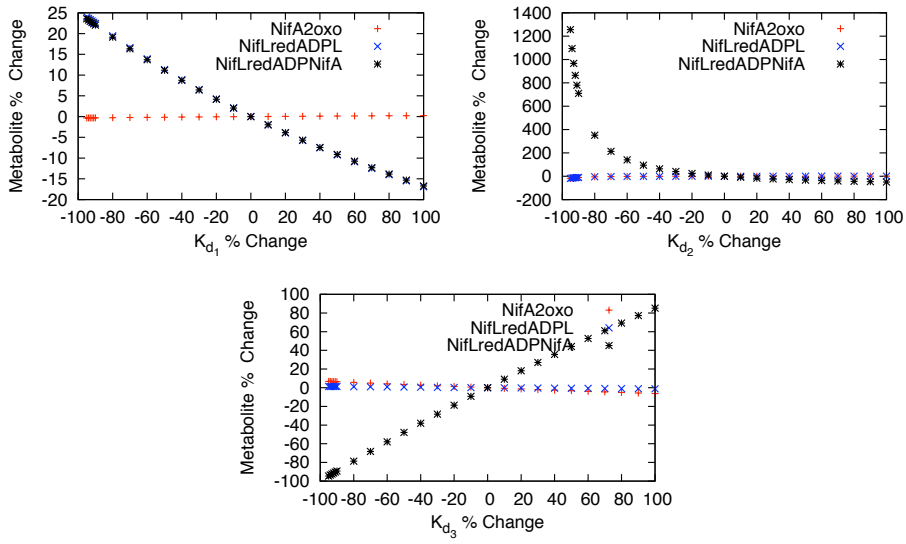


Figure 78: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 13.

Parameter Estimation

Table 41 presents the result of parameter estimation in this steady state point. The estimated parameters are close to the original ones.

Table 41: Estimated Parameters ($K_d/\mu M$) for steady state point 14.

K_d	Estimation	Experiment
K_{d_1}	18.08	13
K_{d_2}	0.019	0.03
K_{d_3}	115.64	57

Forward Modeling

Figure 79 presents the simulation of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ in this steady state point. The simulations are in agreement with the data.

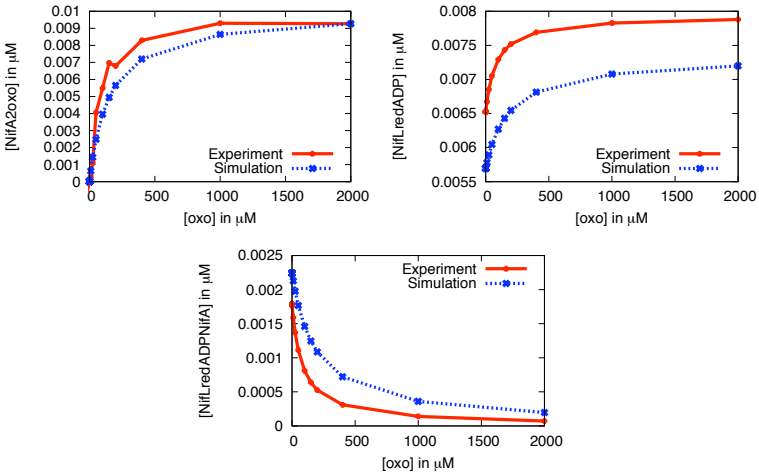


Figure 79: Simulations of $NifL_{red}ADPNifA$, $NifL_{red}ADP$, and $NifA(2OG)$ using estimated parameters from steady state poin 14. $NifA(2OG)$ is compared with the experimental data, and $NifL_{red}ADP$ and $NifL_{red}ADPNifA$ are compared to the simulation with the original parameters.

Steady-state analysis

The stability analysis of $2-oxoglutarate$ system in steady state point 14 shows that at the corresponding state this system is asymptotically stable using the estimated parameters. Table 42 presents the concentration of metabolites of $2-oxoglutarate$ system in this state.

Table 42: Concentration values $(C/\mu M)$ in the steady state conditions for steady state point 14.

metabolite	value
$NifL_{red}$	0.0026
ADP	49.99
$NifL_{red}ADP$	0.0071
NifA	0.0005
$NifL_{red}ADPNifA$	0.00019
$2-oxoglutarate$	1999.9
NifA2oxo	0.0092

Sensitivity analysis

Figure 80 presents the results of the sensitivity analysis in steady state point 14. The result is similar to the last point.

