QM/MM Studies of Nitrogenase

Jiang, Hao

2024

QM/MM Studies of Nitrogenase

HAO JIANG | COMPUTATIONAL CHEMISTRY | LUND UNIVERSITY

QM/MM Studies of Nitrogenase

QM/MM Studies of Nitrogenase

by Hao Jiang

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy

Thesis advisors: Prof. Ulf Ryde

Faculty opponent: Prof. Simone Raugei, Pacific Northwest National Laboratory

To be presented, with the permission of the Division of Computational Chemistry of Lund University, for public defense in lecture hall A, Center for Chemistry and Chemical Engineering, Lund, on Friday, the 27th of September 2024 at 09:00.

Organization: LUND UNIVERSITY

Document name: DOCTORAL DISSERTATION

Date of issue: 2024-09-03

Author(s): Hao Jiang

Sponsoring organization:

Title and subtitle: QM/MM Studies of Nitrogenase

Abstract:

Nitrogenase is the only enzyme that can cleave the strong triple bond in N_2, making nitrogen available for biological life. Despite extensive research, the mechanism of nitrogen fixation by nitrogenase is not fully understood, partly due to the enzyme's complex structure, which includes the largest iron–sulfur cluster known in metalloenzymes, the FeMo cluster. Understanding this process requires the integration of various scientific disciplines, such as inorganic chemistry, biochemistry, crystallography, spectroscopy, and computational chemistry. This thesis employs combined quantum-mechanics and molecular-mechanics (QM/MM) computational methods to investigate the structure and function of nitrogenase, focusing on its reaction mechanisms, intermediates, and electronic states.

The thesis comprises eight papers that explore different parts of the nitrogenase catalytic cycle. **Papers I** and **II** focus on the early part (E_0–E_4), revealing variability in the predictions of different DFT functionals regarding H_2 formation and N_2 binding. **Papers III** and **IV** examine the latter part (E_4–E_8), emphasizing the role of the S2B ligand and proton-transfer mechanisms, which suggest that S2B should remain bound to ensure lower energy barriers for proton transfers. **Papers V** and **VI** investigate the possibility of S2B ligand dissociation, and the results indicate that the stability of various structures is highly dependent on the DFT method used. **Paper VII** presents redox-potential calculations that provide insights into the redox properties of nitrogenase. **Paper VIII** analyzes the first Fe-nitrogenase structure, focusing on protonation states and the electronic structure in the E_0 and E_1 states.

These studies highlight the complexities of nitrogenase catalysis and underscore the limitations of current experimental techniques in capturing reaction intermediates. Computational methods have proven invaluable for studying these intermediates, offering insights that are difficult to obtain experimentally. While challenges remain, particularly in determining the exact structure of the E_4 intermediate, this research advances our understanding of nitrogenase.

Key words: Nitrogenase, QM/MM, DFT, nitrogen fixation, reaction mechanism, H_2 formation, N_2 binding, proton transfer, E_2 intermediate, redox potential, broken-symmetry state.

Classification system and/or index terms (if any)Supplementary bibliographical information

Language ISSN and key title:

ISBN:

978-91-8096-058-8 (print)

978-91-8096-059-5 (pdf)

Recipient's notes

Number of pages:

Price Security classification

Signature Date 2024-08-12

QM/MM Studies of Nitrogenase

Hao Jiang

蒋 浩

LUND
UNIVERSITY

Cover illustration front: Front layer created by the author using Blender and the background generated by QiyuAi.

Cover illustration back: Created by the author using PowerPoint.

Division of Computational Chemistry, Department of Chemistry

ISBN: 978-91-8096-058-8 (print)

ISBN: 978-91-8096-059-5 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2024

“人生天地之间，若白驹之过隙，忽然而已”

庄子

“Life in the universe is like a white pony passing through a crevice—
it passes in the blink of an eye.”

ZhuangZi

Contents

List of Publications

I. H_2 formation from the E_2–E_4 states of nitrogenase
H. Jiang and U. Ryde
Physical Chemistry Chemical Physics, **2024**, *26*, 1364–1375.

II. N_2 binding to the E_0–E_4 states of nitrogenase
H. Jiang and U. Ryde
Dalton Transactions, **2023**, *52*, 9104–9120.

III. Thermodynamically Favourable States in the Reaction of Nitrogenase without Dissociation of any Sulfide Ligand
H. Jiang and U. Ryde
Chemistry – A European Journal, **2022**, *28*, e202103933.

IV. Proton Transfer Pathways in Nitrogenase with and without Dissociated S2B
H. Jiang, O. K. G. Svensson, L. Cao and U. Ryde
Angewandte Chemie International Edition, **2022**, *61*, e202208544.

V. QM/MM Study of Partial Dissociation of S2B for the E_2 Intermediate of Nitrogenase
H. Jiang, O. K. G. Svensson and U. Ryde
Inorganic Chemistry, **2022**, *61*, 18067–18076.

VI. Putative reaction mechanism of nitrogenase with a half-dissociated S2B ligand
H. Jiang and U. Ryde
Dalton Transactions, **2024**, *53*, 11500–11513.

VII. Quantum Mechanical Calculations of Redox Potentials of the Metal Clusters in Nitrogenase
H. Jiang, O. K. G. Svensson and U. Ryde
Molecules, **2022**, *28*, 65.

VIII. Protonation of Homocitrate and the E_1 State of Fe-Nitrogenase Studied by QM/MM Calculations
H. Jiang, K. J. M. Lundgren and U. Ryde
Inorganic Chemistry, **2023**, *62*, 19433–19445.

Publications not included in this thesis

I. Assessment of DFT functionals for a minimal nitrogenase $[Fe(SH)_4H]^-$ model employing state-of-the-art *ab initio* methods

P. Vysotskiy, M. Torbjörnsson, **H. Jiang**, E. D. Larsson, L. Cao, U. Ryde, H. Zhai, S. Lee and G. K.-L. Chan

Journal of Chemical Physics, **2023***, 159,* 044106.

II. Reaction Mechanism for CO Reduction by Mo-Nitrogenase Studied by QM/MM

H. Jiang and U. Ryde

Inorganic Chemistry, in press, DOI: 10.1021/acs.inorgchem.4c02323.

III. Understanding the CO_2 Reduction Mechanism of Fe-Nitrogenase: A QM/MM Study

H. Jiang and U. Ryde

Manuscript.

Acknowledgements

As I finally submit my thesis, I can't believe how quickly these four years have passed—it feels like just yesterday that it all began. First and foremost, I want to express my deepest thanks to my supervisor, **Ulf**. I will never forget that rainy night four years ago when you picked me up from Lund C and brought me to Systervägen. Without your support, I wouldn't have been able to complete this dissertation. Thank you so much for your guidance over these four years. I am especially thankful for the freedom you gave me to study the projects I was interested in and for all the help you provided. These four years have been the most important period of growth for me.

I would also like to thank all my colleagues in the Division of Computational Chemistry, especially the group members, both past and present. Special thanks to **Lili** for her previous study on nitrogenase, which made it easier for me to continue the research. I am very grateful to my current office mate, **Kristoffer**, and my previous office mate, **Sara**. I had a great time at the conference with **Vilhelm** at AstraZeneca in Gothenburg. Thank you to **Oskar** for working together on nitrogenase calculations. I appreciate the teamwork on the CAS-N2ase project with **Ernst**, **Magne**, and **Victor**. Many thanks to **Meiting** for teaching me about free energy calculations, especially thermodynamic integration. **Marcos**, your enthusiasm and humor were much appreciated, and I enjoyed our discussions on QM/MM calculations using ORCA with you and **Xiaoli**. **Simon**, your introduction to Solar Fuel Simulations at the Winter meeting 2024 in Compute was very insightful. Sincere thanks to **Mikael** and **Valera** for the technical support, and to **Maria Lövgren** and **Maria Södergren** for the help with daily tasks. I am also very grateful to **Justin**, **Joel**, **Mickael**, **Iria**, **Kosala**, **Gayathri**, **Parisa**, and **Gaia** for sharing articles from various fields in our Journal Club. **Isabel**, **David**, **Alex** and all other members of Computational Chemistry, thank you for sharing the enjoyable coffee and Friday cakes.

Finally, I want to thank my mom and my younger sister. Despite the thousands of miles and six-hour time difference between us, our hearts have always remained close.

Thank you.

Abbreviations

ADP	Adenosine Diphosphate
AMBER	Assisted Model Building with Energy Refinement
AO	Atomic Orbital
ATP	Adenosine Triphosphate
BS	Broken Symmetry
CHARMM	Chemistry at Harvard Macromolecular Mechanics
Cryo-EM	Cryogenic Electron Microscopy
DFT	Density Functional Theory
ENDOR	Electron Nuclear Double Resonance
EPR	Electron Paramagnetic Resonance
ESEEM	Electron Spin Echo Envelope Modulation
FF	Force Field
GAFF	General Amber Force Field
GGA	Generalized Gradient Approximation
GTO	Gaussian-Type Orbital
HCA	Homocitrate
HERFD-XAS	High-Energy Resolution Fluorescence-Detected X-Ray Absorption Spectroscopy
HF	Hartree–Fock
LCAO	Linear Combination of Atomic Orbital
LDA	Local Density Approximation
LT	Lowe–Thorneley
MD	Molecular Dynamics
MM	Molecular Mechanics
MO	Molecular Orbital
PDB	Protein Data Bank
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics

re/oa	Reductive Elimination/Oxidative Addition
SAM	S-Adenosylmethionine
SCF	Self-Consistent Field
SD	Slater Determinant
STO	Slater-Type Orbital
TST	Transition State Theory
UFF	Universal Force Field
WT	Wild Type
XES	X-Ray Emission Spectroscopy

Popular science in English

Nitrogenase: Cleaving the Nitrogen–Nitrogen Triple Bond

Nitrogenase is an enzyme that transforms atmospheric nitrogen into ammonia, a critical process known as nitrogen fixation, which is vital for sustaining life on Earth. Nitrogenase is the only known biological catalyst capable of breaking the strong nitrogen–nitrogen triple bond. Nitrogen is a fundamental element of life, essential for proteins, nucleic acids, and other biomolecules. It plays a crucial role in plant growth, biomolecule synthesis, and metabolism. Despite nitrogen's abundance in the atmosphere (nearly 80%), only a few prokaryotes—such as certain bacteria and cyanobacteria—can convert atmospheric nitrogen into a usable form through nitrogen fixation.[1–5]

The Process of Nitrogen Fixation: Nitrogen fixation converts nitrogen gas (N_2) into ammonia (NH_3), a form that plants can utilize. There are three primary pathways for nitrogen fixation on Earth:

- **Biological Nitrogen Fixation**: Carried out by microorganisms using nitrogenases.
- **Industrial Nitrogen Fixation**: Achieved through the Haber–Bosch process, which uses high temperature, high pressure, and chemical catalysis.
- **High-Energy Nitrogen Fixation**: Occurs naturally through lightning and other high-energy atmospheric discharges, combining nitrogen and water to form ammonia and nitric acid, which rain brings to the ground.

Among these, biological nitrogen fixation is vital for the nitrogen cycle, making inert N_2 accessible to living organisms. Scientists are particularly interested in this process to develop sustainable methods for ammonia production, as the current industrial method, the Haber–Bosch process, is energy-intensive and environmentally harmful.

Types of Nitrogenases: There are three known types of nitrogenases, each with distinct clusters: Mo-nitrogenase, V-nitrogenase and Fe-nitrogenase.[1,4,6,7]

The Mechanism of Nitrogen Fixation: The mechanism of nitrogen fixation by nitrogenases is complex and not yet fully understood. It involves multiple electron and proton transfers, ultimately leading to the cleavage of the N_2 bond and the formation of ammonia. Current research focuses on identifying and characterizing the reaction intermediates to better understand this intricate process.[1,4,7]

Evolution and Importance: The evolution of nitrogenases has been pivotal in shaping life on Earth. The ability to fix nitrogen allowed early microorganisms to thrive in nitrogen-poor environments. The emergence of molybdenum nitrogenase

is particularly linked to the transition from anaerobic to aerobic metabolism, marking a significant evolutionary advancement.

Computational studies: Computational methods, such as density functional theory, have been crucial in understanding the electronic structure and properties of nitrogenase clusters. These approaches complement experimental findings and provide valuable insights into nitrogenase mechanisms, despite challenges in accuracy and model system limitations.[1,8–10]

Nitrogenases are remarkable enzymes essential for the global nitrogen cycle. Their ability to convert atmospheric nitrogen into a usable form for other organisms is vital for life. As research continues, we can expect deeper insights into their mechanisms, evolution, and potential applications, paving the way for a more sustainable future.

科学摘要

固氮酶: 一把打开氮氮三键的钥匙

固氮酶是一种能够将大气中的氮气转化为氨的酶，这一过程称为固氮过程，对维持地球上的生命活动至关重要。同时固氮酶是唯一已知的能够打破强氮氮三键(N≡N)的生物催化剂。氮是生命的基本元素，对于蛋白质、核酸和其他生物分子至关重要。它在植物生长、材料合成和代谢中发挥着重要作用。尽管氮在大气中非常丰富（接近 80%），只有少数原核生物——如某些细菌和蓝藻,能够通过固氮作用将大气中的氮转化为可利用的形式。[1–5]

固氮过程: 固氮作用将氮气（N_2）转化为氨（NH_3），以便植物可以利用这种形式的氮。地球上有三种主要的固氮途径:

- **生物固氮**：由利用含有固氮酶的微生物进行。
- **工业固氮**：通过哈伯-博施法实现，该方法使用高温、高压和金属催化剂。
- **高能固氮**：通过闪电和其他高能大气放电自然发生，结合氮和水形成氨和硝酸，并通过降雨带到地面。

在这些途径中，生物固氮对于氮循环至关重要，使不活泼的氮气（N_2）转化成为对生物体可利用的形式（NH_3）。科学家特别关注这一过程，以开发可持续的氨生产的方法，因为当前的工业方法——哈伯-博施法——耗能巨大且对环境有害。

固氮酶的类型: 已知有三种类型的固氮酶，每种酶都有不同的辅因子: 钼固氮酶, 钒固氮酶和铁固氮酶。[1,4,6,7]

固氮酶反应机理: 固氮酶的反应机制复杂且尚未被完全理解。该过程涉及多个电子和质子的转移，最终导致氮氮三键键的裂解和氨的形成。目前的研究集中在识别和表征反应中间体，以更好地理解这一复杂过程。[1,4,7]

进化与重要性: 固氮酶的进化对地球上的生命产生了深远影响。拥有固氮能力使早期微生物在贫氮环境中得以繁衍生息。特别是钼固氮酶的出现与微生物从厌氧代谢转向有氧代谢有关，这标志着重要进化过程的发生。

计算研究：计算研究（如密度泛函理论）在理解固氮酶辅因子的电子结构和性质方面起到了至关重要的作用。这些方法补充了实验结果，尽管在准确性和模型系统方面存在挑战，但它们为固氮酶反应机理提供了宝贵的见解。[1,8–10]

固氮酶是对全球氮循环至关重要的酶。它们将大气中的氮转化为其他生物可利用的形式，这对于生命的生存至关重要。随着研究的不断深入，我们可以期待对其机制、进化和潜在应用有更深入的了解，以便为更可持续的未来铺平道路。

1 Introduction

Nitrogenase is the only enzyme that can cleave the N_2 triple bond in nature, making nitrogen available for biological lifeforms. This is because nitrogen gas (N_2) is a very stable molecule, and it requires much energy to cleave the strong triple bond between the two nitrogen atoms. Nitrogenase is a large and complex protein. Crystallographic studies have revealed that the most active type of nitrogenase consists of two proteins. The MoFe protein contains the catalytic [Mo-7Fe-9S-C]-homocitrate (FeMo) cluster and the electron-transfer [8Fe-7S]-6Cys (P) cluster, whereas the Fe protein contains a [4Fe-4S] cluster. The mechanism of nitrogen fixation is not fully understood.[1,4,11]

The mechanism of N_2 reduction by nitrogenase is a complex process, the understanding of which requires the integration of diverse scientific disciplines, including inorganic chemistry, biochemistry, crystallography, spectroscopy, and computational chemistry.[12–16] Each approach has its own strengths and limitations. One of the most crucial aspects of comprehending nitrogenase function is the determination of its three-dimensional structure. X-ray crystallography remains a primary method for determining the three-dimensional structure of enzymes, allowing the analysis of scattered X-ray patterns from crystalline samples.[17–20] However, this technique has limitations in visualizing hydrogen atoms and directly measuring reaction dynamics.

One of the reasons why the mechanism of nitrogenase is so difficult to understand is that it contains the largest iron–sulfur cluster known in metalloenzymes, the FeMo cluster. This unique structure means that the active site of catalysis has flexible and variable electronic states of transition metals. Given the rapid nature of biological enzyme reactions, it is often challenging to capture reaction intermediates using existing experimental techniques. Therefore, computational methods provide powerful tools for investigating the reaction mechanisms of enzymes. Based on atomic structures, computational methods can be employed to predict possible structures, including those of different protonation states of residues, intermediates, and transition states of the reaction that are not accessible through current experimental techniques. By comparing the relative energy, the most stable species in each step can be identified. Moreover, quantum mechanical methods can describe the intricate electronic structure of transition metal systems.[21–27]

Combined quantum-mechanics and molecular-mechanics (QM/MM) approaches perform well in modelling reactions in biomolecular systems.[9,10,28–33] While quantum-mechanical (QM) methods are essential for describing chemical reactions, they are limited to relatively small systems. The size and complexity of biomolecules require molecular-mechanics (MM) methods, which can handle large systems and long time-scales. QM/MM methods combine the strengths of the two approaches, using QM to handle chemically active regions and MM to handle the surrounding protein and solvent.

In this thesis, we used models from accurate crystal structures and performed systematic studies of nitrogenase using the QM/MM approach. We have studied N_2 binding in **Paper II** [34] and H_2 formation in **Paper I**,[35] as well as the partial dissociation of the S2B ligand in **Papers V** and **VI**.[36,37] Furthermore, we calculated the redox potentials in **Paper VII**[38] and proposed putative reaction mechanisms of the cluster with the S2B ligand still bound in **Paper III**, along with a study of the proton-transfer process for both with and without S2B ligand in **Paper IV**.[39,40] The electronic structure and protonation states of the FeFe cluster within Fe-nitrogenase were examined in **Paper VIII**[41], based on recent crystallographic data.

2 Methods

With the development of computational methods and advancements in computer hardware, computational chemistry has been widely applied in multiple disciplines and has played an increasingly important role in chemical research. It has solved problems that cannot be studied by experimental methods, especially the study of enzyme reaction mechanisms. Due to the fast reaction rate of enzyme-catalyzed reactions, it is difficult to capture reaction intermediates experimentally. Theoretical calculations, especially multiscale simulations of enzyme systems using QM/MM methods, have become an important method for studying enzyme reaction mechanisms. This chapter will briefly introduce the theory and methods used in this thesis.

2.1 Quantum Mechanics

Quantum chemistry is the application of quantum mechanical principles to solve chemical problems. The improvement of computer performance and the development of new methods for molecular calculations have made quantum chemistry a practical tool in all fields of chemistry. The main objects of study of quantum chemistry methods are isolated systems containing a few to hundreds of atoms, such as molecules and clusters. The core problem of quantum chemistry is to solve the **Schrödinger equation** of the system.

$$\mathbf{H}\Psi = \boldsymbol{E}\Psi \tag{2.1}$$

Here, **H** is the Hamiltonian operator representing the total energy of the system, E is the energy eigenvalue, and Ψ is the wavefunction, an eigenfunction of the Hamiltonian. The wavefunction Ψ itself does not have direct physical meaning, but its square, $|\Psi|^2$, represents the probability density of finding a particle at a particular position.

The Hamiltonian operator for a molecular system consists of five terms that are sums of potential (**V**) and kinetic (**T**) energy operators:

$$\mathbf{H} = \mathbf{V} + \mathbf{T} = \mathrm{V}_{ee} + \mathrm{V}_{nn} + \mathrm{V}_{en} + \mathrm{T}_e + \mathrm{T}_n \tag{2.2}$$

The potential energy terms are given by

$$\mathbf{V}_{ee} = \frac{e^2}{4\pi\varepsilon_0}\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.3}$$

$$\mathbf{V}_{nn} = \frac{e^2}{4\pi\varepsilon_0}\sum_{A=1}^{n}\sum_{B>A}^{n}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{2.4}$$

$$\mathbf{V}_{en} = -\frac{e^2}{4\pi\varepsilon_0}\sum_{i=1}^{N}\sum_{A=1}^{n}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \tag{2.5}$$

Here, $\mathbf{V}_{ee}$ represents the repulsive potential energies between electrons, $\mathbf{V}_{nn}$ represents the repulsive potential energies between nuclei, and $\mathbf{V}_{en}$ represents the attractive potential energy between electrons and nuclei. e is the proton charge, ε_0 is the permittivity of vacuum, N is the number of electrons and n is the number of nuclei, $\mathbf{r}_i$ and $\mathbf{R}_A$ are the coordinates of electrons and nuclei, respectively, and Z_A and Z_B are the atomic numbers of nuclei A and B. The kinetic energy terms are

$$\mathbf{T}_e = -\frac{\hbar}{2m_e}\sum_{i=1}^{N} \nabla_i^2 \tag{2.6}$$

$$\mathbf{T}_n = -\frac{\hbar}{2M_A}\sum_{A=1}^{n} \nabla_A^2 \tag{2.7}$$

$\mathbf{T}_e$ and $\mathbf{T}_n$ are the kinetic energy operators for electrons and nuclei, respectively, $\hbar$ is Planck's constant, m_e is the electron mass, ∇_i^2 and ∇_A^2 are the Laplacian operator for electron and nuclei, respectively.

For systems involving more than two particles, solving the Schrödinger equation analytically is generally impossible, necessitating approximations. One fundamental approximation is the **Born–Oppenheimer approximation**, which separates the motions of electrons and nuclei based on their mass difference. Electrons, being much lighter than nuclei, move much faster. This allows us to treat the nuclei as stationary from the perspective of the electrons. Under this approximation, the Hamiltonian simplifies to the electronic Hamiltonian:

$$\boldsymbol{H_e} = \mathbf{V}_{ee} + \mathbf{V}_{\boldsymbol{nn}} + \mathbf{V}_{\boldsymbol{en}} + \mathbf{T}_e \tag{2.8}$$

Here, the nuclear kinetic energy operator $\mathbf{T}_{\boldsymbol{n}}$ is omitted and the V_{nn} term becomes a constant because the nuclei are fixed. The electronic Hamiltonian depends only on the positions of the nuclei and the resulting electronic wavefunction depends parametrically on these nuclear coordinates. This approximation significantly simplifies solving the Schrödinger equation for complex molecular systems, making QM methods more feasible for practical applications in computational chemistry.[42,43]

2.1.1 Hartree–Fock Theory

The Hartree–Fock (HF) method is a fundamental method to approximating solutions for the Schrödinger equation in many-electron systems. It transforms the complex multi-electron problem into single-electron problems by using single-electron wavefunctions and employs the self-consistent field (SCF) method to achieve a stable wavefunction and energy. The HF method simplifies the Schrödinger equation through several key approximations, the primary one being the orbital approximation. This approximation suggests that the total wavefunction of a system can be expressed as a product of one-electron wavefunctions or orbitals:

$$\Psi(r_1, r_2, .., r_M) = \psi_1(r_1)\psi_2(r_2)\cdots\psi_M(r_M) \tag{2.9}$$

Each electron has a spin quantum number of 1/2 and in the presence of a magnetic field, there are two possible states, corresponding to alignment along or opposite to the field. The HF method employs the concept of a Slater determinant to ensure the wavefunctions are antisymmetric under the exchange of any two electrons. The Slater determinant inherently satisfies the requirement that no two electrons can occupy the same quantum state simultaneously, a fundamental property of fermions like electrons (the Pauli principle). For the general case of *M* electrons and *M* spin-orbitals, the Slater determinant is given by

$$\Psi = \frac{1}{\sqrt{M!}}\begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_M(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_M(\mathbf{r}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_1(\mathbf{r}_M) & \psi_2(\mathbf{r}_M) & \dots & \psi_M(\mathbf{r}_M) \end{vmatrix} \tag{2.10}$$

To further refine the approximation, the one-electron molecular orbitals (MOs) are expressed as linear combinations of atomic orbitals (AOs), a technique known as the Linear Combination of Atomic Orbitals (LCAO) approximation. This approach allows for the description of molecular orbitals in terms of simpler, well-understood atomic functions.

