Insights into the Mechanism of Biological Nitrogen Fixation through Characterization of the Nitrogenase Molybdenum-Iron Protein

Thesis by

Christine N. Morrison

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemistry

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2017

(Defended 1 May 2017)

Christine N. Morrison
ORCID: 0000-0002-4180-8407

ACKNOWLEDGMENTS

"Science does not happen in a vacuum", as the saying goes. The work presented in this thesis would not have been possible without the support, time, and energy of many people. My thesis committee, including Professor Douglas C. Rees (my research advisor), Professor Jonas C. Peters (the chair of my committee), Professor William M. Clemons, and Professor Harry B. Gray, provided unwavering support and constructive feedback throughout my Ph.D. Many skills I learned from Dr. Limei Zhang, who mentored me during the first few years of my Ph.D. I also worked with Dr. Julie Hoy and Dr. Jens Kaiser to learn X-ray crystallography – a core technique of my thesis. "Team Nitro" (Dr. Limei Zhang, Dr. Thomas Spatzal, Dr. Kathryn Perez, Dr. Helen Segal, Belinda Wenke, Renee Arias, and Camille McAvoy) assisted with experimental design and execution, and they made lab a fun place to be. The Rees lab was a great group to be a part of as everyone was positive and quick to help when possible. There were many other people at Caltech who supported me in various ways during my Ph.D., including Agnes Tong, Felicia Hunt, Sue Chiarchiaro, Joe Drew, Anne Penney, Jeff Groseth, Dr. Angelo Di Bilio, and Dr. Paul Oyala. My work was financially supported by fellowships from the National Science Foundation Graduate Research Fellowship Program, the American Association of University Women, and the Philanthropic Educational Organization.

My family, including my partner, parents, brother, grandparents, cousins, aunts and uncles, Deborah Castleman, and friends, provided tremendous support over the years. There were also several communities that became like family to me, including my speed skating team, my scuba diving buddies, my student affairs team, the Chemistry Graduate Studies Committee, and my housemates.

I would like to thank everyone for their support, time, and energy.

You have my deepest gratitude.

ABSTRACT

Nitrogen fixation, the process of converting dinitrogen to ammonia, is performed industrially and biologically by the Haber-Bosch process and nitrogenase, respectively. The resulting ammonia is largely used as fertilizer. Since there is a finite amount of ammonia produced by nitrogenase, we are heavily dependent on the Haber-Bosch process – only two-fifths of the world's population could be fed without it. Although the importance of the Haber-Bosch process cannot be overstated, our dependence on it has several drawbacks, including significant energy costs (~5% of the annual natural gas consumption), greenhouse gas emissions, and nitrate runoffs. By understanding the biological mechanism of nitrogen fixation, we may be able to (1) develop more efficient nitrogen fixing catalysts to replace those in the Haber-Bosch process or (2) express *de novo* nitrogen fixing proteins in plants so crops can essentially fertilize themselves. The projects described in this thesis aim to contribute to our understanding of the mechanism of biological nitrogen fixation through structural studies of nitrogenase. Nitrogenase consists of the iron and molybdenum-iron (MoFe) proteins, the latter of which contains the active site, the FeMo-cofactor. Throughout my work, I compare the MoFe proteins from A*zotobacter vinelandii* (Av1) and *Clostridium pasteurianum* (Cp1), the two most structurally divergent molybdenum nitrogenases known. Determining the similarities and differences between these proteins may aid our understanding of biological nitrogen fixation. My first project (Chapter III) compares a 1.08 Å Cp1 X-ray structure to a previously published 1.0 Å Av1 structure. I determined that the center atom of the Cp1 FeMo-cofactor is carbon, showing conservation of cofactor structure among molybdenum nitrogenases. Next, I compared substrate pathways in Av1 and Cp1 via Xe pressurization and identification of small molecule binding sites (Chapter IV). My most significant results include the structural and electronic characterization of a reversible protonated resting state of Av1 and Cp1 (Chapter VII).

Published Content and Contributions

Zhang, L.-M., Morrison, C. N., Kaiser, J. T. & Rees, D. C. Nitrogenase MoFe Protein from *Clostridium pasteurianum* at 1.08 Å Resolution: Comparison with the *Azotobacter vinelandii* MoFe Protein. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71,** 274–282 (2015). doi: 10.1107/S1399004714025243.