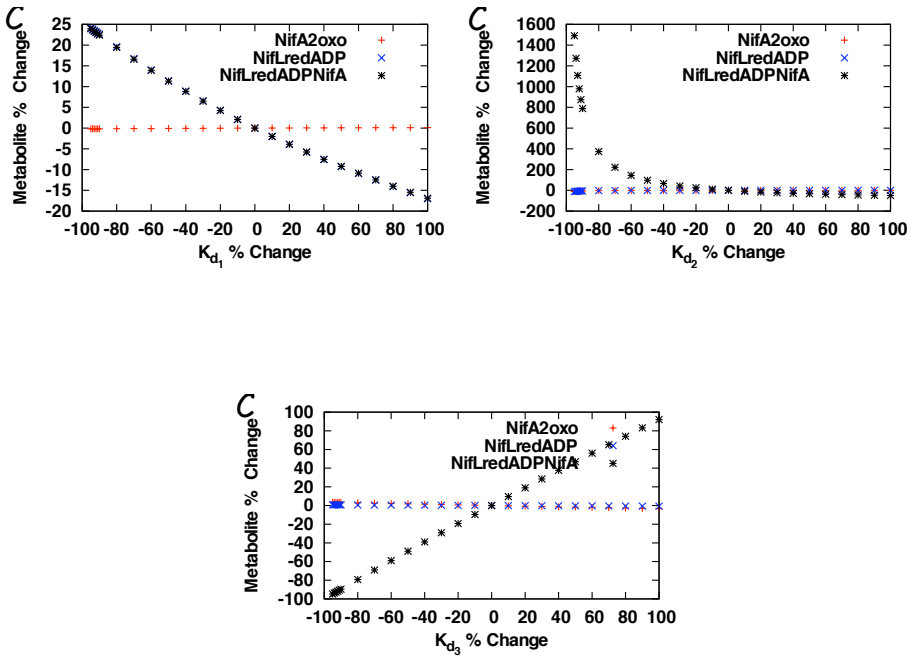


Figure 80: Sensitivity analysis of K_{d_1}, K_{d_2} , K_{d_3} in steady state point 14.

C

AVAILABLE DATA

Table 43

Reaction	Forward rate constant	Backward	Dissociation constant
A.1	5×10^7 $M^{-1}s^{-1}$	15 s^{-1}	10 μM
A.2	NA	NA	NA
A.3	NA	NA	$13 - 25$ μM
A.4	4.5×10^5 $M^{-1}s^{-1}$	0.02 s^{-1}	$0.1 - 0.23$ μM
A.5	NA	NA	57 μM
A.6	NA	NA	NA
A.7	NA	NA	NA
A.8	NA	NA	NA
A.9	NA	NA	NA
A.10	NA	NA	NA
A.11	NA	NA	NA
A.12	NA	NA	NA
A.13	NA	NA	NA
A.14	NA	NA	NA
A.15	NA	NA	NA
A.16	NA	NA	NA
A.17	NA	NA	NA
A.18	NA	NA	NA
A.19	NA	NA	NA
A.20	NA	NA	NA
A.21	NA	NA	NA
A.22	NA	NA	NA
A.23	NA	NA	NA
A.24	NA	NA	NA
A.25	NA	NA	NA
A.26	NA	NA	NA
A.27	NA	NA	NA
A.28	NA	NA	NA
A.29	NA	NA	NA
A.30	NA	NA	NA
A.31	NA	NA	NA
A.32	NA	NA	NA
A.33	NA	NA	NA
A.34	NA	NA	NA
A.35	NA	NA	NA
A.36	NA	NA	NA

BIBLIOGRAPHY

[1] U Alon. *An introduction to systems biology: design principles of biological circuits*, volume 10. Chapman and HallCRC, 2007. ISBN 1584886420. (Cited on pages 50, 51, 52, and 95.)

[2] R Alves, F Antunes, and A Salvador. Tools for kinetic modeling of biochemical networks. *Nat Biotechnol*, 24(6):667–672, 2006. ISSN 1087-0156 (Print). doi: 10.1038/nbt0606-667. (Cited on page 56.)

[3] B K Burgess and D J Lowe. Mechanism of molybdenum nitrogenase. *Chem Rev*, 96(7):2983–3012, 1996. ISSN 0009-2665 (Print). (Cited on pages 32, 33, 34, and 63.)

[4] C D Butler. Food security in the asia-pacific: Malthus, limits and environmental challenges. *Asia Pac J Clin Nutr*, 18(4):577–84, 2009. (Cited on page 29.)