The HF method is based on the variational principle, which asserts that the optimal approximation to the ground-state energy of a system can be achieved by minimizing the total energy with respect to the parameters of the selected wavefunction. This results in the derivation of the Fock equations that are iteratively solved to identify the optimal molecular orbitals (MOs) and their corresponding energies.

Despite its considerable strengths, the HF method is not without limitations. The method approximates the electron–electron repulsion by considering that each electron moves in an average field created by all other electrons, thus neglecting electron correlation, the instantaneous interactions between electrons. As a result,

while the HF method provides a good initial approximation, more advanced post-HF methods are often used to account for these correlation effects for more accurate results. Although the HF method accurately represents the exchange interaction between electrons of the same spin, it does not fully consider the Coulomb interactions. Consequently, while the HF method provides qualitative insights, its quantitative accuracy is limited.[42,43]

2.1.2 Basis Set

A basis set in quantum chemistry is a collection of mathematical functions used to represent the atomic orbitals. The wavefunction is expressed as linear combinations of these basis functions. The choice and quality of the basis set significantly influence the accuracy and efficiency of quantum chemical calculations. There are two types of basis functions: Gaussian-Type Orbitals (**GTOs**) and Slater-Type Orbitals (**STOs**). The GTOs are given by

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\zeta r^2} \tag{2.11}$$

and the STOs are given by

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\zeta r} \tag{2.12}$$

Here, n, l, m are quantum numbers. r, θ, and φ are spherical coordinates, where r is the distance between the electron and the nucleus. N is a normalization constant, $Y_{l,m}$ is a spherical harmonic function, ζ is a constant. GTOs are favored in many quantum-chemical calculations due to their computational efficiency, particularly in the evaluation of integrals. STOs more closely resemble the actual shape of atomic orbitals but are less commonly used due to their more complex integral evaluations.[42–44]

Minimal Basis Sets: These basis sets use one basis set for each atomic orbital. An example is the STO-3G[45] basis set. Minimal basis sets are typically used for preliminary studies and teaching purposes.

Split-Valence Basis Sets: These basis sets use two basis functions for the valence electrons (but still one for the core orbitals), allowing for more flexibility and accuracy. Examples include 3-21G[46] and 6-31G.[47] Double-zeta and triple-zeta basis sets use two and three basis functions for all electrons, respectively.

Polarization functions (i.e. basis functions with the l quantum number one step higher than the electron it should describe) are added to the basis sets to get more accurate molecular geometries and electronic distributions (e.g., 6-31G(d) or 6-31G(d,p)[47,48]). Diffuse functions (i.e. with smaller values of ζ than normal) are used to describe the anions, e.g., 6-31++G,[47,49] where the '+' symbols denote the inclusion of diffuse functions for heavy atoms and hydrogen atoms, respectively.

In this thesis, we use the Karlsruhe basis sets[50]. For geometry optimization, we utilize the def2-SV(P)[50] basis set, which is a split-valence basis set with polarization functions on heavy atoms (excluding hydrogen). For single-point calculations to achieve more accurate energy values, we sometimes use the larger def2-TZVPD[50,51] basis set, which is a valence triple-zeta basis set with polarization and diffuse functions.

2.1.3 Density Functional Theory

Density Functional Theory (DFT) is a QM method used to study the electronic structure of many-body systems, which is based on Hohenberg–Kohn theorems, rather than the Schrödinger equation: [42,43]

First Theorem: The ground state energy E is a unique functional of the electron density ρ.

$$E = E[\rho] \tag{2.13}$$

Second Theorem: The electron density that minimizes the energy of the overall functional is the true ground-state electron density.

$$E[\rho] > E_0[\rho_0] \tag{2.14}$$

These theorems indicate that the ground-state electron density ρ unambiguously describes the system.[52] In 1965, Kohn and Sham introduced a practical way to implementing DFT.[53] They proposed a set of self-consistent field (SCF) equations, known as the Kohn–Sham equations, which describe the behavior of non-interacting electrons in a fictitious system that has the same electron density as the real interacting system. The total energy functional is decomposed into several parts:

$$E[\rho] = T_s[\rho] + E_{\text{ne}}[\rho] + J[\rho] + E_{\text{xc}}[\rho] \tag{2.15}$$

where $T_s[\rho]$ is the kinetic energy of the non-interacting electrons, $E_{\text{ne}}[\rho]$ is the attractive potential between nuclei and elections, $J[\rho]$ is the Coulomb repulsion between electrons and $E_{\text{xc}}[\rho]$ is the exchange–correlation energy, which includes all many-body effects beyond the Hartree approximation.

The exchange–correlation energy is the most difficult term to calculate accurately and needs to be approximated. Various levels of approximations have been developed for $E_{\text{xc}}[\rho]$, including *local density approximation* (LDA),[53] *generalized gradient approximation* (GGA),[54] meta-GGA and hybrid functionals.[55]

In 1998, Kohn was awarded the Nobel Prize in Chemistry (shared with Pople) for his contributions to DFT.[56] Due to its relatively high computational accuracy and efficiency, DFT has developed rapidly. However, it still faces challenges, such as poor description of weak interactions. Most traditional functionals, like B3LYP, completely fail to describe dispersion interactions. However, the DFT-D[57] series of empirical dispersion corrections proposed by Grimme have addressed this issue.

In this thesis, we have used TPSS,[55] r^{2}SCAN,[58,59] TPSSh,[60] B3LYP[61–63] functionals. The former two are meta-GGA functionals, while the other two are hybrid functionals with 10 and 20% Hartree–Fock exchange, respectively. For all methods, we have employed the DFT-D4 dispersion correction.[57]

2.2 Classical Mechanics

Molecular mechanics (MM) methods are computational techniques that employ classical Newtonian mechanics to simulate molecular systems. MM methods use molecular force fields to calculate the system energy, optimize the geometric structure, perform vibrational analysis, and conduct dynamic simulations. A molecular force field is a collection of empirical functions that describe the energy as a function of the coordinates.[42,64]

2.2.1 Force Field Methods

The general form of a molecular force field is:

$$E_{total} = E_{covalent} + E_{noncovalent} \tag{2.16}$$

The total energy of the system includes covalent and non-covalent interactions. Covalent interactions consist of bond stretching energy, bond angle bending energy, and dihedral angle torsion energy:

$$E_{covalent} = E_{bonds} + E_{angles} + E_{dihedrals} \tag{2.17}$$

Non-covalent interactions involve van der Waals forces and electrostatics:

$$E_{noncovalent} = E_{VdW} + E_{el} \tag{2.18}$$

MM methods play a crucial role in enzyme simulations, despite they cannot simulate chemical reactions. MM methods can simulate protein dynamics over extended timescales (e.g., nanoseconds to microseconds).

The MM method treats each atom as a point mass, ignoring the electrons. Consequently, MM methods cannot provide information on electron structures, nor can they describe chemical reactions involving bond formation or breaking. Unlike quantum chemistry methods, MM methods cannot identify reaction intermediates and transition state structures.

Employing a molecular force field requires prior knowledge of parameters for specific atom types within a given bond. These empirical parameters include atomic mass, van der Waals radius and energies, partial atomic charges, and bond lengths, bond angles, and force constants. These parameters may be obtained from experimental and theoretical (typically QM) studies of small organic molecules. As a result, different force field parameters are only applicable to their corresponding systems. The accuracy of MM methods is lower than that of quantum mechanical

methods, but they offer significantly faster computational speeds due to the absence of electron integral calculations.

MM methods are widely used to study large molecular systems such as proteins and nucleic acids. Currently, some of the most well-established and widely used force fields include AMBER[65] (proteins, nucleic acids), CHARMM[66] (small molecules, lipids, nucleic acids, proteins), GAFF[67] (small molecules), and UFF[68] (universal force field).

2.2.2 Molecular Dynamics

Molecular Dynamics (MD) simulations have become an essential tool in biological research, providing insights into the complex dynamics and thermodynamics of biomolecular systems at the atomic level. MD simulations allow to model the behavior of molecules as collections of interacting classical particles governed by Newtonian mechanics. This computational technique has enhanced our understanding of various biochemical processes, from protein folding to drug binding. These simulations provide dynamic trajectories that reveal the complex motions and interactions of atoms and molecules over time.[42,64]

The core of MD simulations is based on Newton's second law of motion, which states that the force acting on an atom equals the mass of the atom multiplied by its acceleration:

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m_i \frac{d^2 x_i}{dt^2} \tag{2.19}$$

Where F_i is the force on atom i, m_i is the mass, and a_i is the acceleration of atom i, x_i is the position of the atom. The acceleration can be obtained from the force, which in turn is derived from the potential energy function U of the system:

$$F = -\frac{dU(x)}{dx_i} \tag{2.20}$$

The potential energy can be computed using different approaches: classical MD through a force field, ab initio MD by solving the Schrödinger equation, or QM/MM MD by combining both methods. The position of a particle at time $t + \Delta t$, starting from an initial guess, can be expressed using a Taylor expansion:

$$x(t + \Delta t) = x(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 + .. \tag{2.21}$$

with

$$x(t + \Delta t) \approx x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \tag{2.22}$$

then with Eq.2.19, we get

$$v(t + \Delta t) \approx v(t) + a(t)\Delta t \tag{2.23}$$

The MD algorithm starts with initial guesses of $x(0)$ and $v(0)$. Subsequently, F and a are calculated, allowing the atoms to be moved to $x(t + \Delta t)$ and the velocities updated to $v(t + \Delta t)$. This procedure is repeated many time steps to simulate the dynamic behavior of the system.

With the growth of computational power, the scope and accuracy of MD simulations are expected to increase. New algorithms and machine-learning techniques are integrated to handle larger systems and longer timescales, providing deeper insights into biomolecular dynamics.[69] The combination of MD simulations with experimental techniques will further accelerate discoveries, making MD an essential tool in biological research. By providing detailed atomic-level insights into the dynamics of biological systems, MD simulations have transformed our understanding of molecular interactions and are instrumental in the design of new therapeutics.[70]

2.3 QM/MM

The hybrid QM/MM (quantum mechanics combined with molecular mechanics) method is a widely used tool for studying reactions in large biomolecules, combining the strengths of QM (accuracy) and MM (speed). This method enables the study of chemical processes in solutions and proteins by applying QM to a small, critical part of the system—where chemical bonds break or electronic structures change—while using MM for the larger, surrounding environment. The main advantage of QM/MM is its computational efficiency, which makes it possible to perform large-scale molecular simulations with minimal accuracy loss. This concept was initially proposed by Warshel and Levitt in 1976,[71] and they, along with Karplus, received the Nobel Prize in Chemistry in 2013 for their development of multiscale models for complex chemical systems.[72]

The core of QM/MM involves partitioning the system into three regions: the QM region, the MM region, and the boundary region. The QM region, usually the reaction site or a specific ligand, is treated by QM methods for accurate bond breaking and electronic structure changes. The MM region, encompassing the surrounding solvent or non-reactive parts of the biomolecule, is treated with MM methods, offering a realistic description of the broader environment.

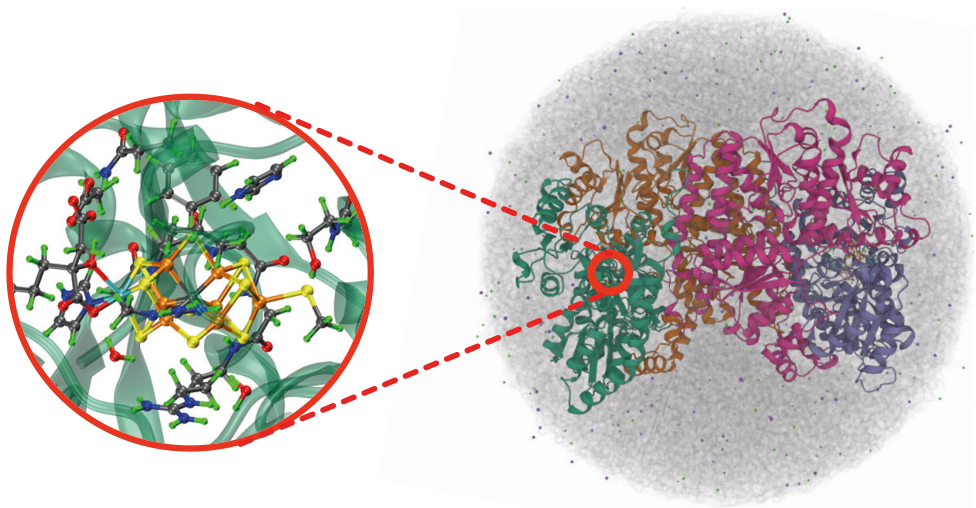


Figure 2.1: The systems in a QM/MM calculation.The QM system is shown in a balls-and-sticks model.

In the QM/MM framework, the boundary region is crucial for interfacing the QM and MM areas, with electrostatic interactions being key to their coupling. There are several methods for handling these interactions, each with varying degrees of complexity and accuracy:

- **Mechanical Embedding:** Uses classical point charges to represent electrostatic interactions between the QM and MM systems. It is computationally efficient but does not involve any polarization effects.
- **Electrostatic Embedding:** Considers the polarization of the QM region by the MM region, using a point-charge model of the MM region in the QM calculations. This typically leads to improved accuracy. This approach is widely used in biomolecular simulations where polarization effects are important. It is also used in this thesis.
- **Polarized Embedding:** Extending upon electrostatic embedding, polarized embedding additionally accounts for the polarization of the MM region by the QM region. This method offers the highest level of accuracy but is also the most computationally demanding. Moreover, it requires polarizable MM model and a QM software that may treat both MM charges and polarizabilities in the SCF calculations.

A significant challenge in QM/MM methods is how to deal with covalent bonds across the boundary between the QM and MM regions. This is essential to prevent significant distortion of the electronic structure for QM region. Various approaches to link the QM and MM regions typically considers charged groups near the QM system and often focus on carbon–carbon bonds. Three methods address the atomic treatment at the QM/MM boundary:

- **Link-Atom Method:** Introduces an additional atom, typically a hydrogen atom, that forms a covalent bond with the QM atoms, effectively saturating their valence by replacing the cut bonds.[73–77] This method is simple and computationally efficient, and it is therefore used in this thesis.
- **Boundary Atom Method:** Replaces MM atoms that bond with QM atoms across the boundary with specialized boundary atoms. These atoms function as MM atoms in MM calculations but simulate the electronic properties of MM atoms in the QM region.[78–81]
- **Localized Orbital Method:** Avoids introducing new atoms by placing hybrid orbitals at the boundary and freezing some of them to cover the QM region and substitute the broken bond.[82–86]

The total energy computed using QM/MM methods encompasses three types of interactions: interactions within the QM region, interactions within the MM region, and interactions between QM and MM regions. Two main approaches exist for calculating the total energy:

- **Additive Scheme:** This method expresses the total energy as the sum of three terms: the energy of the QM region, the energy of the MM region, and the QM/MM coupling energy. The latter term accounts for the interactions at the

boundary region, including bond formation, electrostatic interactions, and van der Waals forces.[87]

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM(QM)}} + E_{\mathrm{MM(MM)}} + E_{\mathrm{QM-MM}} \tag{2.24}$$

- **Subtractive Scheme:** This method, where the total energy is first calculated using molecular mechanics (MM), is followed by adding the quantum mechanics (QM) energy of the QM region and subtracting the MM energy of the same region, ensuring that no interactions are double counted. It can easily be extended to more than two computational methods and regions. While ONIOM[88] is the typical example, other software like ComQum[89] also employs similar methods. It has the advantage that it is simple to implement and does not require any modified MM code.

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM(QM)}} + E_{\mathrm{MM,total}} - E_{\mathrm{MM(QM)}} \tag{2.25}$$

Hybrid QM/MM methods have revolutionized our understanding of biomolecular reactions, providing a powerful tool for unravelling the intricate dynamics and thermodynamics of these processes at the atomic level. Their ability to bridge the gap between quantum and classical mechanics, coupled with their computational efficiency, has made them an indispensable tool in modern biological research. As computational power continues to increase, QM/MM methods will undoubtedly play an even more prominent role in shaping our understanding of the intricate workings of the living world.[33]

2.4 Transition State Theory

Transition state theory (TST) is a fundamental framework in chemical kinetics that explains how reactions occur and predicts their rates. According to TST, for a reaction to occur, the reactant molecules must pass through a high-energy structure, the transition state or activated complex. This transition state exists at the peak of the energy barrier along the reaction coordinate, and molecules need sufficient energy, called activation energy, to reach this state.

TST assumes an equilibrium between reactants and the transition state. The reaction rate is determined by how frequently reactant molecules achieve the transition state and successfully convert into products. The theory employs principles from statistical mechanics, treating the transition state as a type of molecule describable using thermodynamic concepts. TST provides valuable insights into factors affecting reaction rates, such as temperature and catalysts, and helps elucidate reaction mechanisms. However, it has limitations, including the assumption of a single reaction pathway and its dependence on accurate potential energy surfaces.[64]

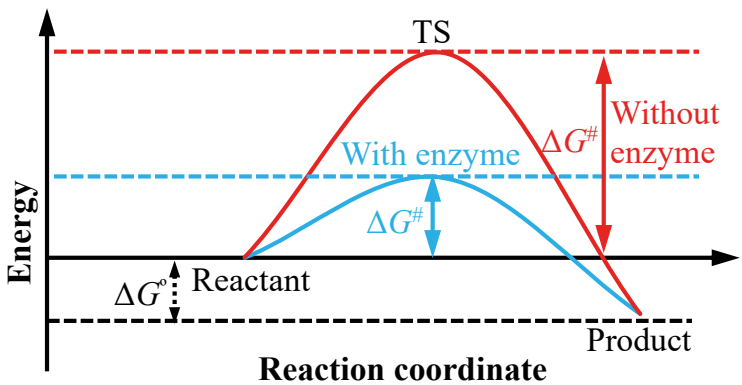


Figure 2.2: Energy profile of a chemical reaction showing the effect of an enzyme on activation energy. The red curve represents the reaction pathway without the enzyme, indicating a higher activation energy. The blue curve shows the reaction pathway with the enzyme, demonstrating a lower activation energy. Enzymes facilitate reactions by stabilizing the transition state, thereby reducing the activation energy.

Transition State Theory (TST) in enzymology explains how enzymes catalyze reactions by stabilizing the transition state. Enzymes lower the activation energy ($\Delta G^{\#}$) required for the reaction by binding to substrates and forming an enzyme–substrate complex, which then transitions to the enzyme–transition-state complex. This stabilization is achieved through precise interactions at the enzyme's active site, which reduce the energy barrier and increases the reaction rate. Consequently, enzymes enhance catalytic efficiency without altering the overall free energy change (ΔG^{o}) of the reaction, making them vital for various biochemical processes.[42]

3 Nitrogenase

Nitrogen is crucial for sustaining life on Earth, being a component of all amino acids and nucleic acids. Although N_2 constitutes 78% of the Earth's atmosphere, nitrogen remains a limiting factor for plant growth and is a main component in artificial fertilizers,[1] because plants cannot metabolize N_2 due to its strong and inert triple bond.

Nitrogenase is the only enzyme that can convert N_2 to NH_3 in nature. This process is called nitrogen fixation which is part of the global nitrogen cycle. The fixation of nitrogen also occurs in the industrial conversion through the energy-intensive Haber–Bosch process and through lightning induced chemical conversion in the Earth's atmosphere. The Haber–Bosch process requires high temperatures and pressures, as well as a metal catalyst, and accounts for nearly 2% of the world's total energy consumption.[5] This process is considered one of the important reasons for the explosive population growth after World War II.

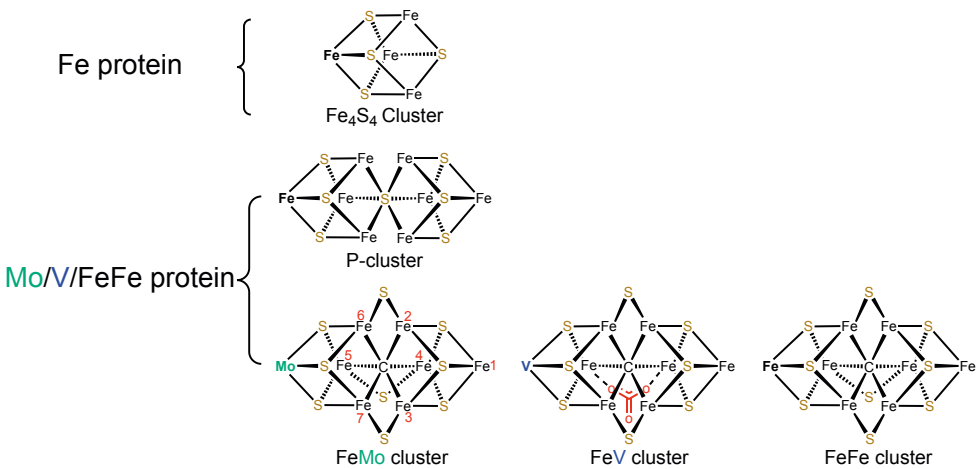


Figure 3.1: Structures of the FeS clusters in nitrogenases.

Nitrogenase exists in three variants: Mo-nitrogenase, V-nitrogenase, and Fe-nitrogenase. Mo-nitrogenase is the most abundant form with the highest N_2-reducing activity, and it is also the most studied type.[1,4,7,8,19,90–92] Crystal structures of Mo-nitrogenase have been known since 1992,[20,93] V-nitrogenase since 2017,[94]

and Fe-nitrogenase since 2023.[95] All nitrogenases contain two proteins, the Fe protein with a [Fe_4S_4] cluster and the Mo/V/FeFe protein. The Mo/V/FeFe protein contains an electron-transfer [Fe_8S_7] (P) cluster and the FeMo/FeV/FeFe cluster, which is one of the most complex clusters known in nature. Mo-nitrogenase contains a catalytic $MoFe_7S_9C$(homocitrate) cluster, known as the FeMo cluster, V-nitrogenase contains a $VFe_7S_8C(CO_3)$(homocitrate) cluster (FeV cluster), while Fe-nitrogenase contains a Fe_8S_9C(homocitrate) cluster (the FeFe cluster). In all three cases, the active-site cluster is coordinated to the protein via a cysteine and a histidine residue.[17,20,93,96,97]

3.1 Atomic Structure

The most studied nitrogenase has been the Mo-dependent form (Mo nitrogenase), and this will also be the main studied this thesis. The Fe protein for these three homologous nitrogenase are encoded by the *nifH*, *vnfH*, and *anfH* genes.[2]

3.1.1 Mo-Nitrogenase

Mo-nitrogenase is a two-component system composed of the MoFe protein and the electron-transfer Fe protein. The Figure 3.2 illustrates the crystal structure of the MoFe protein. The MoFe protein is encoded by *nifDK* and it is a $\alpha_2\beta_2$ heterotetramer that contains two iron–sulfur clusters: the FeMo cluster and P-cluster.[11]

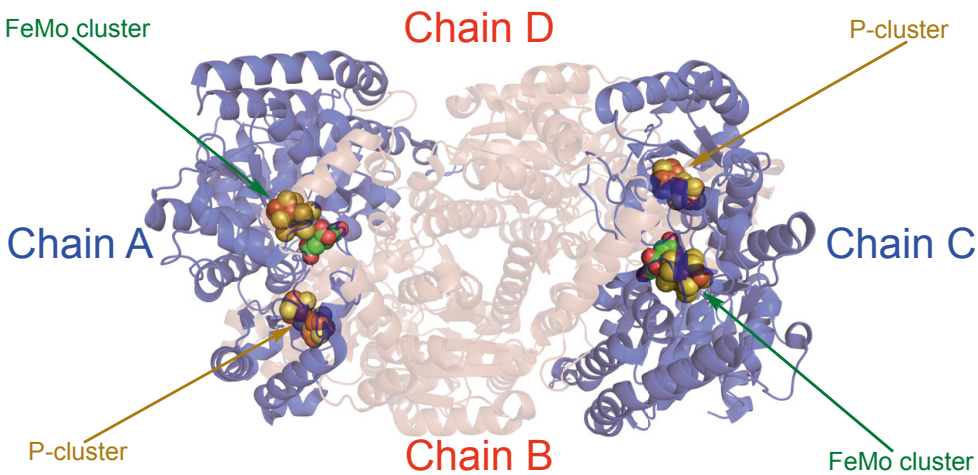


Figure 3.2: Structure of MoFe protein forms an $\alpha_2\beta_2$ heterotetramer in which two αβ units (NifDK) are connected solely via the NifK peptides. Each $\alpha\beta$ unit holds a FeMo cluster and a P-cluster (PDB 3U7Q).

