QM/MM Studies of Nitrogenase

Jiang, Hao

2024

Total number of authors:
1

QM/MM Studies of Nitrogenase

HAO JIANG | COMPUTATIONAL CHEMISTRY | LUND UNIVERSITY
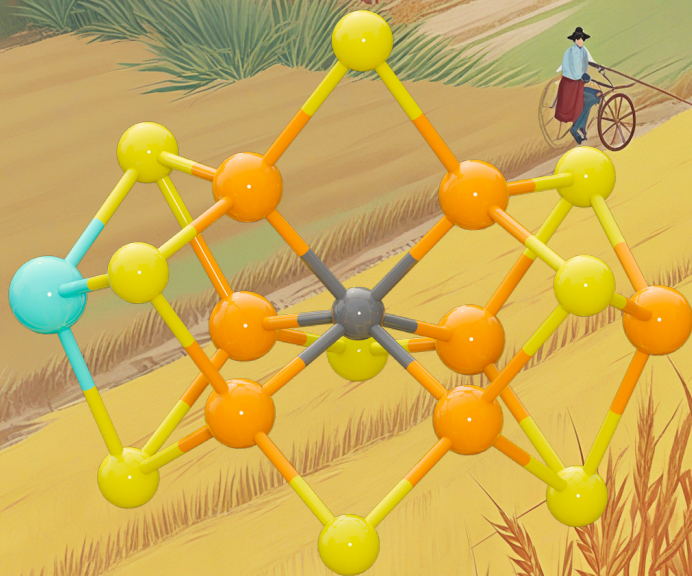
QM/MM Studies of Nitrogenase

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

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.

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature   Date 2024-08-12

QM/MM Studies of Nitrogenase

Hao Jiang

蒋 浩

LUND UNIVERSITY

"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.

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

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:

$$\mathbf{H} = \mathbf{V} + \mathbf{T} = V_{ee} + V_{nn} + V_{en} + T_e + T_n \tag{2.2}$$

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

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

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

$\mathbf{T}_e$ 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:

$$\boldsymbol{H_e = V_{ee} + V_{nn} + V_{en} + T_e} \qquad (2.8)$$

Here, the nuclear kinetic energy operator $\mathbf{T}_n$ 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:

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

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

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

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:

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

where $T_s[\rho]$ is the kinetic energy of the non-interacting electrons, $E_{ne}[\rho]$ is the attractive potential between nuclei and elections, $J[\rho]$ is the Coulomb repulsion between electrons and $E_{xc}[\rho]$ 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 $E_{xc}[\rho]$, 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] r²SCAN,[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:

$$E_{total} = E_{covalent} + E_{noncovalent} \qquad (2.16)$$

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:

$$E_{covalent} = E_{bonds} + E_{angles} + E_{dihedrals} \qquad (2.17)$$

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

$$E_{noncovalent} = E_{VdW} + E_{el} \qquad (2.18)$$

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

$$v(t + \Delta t) \approx v(t) + a(t)\Delta t \qquad (2.23)$$

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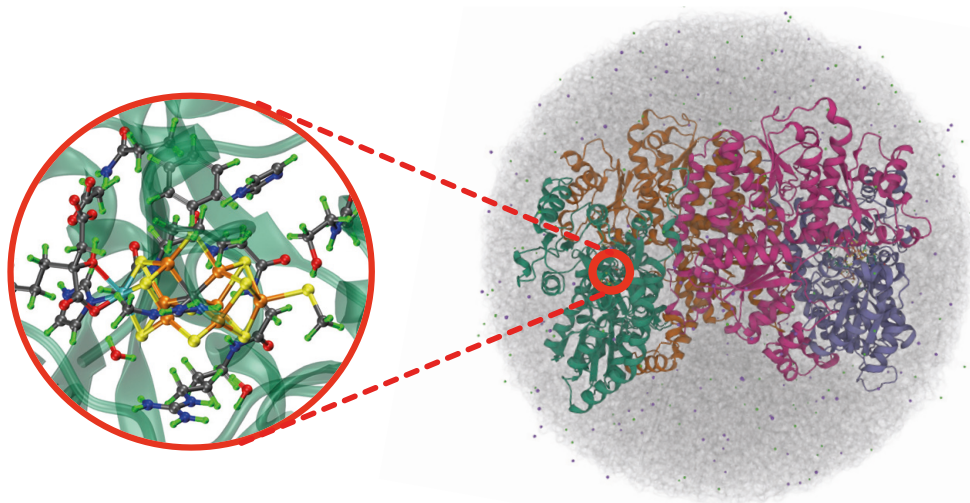
Figure 2.1: The systems in a QM/MM calculation. The QM system is shown in a balls-and-sticks model.