[5] P Cabello, M D Roldán, and C Moreno-Vivián. Nitrate reduction and the nitrogen cycle in archaea. *Microbiology*, 150(Pt 11):3527–46, Nov 2004. doi: 10.1099/mic.0.27303-0. (Cited on page 29.)

[6] Q Cheng. Perspectives in biological nitrogen fixation research. *J Integr Plant Biol*, 50(7):786–98, Jul 2008. doi: 10.1111/j.1744-7909.2008.00700.x. (Cited on page 29.)

[7] Z Cui, A J Dunford, M C Durrant, R A Henderson, and B E Smith. Binding sites of nitrogenase: kinetic and theoretical studies of cyanide binding to extracted femo-cofactor derivatives. *Inorg Chem*, 42(20):6252–6264, 2003. ISSN 0020-1669 (Print). doi: 10.1021/ic030108q. (Cited on page 63.)

[8] D D Mendoza and R Dixon. Getting the signals across: networking inside and out. *Curr Opin Microbiol*, 8(2):113–115, 2005. ISSN 1369-5274 (Print). doi: 10.1016/j.mib.2005.02.017. (Cited on page 35.)

[9] R Dixon. The oxygen-responsive nifl-nifa complex: a novel two-component regulatory system controlling nitrogenase synthesis in gamma-proteobacteria. *Arch Microbiol*, 169(5):371–380, 1998. ISSN 0302-8933 (Print). (Cited on pages 39, 42, 91, and 101.)

[10] R Dixon and D Kahn. Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol*, 2(8):621–631, 2004. ISSN 1740-1526 (Print). doi: 10.1038/nrmicro954. (Cited on pages 32, 35, 39, 40, 91, 145, and 149.)

[11] Jeremy A Dodsworth and John A Leigh. Nifi inhibits nitrogenase by competing with fe protein for binding to the mofe protein. *Biochem Biophys Res Commun*, 364(2):378–82, Dec 2007. doi: 10.1016/j.bbrc.2007.10.020. (Cited on page 32.)

[12] M G Duyvis, H Wassink, and H Haaker. Nitrogenase of azotobacter vinelandii: kinetic analysis of the fe protein redox cycle. *Biochemistry*, 37(50):17345–17354, 1998. ISSN 0006-2960 (Print). doi: 10.1021/bi981509y. (Cited on pages 31 and 35.)

[13] M Dworkin, S Falkow, E Rosenberg, and K H Schleifer. *The Prokaryotes*, volume 2: Ecophysiology and Biochemistry. Springer, 2006. (Cited on pages 30 and 32.)

[14] M Dworkin, S Falkow, E Rosenberg, and K H Schleifer. *The Prokaryotes*, volume 1: Symbiotic Associations, Biotechnology, Applied Microbiology. Springer, 2006. (Cited on page 31.)

[15] M Dworkin, S Falkow, E Rosenberg, K H Schleifer, and E Stackebrandt. *The Prokaryotes*, volume 6 Proteobacteria - Gamma Subclass. Springer, 2006. (Cited on pages 31 and 35.)

[16] R R Eady and J R Postgate. Nitrogenase. *Nature*, 249(460):805–810, 1974. ISSN 0028-0836 (Print). (Cited on page 31.)

[17] R R Eady, D J Lowe, and R N Thorneley. Nitrogenase of klebsiella pneumoniae: a pre-steady state burst of atp hydrolysis is coupled to electron transfer between the component proteins. *FEBS Lett*, 95(2):211–213, 1978. ISSN 0014-5793 (Print). (Cited on page 34.)

[18] O V Eberhard. *Computational Analysis of Biochemical Systems*. Cambridge University Press, 2000. (Cited on page 46.)

[19] D Fell. *Understanding the Control of Metabolism*. Portland Press, 1997. (Cited on pages 46, 47, 48, and 49.)

[20] S J Ferguson. Nitrogen cycle enzymology. *Curr Opin Chem Biol*, 2(2): 182–93, Apr 1998. (Cited on page 29.)

[21] A Finney and M Hucka. Systems biology markup language: Level 2 and beyond. *Biochem Soc Trans*, 31(Pt 6):1472–1473, 2003. ISSN 0300-5127 (Print). doi: 10.1042/. (Cited on page 56.)

[22] K Fisher, D J Lowe, and R N Thorneley. klebsiella pneumoniae nitrogenase. the pre-steady-state kinetics of mofe-protein reduction and hydrogen evolution under conditions of limiting electron flux show that the rates of association with the fe-protein and electron transfer are independent of the oxidation level of the mofe-protein. *Biochem J*, 279 (Pt 1):81–85, 1991. ISSN 0264-6021 (Print). (Cited on page 63.)