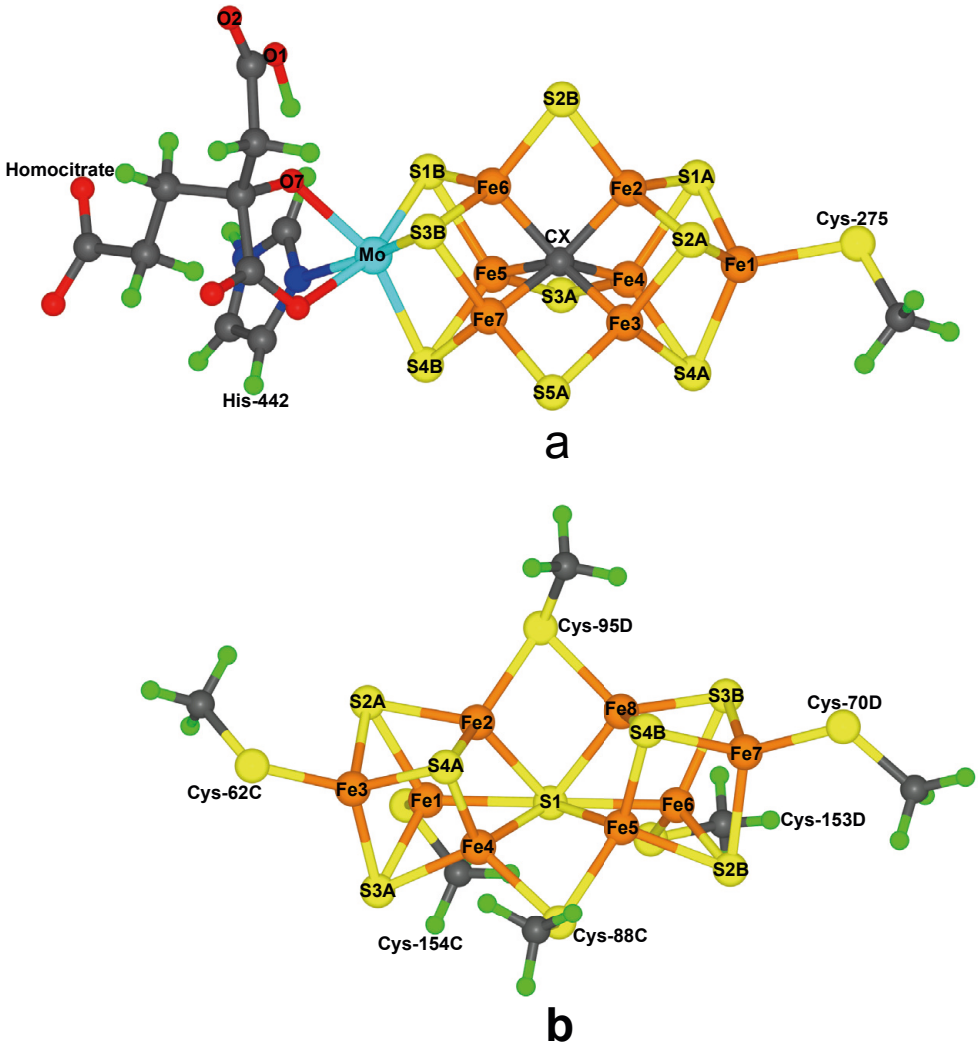


Figure 3.3: The atomic structure of a) FeMo cluster and b) P-cluster with atom and residue names taken from the 3U7Q crystal structure.

The FeMo cluster consists of one molybdenum (Mo) atom, seven iron (Fe) atoms, nine sulfur (S) atoms, one central carbon (C) atom, a homocitrate ligand bound to the Mo atom, and one cysteine (Cys) and one histidine (His) residue that bind the cluster to the protein. It is believed to be the active site for substrate binding and reduction (Figure 3.3 a). The P-cluster contains eight iron (Fe) atoms and seven sulfur (S) atoms, with six cysteine residues from the protein. It is assumed that the P-cluster transfers electrons from Fe protein to FeMo cluster (Figure 3.3 b). Despite

extensive research into the cluster's complexities, including its functionality, reactivity, and electronic structure, many aspects remain subjects of debate.[1,2,4,11]

The first atomic crystal structure of MoFe protein was solved by Kim and Rees in 1992 at 2.7 Å resolution.[20] In this structure, the FeMo-co was described as [4Fe:3S] and [Mo:3Fe:3S] clusters bridged by three sulfide ligands. In 2002, a higher resolution crystal structure at 1.16 Å was solved by Rees, Einsle and coworkers ,[96] showing a light atom in the center of the cluster It was designated "X" (C, N or O) and was debated for many years. Many DFT calculation suggested that it is a N^{3-} ion.[98,99] In 2011, Rees, Einsle and coworkers used atomic-resolution X-ray diffraction data and an electron spin echo envelope modulation (ESEEM) analysis provide direct evidence that the ligand is C rather than N.[17] Almost at the same time, DeBeer and coworkers used X-ray emission spectroscopy (XES) to confirmed that the central was C.[19] A year later, Ribbe and coworkers demonstrated that the central carbide of the FeMo cluster originates from S-adenosylmethionine (SAM) and is inserted by the radical SAM enzyme NifB.[100]

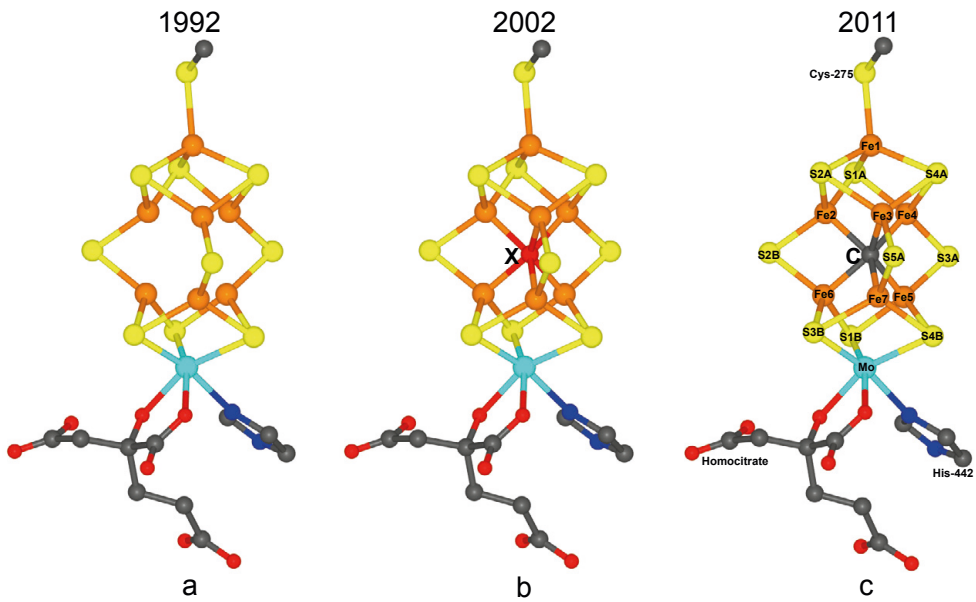


Figure 3.4: The structure of the FeMo cluster based on data from a) 1992 (PDB: 1N2C), b) 2002 (PDB: 1M1N), and c) 2011(PDB: 3U7Q).

3.1.2 Fe-Nitrogenase

Compared to Mo-nitrogenase, V-nitrogenase and Fe-nitrogenase differ not only by the replacement of the Mo ion with V or Fe but also by the presence of two additional G subunits in the VFe and FeFe proteins. These two alternative nitrogenases are much less stable than Mo-nitrogenase, making VFe and FeFe proteins less likely to crystallize. It was not until 2023 that Einsle and coworkers reported the first crystal structure of Fe-nitrogenase and its FeFe cluster from *Azotobacter vinelandii*.[95] Almost the same time, another study on the structure of Fe-nitrogenase, utilizing cryo-EM, was reported by Rebelein and colleagues.[101] The FeFe protein is encoded by *anfDGK*. Chains C and F in Figure 3.5 are the G subunit, the role and function of which remain unclear.

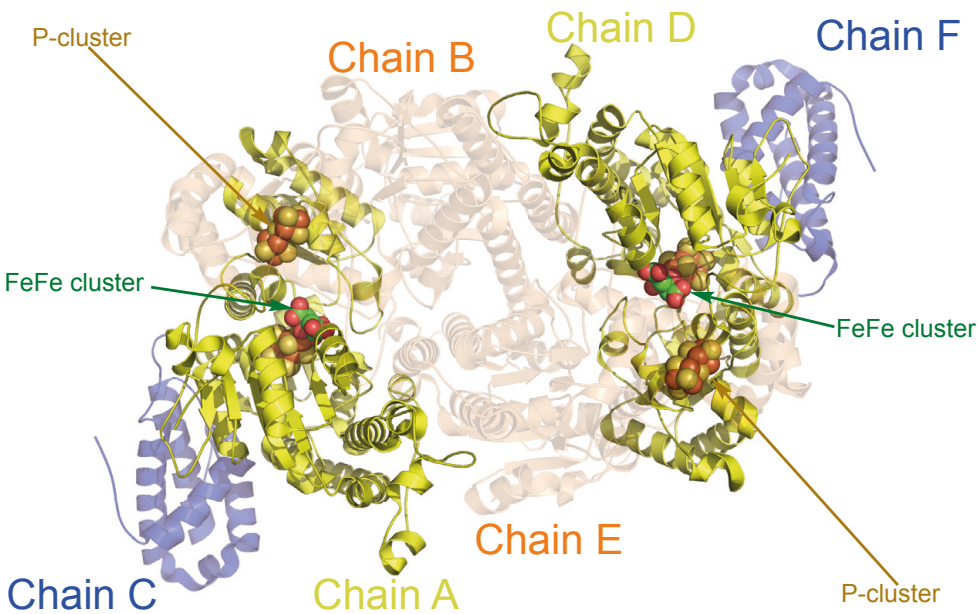


Figure 3.5: Structure of Fe-nitrogenase. Chains A and D are the D subunits and are shown in yellow. Chains B and E are the K subunits and are shown in transparent orange. Chains C and F are the G subunits and are shown in transparent blue (PDB 8BOQ).

The FeFe cluster (Figure 3.6 a) is a [8Fe-9S-C] cluster with an interstitial carbide ion and an organic homocitrate ligand at the apical iron that substitutes for Mo or V in the other isoforms and is the active site of substrate reduction. The P-cluster (Figure 3.6 b) has the same structure as in Mo and V-nitrogenase.[95,101,102]

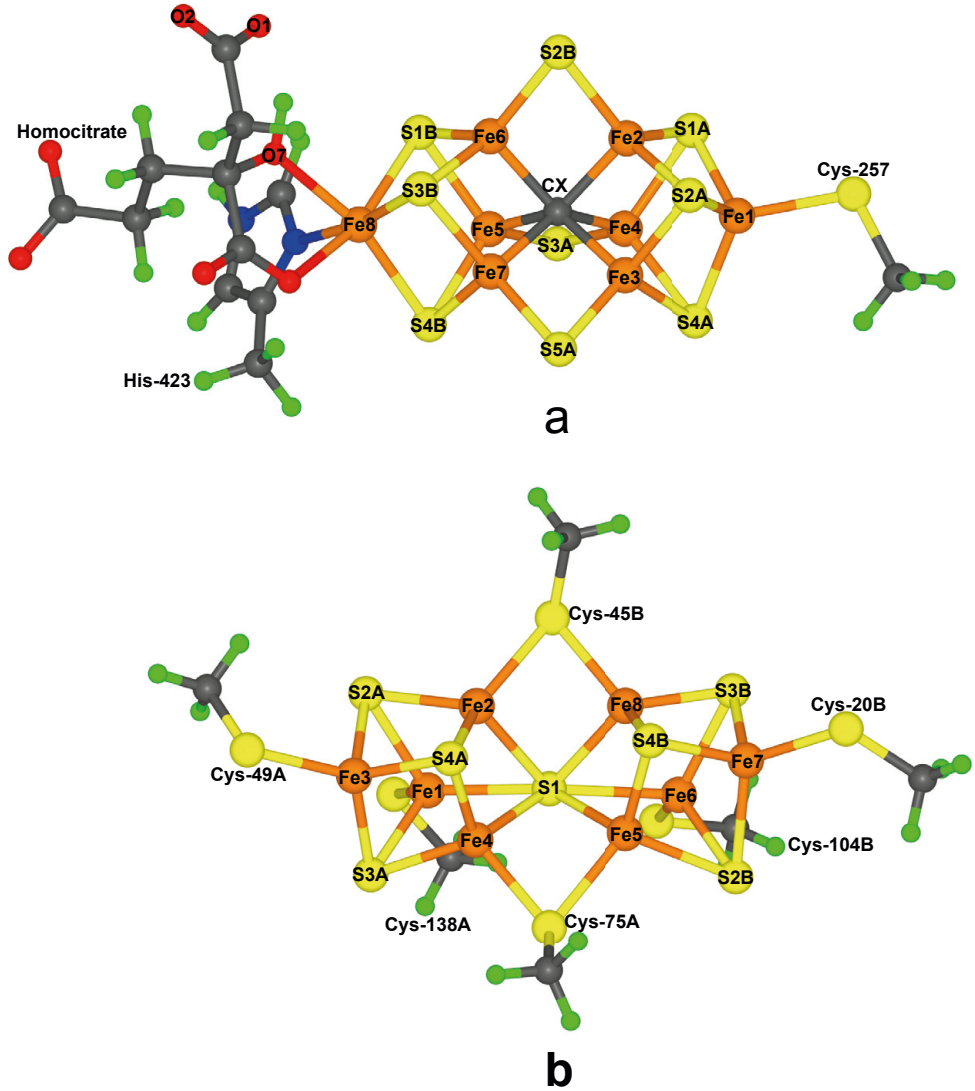


Figure 3.6: The atomic structures of the a) FeFe cluster and b) P-cluster in Fe-nitrogenase. The figure shows atom and residue names from the 8BOQ crystal structure.

3.2 Mechanism

Mo-nitrogenase is the most extensively studied, most active, and best characterized form of nitrogenase, and it is also the main enzyme studied of this thesis. The Mo-nitrogenase reaction requires 16 ATP molecules, eight electrons, and eight protons to convert one molecule of N_2 into two molecules of NH_3:

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \qquad (3.1)$$

Nitrogenase has been studied extensively using spectroscopic, biochemical, and kinetic methods. [1,4,7,14,17,20,93,97] The reaction is commonly described by the Lowe–Thorneley cycle (Figure 3.7) for dinitrogen reduction, developed in the 1970s and 1980s. It describes the kinetics of transformations among catalytic intermediates of nitrogenase, which involves nine intermediates (E_0–E_8) that differ in the number of added electrons and protons. Although nitrogenase has been studied for decades and the structures of various forms of the MoFe and Fe proteins and their complexes have been determined, many mechanistic questions remain unanswered.

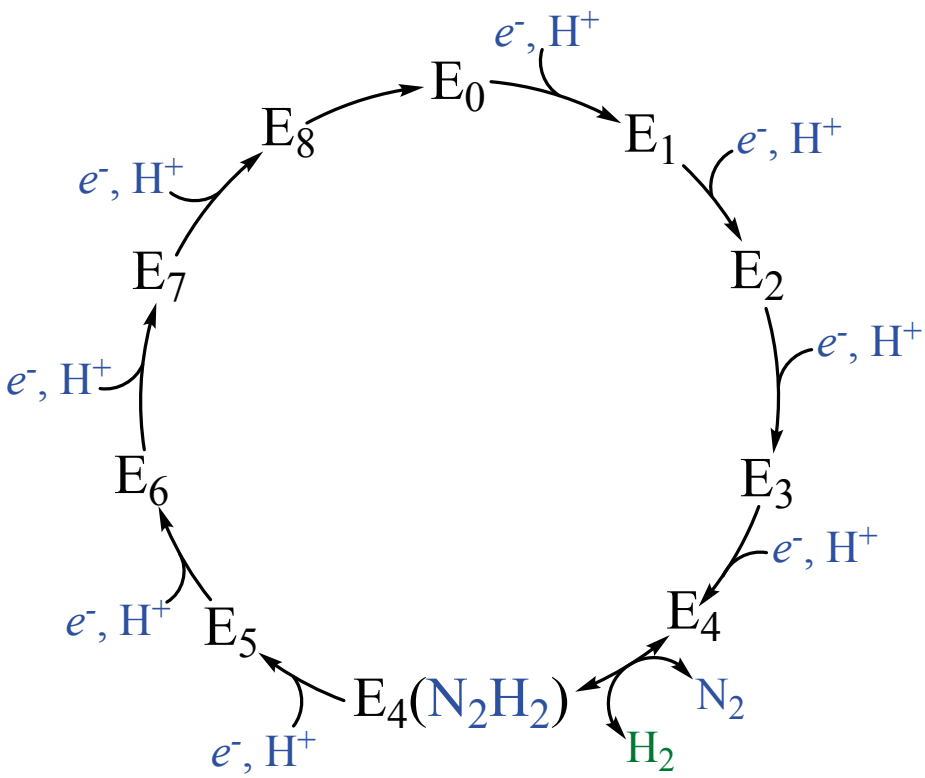


Figure 3.7: Simplified kinetic Lowe–Thorneley scheme.

3.2.1 E_0–E_4 States

The E_0–E_4 states represent the first part of the Lowe–Thorneley (LT) kinetic model.[103,104] Some of these intermediates release H_2. Recent advancements in nitrogenase research have significantly enhanced our understanding of the early catalytic states (E_0 to E_4) and the mechanism of biological nitrogen fixation. This progress has been made possible through a combination of advanced spectroscopic techniques, computational methods, and kinetic studies.

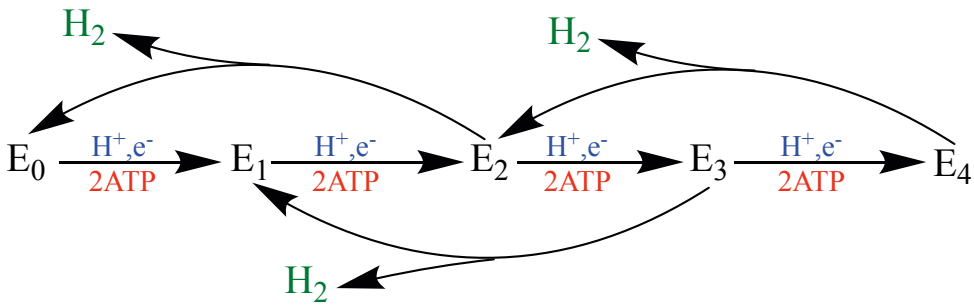


Figure 3.8: Early stages of the Lowe-Thorneley kinetic model for N_2 reduction.

The resting E_0 state is well-documented through accurate crystal structures[17–19] and QM/MM calculations.[105,106] Quantum refinement has shown that this state does not contain any additional protons.[105] The E_0 state has an odd number of unpaired electrons, which allows for EPR (electron paramagnetic resonance) spectroscopic studies, including pulsed methods like electron nuclear double resonance (ENDOR). The E_1 state arises after the first electron and proton transfer to the FeMo cluster. Due to its even spin state, E_1 is challenging to study using spin-selective techniques like EPR. However, XAS and Mössbauer spectroscopy suggest that the additional electron in E_1 primarily resides on an Fe atom within the FeMo cluster, rather than on the Mo atom.[15,107] Computational studies indicate that the most energetically favorable protonation site in E_1 is the S2B atom of FeMo cluster.[31,41] The E_2 state arises after the accumulation of two electrons and two protons by the FeMo cluster. Although EPR studies confirm the existence of E_2, it remains unclear whether this state contains a hydride or not. Cryoannealing studies suggest that E_2 possesses a high-spin ($S = 3/2$) state.[108–110]

The E_4 state arises after the accumulation of four electrons and four protons. The E_4 state is regarded as the reactive state to which N_2 is proposed to bind and from which the reduction begins.[4,6] An additional feature of the LT cycle is that the E_2, E_3, and E_4 states can convert back to the E_0, E_1, and E_2 states by generating H_2.[111] This represents nonproductive hydrogen release that competes with N_2 reduction. According to ^{95}Mo ENDOR spectroscopy, the four electrons and two of the protons in E_4 are formally assigned to two bridging hydrides ([Fe–H–Fe]).[91,112] This state

is short-lived, so the enzyme must ensure that N_2 is readily available for binding before H_2 is eliminated from the E_4 state.[4,6]

The location of the added protons in the various E_n states is hard to determine. The FeMo cluster comprises one molybdenum, seven iron, one carbide, nine sulfur ions, and various ligands such as cysteine, histidine, and homocitrate, which sum up to at least 21 sites capable of protonation. Each of these sites typically allows for protonation at two or three distinct positions. Additionally, a hydride ion can bind to a single metal ion or bridge between two ions, resulting in over 50 potential positions for each additional proton. Consequently, for the E_4 state with its four added protons, there are theoretically over $50^4 = 6.25 \times 10^6$ distinct protonation configurations. Each configuration can further adopt 35 possible broken-symmetry states and two to four spin states, resulting in approximately 10^9 potential states. This vast complexity makes it exceedingly challenging to systematically study all these states comprehensively.[30,31]

Hoffman and coworkers have proposed that N_2 binds to the E_4 state via a *reductive elimination* (re) of H_2.[1,4,6,113] After the addition of four electrons and protons to the FeMo cluster, nitrogenase reaches the E_4 state, which can bind N_2 via *reductive elimination* (*re*) of H_2, forming $E_4(N_2H_2)$. This mechanism requires the loss of one H_2 molecule for each N_2 molecule, which explains the requirement of $8e^-$ and why H_2 is a compulsory byproduct. By identifying the EPR spectra for all states during the relaxation of $E_4(4H)$ and $E_4(N_2H_2)$, it has been shown that WT $E_4(N_2H_2)$ can relax back to the E_4 state via *oxidative addition* (*oa*; Figure 3.9).[114,115] The *re/oa* mechanism is reversible, so an increase in H_2 partial pressure can push the $E_4(N_2H_2)$ state back to $E_4(4H)$ through *oxidative addition*.[116]

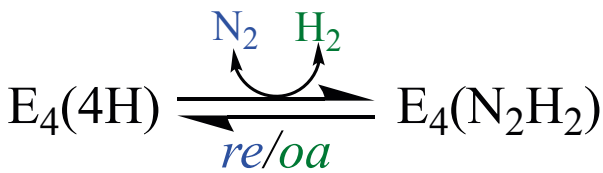


Figure 3.9: Scheme of *re/oa* mechanism.