C. N. M. participated in the conception of the project, solved and analyzed the crystal structures, prepared some of the data, and participated in the writing of the manuscript.

Morrison, C. N., Hoy, J. A., Zhang, L., Einsle, O. & Rees, D. C. Substrate Pathways in the Nitrogenase MoFe Protein by Experimental Identification of Small Molecule Binding Sites. *Biochemistry* **54,** 2052–2060 (2015). doi: 10.1021/bi501313k.

C. N. M. led the project, collected some of the X-ray data, solved and analyzed the crystal structures, prepared the data, and wrote the manuscript.

Morrison, C. N., Spatzal, T., & Rees, D. C. Reversible Protonated Resting State of the Nitrogenase Active Site. Reproduced with permission from *Journal of the American Chemical Society*. Copyright 2017 American Chemical Society.

C. N. M. led the project, collected the X-ray data, solved and analyzed the crystal structures, prepared EPR samples, collected and analyzed EPR data, prepared the data, and wrote the manuscript.

TABLE OF CONTENTS

ii Copyright

iii Acknowledgments

iv Abstract

v Published Content

vi Table of Contents

vii Key Nomenclature

1 **Chapter I**: Conceptual Framework

9 **Chapter II**: Methodology

38 **Chapter III**: Nitrogenase MoFe Protein from *Clostridium pasteurianum* at 1.08 Å Resolution: Comparison to the *Azotobacter vinelandii* MoFe Protein

57 **Chapter IV**: Substrate Pathways in the Nitrogenase MoFe Protein by Experimental Identification of Small Molecule Binding Sites

97 **Chapter V**: Thiocyanate and Methylamine Binding Sites in the MoFe Protein

105 **Chapter VI**: Progress toward determining the Location of Hydrogen Atoms in the MoFe Protein using Neutron Diffraction

115 **Chapter VII**: Reversible Protonated Resting State of the Nitrogenase Active Site

139 Concluding Perspectives

140 References

KEY NOMENCLATURE

ADP: Adenosine diphosphate, a nucleotide, is the product of ATP hydrolysis. ADP temporarily exists in two binding sites in the Fe protein until it can be replaced with ATP.

ATP: Adenosine triphosphate is a nucleotide. There are two ATP binding sites in the Fe protein. During nitrogen reduction, 16 ATP molecules are hydrolyzed to ADP. The resulting energy drives electron transfer between the Fe and MoFe proteins.

Av1: The MoFe protein from *Azotobacter vinelandii.*

Av2: The Fe protein from *Azotobacter vinelandii.*

Azotobacter vinelandii: An aerobic soil bacterium that expresses nitrogenase; the most commonly studied species of nitrogenase in structural, spectroscopic, and biosynthetic investigations.

***Clostridium pasteurianum*:** An anaerobic soil bacterium that expresses nitrogenase; this species of C*lostridium* is nonpathogenic.

Cp1: The MoFe protein from *Clostridium pasteurianum.*

Cp2: The Fe protein from *Clostridium pasteurianum.*

EPR: In this work, electron paramagnetic resonance (EPR) spectroscopy is used to probe the electronic structure of the FeMo-cofactor.

Fe protein: The iron protein of nitrogenase; a homodimeric protein containing one [4Fe:4S] cluster and two nucleotide binding sites (one in each subunit).

FeMo-cofactor: A [7Fe:Mo:9S:C:*R*-homocitrate] cluster that resides in the alpha subunit of the MoFe protein and is the active site of nitrogenase.

Klebsiella pneumoniae: An aerobic soil bacterium that expresses nitrogenase.

Kp1: The MoFe protein from *Klebsiella pneumoniae.*

Kp2: The Fe protein from *Klebsiella pneumoniae.*

MoFe protein: The molybdenum-iron protein of nitrogenase; a heterodimeric protein in which each dimer contains one [8Fe:7S] P-cluster and one [7Fe:Mo:9S:C:*R*-homocitrate] FeMo-cofactor.

Neutron crystallography: A technique used to determine the atomic structure of crystallized molecules based on the diffraction neutrons from the nuclei of atoms in the molecules.