[23] A Funahashi, A Jouraku, Y Matsuoka, and H Kitano. Integration of celldesigner and sabio-rk. *In Silico Biol*, 7(2 Suppl):S81–90, 2007. ISSN 1386-6338 (Print). (Cited on page 58.)

[24] J Gagneur and S Klamt. Computation of elementary modes: a unifying framework and the new binary approach. *BMC Bioinformatics*, 5:175, 2004. ISSN 1471-2105 (Electronic). doi: 10.1186/1471-2105-5-175. (Cited on page 54.)

[25] J N Galloway and E B Cowling. Reactive nitrogen and the world: 200 years of change. *Ambio*, 31(2):64–71, Mar 2002. (Cited on page 30.)

[26] C S Gillespie, D J Wilkinson, C J Proctor, D P Shanley, R J Boys, and T B L Kirkwood. Tools for the sbml community. *Bioinformatics*, 22(5):628–629, 2006. ISSN 1367-4803 (Print). doi: 10.1093/bioinformatics/btk042. (Cited on page 56.)

[27] Jens Glöer, Robert Thummer, Heike Ullrich, and Ruth A Schmitz. Towards understanding the nitrogen signal transduction for nif gene expression in klebsiella pneumoniae. *FEBS J*, 275(24):6281–94, Dec 2008. doi: 10.1111/j.1742-4658.2008.06752.x. (Cited on page 148.)

[28] C M Halbleib and P W Ludden. Regulation of biological nitrogen fixation. *J Nutr*, 130(5):1081–1084, 2000. ISSN 0022-3166 (Print). (Cited on pages 29 and 30.)

[29] T R Hawkes, D J Lowe, and B E Smith. Nitrogenase from klebsiella pneumoniae. an e.p.r. signal observed during enzyme turnover under ethylene is associated with the iron-molybdenum cofactor. *Biochem J*, 211(2):495–497, 1983. ISSN 0264-6021 (Print). (Cited on page 34.)

[30] M H Hefti, K Francoijs, S C D Vries, R Dixon, and J Vervoort. The pas fold. a redefinition of the pas domain based upon structural prediction. *Eur J Biochem*, 271(6):1198–1208, 2004. ISSN 0014-2956 (Print). (Cited on page 37.)

[31] S Hill, S Austin, T Eydmann, T Jones, and R Dixon. azotobacter vinelandii nifl is a flavoprotein that modulates transcriptional activation of nitrogen-fixation genes via a redox-sensitive switch. *Proc Natl Acad Sci U S A*, 93(5):2143–2148, 1996. ISSN 0027-8424 (Print). (Cited on pages 35 and 37.)

[32] Brian M Hoffman, Dennis R Dean, and Lance C Seefeldt. Climbing nitrogenase: toward a mechanism of enzymatic nitrogen fixation. *Acc Chem Res*, 42(5):609–19, May 2009. doi: 10.1021/ar8002128. (Cited on page 31.)

[33] S Hoops, S Sahle, R Gauges, C Lee, J Pahle, N Simus, M Singhal, L Xu, P Mendes, and U Kummer. Copasi–a complex pathway simulator. *Bioinformatics*, 22(24):3067–3074, 2006. ISSN 1460-2059 (Electronic). doi: 10.1093/bioinformatics/btl485. (Cited on pages 57 and 66.)

[34] J B Howard and D C Rees. Structural basis of biological nitrogen fixation. *Chem Rev*, 96(7):2965–2982, Nov 1996. ISSN 0009-2665 (Print). (Cited on pages 29 and 34.)

[35] J B Howard and D C Rees. How many metals does it take to fix n2? a mechanistic overview of biological nitrogen fixation. *Proc Natl Acad Sci U S A*, 103(46):17088–17093, 2006. ISSN 0027-8424 (Print). doi: 10.1073/pnas.0603978103. (Cited on page 30.)

[36] B D Howes, K Fisher, and D J Lowe. Nitrogenase of klebsiella pneumoniae: electron nuclear double resonance (endor) studies on the substrate reduction site. *Biochem J*, 297 (Pt 2):261–264, 1994. ISSN 0264-6021 (Print). (Cited on page 34.)

[37] M Hucka, A Finney, H M Sauro, H Bolouri, J Doyle, and H Kitano. The erato systems biology workbench: enabling interaction and exchange between software tools for computational biology. *Pac Symp Biocomput*, pages 450–461, 2002. ISSN 1793-5091 (Print). (Cited on pages 57 and 66.)