The relative energies of the various BS states have been shown to be sensitive to the specific exchange and correlation functional used.[10] Moreover, different functionals can lead to variations in the predicted ordering of possible E_4 isomers, creating ambiguity in identifying the most stable configuration.[30] This sensitivity highlights the importance of considering the limitations of DFT and the potential for variations in results based on the chosen functional.

3.2.2 $E_4(N_2H_2)$–E_8

Following the formation of the $E_4(N_2H_2)$ intermediate, the catalytic cycle continues through four more states (E_5–E_8), which involve partially reduced intermediates of N_2.[1–4,6,11] Two possible pathways have been proposed for this latter part of the mechanism: the alternating pathway and the distal pathway. The two pathways are illustrated in Figure 3.10.

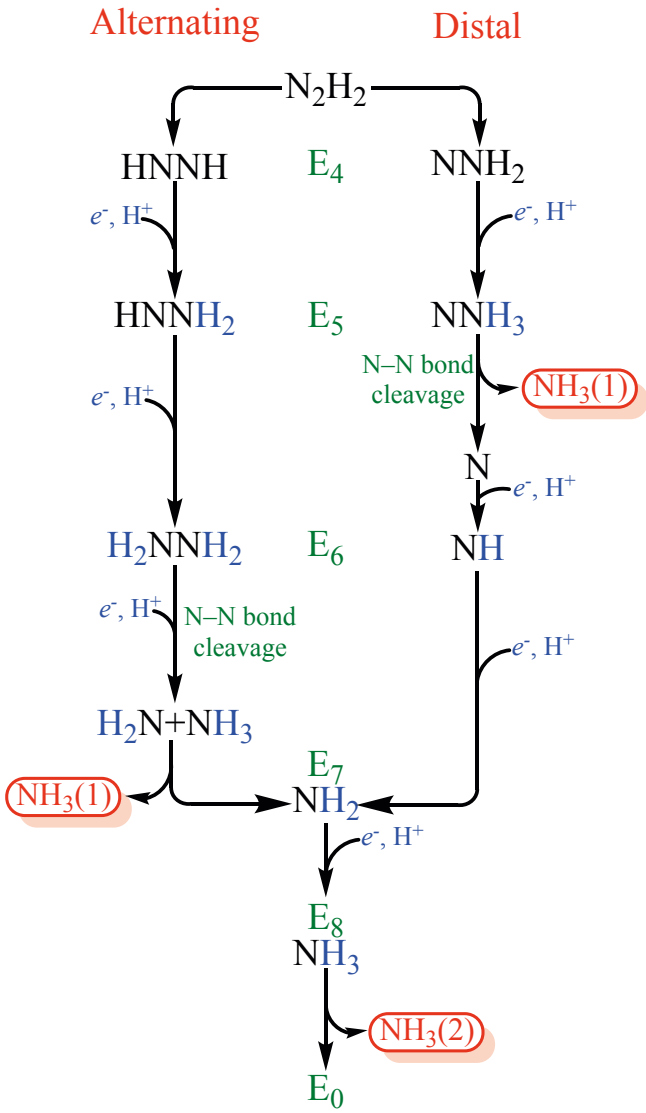


Figure 3.10: The ***alternating*** and ***distal*** mechanisms of N_2 reduction for the later parts of the Lowe–Thorneley cycle of nitrogenase.

Alternating Pathway: In the alternating pathway, the two nitrogen atoms are hydrogenated alternately.[13,117] This pathway involves the formation of hydrazine (N_2H_4), at the E_6 intermediate stage. The first NH_3 is released during the subsequent hydrogenation step, leading to the formation of E_7.[118,119] The alternating pathway is supported by inorganic nitrogenase model complexes,[120] as well as by some computational and crystallographic studies of nitrogenase.[28,39,40,121] The experimental evidence includes the observation of hydrazine as a product when Mo-nitrogenase is quenched with acid during N_2 reduction and the fact that hydrazine reduction is not inhibited by H_2, unlike diazene and N_2 reduction. Additionally, it has been shown that N_2, N_2H_2, CH_3NH_2, and N_2H_4 react through a common intermediate.[12,122–124]

Distal Pathway: In the distal pathway, the distal nitrogen atom of Fe-bound N_2 undergoes three sequential hydrogenation steps. After the addition of the fifth electron and proton, the first NH_3 molecule is released, resulting in a nitrido (E_5) species. The remaining nitrogen atom, now a nitrido-N, undergoes three more hydrogenations to produce the second NH_3. This mechanism was originally suggested by Chatt and has gained support from N_2-fixing inorganic model complexes.[1–4,6]

3.3 Electronic Configurations

3.3.1 FeMo Cluster

FeMo cluster is the catalytic site of MoFe protein in nitrogenase. The resting state of the FeMo cluster, often denoted as E_0, is characterized by a quartet ground spin state ($S = 3/2$) according to EPR spectroscopy.[125,126] This state has been extensively studied both experimentally and computationally, and a generally accepted electronic structure model has emerged.

Early Mo K-edge XAS[127,128] studies and ^{95}Mo ENDOR.[129–131] investigations suggested that the molybdenum ion in the FeMo cluster was in the Mo^{4+} oxidation state. However, high-energy resolution fluorescence-detected X-ray absorption spectroscopy (HERFD-XAS) studies[8] showed that the molybdenum is best described as Mo^{3+}. This Mo^{3+} ion has an unusual electronic configuration, with two electrons in the *d* orbital with down spin and one electron in the *d* orbital with up spin, which does not follow Hund's rule (Figure 3.11). This non-Hund configuration was proved by a time-dependent density functional theory study.[8]

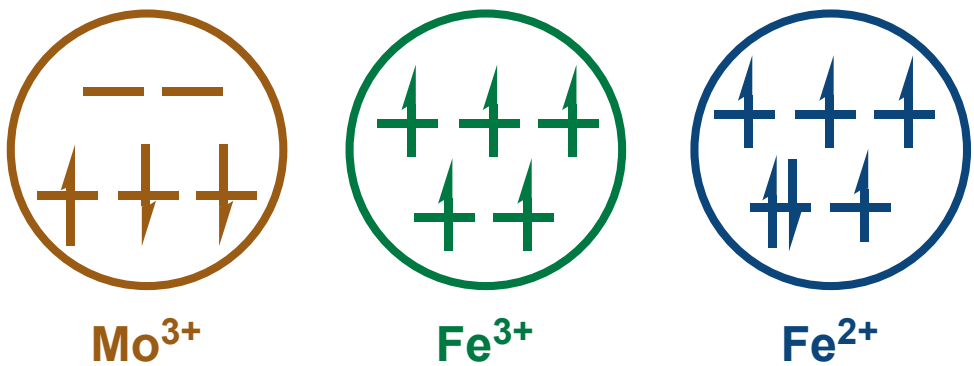


Figure 3.11: The electronic configurations for Mo^{3+}, Fe^{3+}, and Fe^{2+} in the FeMo cluster.

In early studies of nitrogenase, three different oxidation states were considered: $Mo(IV)Fe(III)Fe(II)_6$,[132] $Mo(IV)Fe(III)_3Fe(II)_4$,[107] and $Mo(IV)Fe(III)_5Fe(II)_2$.[133] Spatzal and colleagues applied spatially resolved anomalous dispersion to the MoFe protein, obtaining site-specific Fe K-edges for each Fe atom in FeMo cluster and the P-cluster. Their data supported the $Mo(III)Fe(III)_4Fe(II)_3$ oxidation state.[134] A study by Siegbahn's also supported this charge for the resting state, as other assignments gave unreasonable energetics for the mechanism.[135] DeBeer and colleagues demonstrated that the $Mo(III)Fe(III)_4Fe(II)_3$ state is best supported by DFT calculations and analysis of predicted Mössbauer isomer shifts.[136] This

oxidation state distribution of FeMo cluster gives an overall charge of –1 for the FeMo cluster.

Noodleman and coworkers introduced broken-symmetry (BS) DFT to study the electronic properties of the FeMo cluster.[137–139] The seven Fe ions in the FeMo cluster are in their high-spin states with five unpaired electrons for Fe(III) and four for Fe(II). These combine antiferromagnetically to a net quartet state for the E_0 resting state. This can be done in 35 different ways (Figure 3.12).Subsequent studies have shown that a particular category of BS solutions, referred to as BS7, consistently emerges as energetically most favorable across various studies, regardless of the specific charge state or the identity of the interstitial atom within the FeMo cluster.[105,106,140] The BS7 solutions are characterized by the maximization of antiferromagnetic coupling between neighboring Fe atoms within the cluster. They consist of four iron sites with spin-up (α) configuration coupled to three iron sites with spin-down (β) configuration. There are three specific BS7 solutions, often labeled according to the iron ions with spin-down configurations: BS7-235, BS7-346, and BS7-247. These three solutions are energetically very similar, typically falling within 4 kJ/mol of each other. The primary distinction between these BS7 spin isomers lies in the specific arrangement of Fe atoms with spin-down configurations. In the BS7 configuration, Fe3–Fe4 and Fe5–Fe7 form ferromagnetically coupled pairs within the cluster. Therefore, in the E_0 state of FeMo cluster, the iron sites Fe2 and Fe6 are the most highly oxidized positions in the cluster.[2,3]

In this thesis, we used this BS approach in all QM calculations.[137] The various BS states were obtained either by swapping the coordinates of the Fe ions[108] or with the fragment approach by Szilagyi and Winslow.[141] We usually start from the BS10-147 state (i.e. with minority spin on Fe1, Fe4 and Fe7). Then, we did a comprehensive study of all BS states for the lowest structure and used the best BS state found for all structures. All QM calculations were performed with the Turbomole software.[142]

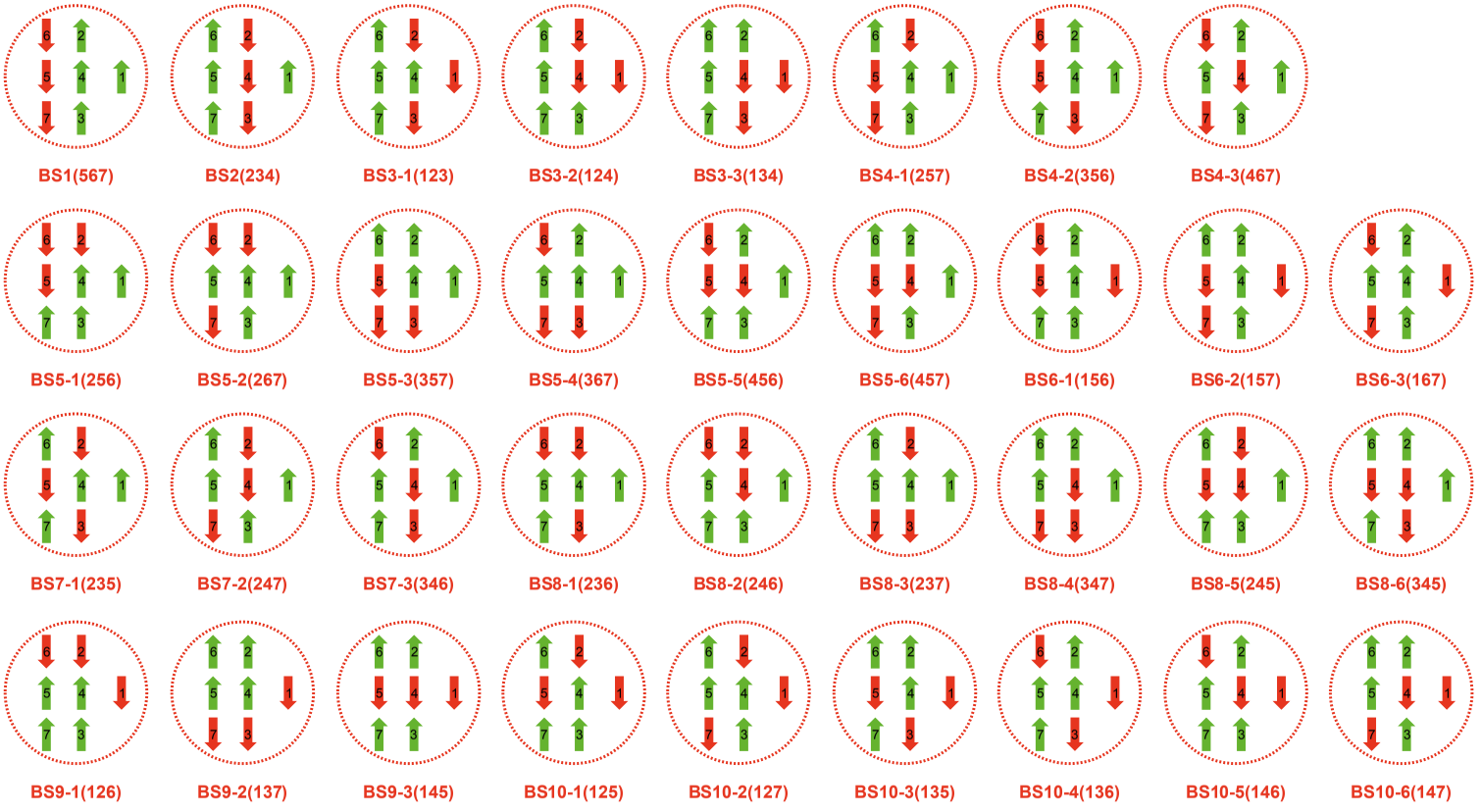
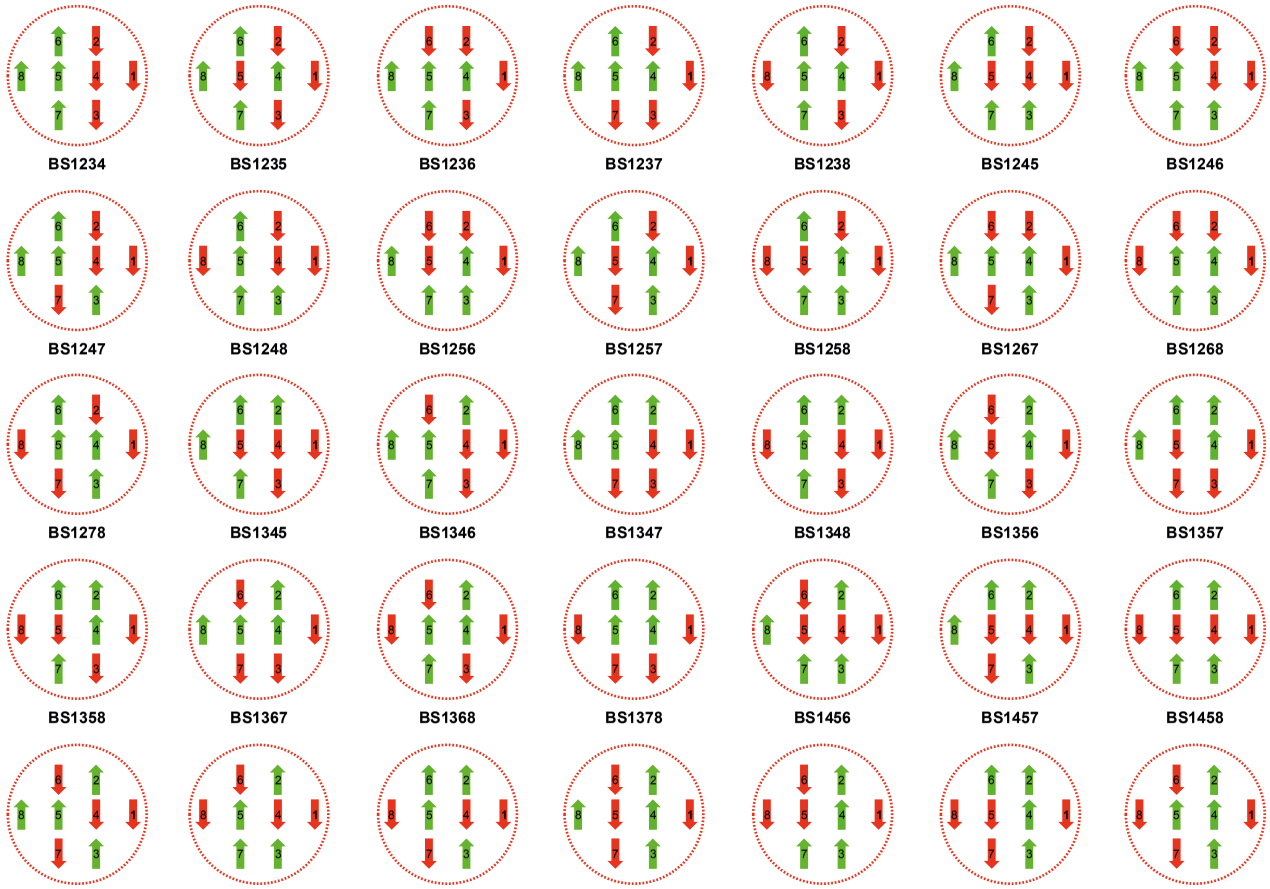


Figure 3.12: The 35 possible BS states of the FeMo cluster in nitrogenase.

3.3.2 FeFe Cluster

The FeFe cluster, found in FeFe protein, is less well studied and has a [8Fe-9S:C] chemical composition.[95,101] Like the FeMo cluster, the FeFe cluster also has a μ_2-sulfide-bridged dicubane architecture with a central carbide atom. In the resting state (E_0), which is the state typically isolated in the presence of a chemical reductant, the FeFe cluster is diamagnetic, implying an even number of ferrous (Fe^{2+}) and ferric (Fe^{3+}) sites. Mössbauer spectroscopy suggests that the average isomer shift of FeFe cluster is ~0.4 mm/s, consistent with an even number of Fe^{2+} and Fe^{3+} sites.[143]

For the FeFe cluster in its resting E_0 state, the oxidation states are assigned as $4Fe^{3+}4Fe^{2+}$ with a singlet spin state ($S = 0$) according to EPR spectroscopy.[144] These spins are coupled antiferromagnetically to result in a singlet state. Thus, the cluster can be described with four Fe ions having a surplus of α spin and the other four having a surplus of β spin. There are multiple ways to arrange these spins among the Fe ions, resulting in 70 different BS states, shown in Figure 3.13.


BS1234
BS1235
BS1236
BS1237
BS1238
BS1245
BS1246
BS1247
BS1248
BS1256
BS1257
BS1258
BS1267
BS1268
BS1278
BS1345
BS1346
BS1347
BS1348
BS1356
BS1357
BS1358
BS1367
BS1368
BS1378
BS1456
BS1457
BS1458

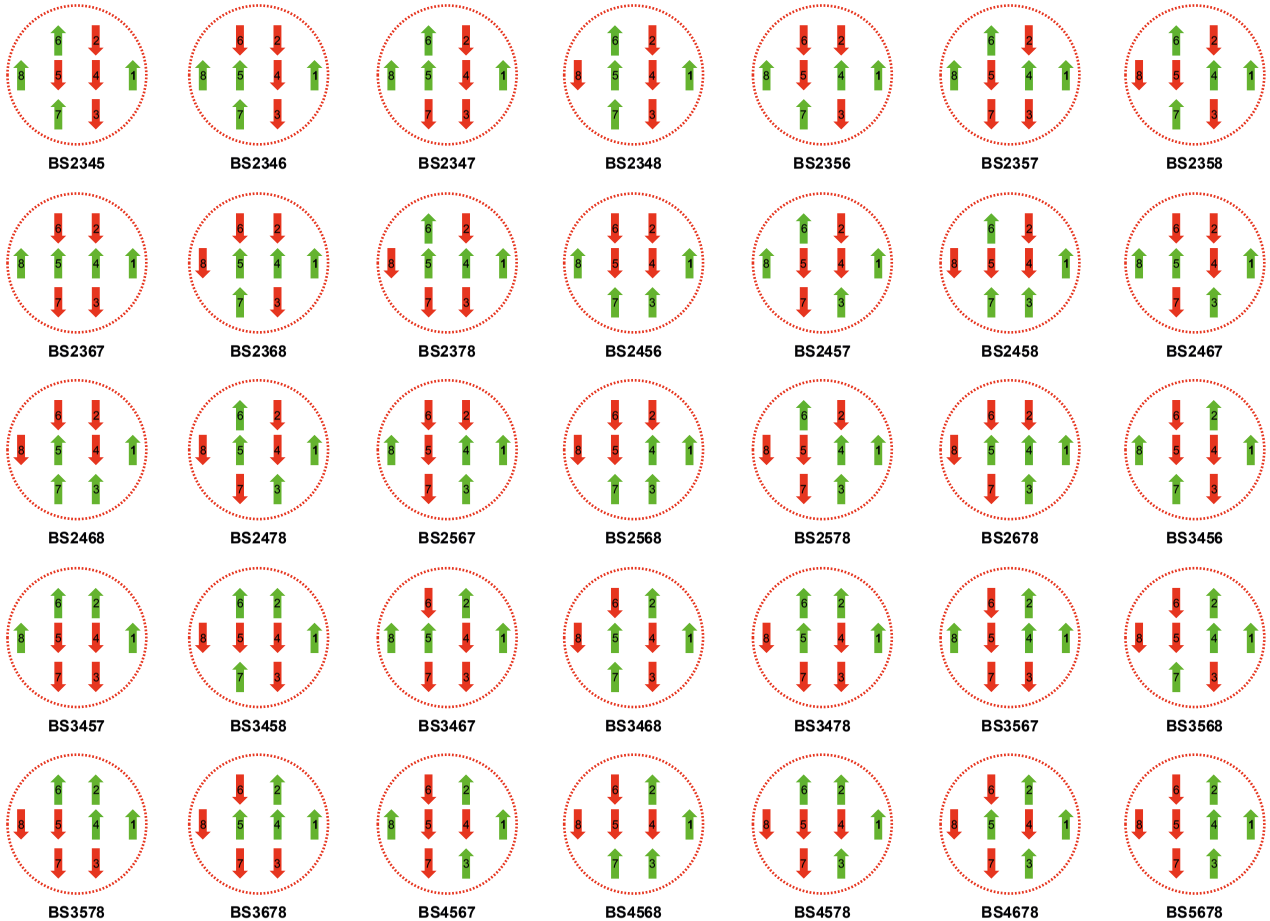


Figure 3.13: The 70 possible BS states of the FeFe cluster in Fe-onlu nitrogenase.

3.4 Critical Components

3.4.1 His195

His195, a residue on the α-subunit of the MoFe protein in nitrogenase, is critical for the functionality of nitrogenase. This residue forms a hydrogen bond with a bridging sulfide ion (S2B) in the FeMo cluster,[18,19] which may serve as a proton donor during the nitrogen reduction process.[145,146] Substituting His195 with glutamine (His→Gln) drastically reduces nitrogen fixation activity to less than 1% of the wild-type, while acetylene reduction remains unaffected.[146,147] This indicates a specific role of His195 in nitrogen reduction. The His195→Gln195 substituted MoFe protein still exhibits N_2 inhibition of acetylene and proton reduction, suggesting that N_2 and acetylene share a binding site. Although it has been proposed that the Nε–H side chain of His195 directly donates protons to FeMo cluster, the mechanism by which the proton is replenished on His195 remains unclear. Rotating the imidazole ring for multiple proton donations could disrupt this process.[148] Photolysis studies on His195→Asn195 and His195→Gln195 nitrogenase variants under high-CO conditions reveal changes in CO binding to FeMo cluster, highlighting the influence of His195 on CO binding.[149,150] In summary, His195 is essential for nitrogen reduction in nitrogenase, with its sensitivity to substitution and impact on CO binding underscoring its critical role.