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]

$$E_{\text{QM/MM}} = E_{\text{QM(QM)}} + E_{\text{MM(MM)}} + E_{\text{QM-MM}} \qquad (2.24)$$

- **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.

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

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]
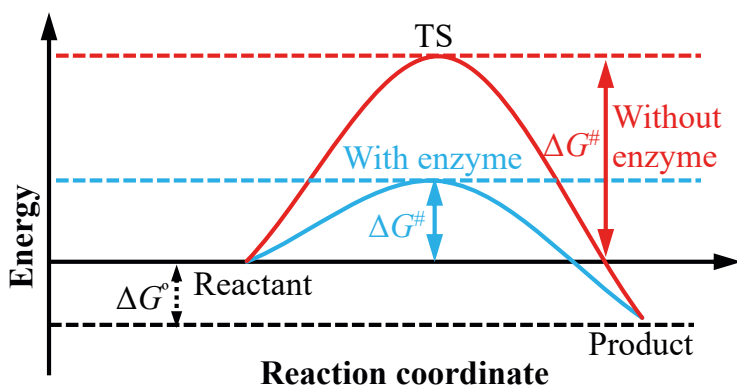
Figure 2.2: Energy profile of a chemical reaction showing the effect of an enzyme on activation energy. The red curve represents the reaction pathway without the enzyme, indicating a higher activation energy. The blue curve shows the reaction pathway with the enzyme, demonstrating a lower activation 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.
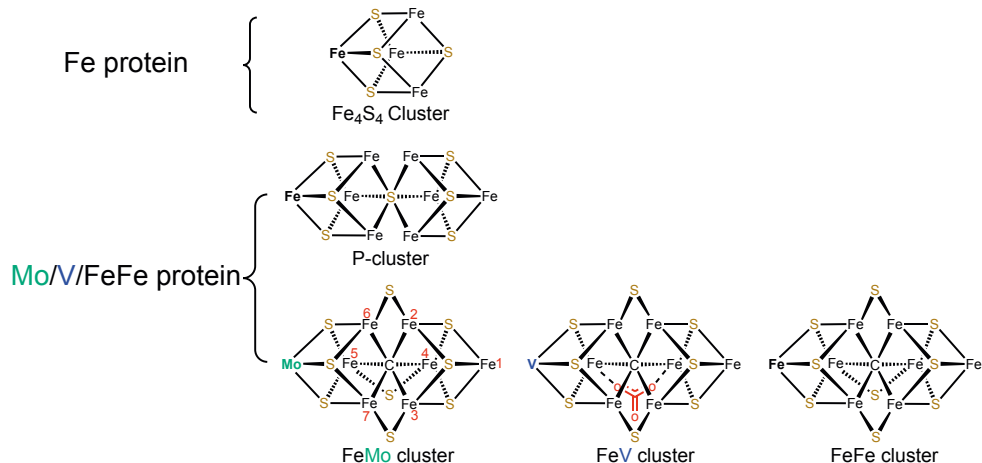
Figure 3.1: Structures of the FeS clusters in nitrogenases.