[38] M Hucka, A Finney, H M Sauro, H Bolouri, J C Doyle, H Kitano, A P Arkin, B J Bornstein, D Bray, A Cornish-Bowden, A A Cuellar, S Dronov, E D Gilles, M Ginkel, V Gor, I I Goryanin, W J Hedley, T C Hodgman, J-H Hofmeyr, P J Hunter, N S Juty, J L Kasberger, A Kremling, U Kummer, N L Novere, L M Loew, D Lucio, P Mendes, E Minch, E D Mjolsness, Y Nakayama, M R Nelson, P F Nielsen, T Sakurada, J C Schaff, B E Shapiro, T S Shimizu, H D Spence, J Stelling, K Takahashi, M Tomita, J Wagner, and J Wang. The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003. ISSN 1367-4803 (Print). (Cited on page 56.)

[39] K J Kauffman, P Prakash, and J S Edwards. Advances in flux balance analysis. *Curr Opin Biotechnol*, 14(5):491–496, 2003. ISSN 0958-1669 (Print). (Cited on page 54.)

[40] D B Kell. Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discov Today*, 11(23-24):1085–1092, 2006. ISSN 1359-6446 (Print). doi: 10.1016/j.drudis.2006.10.004. (Cited on page 68.)

[41] S Klamt and E D Gilles. Minimal cut sets in biochemical reaction networks. *Bioinformatics*, 20(2):226–234, 2004. ISSN 1367-4803 (Print). (Cited on page 54.)

[42] S Klamt and J Stelling. Two approaches for metabolic pathway analysis? *Trends Biotechnol*, 21(2):64–69, 2003. ISSN 0167-7799 (Print). (Cited on page 54.)

[43] S Klamt, J Stelling, M Ginkel, and E D Gilles. Fluxanalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps. *Bioinformatics*, 19(2):261–269, 2003. ISSN 1367-4803 (Print). (Cited on page 58.)

[44] S Klamt, J Gagneur, and A V Kamp. Algorithmic approaches for computing elementary modes in large biochemical reaction networks. *IEE Proc Syst Biol*, 152(4):249–255, 2005. (Cited on page 54.)

[45] S Klamt, J Saez-Rodriguez, J A Lindquist, L Simeoni, and E D Gilles. A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics*, 7:56, 2006. ISSN 1471-2105 (Electronic). doi: 10.1186/1471-2105-7-56. (Cited on page 54.)

[46] S Klamt, J Saez-Rodriguez, and E D Gilles. Structural and functional analysis of cellular networks with cellnetanalyzer. *BMC Syst Biol*, 1:2, 2007. ISSN 1752-0509 (Electronic). doi: 10.1186/1752-0509-1-2. (Cited on page 58.)

[47] C Kneip, P Lockhart, C Voss, and U Maier. Nitrogen fixation in eukaryotes–new models for symbiosis. *BMC Evol Biol*, 7:55, 2007. ISSN 1471-2148 (Electronic). doi: 10.1186/1471-2148-7-55. (Cited on page 30.)

[48] H Kurata, N Matoba, and N Shimizu. Cadlive for constructing a large-scale biochemical network based on a simulation-directed notation and its application to yeast cell cycle. *Nucleic Acids Res*, 31(14):4071–4084, 2003. ISSN 1362-4962 (Electronic). (Cited on page 58.)

[49] H Kurata, K Masaki, Y Sumida, and R Iwasaki. Cadlive dynamic simulator: direct link of biochemical networks to dynamic models. *Genome Res*, 15(4):590–600, 2005. ISSN 1088-9051 (Print). doi: 10.1101/gr.3463705. (Cited on page 58.)

[50] N L Novere. Model storage, exchange and integration. *BMC Neurosci*, 7 Suppl 1:S11, 2006. ISSN 1471-2202 (Electronic). doi: 10.1186/1471-2202-7-S1-S11. (Cited on page 56.)

[51] N L Novere, B Bornstein, A Broicher, M Courtot, M Donizelli, H Dharuri, L Li, H Sauro, M Schilstra, B Shapiro, J L Snoep, and M Hucka. Biomodels database: a free, centralized database of curated, published, quanti-

tative kinetic models of biochemical and cellular systems. *Nucleic Acids Res*, 34(Database issue):D689–91, 2006. ISSN 1362-4962 (Electronic). doi: 10.1093/nar/gkj092. (Cited on page 68.)

[52] S Y Lee. Systems biotechnology. *Genome Inform*, 23(1):214–6, Oct 2009. (Cited on page 165.)