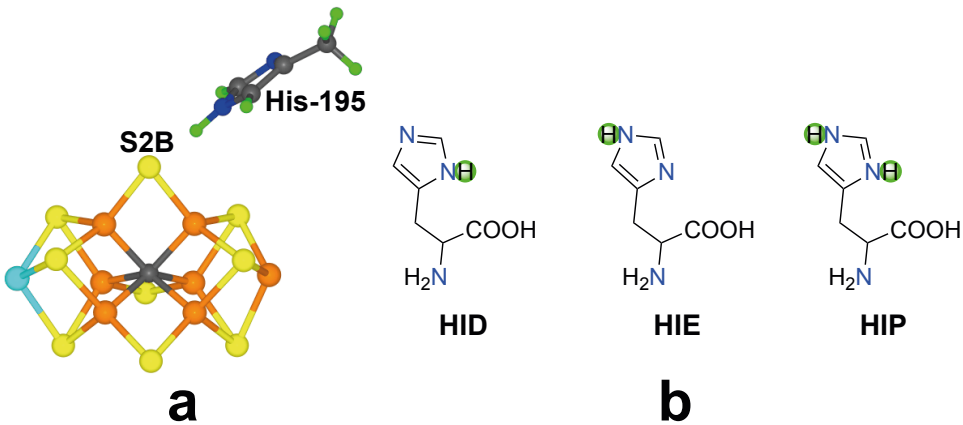


Figure 3.14: (a) The relative position of His195. (b) The three possible protonation states of histidine.

3.4.2 Homocitrate

Homocitrate is essential for the FeMo cluster in nitrogenase, coordinating to the Mo atom through its 2-hydroxy and 2-carboxy groups, which are crucial for the enzyme's function and structure.[11] Its role in nitrogen fixation, compared to similar molecules like citrate, is not fully understood but may involve facilitating proton transfer to the active site, potentially through a network of water molecules.[148,151] Computational models suggest that homocitrate helps transfer a proton from its hydroxyl group to a nearby sulfide, S2B, during hydride intermediate formation.[28,37,39,40] Additionally, homocitrate is vital for the correct assembly and insertion of the FeMo cluster into the nitrogenase enzyme, as FeMo cluster precursors lacking homocitrate cannot be integrated into the MoFe protein. Substituting homocitrate with other carboxylic acids severely impairs nitrogen reduction activity, especially for reactions needing multiple protonation steps, highlighting homocitrate's critical role in proton transfer. The longer "arm" of homocitrate may influence water molecule positioning and proton transfer efficiency, with crystallographic studies showing that citrate's shorter length results in a less effective proton transfer pathway and reduced activity. Further research is needed to fully understand homocitrate's mechanistic role in nitrogenase, particularly its impact on proton transfer and the FeMo cluster electronic structure.[11,152] In our studies (**Paper III**, **IV**, **VI**), we show that homocitrate may play a role as proton buffer, stabilizing certain intermediates, e.g. H_2NNH_2 and NH_3.

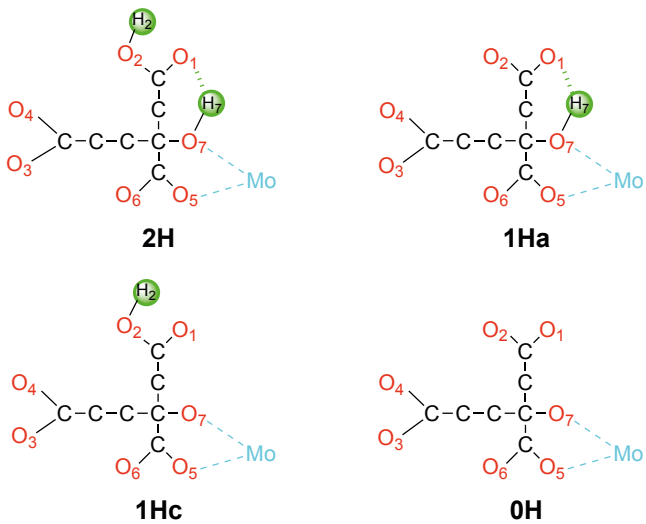


Figure 3.15: Four considered protonation states of homocitrate, 2H, 1Ha, 1Hc, and 0H. Atom numbers are also shown. Nonpolar H atoms are omitted. The charge of homocitrate is −2, −3, −3, and −4, respectively, in these four protonation states.

3.4.3 S2B

S2B is a bridging μ_2 sulfide ligand in the FeMo cluster of nitrogenase, which plays an important role in substrate and inhibitor binding. Crystallographic studies of the MoFe, VFe, and FeFe proteins have consistently demonstrated that S2B is bridging between Fe2 and Fe6, beneath the conserved His195 residue. During ligand binding and exchange, S2B can be reversibly displaced by CO,[97] forming a bridging carbonyl in the "low-CO" state under turnover conditions. Similarly, selenide can replace S2B during turnover with selenocyanide, and extended turnover can replace also other sulfides in the FeMo cluster, indicating the structural flexibility of the cluster.[153] In VFe protein structures, a light atom, potentially NH^{2-} or OH^{-}, replaces S2B, underscoring its reversible exchangeability and its function in creating a coordination site for substrates or inhibitors absent in the resting state.[154–156] However, while crystallographic data emphasize the lability and displacement of S2B, theoretical studies suggest it might undergo protonation and destabilize an Fe–S bond, forming a dangling thiol that remains associated with the cluster.[156,157] This discrepancy between experimental and theoretical findings underscores the need for further research to fully understand the behavior of S2B during nitrogenase catalysis.

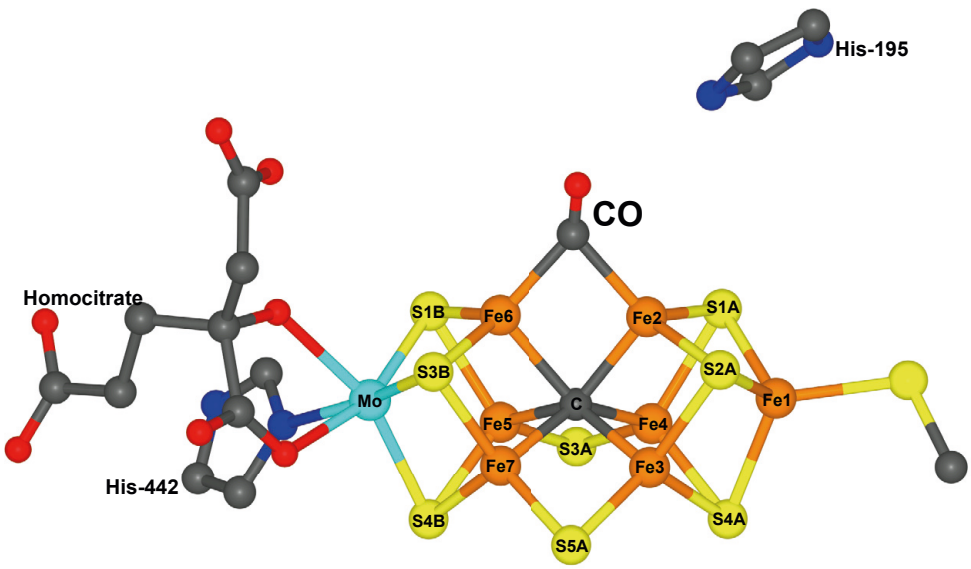


Figure 3.16: The active sites of CO-inhibited Mo-nitrogenase (PDB: 4TKV)

4 Summary of papers

This thesis is based on eight publications, Papers I–VIII. In this thesis, we performed systematic studies of nitrogenase using the combined quantum and molecular mechanics (QM/MM) approach:

- **E_0–E_4 States:** We investigated the formation of H_2 from the E_2–E_4 states (**Paper I**) and the binding of N_2 to the E_0–E_4 states (**Paper II**).
- **E_5–E_8 States:** We proposed putative reaction mechanisms of the cluster without the S2B ligand dissociated (**Paper III**), along with study the proton transfer process both with and without the S2B ligand (**Paper IV**).
- **S2B Half-Dissociated Process:** We examined half-dissociation of the S2B ligand for the E_2 intermediate (**Paper V**) and the latter part of the reaction mechanism of nitrogenase with a half-dissociated S2B ligand (**Paper VI**).
- **Redox Potentials:** We calculated redox potentials of the metal clusters in Mo-nitrogenase (**Paper VII**).
- **Fe-Nitrogenase:** We set up QM/MM calculations for Fe-nitrogenase and examined the electronic structure and protonation states of the FeFe cluster in the E_0 and E_1 states (**Paper VIII**).

Paper I: H_2 Formation in States E_2–E_4

The aim of this project is to investigate the formation of H_2 from the E_2–E_4 states of nitrogenase. The enzyme must be loaded with four electrons and protons (reaching the E_4 state) before N_2 can bind.[1,4] Current research suggests that N_2 binding is facilitated by the dissociation of H_2, which forms through the reductive elimination of two hydride ions bridging two Fe ions each.[91,92,113,116,158] This explains why H_2 is a compulsory byproduct in the reaction mechanism.

Despite numerous studies, both experimental and theoretical, we still do not fully understand many aspects of the reaction mechanisms of nitrogenase.[1,4,7,8,14,17–20,93,97,152,159–161] One important reason is that different DFT methods give widely different predictions of the relative stability of various models of the active site of nitrogenase.[10,31,36,159]

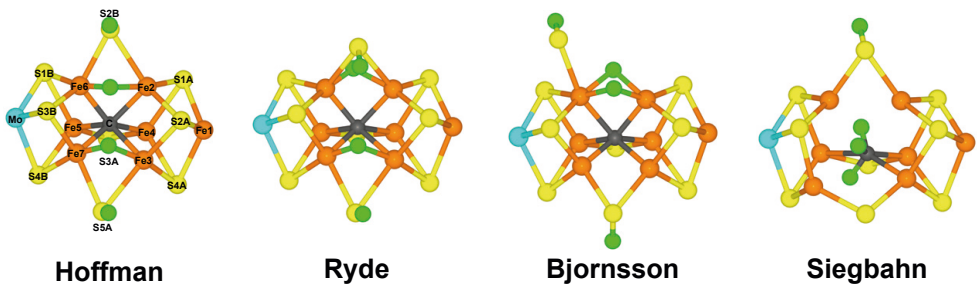


Figure 4.1: Previously suggested E_4 structures.

Recent research has examined the formation and dissociation of H_2 formation and dissociation within the FeMo cluster. Different DFT methods give different results for the E_2 state[31]: TPSS suggests a structure with a proton on S2B and a bridging hydride ion between Fe2 and Fe6,[36] while other functionals like r^{2}SCAN and TPSSh propose similar structure but with half-dissociated protonated S2B ions.[34] In contrast, B3LYP favors a doubly protonated central carbide ion. For the E_3 state, B3LYP also prefers a triply protonated carbide ion, contrasting with other functionals suggesting a structure involving protonated S2B and hydride ions bridging Fe2 and Fe6.[34] Likewise, several hypotheses for the E_4 structure have been suggested. ENDOR experiments show that E_4 should contain two bridging hydride ions.[92,115] **Hoffman** and coworkers have suggested a structure in which S2B and S5A are protonated and there are two hydride ions bridging the Fe2/6 and the Fe3/7 pairs of ions.[157] **Ryde** and coworkers studied structures with two bridging hydride ions and showed that a structure with the protons and the hydride ions at the same positions as suggested by Hoffman is most stable, but three of them are pointing

towards to other faces of the FeMo cluster.[30] **Björnsson** and coworkers have advocated for structures in which the two hydride ions both bridge Fe2/6, S2B is protonated and dissociated from one Fe and the second proton is on S5A. **Siegbahn** suggested a structure with a triply protonated carbide ion and the fourth proton on S2B with B3LYP.[135,162]

Another related question is the formation and dissociation of H_2 from the FeMo cluster which is crucial for nitrogenase's catalytic cycle. If H_2 forms in the E_2 or E_3 states, the enzyme will go back to E_0 or E_1 states. However, H_2 formation in the E_4 state is beneficial only in combination with N_2 binding.[4] **Thorhallsson** and coworkers studied the formation of H_2 from the E_2 state calculations and they found that activation barriers of 86–95 kJ/mol from a structure with S2B dissociated from one of the Fe ions.[163]

In this study, we extend previous investigations by examining the formation and dissociation of H_2 from the FeMo cluster in the E_2–E_4 states because it is not possible until the E_2 state when two protons have accumulated on the cluster. The investigation has three primary objectives: **First**, to explore the H_2 formation reaction across various structures, including the combination of either two hydride ions or one proton and one hydride ion, and to study reactions at different cluster positions, such as terminal and bridging hydride ions located on different Fe ions or the same Fe ion. **Second**, to investigate the ease of interconversion between different protonation states, which may elucidate whether the enzyme can prevent H_2 formation by positioning the proton and hydride ions far apart. **Third**, to compare the results obtained from four different DFT functionals. It is known that the enzyme must prevent H_2 loss in the E_2 and E_3 states, while H_2 formation is essential in the E_4 state, but only when accompanied by N_2 binding. If H_2 formation occurs too easily in the E_2 and E_3 states, it may indicate issues with the DFT method used or suggest that the correct structure has not yet been identified.

Our studies find large differences in the predictions of the different methods. B3LYP strongly favors protonation of the central carbide ion and H_2 cannot form from such structures. On the other hand, with TPSS, r^{2}SCAN and TPSSh, H_2 formation is strongly exothermic for all structures and therefore needs strict kinetic control to be avoided. For the E_2 state, the kinetic barriers for the low-energy structures are high enough to avoid H_2 formation. However, for both the E_3 and E_4 states, all three methods predict that the best structure has two hydride ions bridging the same pair of Fe ions (Fe2 and Fe6) and these two ions can combine to form H_2 with an activation barrier of only 29–57 kJ/mol, corresponding to rates of 7×10^2 to 5×10^7 s^{-1}, i.e. much faster than the turnover rate of the enzyme (1–5 s^{-1}). We have also studied H-atom movements within the FeMo cluster, showing that the various protonation states can quite freely be interconverted (activation barriers of 12–69 kJ/mol).

Paper II: N_2 Binding to the E_0–E_4 States

In this project, we aimed at investigating the binding of nitrogen (N_2) to the various states of nitrogenase. Experimentally, it is known that three or four electrons must be added before nitrogenase can bind N_2. Our aim was to study N_2 binding to the E_0–E_4 states by using QM/MM.

Despite extensive studies of nitrogenase using spectroscopic, biochemical, and kinetic methods, the details of the reaction remain poorly understood.[1,4,7,14,17,20,93,97] The E_0 resting state has been thoroughly characterized by crystallography, spectroscopy, and computational studies.[4,17,106,134] The E_1 state, examined using X-ray absorption and Mössbauer spectroscopy,[15,107] likely contains a proton on the S2B μ_2 bridging sulfide ion.[31,41] The E_2 state is known to involve two conformers, with at least one containing an iron-bound hydride ion.[109–111,164] The E_3 state is less studied experimentally due to its EPR-silent nature.[1,4,14] The E_4 state, characterized by EPR and ENDOR spectroscopy, has been shown to contain two hydride ions that bridge between two Fe ions of the FeMo cluster.[91,92,115] It has been established that N_2 binds to the E_3 and E_4 states, but not the E_0–E_2 states.[7,103,104,114,115] In connection with N_2 binding, H_2 is released by reductive elimination, i.e., by forming H_2 from two hydride ions.[92,113,115,158] Subsequently, N_2 is progressively reduced and protonated to form two molecules of NH_3. Based on mutational studies of Val70, the Fe2–Fe3–Fe6–Fe7 face of the FeMo cluster is proposed as the primary site for N_2 reduction, with Fe2 or Fe6 being the most likely N_2 binding sites.[165,166]

Nitrogenase has also been extensively studied by DFT methods.[159] **Blöchl**, **Kästner**, and coworkers proposed that N_2 binds to Fe7 following the dissociation of S5A.[165,167] Other groups have also suggested that such half-dissociation of the μ_2-bridging sulfide ions may enhance N_2 binding, but mainly for S2B and N_2 binding to Fe2 or Fe.[168–170] Likewise, crystallographic studies have indicated that S2B may sometimes be replaced by other ligands,[97,154,171] indicating that sulfide lability may be mechanistically relevant.[172,173] **Björnsson** and coworkers indicated N_2 binding to Fe2 or Fe6 in the E_4 state with favorable binding energies (56 or 43 kJ/mol), and unfavorable N_2-binding energies to E_0, E_1, and E_2, but slightly favorable to E_4 (17 kJ/mol), emphasizing the role of two doubly occupied 3*d* orbitals on Fe binding N_2.[174] **Hoffman** and coworkers suggested that reductive elimination of H_2 from E_4 is necessary for N_2 binding and proposed a structure for E_4 with two protons on S2B and S5A and two hydride ions bridging Fe2/6 and Fe3/7.[157,175] DFT calculations showed a favorable binding free energy (13 kJ/mol) with an endergonic formation of H_2 (20 kJ/mol) and a barrier of 49 kJ/mol from E_4. **Dance** showed that side-on binding of N_2 is less favorable than end-on binding, and bridging N_2 between two Fe ions is unfavorable. He initially suggested binding to Fe6 and later proposed a two-step binding process with a promotional N_2 binding to Fe2 (exo-position) and a reactive N_2 binding to Fe6 (endo-position), reporting favorable binding energies up to 38 kJ/mol, enhanced with a bound H_2 molecule (up to 59

kJ/mol).[16,176–178] **Siegbahn** argued that N_2 binding to E_0–E_4 states is endergonic and suggested that nitrogenase requires four additional electrons for N_2 binding, positioning the E_4 state as the E_0 state in his catalytic cycle. Early studies suggested N_2 binding between Fe4 and Fe6 (endergonic by 13 kJ/mol), while later studies indicated S2B dissociation with N_2 binding to Fe4 (less endergonic by 10 kJ/mol), highlighting the dependence of binding energy on the Hartree–Fock exchange in the functional.[179,180]

Thus, there is no consensus on how N_2 binds to the FeMo cluster, partly due to disagreements about the structure of the E_4 state and significant variations in the structures and energies obtained with different DFT functionals. To address this, we investigated the binding of N_2 to nitrogenase using four different DFT methods. We examined the binding of N_2 to the five E_0–E_4 states and evaluated how well different DFT functionals reproduce the experimental observation that N_2 binds only to the E_3 and E_4 states.[7,104,114,115] For the E_0–E_2 states, there is reasonable consensus on the preferred protonation states.[15,31,36,159,163,181] For the E_3 and E_4 states, we expanded previous studies on the preferred protonation states,[10,30,31,162,182,183] particularly focusing on structures where S2B has dissociated from either Fe2 or Fe6.

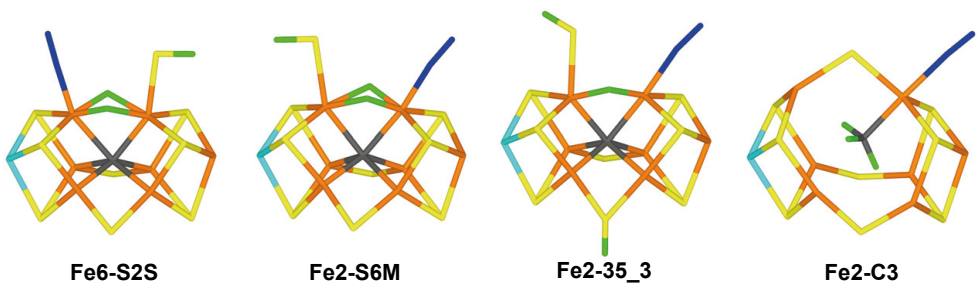


Figure 4.2: The best E_3 structures with N_2 bound, Fe6-S2S, Fe2-S6S, Fe2-3523 and Fe2-C3. The first three structures were optimised with TPSS, whereas Fe2-C3 was optimised with B3LYP.

We provide insights into the stability of various E_3 and E_4 state structures, showing that the S6M and S6S states (with two hydride ions both bridging Fe2/6, cf. Figure 4.2) are the best models for E_3 with TPSS, TPSSh, and r^{2}SCAN functionals, while B3LYP favors a triply deprotonated carbide ion. For E_4, the S6S structure is most stable with TPSS, TPSSh, and r^{2}SCAN, although the 3323 and 3523 structures are close in energy. N_2 binding is observed in the E_2–E_4 states, occurring end-on in the exo position of Fe2 or Fe6, with Fe–N bond lengths of 1.80–1.98 Å. Half-dissociation of S2B enhances N_2 binding, especially to Fe2, with TPSS showing less preference for half-dissociation than TPSSh, and r^{2}SCAN. TPSS generally favors N_2 binding to Fe6, while the other functionals prefer binding to Fe2.

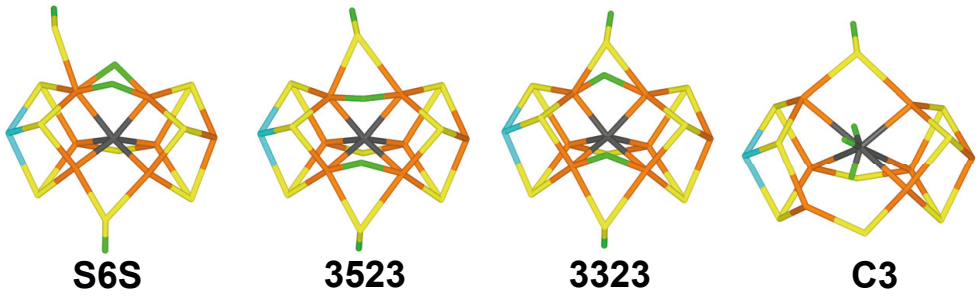


Figure 4.3: The best E_4 structures without N_2 bound.

Binding free energy varies with the DFT functional, entropy correction, and binding definition. Various groups have suggested different sizes of the entropy correction.[177,179,180,184] Using Björnsson and Siegbahn's large entropy correction (41–45 kJ/mol), no functional shows favorable binding; however, using Dance's lower entropy penalty (17 kJ/mol), TPSS favors binding to E_3 and E_4, and r^2SCAN to E_4. Hybrid functionals in general give weaker binding, favoring Fe2 binding and central carbide protonation. Our results indicate that structures with S2B dissociated from Fe2 or Fe6 and those with hydride ions bridging Fe2 and Fe6 are likely involved in the reaction mechanism. We find no support for the suggestion that reductive elimination of hydride ions in E_4 enhances N_2 binding (the formation of H_2 from E_4 will move the cluster back to a reactive state of E_2, but N_2 binding to such a E_2 state is still unfavorable). Further studies on H_2 dissociation from the FeMo cluster and its effect on N_2 binding are needed.

Paper III: Second Part of the Reaction Mechanism with S2B Bound

The goal of this project was to study the reaction mechanism of nitrogenase without the dissociation of any sulfide ligand QM/MM calculations. We started from a state where N_2 is bound to the cluster and protonated to N_2H_2, following the dissociation of H_2.