P-cluster: A [8Fe:7S] cluster that resides at the interface of the alpha and beta subunits in the MoFe protein. This cluster receives electrons from the [4Fe:4S] cluster in the Fe protein and transfers them to the FeMo-cofactor.

X-ray crystallography: A technique used to determine the atomic structure of crystallized molecules based on the diffraction of X-rays off of electrons in the molecule.

Chapter I

Conceptual Framework

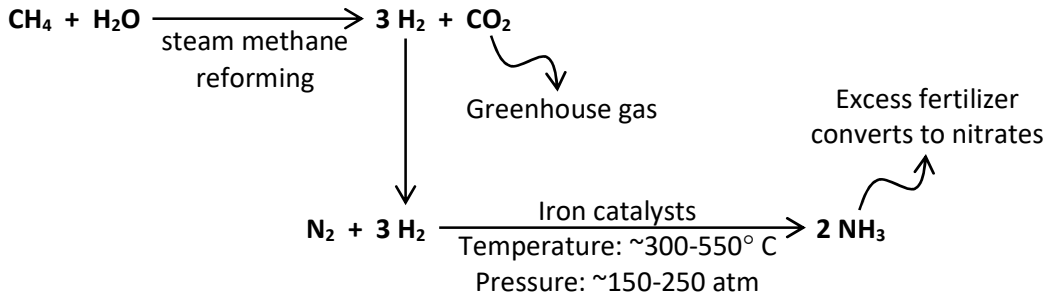
I.A. Motivation

The first mention of nitrogenase in published scientific work is in 1934.[1,2] Over 80 years later, its mysteries are being probed with ever increasing vigor. This enzyme catalyzes the reduction of atmospheric dinitrogen to its bioavailable form, ammonia, in a process called nitrogen fixation. Nitrogenase has intrigued so many scientists over the years because of its role in the nitrogen cycle, its complex structure and function, and its potential impact on sustainable agriculture. Its complexity is illustrated by the fact that after decades of work, the mechanism of biological nitrogen fixation remains enigmatic. The work described in this thesis aims to contribute to our understanding of the mechanism of biological nitrogen fixation through studies of the nitrogenase component proteins that (1) compare and contrast the structure of two different species of nitrogenase to understand what structural features are conserved and may therefore be mechanistically important, (2) identify small molecule binding sites within the proteins as a means of elucidating potential pathways and other functionally relevant aspects of the enzyme, and (3) investigate the atomic and electronic structure of a protonated resting state of the active site to learn about the protonation events on and near the active site during catalytic turnover.

A complete understanding of the biological nitrogen fixation mechanism has the potential to greatly impact sustainable agriculture. Quoting Vaclav Smil, environmental scientist, policy analyst, and author of Enriching the Earth: Fritz Haber, Carl Bosch, and the

Transformation of World Food Production, "Agriculture's principal objective is the production of digestible nitrogen".[3] Digestible nitrogen – ammonia – has been manufactured by the Haber-Bosch process since its invention in the early 1900's. Today, the Haber-Bosch process (Scheme I-1) yields about 275 billion pounds of ammonia per year,[4] the vast majority of which is used in synthetic fertilizers. Less than two-fifths of the current world population could be fed without synthetic nitrogen fertilizers, making the Haber-Bosch process a chief component of agriculture.[3] It is estimated that roughly half of the nitrogen atoms in a human body originate from biological nitrogen fixation and the other half from synthetic fertilizers.[3] While the importance of synthetic fertilizers can be overstated, we face the problem of the unsustainable energetic and environmental costs of this process. For example, the Haber-Bosch process is responsible for ~5% of the world's annual consumption of natural gas, whose steam reforming releases greenhouse gases.[3,4] Also, excess fertilizer that is not incorporated into crops leads to nitrate runoffs resulting in algal blooms that can be toxic.[5,6]