[53] R Little, F Reyes-Ramirez, Y Zhang, W C V Heeswijk, and R Dixon. Signal transduction to the azotobacter vinelandii nifl-nifa regulatory system is influenced directly by interaction with 2-oxoglutarate and the pii regulatory protein. *EMBO J*, 19(22):6041–6050, 2000. ISSN 0261-4189 (Print). doi: 10.1093/emboj/19.22.6041. (Cited on pages 9, 40, 42, 83, 86, 92, 94, 101, 105, 119, 120, and 145.)

[54] R Little, V Colombo, A Leech, and R Dixon. Direct interaction of the nifl regulatory protein with the glnk signal transducer enables the azotobacter vinelandii nifl-nifa regulatory system to respond to conditions replete for nitrogen. *J Biol Chem*, 277(18):15472–15481, 2002. ISSN 0021-9258 (Print). doi: 10.1074/jbc.M112262200. (Cited on page 35.)

[55] 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. *J Biol Chem*, 278(31):28711–8, Aug 2003. doi: 10.1074/jbc.M301992200. (Cited on page 148.)

[56] D J Lowe and R N Thorneley. The mechanism of klebsiella pneumoniae nitrogenase action pre-steady-state kinetics of h2 formation. *Biochem J*, 224(3):877–886, 1984. ISSN 0264-6021 (Print). (Cited on pages 34, 63, 64, and 66.)

[57] D J Lowe and R N Thorneley. The mechanism of klebsiella pneumoniae nitrogenase action the determination of rate constants required for the simulation of the kinetics of n2 reduction and h2 evolution. *Biochem J*, 224(3):895–901, 1984. ISSN 0264-6021 (Print). (Cited on pages 8, 34, 63, 64, 65, and 75.)

[58] D J Lowe, K Fisher, and R N Thorneley. klebsiella pneumoniae nitrogenase mechanism of acetylene reduction and its inhibition by

carbon monoxide. *Biochem J*, 272(3):621–625, 1990. ISSN 0264-6021 (Print). (Cited on page 63.)

[59] D J Lowe, K Fisher, and R N Thorneley. klebsiella pneumoniae nitrogenase: pre-steady-state absorbance changes show that redox changes occur in the mofe protein that depend on substrate and component protein ratio; a role for p-centres in reducing dinitrogen? *Biochem J*, 292 (Pt 1):93–98, 1993. ISSN 0264-6021 (Print). (Cited on page 63.)

[60] I Martinez-Argudo, R Little, N Shearer, P Johnson, and R Dixon. The nifl-nifa system: a multidomain transcriptional regulatory complex that integrates environmental signals. *J Bacteriol*, 186(3):601–610, 2004. ISSN 0021-9193 (Print). (Cited on pages 31, 35, 37, 38, 39, 40, 42, 101, and 145.)

[61] P Mendes. Gepasi: a software package for modelling the dynamics, steady states and control of biochemical and other systems. *Comput Appl Biosci*, 9(5):563–571, 1993. ISSN 0266-7061 (Print). (Cited on page 56.)

[62] Z Michalewicz and D B Fogel. *How to solve it: modern heuristics*. Springer, Berlin, 2nd ed., rev. and extended ed edition, 2004. ISBN 3540224947 (acid-free paper). URL http://www.loc.gov/catdir/enhancements/fy0818/2004094686-d.html. (Cited on page 88.)

[63] T Money, J Barrett, R Dixon, and S Austin. Protein-protein interactions in the complex between the enhancer binding protein nifa and the sensor nifl from azotobacter vinelandii. *J Bacteriol*, 183(4):1359–1368, 2001. ISSN 0021-9193 (Print). doi: 10.1128/JB.183.4.1359-1368.2001. (Cited on pages 37 and 38.)

[64] W R Moomaw. Energy, industry and nitrogen: strategies for decreasing reactive nitrogen emissions. *Ambio*, 31(2):184–189, 2002. ISSN 0044-7447 (Print). (Cited on page 30.)

[65] T Ogura and A J Wilkinson. Aaa+ superfamily atpases: common structure–diverse function. *Genes Cells*, 6(7):575–597, 2001. ISSN 1356-9597 (Print). (Cited on page 37.)