The debate on whether nitrogenases follow alternating or distal pathway has persisted for a long time. (cf. Section **3.2.2**) The nitrogenases have been thoroughly studied also by computational methods,[1,8,16,135,157,167,184–187] but such studies have produced divergent and disparate suggestions. Notably, there is no consensus on the structure of the key E_4 intermediate, partly because different DFT methods predict widely different relative stabilities of various protonation states, with discrepancies up to 600 kJ/mol.[10] The numerous potential structures and electronic states of the intermediates further complicate the matter.[29,31]

For example, **Hoffman** and colleagues suggested a structure with two hydride ions bridging Fe2–Fe6 and Fe3–Fe7, and protons on sulfides S2B and S5A, all positioned on the same face of the FeMo cluster.[92,175] **Dance** proposed an E_4 state with terminal hydride ions on Fe2 and Fe6, and protons on S2B and S3B. Here, N_2 binds side-on to Fe6 without H_2 dissociation, and is protonated to H_2NNH_2, leading to the cleavage of the N–N bond and the formation of NH_2 fragments on Fe2 and Fe6.[117,159,188] **Nørskov** and co-workers suggested that the E_0 state is doubly protonated, and a sulfide ligand dissociates from the cluster, creating a binding site for N_2.[185]

This mechanism of sulfide dissociation was inspired by crystallographic studies indicating that the S2B group in Mo and V-nitrogenases can be replaced by ligands such as CO, OH^-, and Se.[97,154,155,171,189,190] Recently, our group used QM/MM to investigate a similar mechanism involving the dissociation of S2B using a larger and more realistic model system than the one used by Nørskov and coworkers.[28] The results indicated that the conversion of N_2H_2 to two NH_3 molecules is thermodynamically favorable and primarily follows an alternating pathway, although the first intermediate involves a bridging NNH_2 group, typically associated with a distal mechanism..

However, this study did not prove that the nitrogenase mechanism involves a dissociated S2B group. Therefore, we decided to study the later part of the reaction mechanism of nitrogenase assuming that the S2B ligand does not dissociate. To avoid issues arising from the lack of consensus on the structure and protonation of the E_4 state, we started our study from a state where N_2 is already bound to the cluster and protonated to N_2H_2, following H_2 dissociation.

For each intermediate (E_4–E_8), we evaluated all structures potentially involved in the alternating or distal mechanisms, considering various protonation states of His195, which may form hydrogen bonds with the substrate and intermediates. Based on the best structure for each E_n state, we suggested a mechanism that is primarily alternating and the substrate binds to Fe6, as shown in Figure 4.4.

In the E_5 state, the substrate is protonated to H_2NNH_2 (hydrazine), abstracting a proton from homocitrate. In E_6, a proton is added back to homocitrate, maintaining H_2NNH_2 as the ligand. In E_7, the substrate is protonated to H_2NNH_3, the N–N bond is cleaved, and the first NH_3 dissociates. The resulting NH_2 group remains bound to Fe6 and is protonated to NH_3, again using a proton from homocitrate. In the E_8 state, NH_3 dissociates, forming the resting E_0 state. Our findings suggest that the enzyme follows an alternating mechanism, with HNNH and H_2NNH_2 as intermediates, and the N–N bond is cleaved in the E_7 state with NH_3 products dissociating at the E_7 and E_8 levels.

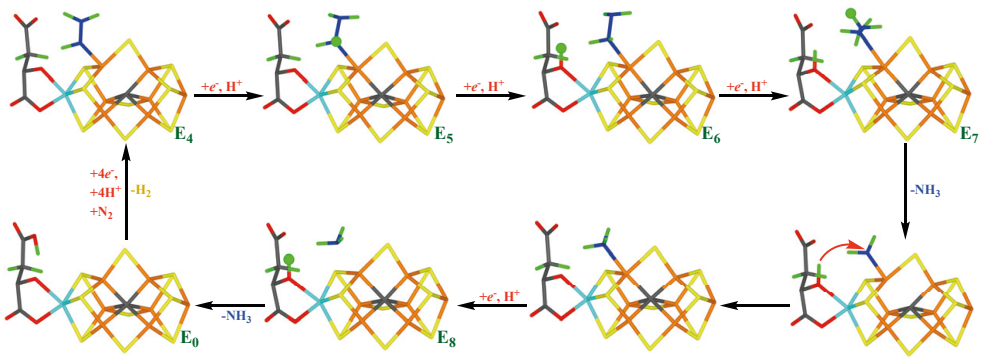


Figure 4.4: The suggested reaction mechanism for nitrogenase, assuming that S2B remains bound to the cluster.

To explore proton delivery, we considered Dance's hypothesis involving proton transfer from the solvent to the FeMo cluster via a chain of water molecules and homocitrate.[148,191,192] His-195 may also provide protons,[135,179] though our calculations suggest its contribution is limited. Interestingly, our results indicated that homocitrate might act as a proton buffer, stabilizing intermediates like H_2NNH_2 and NH_3, which aligns with its essential role in nitrogenase function.[4,184,193]

Our results also show that bridging intermediates are less stable due to spatial constraints. Instead, substrate binding to Fe6 is favored, allowing a significant role of homocitrate.

Paper IV: Proton Transfer in the E_4–E_8 States

Crystallographic studies have demonstrated that the sulfide ligand S2B, which bridges the Fe2 and Fe6 ions in the FeMo cluster, can be reversibly replaced by inhibitors such as CO, OH^-, and Se.[97,154,155,171,189,190] A potential storage site for the dissociated SH^- ion has been identified, suggesting that S2B may reversibly dissociate also during the normal reaction mechanism, thereby providing a binding site for the substrate.[28,185] A recent crystal structure proposed the replacement of S2B (and the S3A and S5A sulfide ions) by N_2,[171] although this interpretation has been questioned.[194,195]

Dance explored possible proton-delivery pathways from the solvent to the FeMo cluster, identifying a conserved chain of water molecules terminating near the S3B ion of the FeMo cluster.[148,151,191,192] He proposed that protons are delivered to the FeMo cluster at the S3B atom and are subsequently transferred to the substrate via various Fe and sulfide ions. He identified six local minima for proton binding on S3B and demonstrated proton movement between these sites with barriers of 10–60 kJ/mol.[192,196] However, his studies primarily addressed the initial four steps of the reaction mechanism (E_0 to E_4) and utilized a somewhat outdated model of the FeMo cluster.

Our group has studied putative reaction mechanisms of nitrogenase, beginning from bound and protonated N_2H_2 and progressing to the formation of two NH_3 molecules, either with S2B still bound (**Paper III**)[39] or dissociated[28] from the cluster. Both scenarios produced reasonable pathways, predominantly following alternating mechanisms where the two N atoms are protonated alternately, and products do not dissociate until reaching the E_7 intermediate. Therefore, these studies could not definitively favor one scenario over the other. Furthermore, protons were hypothetically added to all possible sites on the substrates between each E_n state, assuming free proton movement within the FeMo cluster. These studies primarily focused on determining the thermodynamically most stable protonation states and substrate binding conformations at each E_n level.

In this study, we studied proton transfers within the FeMo cluster, assuming that the proton enters at S3B, S4B, or S5A and is then transported to the substrate via the sulfide and Fe ions. Our results indicate that the net barriers for proton transfers are generally higher when S2B has dissociated from the cluster compared to when it remains bound. In the dissociated case, the maximum barriers are prohibitively high (107–213 kJ/mol) for the E_5–E_7 levels, whereas with S2B bound, the barriers are lower (69–83 kJ/mol). These results strongly argue against the dissociation of S2B. Figure 4.5 compares the energies with and without S2B.

For all E_n levels, protonation of S5A is consistently more favorable (29–98 kJ/mol) than protonation of S4B and S3B. States with the proton on Fe7 and Fe2 are also less stable (16–74 kJ/mol). Even if the proton initially binds to S3B, it would rapidly

be transferred to S5A, which is thermodynamically more stable. However, the stable protonation of S5A poses a problem as it becomes a thermodynamic sink, increasing the effective barriers for the proton-transfer reactions. The individual barriers for proton-transfer and proton-rotation reactions range from 6–67 kJ/mol, typically highest at the initial step (S5A to Fe7) or the final step (to the substrate). These barriers, while lower than the experimental reaction rate, suggest that S5A may always be protonated during the nitrogenase reaction mechanism, as this would reduce the proton-transfer barriers by ~20 kJ/mol.

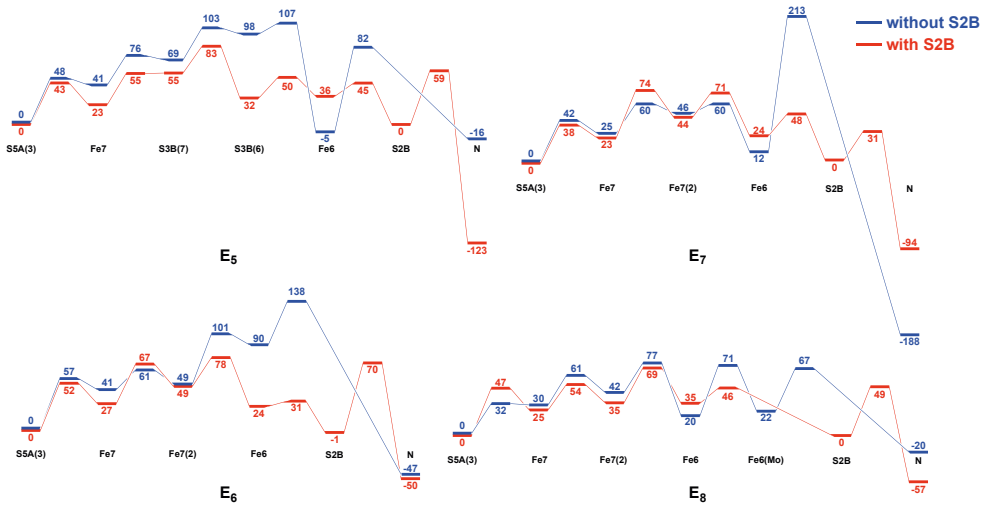


Figure 4.5: Relative energies for the proton-transfer reactions at the various E_n states with and without S2B bound.

Dance also studied proton-transfer reactions within the FeMo cluster but did not consider transfers to or from S5A, resulting in an underestimation of the barriers.[148,151,191,192] In our mechanisms, protonation of S3B is 29–69 kJ/mol less stable than protonation of S5A, and backward barriers to S5A are always lower than forward barriers towards the substrate. With S2B bound, the energy variation of intermediates and transition states is minimal (4–20 kJ/mol) between the E_5 to E_8 states, except for the final step. Two pathways are observed: one via S3B and another via Fe7, with most pathways involving the transfer of the proton from S2B to the substrate. When S2B is dissociated, the energy variation is larger, and barriers are higher, requiring the proton to transfer directly from Fe6 to the substrate, often leading to prohibitively large barriers. It should be noted that Siegbahn has argued that the proton transfer within the cluster can be significantly accelerated by using the surrounding water molecules.[197,198] However, this is based on QM-cluster calculations where the surrounding protein is ignored.

Further analysis of how the surrounding protein affects proton transfers revealed that steric and electrostatic effects from the protein and solvent outside the QM system are significant. The MM energy correction is small (–11 to 15 kJ/mol) but slightly biased to positive values, while the point-charge model has a larger influence (–18 to 84 kJ/mol). Protonation of S2B is more favored by the surrounding electrostatics than other protonation states.

Paper V: E_2 States with S2B Half-Dissociated

Recent studies have suggested that the protonated S2B ligand in the E_2 state of the FeMo cluster may dissociate from one of its two Fe ions (Fe2 or Fe6).[167,169,184,199] **Thorhallsson & Björnsson** (T&B) conducted a comparative study of 18 different states, considering protonation of the three μ_2 bridging sulfide ions, Fe ions, or the central carbide using QM/MM calculations.[163] They found that the most favorable structures had either two protons on S2B and S5A or a bridging hydride between Fe2 and Fe6 and a proton on S2B, which was dissociated from Fe2.

These findings contrast significantly with those of **Cao & Ryde**, who systematically studied approximately 40 different protonation states of E_2, all with a proton on S2B.[31] They found that the most stable states had a bridging hydride ion between Fe2 and Fe6, with the protonated S2B ligand still bound to both Fe2 and Fe6. States with the hydride ion on either side of S2B differed by only 2 kJ/mol, while a state with a terminal hydride ion on Fe5 was only 3 kJ/mol less stable. States with the second proton on S5A (pointing either towards S2B or S3A) were 30 and 37 kJ/mol less stable. No states with the protonated S2B dissociated from either Fe2 or Fe6 were observed, but such states were not systematically explored.

Given the repeated suggestion that half-dissociated S2B states may play a role in the nitrogenase reaction mechanism,[167,169,184,199] it is crucial to resolve these discrepancies. Therefore, we conducted a systematic study of 26 different E_2 structures at various levels of theory to determine whether the inconsistencies are due to the QM model, DFT method, the BS state, the basis set or relativistic effects.

We performed QM/MM calculations with four different functionals and found that the BS state, the size of the QM model, and the relaxation of the surrounding environment influence relative stabilities by up to 12, 20, and 37 kJ/mol, respectively. Additionally, considering all conformations of added protons can change the relative energies by up to 33 kJ/mol, although it does not alter the ranking of different structures. The primary difference between the studies is attributable to the use of different DFT methods.

- TPSS (a pure GGA method) favors structures with both the hydride and S2B bridging Fe2 and Fe6 (B53), which are 15–18 kJ/mol more stable than structures with a half-dissociated S2B (H6S), two terminal hydride ions (D26), or the best structure with no hydride ions (N2353, cf. Figure 4.6).
- B3LYP (a hybrid GGA functional with 20% Hartree–Fock exchange) strongly favors the C2 structure with a doubly protonated carbide ion, making it 101 kJ/mol more stable than N2335, and disfavors all structures with Fe-bound hydride ions.

- TPSSh (a hybrid GGA functional with 10% Hartree–Fock exchange) shows similar tendencies as B3LYP but to a smaller extent. It favors half-dissociated structures, making H6S and H6M 7 kJ/mol more stable than B53.
- r^2SCAN (a modern pure GGA functional) selectively favors half-dissociated structures and C2 (but to a lesser extent than hybrid functionals), making H6S and H6M 16 kJ/mol more stable than B53.

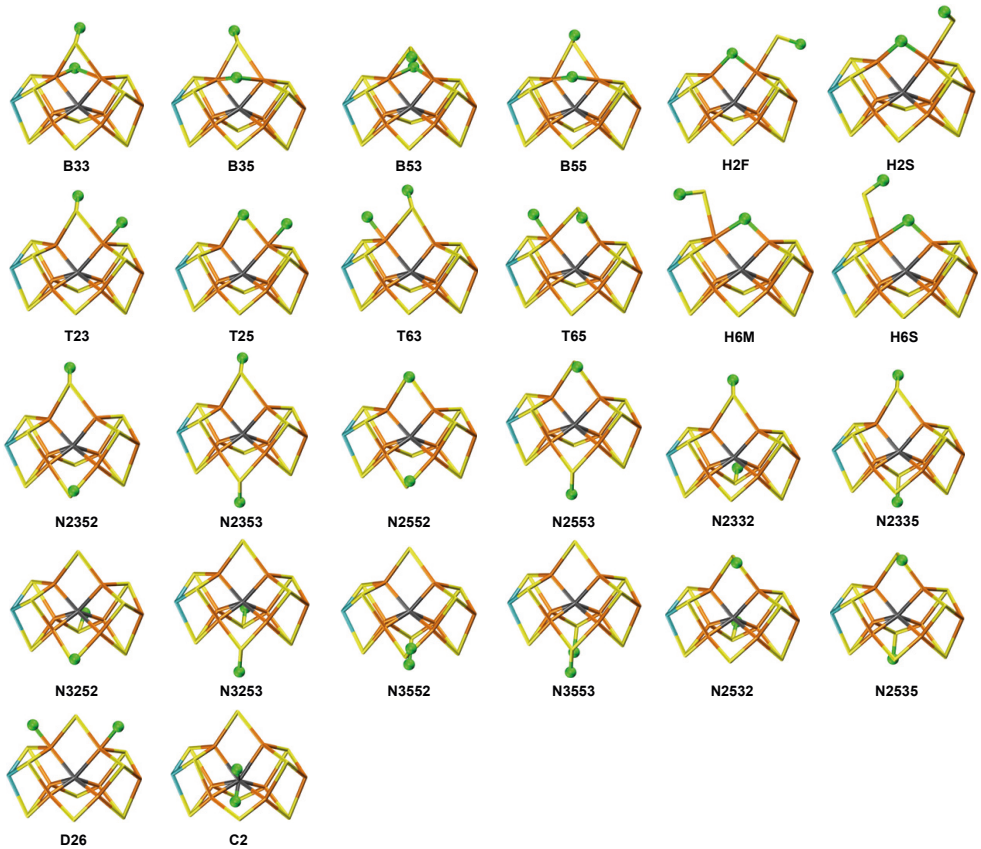


Figure 4.6: The 26 structures of the E_2 state investigated in **Paper V**.

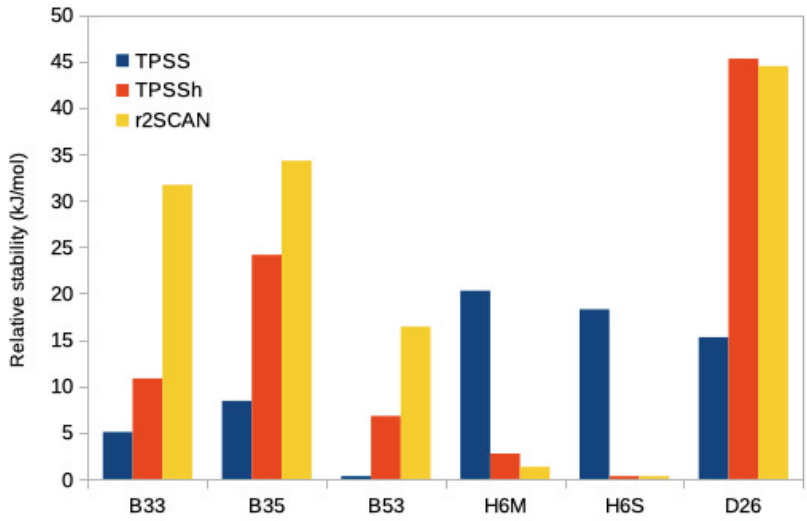


Figure 4.7: Relative stabilities of the best structures containing hydride ions for the TPSS, TPSSh, and r^2SCAN methods.

Our results indicate that several E_2 structures (B53, B33, B35, H6M, H6S, D26, N2353, and C2) are energetically competitive (within 20 kJ/mol). However, two of them (C2 and N2353) do not contain iron-bound hydride ions and are therefore not compatible with EPR data.[109,110] Which structure is most stable depends on the details of the calculations, particularly the DFT method used. Recent studies suggest that r^2SCAN, TPSSh, and B3LYP* (with 15% Hartree–Fock exchange) yield the best structures for Fe2 and FeMo models,[26] while pure GGA functionals, like PBE and PW91, perform better for structures and energies involving H_2 and N_2 binding to small transition-metal models related to nitrogenase.[200] Further research with dispersion-corrected DFT functionals indicates that pure GGA functionals provide better structures, whereas hybrid functionals give more reliable energetic results.[201] Therefore, the relative stabilities of structures with S2B bound to one or two Fe ions are highly sensitive to calculation details, and which structure is most stable remains uncertain and requires more extensive testing of various DFT methods.

Paper VI: Second Part of the Reaction Mechanism with S2B Half-Dissociated

The second part of the nitrogenase reaction mechanism, following N_2 binding, has been widely debated, particularly whether it follows a distal or alternating mechanism.[1,4] In **Paper III**, we thoroughly explored the complete reaction mechanism of nitrogenase, proposing an alternating pathway. We highlighted the role of the homocitrate ligand as a potential proton buffer, crucial for stabilizing intermediates such as H_2NNH_2 and NH_3 at the E_5 and E_7 states, respectively.[39] Furthermore, drawing insights from crystal structures of inhibited nitrogenase,[94,97] we investigated how the dissociation of S2B from the FeMo cluster influences the reaction mechanism.[28,40] Our findings suggested that upon S2B dissociation, N_2H_2 binds as NNH_2, bridging Fe2 and Fe6, with a H_2NNH_2 intermediate at the E_6 state and NH_3 formation occurring in the E_7 state.[28] While both mechanisms appeared equally plausible, our analysis in **Paper IV** of proton-transfer reactions within the cluster suggested that maintaining S2B bound facilitates proton transfers to the substrate.[40]

Recent studies have suggested that S2B may dissociate from only one of the two Fe ions, forming unhooked or half-dissociated structures,[163,167,169,184] which seem to be likely candidates for the E_2–E_4 states of Mo-nitrogenase.[35,36,163] Given this, we aimed to investigate whether such structures are competitive also in the later stages of the nitrogenase reaction, following N_2 binding.

To make half-dissociation possible, we introduced an additional proton to S2B compared to our earlier studies. Furthermore, we employed two functionals, r^{2}SCAN and TPSSh, known for favoring half-dissociation of S2B (**Paper V**) while providing accurate results for the FeMo cluster of nitrogenase.[26] For the E_4 and E_5 states, structures with half-dissociated S2B were found to be significantly less stable (by 16–24 and 9–15 kJ/mol) than the best structures with a bridging S2B. The difference increased dramatically for the E_6 state, for which structures with a half-dissociated S2B were disfavored by 47–52 kJ/mol. On the other hand, our analysis also indicated that in the E_6 state with NH_3 dissociated, the most stable structure showed S2B bound only to Fe6, as the NH ligand occupied the Fe2–Fe6 bridging position. However, our results indicate that such structures are not involved in the reaction mechanism. Specifically, only structures with a bridging S2B ligand were found when the N–N bond remained intact, while after bond cleavage, NH_2 preferred to bridge between Fe2 and Fe6, necessitating S2B dissociation from Fe2 in the most stable state. In the E_8 state, NH_3 may bind either to Fe2 or Fe6, but S2B prefers a bridging position by at least 16 kJ/mol.

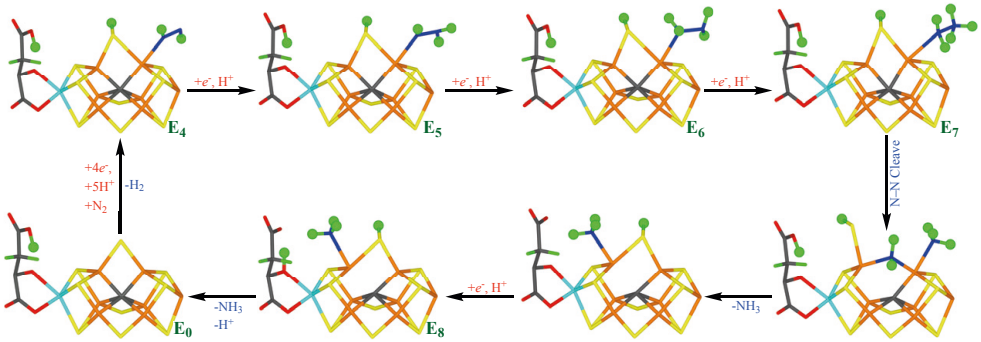


Figure 4.8: Suggested second part of the reaction mechanism for nitrogenase, assuming that the S2B ligand is protonated.