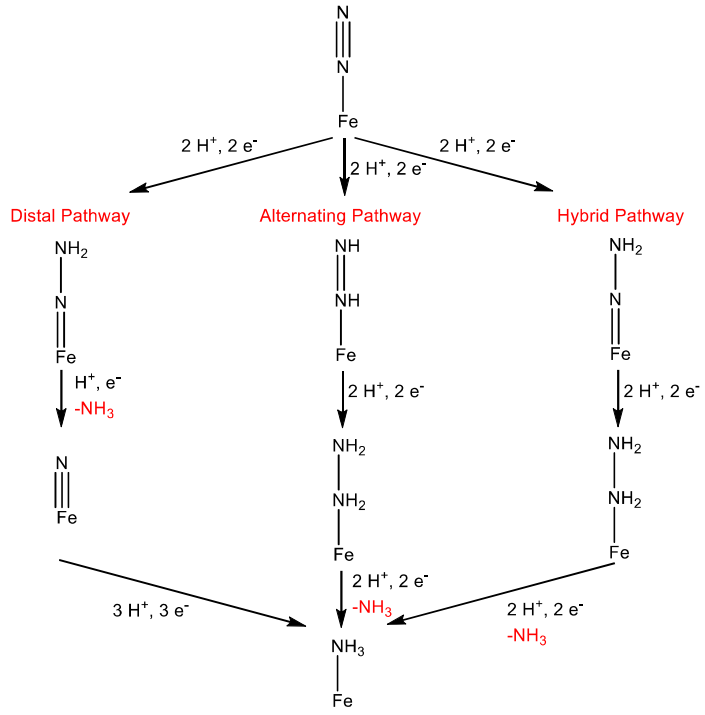
Scheme I-1. Outline of synthetic ammonia production via the Haber-Bosch process



We may be able to lessen our dependence on the Haber-Bosch process by understanding how biological nitrogen fixation works and incorporating this knowledge into new industrial catalysts or genetic engineering of crops to express *de novo* nitrogen-fixing

proteins that allow plants to make their own ammonia. Regarding the former, some synthetic inorganic groups are exploring non-biological nitrogen-fixing catalysts. For example, the formation and isolation of intermediates in small molecule nitrogen fixing catalysts suggests that dinitrogen reduction occurs by a mixed alternating and distal mechanism (Scheme I-2).[7,8] Related studies have explained the importance of a labile bond to the iron, which ecludiates the purpose of the interstitial carbide in the FeMo-cofactor.[9,10]

Scheme I-2. Distal and alternating pathways for dinitrogen reduction as well as the hybrid crossover pathway.



What we currently understand about nitrogenase has already inspired creative approaches to sustainable ammonia production.[11] For example, scientists are using our knowledge of the nitrogenase biosynthetic pathway[12] to engineer the enzyme in cereal crops.[13–15] A hurdle in expressing functional nitrogenase in non-native species is ensuring that all of the

20+ proteins required to yield active nitrogenase are artificially present in the genome and properly expressed in an anaerobic cellular environment.[12,13] As an alternative, imagine an engineered nitrogen-fixing protein with a simpler genome. Such a protein would likely consist of an iron-sulfur cluster similar to the active site of nitrogenase. Work is being done elsewhere on the self-assembly of proteins;[16–18] however, a thorough understanding of the nitrogenase mechanism could inform us on ideal geometry and coordination of metal centers and the nearby protein environment of *de novo* nitrogen-fixing proteins. Engineering such a protein for expression in crops would essentially allow plants to provide themselves with fertilizer; thus, reaching a state of sustainable agriculture.

I.B. Structure of nitrogenase

Nitrogenase (Figure I-1) is an enzyme consisting of two proteins.[19] The simpler of the two is the homodimeric iron (Fe) protein, which has a mass of about 60 kDa and contains a [4Fe:4S] cubane cluster at the interface of the two subunits as well as two nucleotide binding sites (one in each subunit) that bind ATP.[20] The molybdenum-iron (MoFe) protein, about 240 kDa, is a heterodimer containing two copies each of α- and β-subunits. In each αβ-dimer of the MoFe protein, there is one [8Fe:7S] "P-cluster" at the αβ-subunit interface and one [7Fe:Mo:9S:C:*R*-homocitrate] "FeMo-cofactor" buried in the α-subunit, the latter of which is the active site.[21,22] Additionally, there is a mono-nuclear iron site, named Fe16, at the two interfaces of the β- and β'-subunits. The purpose of Fe16 is not known.[23]