[66] B Palsson. *Systems biology: properties of reconstructed networks*. Cambridge University Press, Cambridge, 2006. ISBN

0521859034. URL http://www.loc.gov/catdir/enhancements/fy0659/2005035008-d.html. (Cited on page 165.)

[67] S Perry, N Shearer, R Little, and R Dixon. Mutational analysis of the nucleotide-binding domain of the anti-activator nifl. *J Mol Biol*, 346(4): 935–949, 2005. ISSN 0022-2836 (Print). doi: 10.1016/j.jmb.2004.12.033. (Cited on pages 10, 99, 100, and 101.)

[68] W H Press, W T Vetterling, S A Teukolsky, and B P Flannery. *Numerical Recipes in C++: the art of scientific computing*. Cambridge University Press, v. 2.10 edition, 2002. ISBN 0521750377. URL http://www.loc.gov/catdir/enhancements/fy0668/2002564733-d.html. (Cited on page 53.)

[69] R J Prill, D Marbach, J Saez-Rodriguez, P K Sorger, L G Alexopoulos, X Xue, N D Clarke, G Altan-Bonnet, and G Stolovitzky. Towards a rigorous assessment of systems biology models: The dream3 challenges. *PLoS One*, 5(2):e9202, 2010. doi: 10.1371/journal.pone.0009202. (Cited on page 165.)

[70] C Reder. Metabolic control theory: a structural approach. *J Theor Biol*, 135(2):175–201, 1988. ISSN 0022-5193 (Print). (Cited on page 54.)

[71] D C Rees and J B Howard. Nitrogenase: standing at the crossroads. *Curr Opin Chem Biol*, 4(5):559–566, 2000. ISSN 1367-5931 (Print). (Cited on page 30.)

[72] D C Rees, F T Akif, C A Haynes, M Y Walton, S Andrade, O Einsle, and J B Howard. Structural basis of biological nitrogen fixation. *Philos Transact A Math Phys Eng Sci*, 363(1829):971–984, 2005. ISSN 1364-503X (Print). doi: 10.1098/rsta.2004.1539. (Cited on page 32.)

[73] O Resendis-Antonio, J L Reed, S Encarnación, J Collado-Vides, and B Ø Palsson. Metabolic reconstruction and modeling of nitrogen fixation in rhizobium etli. *PLoS Comput Biol*, 3(10):1887–95, Oct 2007. doi: 10.1371/journal.pcbi.0030192. (Cited on pages 29 and 30.)

[74] F Reyes-Ramirez, R Little, and R Dixon. Mutant forms of the azotobacter vinelandii transcriptional activator nifa resistant to inhi-

bition by the nifl regulatory protein. *J Bacteriol*, 184(24):6777–6785, 2002. ISSN 0021-9193 (Print). (Cited on pages 9, 35, 90, 91, and 92.)

[75] N Rodriguez, M Donizelli, and N L Novere. Sbmleditor: effective creation of models in the systems biology markup language (sbml). *BMC Bioinformatics*, 8:79, 2007. ISSN 1471-2105 (Electronic). doi: 10.1186/1471-2105-8-79. (Cited on page 68.)

[76] Paul Rudnick, Christopher Kunz, Malkanthi K Gunatilaka, Eric R Hines, and Christina Kennedy. Role of glnk in nifl-mediated regulation of nifa activity in azotobacter vinelandii. *J Bacteriol*, 184(3):812–20, Feb 2002. (Cited on page 149.)

[77] M Rudolf and P M H Kroneck. The nitrogen cycle: its biology. *Met Ions Biol Syst*, 43:75–103, 2005. ISSN 0161-5149 (Print). (Cited on page 29.)

[78] F Sánchez, L Cárdenas, and C Quinto. Biological nitrogen fixation and future challenges of agriculture. the endophytic connection. *Adv Exp Med Biol*, 464:107–15, 1999. (Cited on page 29.)

[79] H M Sauro, M Hucka, A Finney, C Wellock, H Bolouri, J Doyle, and H Kitano. Next generation simulation tools: the systems biology workbench and biospice integration. *OMICS*, 7(4):355–372, 2003. ISSN 1536-2310 (Print). doi: 10.1089/153623103322637670. (Cited on pages 57 and 66.)

[80] SBML. System biology markup language, 2007. URL http://www.sbml.org/index.psp. (Cited on pages 56 and 68.)

[81] W H Schlesinger. Carbon sequestration in soils: some cautions amidst optimism. *Agriculture Ecosystems & Environment*, 82(1-3):121–127, December 2000. (Cited on page 30.)

[82] L A Segel. *Biological kinetics*, volume 12. Cambridge University Press, Cambridge, 1991. ISBN 0521391849. URL http://www.loc.gov/catdir/description/cam024/90002465.html. (Cited on pages 46 and 47.)