The two DFT functionals, r^{2}SCAN and TPSSh, gave quite consistent structures and relative energies. Overall, our proposed reaction mechanism (shown in Figure 4.8) follows an alternating pathway, consistent with our previously suggested mechanisms involving S2B either bridging or dissociated from the cluster.[39,202] However, with r^{2}SCAN and TPSSh, the substrate preferentially binds to Fe2, which contrasts with the previously suggested mechanism (obtained with TPSS and without the extra proton on S2B), in which the substrate bound preferentially to Fe6 in the presence of a bridging S2B.[39]

Paper VII: E_0–E_8 Redox Potentials

The Lowe–Thorneley cycle emphasizes the significant roles of electron and proton transfer in the nitrogenase reaction cycle, driven by redox potential. Measuring the redox potentials of the FeMo cluster in nitrogenase is challenging because the reaction cannot be arrested at specific E_n states.[3] The only confirmed redox potential is between the resting E_0 state and a one-electron oxidized state (denoted E_{-1}), which lies outside the Lowe–Thorneley cycle at –0.042 V.[3,90,203,204] For the reduction of the resting state, redox potentials of –0.45 V to –0.49 V have been reported, though they may represent a mixture of reduced states.[3,204–206] Computational methods based on the Poisson–Boltzmann equation or similar approaches yield mean errors of 0.03–0.11 V for relative redox potentials.[22,147,207–211] QM calculations, required for absolute potentials and sites of different types, show larger errors, typically ranging from 0.2 to 0.6 V.[209–212] A prediction of the potential of the FeMo cluster had an error of 1.3 V, leading to incorrect identification of the central carbide ion.[99] Even all-atom QM molecular dynamics and free-energy calculations did not achieve better accuracy than 0.26 V.[213]

Recently, our group calibrated various QM/MM methods to estimate the redox potentials of 13 iron–sulfur clusters containing 1–4 Fe ions.[24] The most accurate results were obtained using QM-cluster calculations in a continuum solvent with a high dielectric constant, employing a large QM model (approximately 300 atoms) based on QM/MM structures. This approach gave a mean absolute error of 0.17 V, after correcting for a systematic error of 0.62 V, with the maximum error among the 13 potentials being 0.44 V. Despite the moderate accuracy, these results are sufficient for making useful predictions, such as identifying the redox couple used by [4Fe–4S] ferredoxins.

In this study, we explored redox potentials of the P-cluster and FeMo cluster within Mo-nitrogenase using QM/MM calculations. For the P-cluster, we analyzed six states and calculated the redox potentials for five transitions. Our results showed acceptable accuracy, with errors within or just slightly outside the range observed in the calibration study. We found that the P^{+1}/P^{N} redox potential was the same regardless if the electron transfer was coupled to proton transfer or not, while the P^{+2}/P^{+1} redox potential indicated that the P-cluster takes up a proton together with the electron. These findings validate our method's reliability for studying large iron–sulfur clusters and redox reactions involving protonation, encouraging its application to the FeMo cluster.

For the FeMo cluster, our calculations examined twelve states to reproduce the experimental redox potentials. The calculations confirmed that the resting E_0 state has the $Mo^{III}Fe_3^{II}Fe_4^{III}$ oxidation state, consistent with previous experimental and theoretical studies.[15,31,99,181] The accuracy of the redox potentials was within the error range observed in our previous P-cluster study. Our results indicated that the

E_0 to E_1 transition involves proton transfer, and the calculated redox potentials of the E_0–E_4 states were within 0.41 V, supporting the hypothesis that these states should have similar potentials to accept electrons from the same donor (the P-cluster).

Table 4-1: Calculated redox potentials for the E_0–E_8 states of the FeMo cluster. The last column (ΔE°_{calc}) reports the difference in the calculated redox potential compared to that of the $E_0 \rightarrow E_1H$ transition. Redox potentials for the most favourable structures of the first four transitions are shown in bold face.

Transition	E°_{calc}	ΔE°_{calc}
$E_0 \rightarrow E_1H$	**−1.28**	**0.00**
$E_1 \rightarrow E_2$	**−1.20**	**0.08**
$E_1 \rightarrow E_2'$	−1.45	−0.17
$E_1 \rightarrow E_2''$	−1.29	0.00
$E_2 \rightarrow E_3$	**−1.47**	**−0.19**
$E_2 \rightarrow E_3'$	−1.81	−0.53
$E_3 \rightarrow E_4$	**−0.87**	**0.41**
$E_3 \rightarrow E_4'$	−1.10	0.18
$E_3 \rightarrow E_4''$	−1.34	−0.06
$E_3 \rightarrow E_4'''$	−1.49	−0.21
With S2B		
$E_4N_2H_2 \rightarrow E_5N_2H_3$	−0.15	1.13
$E_5N_2H_3 \rightarrow E_6N_2H_4$	−0.87	0.41
$E_6NH \rightarrow E_7NH_2$	0.74	2.02
$E_7NH_2 \rightarrow E_8NH_3$	−0.71	0.57
Without S2B		
$E_4N_2H_2' \rightarrow E_5N_2H_3'$	−0.31	0.97
$E_5N_2H_3' \rightarrow E_6N_2H_4'$	−1.07	0.20
$E_6N_2H_4' \rightarrow E_7N_2H_5'$	1.37	2.65
$E_7NH_2' \rightarrow E_8NH_3'$	−1.09	0.19

Additionally, our study investigated the redox potentials of the E_4–E_8 states of the FeMo cluster. The calculated potentials for these states are more positive than those for the E_0/E_1 couple, indicating exothermic electron transfers. The trends in redox potentials are consistent regardless of whether S2B remained bound or dissociated. This suggests that a stronger driving force is not necessary for the reaction mechanism. Our findings imply that the assumption of direct N_2 protonation when bound to the cluster may need reevaluation. Overall, our study demonstrates that calibrated redox potential calculations can provide strong predictive power and identify potential issues in proposed reaction mechanisms despite the inherent limitations in accuracy compared to experimental measurements.

Paper VIII: The E_0 and E_1 States of Fe-Nitrogenase

Crystal structures of Mo-nitrogenase have been known since 1992[20,93] and of V-nitrogenase since 2017,[94] but the first crystal structure of Fe-nitrogenase was published in 2023.[95] The same year, a cryogenic electron microscopy structure of Fe-nitrogenase was also published.[101,102]

It is generally believed that Mo, V and Fe-nitrogenase follow similar reaction mechanisms.[2,214] However, a recent EPR study of the one-electron reduced E_1 state in Fe-nitrogenase (which is EPR active, unlike E_1 in Mo-nitrogenase) suggested that it contains a Fe-bound hydride ion rather than a sulfur-bound proton.[215] In contrast, for Mo-nitrogenase, EXAFS measurements and QM/MM calculations have indicated that the E_1 intermediate most likely involves a protonated μ_2 belt sulfide, probably S2B.[181] Previous QM and QM/MM studies have also identified S2B as the energetically most favorable protonation site in the E_1 state.[31,216] Given these differences, it is of great interest to investigate whether there is a difference in the protonation preferences of Mo and Fe-nitrogenase in the E_1 state. The recent crystal structure of Fe-nitrogenase facilitates such an investigation.

We based our calculations on the crystal structure of Fe-nitrogenase from *Azotobacter vinelandii* (PDB code 8BOQ),[95] which includes the entire $\alpha_2\beta_2\gamma_2$ heterohexamer structure. We carefully determined the protonation states of all residues, including specific assumptions for charged residues and the assignment of protonation for histidine residues based on hydrogen-bond patterns and solvent accessibility. We performed extensive MD simulations to equilibrate the protein structure in a water box with an ionic strength of 0.2 M, restraining the protein's heavy atoms to their crystal structure positions. Using DFT methods, we calculated the electrostatic potentials for the metal sites within the FeFe cluster, P-cluster, and Mg site to obtain charges for the molecular mechanical force field.

Then we performed QM/MM calculations of the E_0 state. We examined all 70 BS states of the cluster, revealing that the relative stabilities of these states are similar to those obtained for the FeMo cluster, though with some notable differences in the order. The most stable BS states were identified as Noodleman's BS7 type. This state also best reproduced the Fe–Fe and Fe–ligand distances of the crystal structure. We also investigated the protonation states of homocitrate and His180, finding that His180 prefers protonation on NE2, and homocitrate is most stable when singly protonated at the alcohol O7 atom, which was also supported by the quantum refinement of the crystal structure.

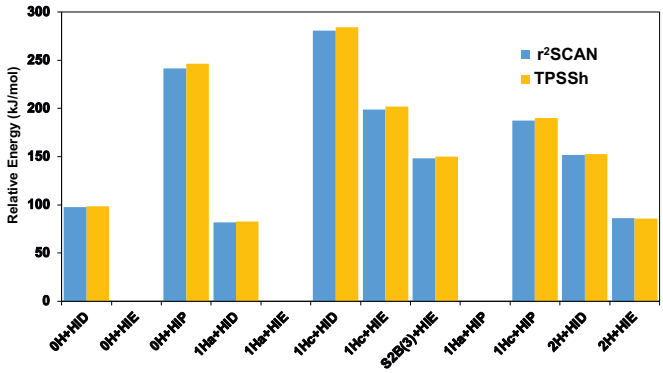


Figure 4.9: Relative energies (kJ/mol) of the various protonation states tested for Fe-nitrogenase in the E_0 state.

For the E_1 state, we also examined the BS states, and the results showed that all 70 BS states are distinct, and many are close in energy. There are significant differences in the preferred BS states for the E_1 state compared to those observed for Mo-nitrogenase. We optimized the structures of 50 different protonation states and found that, across four different DFT functionals (TPSS, r^{2}SCAN, TPSSh, and B3LYP), protonation of the μ_2 belt sulfide ion S2B is more favorable than the formation of a Fe-bound hydride ion. The best hydride-bound structure, with a hydride bound terminally to Fe2, is 14, 26, 32, and 117 kJ/mol less stable than the S2B protonated structure with the TPSS, r^{2}SCAN, TPSSh, and B3LYP functionals, respectively. This stability difference remains even with a larger basis set or relaxed surroundings during geometry optimization. Our results indicate that the E_1 state does not contain a Fe-bound hydride ion, aligning with findings for Mo-nitrogenase but contradicting recent EPR results suggesting a hydride ion in the E_1 state of Fe-nitrogenase.[215]

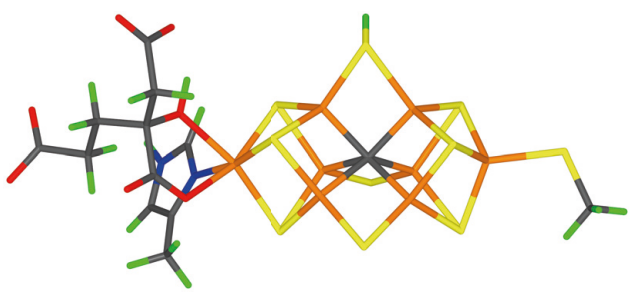


Figure 4.10: Best E_1 structure, protonated on the S2B atom.

5 Conclusions and Outlook

This thesis aims to gain a deeper understanding of the function and mechanism of nitrogenase by using computational methods. We employed QM/MM methods to study molecular structures, electronic states, intermediates, and predict possible reaction mechanisms. The thesis consists of eight papers.

Paper I and **II**: The early stages of the catalytic cycle (E_0–E_4) were examined to understand the formation and dissociation of H_2 and the binding of N_2. The study revealed significant variability in the predictions of different DFT functionals. The B3LYP functional strongly favor protonation of the central carbide ion, preventing H_2 formation, whereas TPSS, r^{2}SCAN, and TPSSh functionals showed exothermic H_2 formation, suggesting that strict kinetic control is required to avoid premature H_2 release. N_2 binding studies indicated that half-dissociation of the S2B ligand enhances N_2 binding, particularly to Fe2 or Fe6, depending on the functional used.

Paper III and **IV**: The later stage of the catalytic cycle (E_4–E_8) was explored, focusing on the role of the S2B ligand and proton-transfer mechanisms. The results supported an alternating mechanism with significant involvement of the homocitrate ligand as a proton buffer. Proton-transfer studies suggested that S2B should remain bound to facilitate lower energy barriers for proton transfers, contradicting the hypothesis of S2B dissociation.

Paper V and **VI**: The possibility that the S2B ligand may dissociate from either Fe2 or Fe6 was thoroughly examined. The studies highlighted that the stability of various E_2 structures depends heavily on the DFT method used. TPSSh and r^{2}SCAN functionals favored structures with S2B half-dissociated from one Fe ion, while B3LYP strongly favors the C2 structure with a doubly protonated carbide ion and TPSS favors structures with both the hydride and S2B bridging Fe2 and Fe6. The investigation was then extended to later stages (E_4–E_8) of the reaction mechanism, finding that half-dissociated structures are less stable than structures in which S2B remains bound to both Fe ions, except when the N–N bond is cleaved.

Paper VII: Redox potential calculations were conducted for the P-cluster and FeMo cluster. The results showed that the E_0 to E_1 transition involves protonation when the electron is taken up. The study provided valuable insights into the redox properties of nitrogenase, suggesting that direct N_2 protonation may need reevaluation.

Paper VIII: The first Fe-nitrogenase structure was analyzed, focusing on the protonation states of homocitrate and His180, as well as the BS states in the E_0 and E_1 states. The results indicate a preference for S2B protonation over Fe-bound hydride formation, aligning with Mo-nitrogenase results but contradicting recent EPR studies suggesting a hydride ion in the E_1 state of Fe-nitrogenase.

This thesis uses QM/MM methods to investigate the reaction mechanism of nitrogenase and examines several significant aspects of the nitrogenase reaction. While the challenges caused by BS states and DFT functionals are often more problematic than commonly supposed, several meaningful results are obtained. While the selection of computational method may influence the results, these computational studies have enhanced our understanding of nitrogenase. Computational methods have proven to be an invaluable tool for studying nitrogenase intermediates that are challenging to observe through conventional experimental techniques.

However, the exact structure of the E_4 intermediate remains controversial.[4,152,160,217] Experimental and computational approaches must continue to be applied to unravel this key intermediate in nitrogenase catalysis. Furthermore, nitrogenase can utilize protons and electrons as reducing equivalents to convert CO, CO_2, and CN^- into hydrocarbons under ambient conditions. These enzymatic Fischer–Tropsch (FT) reactions demonstrate the versatility and importance of the nitrogenase system.[2,152,160,218–220] Understanding the mechanisms of such reactions is not only crucial for understanding the mechanism of nitrogenase, but also has important implications for prebiotic chemistry and biotechnological applications.

In summary, our computational studies have enhanced our understanding of nitrogenase. By revealing the structures of nitrogenase intermediates and employing sophisticated computational methods, we can not only deepen our comprehension of the nitrogenase reaction mechanism but also explore its applications across diverse scientific and practical domains.

References

[1] B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, *Chem. Rev.* **2014**, *114*, 4041–4062.

[2] A. J. Jasniewski, C. C. Lee, M. W. Ribbe, Y. Hu, *Chem. Rev.* **2020**, *120*, 5107–5157.

[3] H. L. Rutledge, F. A. Tezcan, *Chem. Rev.* **2020**, *120*, 5158–5193.

[4] L. C. Seefeldt, Z.-Y. Yang, D. A. Lukoyanov, D. F. Harris, D. R. Dean, S. Raugei, B. M. Hoffman, *Chem. Rev.* **2020**, *120*, 5082–5106.

[5] X. Zhang, B. B. Ward, D. M. Sigman, *Chem. Rev.* **2020**, *120*, 5308–5351.

[6] B. M. Hoffman, D. Lukoyanov, D. R. Dean, L. C. Seefeldt, *Acc. Chem. Res.* **2013**, *46*, 587–595.

[7] B. K. Burgess, D. J. Lowe, *Chem. Rev.* **1996**, *96*, 2983–3012.

[8] R. Bjornsson, F. A. Lima, T. Spatzal, T. Weyhermüller, P. Glatzel, E. Bill, O. Einsle, F. Neese, S. DeBeer, *Chem, Sci,* **2014**, *5*, 3096–3103.

[9] L. Cao, M. C. Börner, J. Bergmann, O. Caldararu, U. Ryde, *Inorg. Chem.* **2019**, *58*, 9672–9690.

[10] L. Cao, U. Ryde, *Phys. Chem. Chem. Phys.* **2019**, *21*, 2480–2488.

[11] O. Einsle, D. C. Rees, *Chem. Rev.* **2020**, *120*, 4969–5004.

[12] B. M. Barney, M. Laryukhin, R. Y. Igarashi, H.-I. Lee, P. C. Dos Santos, T.-C. Yang, B. M. Hoffman, D. R. Dean, L. C. Seefeldt, *Biochemistry (Mosc.)* **2005**, *44*, 8030–8037.

[13] J. Kästner, P. E. Blöchl, *ChemPhysChem* **2005**, *6*, 1724–1726.

[14] C. Van Stappen, L. Decamps, G. E. Cutsail, R. Bjornsson, J. T. Henthorn, J. A. Birrell, S. DeBeer, *Chem. Rev.* **2020**, *120*, 5005–5081.

[15] C. Van Stappen, R. Davydov, Z.-Y. Yang, R. Fan, Y. Guo, E. Bill, L. C. Seefeldt, B. M. Hoffman, S. DeBeer, *Inorg. Chem.* **2019**, *58*, 12365–12376.

[16] I. Dance, *Z. Anorg. Allg. Chem.* **2015**, *641*, 91–99.

[17] T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees, O. Einsle, *Sci* **2011**, *334*, 940–940.

[18] S. Ramaswamy, *Sci* **2011**, *334*, 914–915.

[19] K. M. Lancaster, M. Roemelt, P. Ettenhuber, Y. Hu, M. W. Ribbe, F. Neese, U. Bergmann, S. DeBeer, *Sci* **2011**, *334*, 974–977.

[20] J. Kim, D. C. Rees, *Sci* **1992**, *257*, 1677–1682.

[21] L. C. Davis, M. T. Henzl, R. H. Burris, W. H. Orme-Johnson, *Biochemistry (Mosc.)* **1979**, *18*, 4860–4869.

[22] P. J. Stephens, D. R. Jollie, A. Warshel, *Chem. Rev.* **1996**, *96*, 2491–2514.

[23] J. A. Rees, R. Bjornsson, J. K. Kowalska, F. A. Lima, J. Schlesier, D. Sippel, T. Weyhermüller, O. Einsle, J. A. Kovacs, S. DeBeer, *Dalton Trans.* **2017**, *46*, 2445–2455.

[24] S. Jafari, Y. A. Tavares Santos, J. Bergmann, M. Irani, U. Ryde, *Inorg. Chem.* **2022**, *61*, 5991–6007.

[25] S. Yogendra, D. W. N. Wilson, A. W. Hahn, T. Weyhermüller, C. Van Stappen, P. Holland, S. DeBeer, *Inorg. Chem.* **2023**, *62*, 2663–2671.

[26] B. Benediktsson, R. Bjornsson, *J. Chem. Theory Comput.* **2022**, *18*, 1437–1457.

[27] C. Xiang, W. Jia, W.-H. Fang, Z. Li, *J. Chem. Theory Comput.* **2024**, *20*, 775–786.

[28] L. Cao, U. Ryde, *J. Catal.* **2020**, *391*, 247–259.

[29] L. Cao, U. Ryde, *JBIC J. Biol. Inorg. Chem.* **2020**, *25*, 521–540.

[30] L. Cao, U. Ryde, *J. Chem. Theory Comput.* **2020**, *16*, 1936–1952.

[31] L. Cao, O. Caldararu, U. Ryde, *J. Chem. Theory Comput.* **2018**, *14*, 6653–6678.

[32] X. Lu, D. Fang, S. Ito, Y. Okamoto, V. Ovchinnikov, Q. Cui, *Mol. Simul.* **2016**, *42*, 1056–1078.

[33] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

[34] H. Jiang, U. Ryde, *Dalton Trans.* **2023**, *52*, 9104–9120.

[35] H. Jiang, U. Ryde, *Phys. Chem. Chem. Phys.* **2024**, *26*, 1364–1375.

[36] H. Jiang, O. K. G. Svensson, U. Ryde, *Inorg. Chem.* **2022**, *61*, 18067–18076.

[37] H. Jiang, U. Ryde, *Dalton Trans.* **2024**, *53*, 11500–11513.

[38] H. Jiang, O. K. G. Svensson, U. Ryde, *Molecules* **2022**, *28*, 65.

[39] H. Jiang, U. Ryde, *Chem. – Eur. J.* **2022**, *28*, e202103933.

[40] H. Jiang, O. K. G. Svensson, L. Cao, U. Ryde, *Angew. Chem. Int. Ed.* **2022**, *61*, e202208544.

[41] H. Jiang, K. J. M. Lundgren, U. Ryde, *Inorg. Chem.* **2023**, *62*, 19433–19445.

[42] F. Jensen, *Introduction to Computational Chemistry*, Wiley, Chichester, UK ; Hoboken, NJ, **2017**.

[43] I. N. Levine, *Quantum Chemistry*, Pearson, Boston, **2014**.

[44] B. Nagy, F. Jensen, in *Reviews in Computational Chemistry* (Eds.: A.L. Parrill, K.B. Lipkowitz), Wiley, **2017**, pp. 93–149.

[45] W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.* **1969**, *51*, 2657–2664.

[46] J. S. Binkley, J. A. Pople, W. J. Hehre, *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

[47] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724–728.

[48] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213–222.

[49] T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. V. R. Schleyer, *J. Comput. Chem.* **1983**, *4*, 294–301.

[50] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

[51] D. Rappoport, F. Furche, *J. Chem. Phys.* **2010**, *133*, 134105.

[52] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–B871.

[53] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.

[54] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[55] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **2003**, *91*, 146401.

[56] "The nobel prize in chemistry 1998," can be found under https://www.nobelprize.org/prizes/chemistry/1998/summary/, **1998**.

[57] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2019**, *150*, 154122.

[58] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, J. Sun, *J. Phys. Chem. Lett.* **2020**, *11*, 8208–8215.

[59] S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew, J. G. Brandenburg, *J. Chem. Phys.* **2021**, *154*, 061101.

[60] V. N. Staroverov, G. E. Scuseria, J. Tao, J. P. Perdew, *J. Chem. Phys.* **2003**, *119*, 12129–12137.

[61] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[62] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[63] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[64] K. Dill, S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*, Garland Science, New York, **2010**.

[65] D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M. C. Kaymak, E. King, T. Kurtzman, T.-S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H. M. Nguyen, K. A. O'Hearn, A. V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T. E. I. Cheatham, D. R. Roe, A. Roitberg, C. Simmerling, D. M. York, M. C. Nagan, K. M. Jr. Merz, *J. Chem. Inf. Model.* **2023**, *63*, 6183–6191.

[66] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545–1614.

[67] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[68] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. a. I. Goddard, W. M. Skiff, *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

[69] J. Westermayr, M. Gastegger, K. T. Schütt, R. J. Maurer, *J. Chem. Phys.* **2021**, *154*, 230903.

[70] H. Li, X. Sun, W. Cui, M. Xu, J. Dong, B. E. Ekundayo, D. Ni, Z. Rao, L. Guo, H. Stahlberg, S. Yuan, H. Vogel, *Nat. Biotechnol.* **2024**, *42*, 229–242.

[71] A. Warshel, M. Levitt, *J. Mol. Biol.* **1976**, *103*, 227–249.

[72] "The nobel prize in chemistry 2013," can be found under https://www.nobelprize.org/prizes/chemistry/2013/summary/, **2013**.

[73] P. Amara, M. J. Field, *Theor. Chem. Acc.* **2003**, *109*, 43–52.

[74] D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček, B. R. Brooks, *J. Chem. Phys.* **2002**, *117*, 10534–10547.

[75] N. Ferré, M. Olivucci, *J. Mol. Struct. THEOCHEM* **2003**, *632*, 71–82.

[76] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718–730.

[77] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700–733.

[78] K. Yasuda, D. Yamaki, *J. Chem. Phys.* **2004**, *121*, 3964–3972.

[79] Y. Zhang, *Theor. Chem. Acc.* **2006**, *116*, 43–50.

[80] I. Antes, W. Thiel, *J. Phys. Chem. A* **1999**, *103*, 9290–9295.

[81] G. A. DiLabio, M. M. Hurley, P. A. Christiansen, *J. Chem. Phys.* **2002**, *116*, 9578–9584.

[82] R. B. Murphy, D. M. Philipp, R. A. Friesner, *Chem. Phys. Lett.* **2000**, *321*, 113–120.

[83] V. Kairys, J. H. Jensen, *J. Phys. Chem. A* **2000**, *104*, 6656–6665.