Nitrogenase proteins are named based on their bacterial species and the protein component number, where component 1 and 2 refer to the MoFe and Fe proteins, respectively. For example, the MoFe and Fe proteins expressed in *Clostridium pasteurianum* are called Cp1

and Cp2, respectively. Nitrogenases are divided into six groups based on their structure: (1) MoFe proteins homologous to Av1, (2) MoFe proteins homologous to Cp1, (3 and 4) MoFe proteins with small insertions or deletions in the α- and β-subunits, (Anf) FeFe proteins, and (Vnf) FeV proteins.[24] FeFe and FeV refer to component 1 proteins containing Fe or V at the molybdenum site in the FeMo-cofactor. The focus of this work is on MoFe proteins from groups 1 and 2. Most structural, spectroscopic, and biosynthetic investigations of nitrogenase are performed with Av nitrogenase. Cp1 differs by two insertion/deletion chains of about 50 residues each; it has a sequence identity of ~36% compared to Av1. The cofactors are conserved both in structure and location, as is most of the protein secondary structure elements. The similarities and differences between Cp1 and Av1 may shed light on the features important to the mechanism of biological nitrogen fixation and is fundamental to the work presented herein.

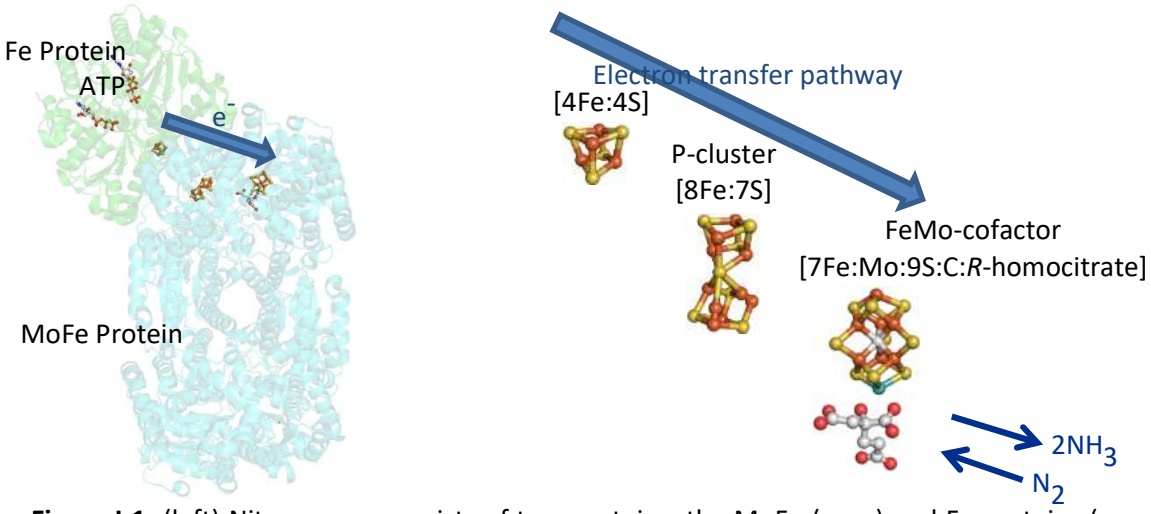


Figure I-1. (left) Nitrogenase consists of two proteins, the MoFe (cyan) and Fe proteins (green). During turnover, ATP hydrolyzes and electrons transfer from the [4Fe:4S] cluster in the Fe protein to the P-cluster and FeMo-cofactor in the MoFe protein (right). Substrates bind and are reduced at the FeMo-cofactor. For simplicity, only one heterodimer of the MoFe protein is illustrated with cofactors and Fe protein binding. The proteins are shown in cartoon representation, and the cofactors are shown in ball-and-stick models and colored by element (yellow: S; orange: Fe; gray: C; cyan: Mo; red: O).