[83] V Smil. *Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production*. MIT Press, Cambridge, Mass., 2001. ISBN 026219449X (hc : alk. paper). (Cited on page 30.)

[84] V Smil. Nitrogen and food production: proteins for human diets. *Ambio*, 31(2):126–131, 2002. ISSN 0044-7447 (Print). (Cited on page 30.)

[85] B E Smith, D J Lowe, and R C Bray. Nitrogenase of klebsiella pneumoniae: electron-paramagnetic-resonance studies on the catalytic mechanism. *Biochem J*, 130(2):641–643, 1972. ISSN 0264-6021 (Print). (Cited on page 34.)

[86] Jessica Stips, Robert Thummer, Melanie Neumann, and Ruth A Schmitz. Glnk effects complex formation between nifa and nifl in klebsiella pneumoniae. *Eur J Biochem*, 271(16):3379–88, Aug 2004. doi: 10.1111/j.1432-1033.2004.04272.x. (Cited on page 149.)

[87] L Stromback and P Lambrix. Representations of molecular pathways: an evaluation of sbml, psi mi and biopax. *Bioinformatics*, 21(24):4401–4407, 2005. ISSN 1367-4803 (Print). doi: 10.1093/bioinformatics/bti718. (Cited on pages 56 and 68.)

[88] L Stromback, V Jakoniene, H Tan, and P Lambrix. Representing, storing and accessing molecular interaction data: a review of models and tools. *Brief Bioinform*, 7(4):331–338, 2006. ISSN 1467-5463 (Print). doi: 10.1093/bib/bbl039. (Cited on pages 56 and 68.)

[89] R N Thorneley and R R Eady. Nitrogenase of klebsiella pneumoniae: evidence for an adenosine triphosphate-induced association of the iron-sulphur protein. *Biochem J*, 133(2):405–408, 1973. ISSN 0264-6021 (Print). (Cited on page 34.)

[90] R N Thorneley and D J Lowe. Nitrogenase of klebsiella pneumoniae. kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: identification of the rate-limiting step for substrate reduction. *Biochem J*, 215(2):393–403, 1983. ISSN 0264-6021 (Print). (Cited on pages 17, 34, 63, 64, and 74.)

[91] R N Thorneley and D J Lowe. The mechanism of klebsiella pneumoniae nitrogenase action pre-steady-state kinetics of an enzyme-bound intermediate in n2 reduction and of nh3 formation. *Biochem J*, 224(3):887–894, 1984. ISSN 0264-6021 (Print). (Cited on pages 34, 63, 64, and 66.)

[92] R N Thorneley and D J Lowe. The mechanism of klebsiella pneumoniae nitrogenase action. simulation of the dependences of h2-evolution rate on component-protein concentration and ratio and sodium dithionite concentration. *Biochem J*, 224(3):903–909, 1984. ISSN 0264-6021 (Print). (Cited on pages 34, 63, and 64.)

[93] D Tilman. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *Proc Natl Acad Sci U S A*, 96(11):5995–6000, May 1999. (Cited on page 29.)

[94] P M Vitousek, S Hattenschwiler, L Olander, and S Allison. Nitrogen and nature. *Ambio*, 31(2):97–101, 2002. ISSN 0044-7447 (Print). (Cited on page 29.)

[95] P E Wilson, A C Nyborg, and G D Watt. Duplication and extension of the thorneley and lowe kinetic model for klebsiella pneumoniae nitrogenase catalysis using a mathematica software platform. *Biophys Chem*, 91(3):281–304, 2001. ISSN 0301-4622 (Print). (Cited on page 79.)

[96] Z Zi, K H Cho, M H Sung, X Xia, J Zheng, and Z Sun. In silico identification of the key components and steps in ifn-gamma induced jak-stat signaling pathway. *FEBS Lett*, 579(5):1101–8, Feb 2005. doi: 10.1016/j.febslet.2005.01.009. (Cited on page 89.)

[97] Z Zi, Y Zheng, A E Rundell, and E Klipp. Sbml-sat: a systems biology markup language (sbml) based sensitivity analysis tool. *BMC Bioinformatics*, 9:342, 2008. doi: 10.1186/1471-2105-9-342. (Cited on pages 59, 90, and 97.)

[98] W G Zumft and L E Mortenson. The nitrogen-fixing complex of bacteria. *Biochim Biophys Acta*, 416(1):1–52, 1975. ISSN 0006-3002 (Print). (Cited on page 30.)

DECLARATION

Norwich, February 2010

Nima Pakseresht