[84] N. Ferré, X. Assfeld, J.-L. Rivail, *J. Comput. Chem.* **2002**, *23*, 610–624.

[85] J. Pu, J. Gao, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 632–650.

[86] V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret, G. G. Ferenczy, *J. Comput. Chem.* **1994**, *15*, 269–282.

[87] L. Cao, U. Ryde, *Front. Chem.* **2018**, *6*, 89.

[88] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J Phys Chem* **1996**, *100*, 19357–19363.

[89] U. Ryde, *J. Comput.-Aided Mol. Des.* **1996**, *10*, 153–164.

[90] B. Schmid, H.-J. Chiu, V. Ramakrishnan, J. B. Howard, D. C. Rees, in *Handbook of Metalloproteins*, John Wiley & Sons, Ltd, **2006**.

[91] R. Y. Igarashi, M. Laryukhin, P. C. Dos Santos, H.-I. Lee, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2005**, *127*, 6231–6241.

[92] V. Hoeke, L. Tociu, D. A. Case, L. C. Seefeldt, S. Raugei, B. M. Hoffman, *J. Am. Chem. Soc.* **2019**, *141*, 11984–11996.

[93] O. Einsle, *JBIC J. Biol. Inorg. Chem.* **2014**, *19*, 737–745.

[94] D. Sippel, O. Einsle, *Nat. Chem. Biol.* **2017**, *13*, 956–960.

[95] C. Trncik, F. Detemple, O. Einsle, *Nat. Catal.* **2023**, *6*, 415–424.

[96] O. Einsle, F. A. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard, D. C. Rees, *Sci* **2002**, *297*, 1696–1700.

[97] T. Spatzal, K. A. Perez, O. Einsle, J. B. Howard, D. C. Rees, *Sci* **2014**, *345*, 1620–1623.

[98] I. Dance, *Chem. Commun.* **2003**, 324–325.

[99] T. Lovell, T. Liu, D. A. Case, L. Noodleman, *J. Am. Chem. Soc.* **2003**, *125*, 8377–8383.

[100] J. A. Wiig, Y. Hu, C. C. Lee, M. W. Ribbe, *Sci* **2012**, *337*, 1672–1675.

[101] F. V. Schmidt, L. Schulz, J. Zarzycki, S. Prinz, N. N. Oehlmann, T. J. Erb, J. G. Rebelein, *Nat. Struct. Mol. Biol.* **2024**, *31*, 150–158.

[102] S. Greed, *Nat. Rev. Chem.* **2023**, *7*, 379–379.

[103] D. J. Lowe, R. N. F. Thorneley, *Biochem. J.* **1984**, *224*, 895–901.

[104] R. N. F. Thorneley, D. J. Lowe, *Biochem. J.* **1984**, *224*, 887–894.

[105] L. Cao, O. Caldararu, U. Ryde, *J. Phys. Chem. B* **2017**, *121*, 8242–8262.

[106] B. Benediktsson, R. Bjornsson, *Inorg. Chem.* **2017**, *56*, 13417–13429.

[107] S. J. Yoo, H. C. Angove, V. Papaefthymiou, B. K. Burgess, E. Münck, *J. Am. Chem. Soc.* **2000**, *122*, 4926–4936.

[108] C. Greco, P. Fantucci, U. Ryde, L. de Gioia, *Int. J. Quantum Chem.* **2011**, *111*, 3949–3960.

[109] D. Lukoyanov, Z.-Y. Yang, S. Duval, K. Danyal, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2014**, *53*, 3688–3693.

[110] D. A. Lukoyanov, N. Khadka, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2018**, *57*, 6847–6852.

[111] D. Lukoyanov, B. M. Barney, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Proc. Natl. Acad. Sci.* **2007**, *104*, 1451–1455.

[112] D. Lukoyanov, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2010**, *132*, 2526–2527.

[113] Z.-Y. Yang, N. Khadka, D. Lukoyanov, B. M. Hoffman, D. R. Dean, L. C. Seefeldt, *Proc. Natl. Acad. Sci.* **2013**, *110*, 16327–16332.

[114] D. Lukoyanov, Z.-Y. Yang, N. Khadka, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2015**, *137*, 3610–3615.

[115] D. Lukoyanov, N. Khadka, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2016**, *138*, 10674–10683.

[116] D. Lukoyanov, N. Khadka, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2016**, *138*, 1320–1327.

[117] I. Dance, *Dalton Trans.* **2008**, 5977–5991.

[118] L. C. Seefeldt, B. M. Hoffman, D. R. Dean, *Annu. Rev. Biochem.* **2009**, *78*, 701–722.

[119] B. M. Hoffman, D. R. Dean, L. C. Seefeldt, *Acc. Chem. Res.* **2009**, *42*, 609–619.

[120] R. R. Schrock, *Acc. Chem. Res.* **2005**, *38*, 955–962.

[121] F. Neese, *Angew. Chem. Int. Ed.* **2006**, *45*, 196–199.

[122] B. M. Barney, D. Lukoyanov, T.-C. Yang, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Proc. Natl. Acad. Sci.* **2006**, *103*, 17113–17118.

[123] B. M. Barney, J. McClead, D. Lukoyanov, M. Laryukhin, T.-C. Yang, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Biochemistry (Mosc.)* **2007**, *46*, 6784–6794.

[124] B. M. Barney, T.-C. Yang, R. Y. Igarashi, P. C. Dos Santos, M. Laryukhin, H.-I. Lee, B. M. Hoffman, D. R. Dean, L. C. Seefeldt, *J. Am. Chem. Soc.* **2005**, *127*, 14960–14961.

[125] R. Zimmermann, E. Münck, W. J. Brill, V. K. Shah, M. T. Henzl, J. Rawlings, W. H. Orme-Johnson, *Biochim. Biophys. Acta* **1978**, *537*, 185–207.

[126] E. Münck, H. Rhodes, W. H. Orme-Johnson, L. C. Davis, W. J. Brill, V. K. Shah, *Biochim. Biophys. Acta (BBA) - Protein Struct.* **1975**, *400*, 32–53.

[127] S. P. Cramer, W. O. Gillum, K. O. Hodgson, L. E. Mortenson, E. I. Stiefel, J. R. Chisnell, W. J. Brill, V. K. Shah, *J. Am. Chem. Soc.* **1978**, *100*, 3814–3819.

[128] S. P. Cramer, K. O. Hodgson, W. O. Gillum, L. E. Mortenson, *J. Am. Chem. Soc.* **1978**, *100*, 3398–3407.

[129] B. M. Hoffman, J. E. Roberts, W. H. Orme-Johnson, *J. Am. Chem. Soc.* **1982**, *104*, 860–862.

[130] R. A. Venters, M. J. Nelson, P. A. McLean, A. E. True, M. A. Levy, B. M. Hoffman, W. H. Orme-Johnson, *J. Am. Chem. Soc.* **1986**, *108*, 3487–3498.

[131] A. E. True, P. McLean, M. J. Nelson, W. H. Orme-Johnson, B. M. Hoffman, *J. Am. Chem. Soc.* **1990**, *112*, 651–657.

[132] H.-I. Lee, B. J. Hales, B. M. Hoffman, *J. Am. Chem. Soc.* **1997**, *119*, 11395–11400.

[133] T. V. Harris, R. K. Szilagyi, *Inorg. Chem.* **2011**, *50*, 4811–4824.

[134] T. Spatzal, J. Schlesier, E.-M. Burger, D. Sippel, L. Zhang, S. L. A. Andrade, D. C. Rees, O. Einsle, *Nat. Commun.* **2016**, *7*, 10902.

[135] P. E. M. Siegbahn, *J. Am. Chem. Soc.* **2016**, *138*, 10485–10495.

[136] R. Bjornsson, F. Neese, S. DeBeer, *Inorg. Chem.* **2017**, *56*, 1470–1477.

[137] T. Lovell, J. Li, T. Liu, D. A. Case, L. Noodleman, *J. Am. Chem. Soc.* **2001**, *123*, 12392–12410.

[138] T. Lovell, J. Li, D. A. Case, L. Noodleman, *JBIC J. Biol. Inorg. Chem.* **2002**, *7*, 735–749.

[139] G. M. Sandala, L. Noodleman, in *Nitrogen Fixation* (Ed.: M.W. Ribbe), Humana Press, Totowa, NJ, **2011**, pp. 293–312.

[140] D. Lukoyanov, V. Pelmenschikov, N. Maeser, M. Laryukhin, T. C. Yang, L. Noodleman, D. R. Dean, D. A. Case, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2007**, *46*, 11437–11449.

[141] R. K. Szilagyi, M. A. Winslow, *J. Comput. Chem.* **2006**, *27*, 1385–1397.

[142] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *WIREs Comput. Mol. Sci.* **2014**, *4*, 91–100.

[143] E. Krahn, B. Weiss, M. Kröckel, J. Groppe, G. Henkel, S. Cramer, A. Trautwein, K. Schneider, A. Müller, *JBIC J. Biol. Inorg. Chem.* **2002**, *7*, 37–45.

[144] K. Schneider, U. Gollan, M. Dröttboom, S. Selsemeier-Voigt, A. Müller, *Eur. J. Biochem.* **1997**, *244*, 789–800.

[145] I. Dance, *J. Inorg. Biochem.* **2017**, *169*, 32–43.

[146] M. J. Dilworth, K. Fisher, C.-H. Kim, W. E. Newton, *Biochemistry (Mosc.)* **1998**, *37*, 17495–17505.

[147] C.-H. Kim, W. E. Newton, D. R. Dean, *Biochemistry (Mosc.)* **1995**, *34*, 2798–2808.

[148] I. Dance, *Dalton Trans.* **2012**, *41*, 7647–7659.

[149] L. Yan, V. Pelmenschikov, C. H. Dapper, A. D. Scott, W. E. Newton, S. P. Cramer, *Chem. - Eur. J.* **2012**, *18*, 16349–16357.

[150] L. Yan, C. H. Dapper, S. J. George, H. Wang, D. Mitra, W. Dong, W. E. Newton, S. P. Cramer, *Eur. J. Inorg. Chem.* **2011**, *2011*, 2064–2074.

[151] I. Dance, *Dalton Trans.* **2015**, *44*, 18167–18186.

[152] O. Einsle, *Curr. Opin. Struct. Biol.* **2023**, *83*, 102719.

[153] I. Dance, *Dalton Trans.* **2016**, *45*, 14285–14300.

[154] D. Sippel, M. Rohde, J. Netzer, C. Trncik, J. Gies, K. Grunau, I. Djurdjevic, L. Decamps, S. L. A. Andrade, O. Einsle, *Sci* **2018**, *359*, 1484–1489.

[155] B. Benediktsson, A. Th. Thorhallsson, R. Bjornsson, *Chem. Commun.* **2018**, *54*, 7310–7313.

[156] J. Bergmann, E. Oksanen, U. Ryde, *J. Inorg. Biochem.* **2021**, *219*, 111426.

[157] S. Raugei, L. C. Seefeldt, B. M. Hoffman, *Proc. Natl. Acad. Sci.* **2018**, *115*, E10521–E10530.

[158] H. Yang, J. Rittle, A. R. Marts, J. C. Peters, B. M. Hoffman, *Inorg. Chem.* **2018**, *57*, 12323–12330.

[159] I. Dance, *ChemBioChem* **2020**, *21*, 1671–1709.

[160] O. Einsle, *Molecules* **2023**, *28*, 7959.

[161] J. G. Rebelein, Y. Hu, M. W. Ribbe, *Angew. Chem. Int. Ed.* **2014**, *53*, 11543–11546.

[162] P. E. M. Siegbahn, *J. Comput. Chem.* **2018**, *39*, 743–747.

[163] A. Th. Thorhallsson, R. Bjornsson, *Chem. – Eur. J.* **2021**, *27*, 16788–16800.

[164] K. Fisher, W. E. Newton, D. J. Lowe, *Biochemistry (Mosc.)* **2001**, *40*, 3333–3339.

[165] B. M. Barney, R. Y. Igarashi, P. C. Dos Santos, D. R. Dean, L. C. Seefeldt, *J. Biol. Chem.* **2004**, *279*, 53621–53624.

[166] R. Sarma, B. M. Barney, S. Keable, D. R. Dean, L. C. Seefeldt, J. W. Peters, *J. Inorg. Biochem.* **2010**, *104*, 385–389.

[167] P. P. Hallmen, J. Kästner, *Z. Anorg. Allg. Chem.* **2015**, *641*, 118–122.

[168] M. L. Mckee, *J. Comput. Chem.* **2007**, *28*, 1342–1356.

[169] I. Dance, *Dalton Trans.* **2019**, *48*, 1251–1262.

[170] I. Dance, *Dalton Trans.* **2022**, *51*, 15538–15554.

[171] W. Kang, C. C. Lee, A. J. Jasniewski, M. W. Ribbe, Y. Hu, *Sci* **2020**, *368*, 1381–1385.

[172] K. L. Skubi, P. L. Holland, *Biochemistry (Mosc.)* **2018**, *57*, 3540–3541.

[173] T. M. Buscagan, D. C. Rees, *Joule* **2019**, *3*, 2662–2678.

[174] Y. Pang, R. Bjornsson, *Inorg. Chem.* **2023**, *62*, 5357–5375.

[175] D. A. Lukoyanov, Z.-Y. Yang, D. R. Dean, L. C. Seefeldt, S. Raugei, B. M. Hoffman, *J. Am. Chem. Soc.* **2020**, *142*, 21679–21690.

[176] I. Dance, *J. Am. Chem. Soc.* **2007**, *129*, 1076–1088.

[177] I. Dance, *Dalton Trans.* **2021**, *50*, 18212–18237.

[178] I. Dance, *Dalton Trans.* **2022**, *51*, 12717–12728.

[179] P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.* **2019**, *21*, 15747–15759.

[180] W. Wei, P. E. M. Siegbahn, *Chem. – Eur. J.* **2022**, *28*, e202103745.

[181] C. Van Stappen, A. T. Thorhallsson, L. Decamps, R. Bjornsson, S. DeBeer, *Chem. Sci.* **2019**, *10*, 9807–9821.

[182] M. Rohde, D. Sippel, C. Trncik, S. L. A. Andrade, O. Einsle, *Biochemistry (Mosc.)* **2018**, *57*, 5497–5504.

[183] D. Lukoyanov, N. Khadka, D. R. Dean, S. Raugei, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2017**, *56*, 2233–2240.

[184] A. Th. Thorhallsson, B. Benediktsson, R. Bjornsson, *Chem. Sci.* **2019**, *10*, 11110–11124.

[185] J. B. Varley, Y. Wang, K. Chan, F. Studt, J. K. Nørskov, *Phys. Chem. Chem. Phys.* **2015**, *17*, 29541–29547.

[186] M. L. McKee, *J. Phys. Chem. A* **2016**, *120*, 754–764.

[187] L. Rao, X. Xu, C. Adamo, *ACS Catal.* **2016**, *6*, 1567–1577.

[188] I. Dance, *Dalton Trans.* **2012**, *41*, 4859.

[189] T. Spatzal, K. A. Perez, J. B. Howard, D. C. Rees, *eLife* **2015**, *4*, e11620.

[190] L. Cao, O. Caldararu, U. Ryde, *JBIC J. Biol. Inorg. Chem.* **2020**, *25*, 847–861.

[191] I. Dance, *Inorg. Chem.* **2013**, *52*, 13068–13077.

[192] I. Dance, *J. Am. Chem. Soc.* **2005**, *127*, 10925–10942.

[193] J. Imperial, T. R. Hoover, M. S. Madden, P. W. Ludden, V. K. Shah, *Biochemistry (Mosc.)* **1989**, *28*, 7796–7799.

[194] W. Kang, C. C. Lee, A. J. Jasniewski, M. W. Ribbe, Y. Hu, *Sci* **2021**, *371*, eabe5856.

[195] J. Bergmann, E. Oksanen, U. Ryde, *JBIC J. Biol. Inorg. Chem.* **2021**, *26*, 341–353.

[196] I. Dance, *Biochemistry (Mosc.)* **2006**, *45*, 6328–6340.

[197] P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.* **2023**, *25*, 23602–23613.

[198] P. E. M. Siegbahn, *J. Phys. Chem. B* **2023**, *127*, 2156–2159.

[199] D. A. Lukoyanov, M. D. Krzyaniak, D. R. Dean, M. R. Wasielewski, L. C. Seefeldt, B. M. Hoffman, *J. Phys. Chem. B* **2019**, *123*, 8823–8828.

[200] I. Dance, *Mol. Simul.* **2018**, *44*, 568–581.

[201] M. Torbjörnsson, U. Ryde, *Electron. Struct.* **2021**, *3*, 034005.

[202] V. P. Vysotskiy, M. Torbjörnsson, H. Jiang, E. D. Larsson, L. Cao, U. Ryde, H. Zhai, S. Lee, G. K.-L. Chan, *J. Chem. Phys.* **2023**, *159*, 044106.

[203] M. J. O'Donnell, B. E. Smith, *Biochem. J.* **1978**, *173*, 831–838.

[204] R. Y. Igarashi, L. C. Seefeldt, *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 351–384.

[205] S. Lough, A. Burns, G. D. Watt, *Biochemistry (Mosc.)* **1983**, *22*, 4062–4066.

[206] G. D. Watt, A. Burns, S. Lough, D. L. Tennent, *Biochemistry (Mosc.)* **1980**, *19*, 4926–4932.

[207] B. S. P. Jr, B. T. Miller, V. Schalk, H. L. Woodcock, B. R. Brooks, T. Ichiye, *PLoS Comput. Biol.* **2014**, *10*, e1003739.

[208] B. S. Perrin Jr., S. Niu, T. Ichiye, *J. Comput. Chem.* **2013**, *34*, 576–582.

[209] L. Noodleman, W.-G. Han, *JBIC J. Biol. Inorg. Chem.* **2006**, *11*, 674–694.

[210] J.-M. Mouesca, J. L. Chen, L. Noodleman, D. Bashford, D. A. Case, *J. Am. Chem. Soc.* **1994**, *116*, 11898–11914.

[211] R. A. Torres, T. Lovell, L. Noodleman, D. A. Case, *J. Am. Chem. Soc.* **2003**, *125*, 1923–1936.

[212] L. Noodleman, T. Lovell, T. Liu, F. Himo, R. A. Torres, *Curr. Opin. Chem. Biol.* **2002**, *6*, 259–273.

[213] J. Cheng, X. Liu, J. VandeVondele, M. Sulpizi, M. Sprik, *Acc. Chem. Res.* **2014**, *47*, 3522–3529.

[214] R. R. Eady, *Chem. Rev.* **1996**, *96*, 3013–3030.

[215] D. A. Lukoyanov, D. F. Harris, Z.-Y. Yang, A. Pérez-González, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *Inorg. Chem.* **2022**, *61*, 5459–5464.

[216] I. Dance, *Inorganics* **2019**, *7*, 8.

[217] R. A. Warmack, D. C. Rees, *Molecules* **2023**, *28*, 7952.

[218] Y. Hu, C. C. Lee, M. Grosch, J. B. Solomon, W. Weigand, M. W. Ribbe, *Chem. Rev.* **2023**, *123*, 5755–5797.

[219] T. M. Buscagan, K. A. Perez, A. O. Maggiolo, D. C. Rees, T. Spatzal, *Angew. Chem. Int. Ed.* **2021**, *60*, 5704–5707.

[220] M. Rohde, K. Laun, I. Zebger, S. T. Stripp, O. Einsle, *Sci. Adv.* **2021**, *7*, eabg4474.

Scientific Publications

Author contributions

Paper I: H_2 formation from the E_2–E_4 states of nitrogenase

I performed all the calculations and participated in writing the manuscript.

Paper II: N_2 binding to the E_0–E_4 states of nitrogenase

I performed all the calculations and participated in writing the manuscript.

Paper III: Thermodynamically Favourable States in the Reaction of Nitrogenase without Dissociation of any Sulfide Ligand

I performed all the calculations and participated in writing the manuscript.

Paper IV: Proton Transfer Pathways in Nitrogenase with and without Dissociated S2B

I performed most of the calculations, I supervised a BSc student to do the remaining calculations and participated in writing the manuscript.

Paper V: QM/MM Study of Partial Dissociation of S2B for the E_2 Intermediate of Nitrogenase

I performed most of the calculations, I supervised a BSc student to do the remaining calculations and participated in writing the manuscript.

Paper VI: Putative reaction mechanism of nitrogenase with a half-dissociated S2B ligand

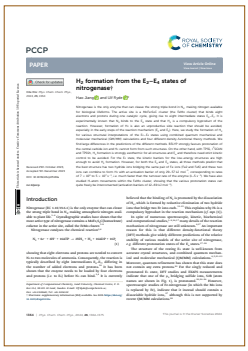
I performed all the calculations and participated in writing the manuscript.

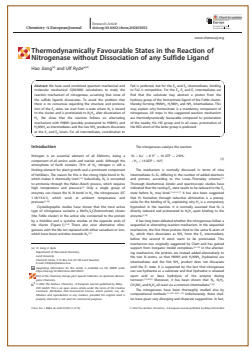
Paper VII: Quantum Mechanical Calculations of Redox Potentials of the Metal Clusters in Nitrogenase

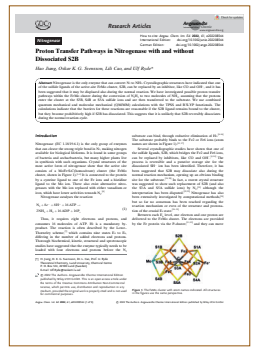
I performed all the geometry optimizations of FeMo cluster and participated in writing the manuscript.

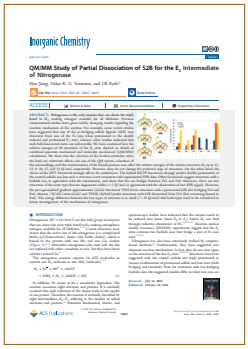
Paper VIII: Protonation of Homocitrate and the E_1 State of Fe-Nitrogenase Studied by QM/MM Calculations

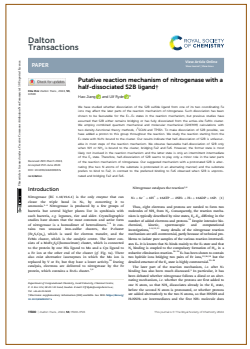
I set up the protein, performed the MD simulations, the QM/MM calculations and participated in writing the manuscript.

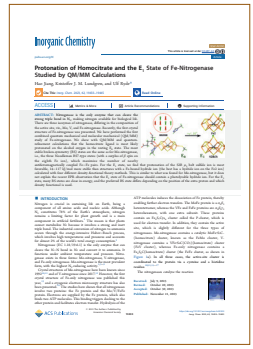












Nitrogenase is the only enzyme that can cleave the N_2 triple bond in nature, enabling nitrogen to be available for biological life. This enzyme features a catalytic FeMo cluster and an electron-transfer P-cluster, as well as a [4Fe:4S] cluster in the Fe protein. The nitrogenase mechanism is not fully understood due to the complicated structure of the enzyme and the challenge of capturing reaction intermediates experimentally. However, computational methods, especially combined quantum-mechanics/molecular-mechanics (QM/MM) method, offer valuable insights. These methods integrate quantum chemistry to handle the active regions and molecular mechanics for the larger protein environment, providing a comprehensive model for studying nitrogenase. This thesis presents systematic QM/MM studies on nitrogenase, focusing on H_2 formation, N_2 binding, S2B ligand half dissociation, redox potentials, and proton transfer processes.

ISBN: 978-91-8096-058-8

Computational Chemistry
Department of Chemistry
Faculty of Science

LUND UNIVERSITY

Printed by Media-Tryck, Lund 2024
NORDIC SWAN ECOLABEL 3041 0903