I.C. Mechanism of biological nitrogen fixation

Scheme I-3. Standard model of biological nitrogen fixation

$$N_2 + 8\,H^+ + 8\,e^- + 16\,ATP \rightarrow 2\,NH_3 + H_2 + 16\,(ADP + P_i)$$

The standard model of biological nitrogen fixation is given in Scheme I-3. For substrate reduction to occur, sufficient buildup of electrons and protons (including bridging hydrides) at the FeMo-cofactor is required. This is achieved through the following steps:[25,26]

(1) The Fe protein binds with the MoFe protein by hinging at the [4Fe:4S] cluster and gripping the MoFe protein so that the shortest possible distance between the [4Fe:4S] cluster and the P-cluster is achieved.[27,28] Structural studies on Av1 and Av2 complexes with no nucleotide, MgADP, and MgAMPPCP suggest that the Fe protein may rock across the Av1-Av2 interface during the process of electron transfer.[28]

(2) One electron transfers from the [4Fe:4S] cluster to the P-cluster upon hydrolysis of two ATP molecules located in the Fe protein.[29–32]

(3) The Fe protein and MoFe protein dissociate from each other.

(4) The Fe protein exchanges two ADP for two ATP and is re-reduced by ferredoxin or flavodoxin.[29]

(5) In the MoFe protein, electrons transfer from the P-cluster to the FeMo-cofactor, and substrates are reduced at the FeMo-cofactor.

This process repeats eight times to reduce dinitrogen.[33] Six of these electrons reduce dinitrogen to ammonia, and two are used for hydrogen evolution that occurs upon substrate binding to the FeMo-cofactor.[34,35] Although it is unknown exactly where substrates are bound and reduced, the nitrogenase scientific community generally favors an iron-based reduction mechanism over a

molybdenum-based mechanism; however, both types are observed using small molecules.[9,36] Recent studies with CO and Se suggest that a substrate molecule may bind to Fe2 and/or Fe6 by displacing a belt sulfur.[37,38] Regarding the flow of electrons during substrate turnover, a "deficit-spending" model has been proposed in which electrons first transfer from the P-cluster to the FeMo-cofactor, which gives the P-cluster a sufficiently low reduction potential to attract electrons from the [4Fe:4S] cluster.[39]

The accepted scheme for the kinetics of biological fixation is given by the Lowe-Thorneley model and reparametrized by the Watt group (Figure I-2).[40–45] The E_n states, in which n varies from 0-7, refers to the number of electrons transferred to the MoFe protein; E_0 is the resting state. During the progression from the resting state to more reduced states of the MoFe protein, molecular hydrogen may be formed at E_n (where $n > 1$) and released from the FeMo-cofactor, causing relaxation to E_{n-2}. Dinitrogen binding occurs at E_3 and E_4. Acetylene, which is commonly used to study the substrate turnover activity of nitrogenase, can bind and be reduced to ethylene in the E_1 and E_2 state.

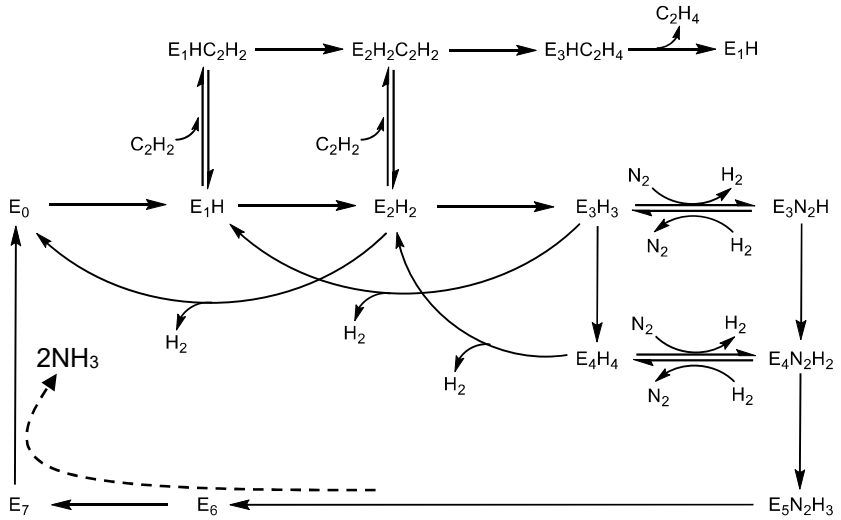


Figure I-2. A summary of the Lowe-Thorneley model.[45] Modeled from Scheme 3 of the following reference.[40]

I hope you enjoy the following chapters, which describe general methodology employed for working with nitrogenase proteins as well as detailed descriptions of several projects that aim to contribute to our understanding of the mechanism of biological nitrogen fixation.