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]
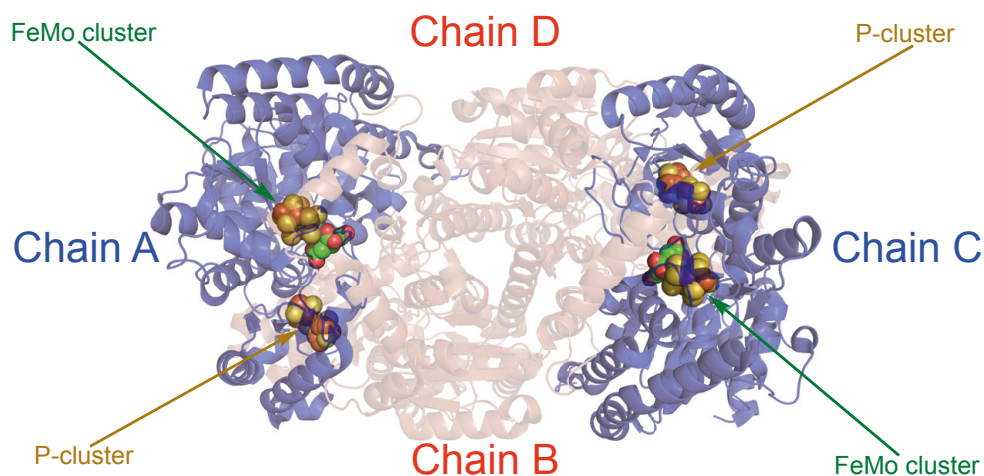
Figure 3.2: Structure of MoFe protein forms an $\alpha_2\beta_2$ heterotetramer in which two αβ units (NifDK) are connected solely via the NifK peptides. Each *αβ* unit holds a FeMo cluster and a P-cluster (PDB 3U7Q).
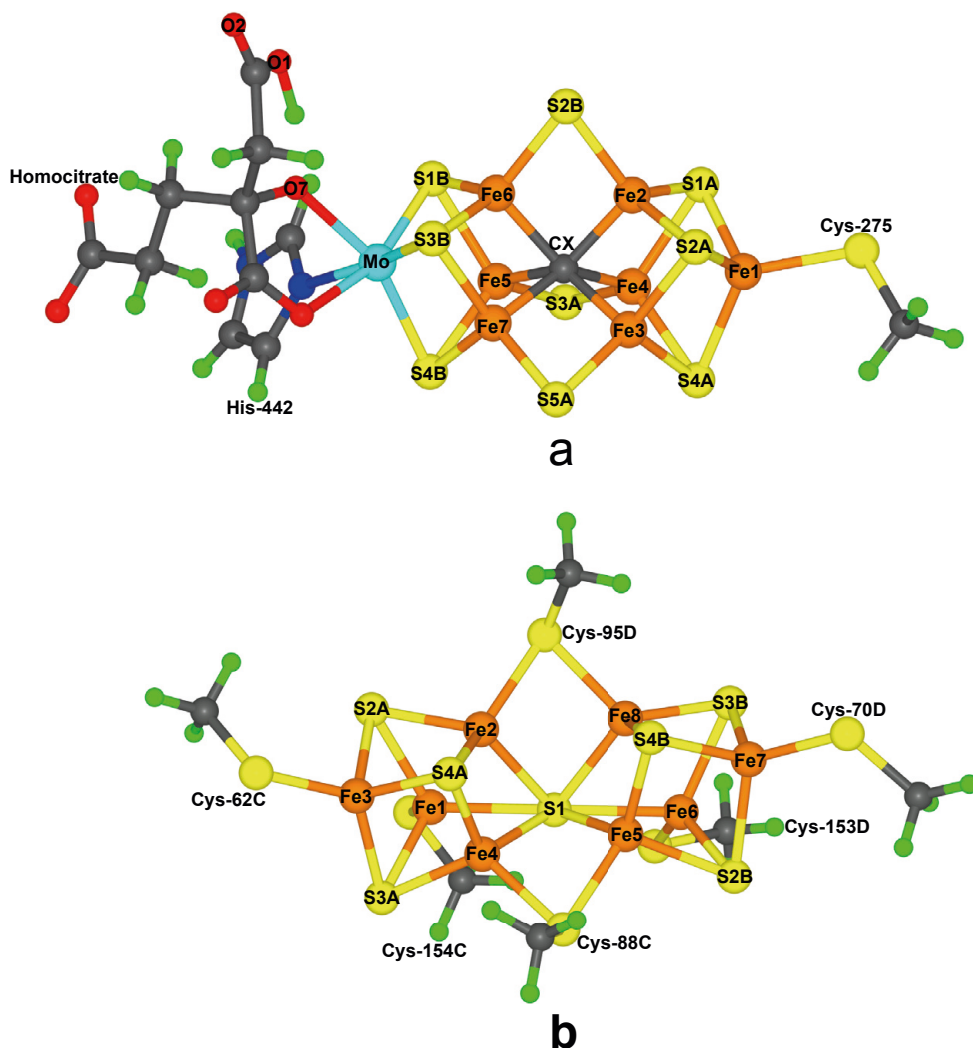
Figure 3.3: The atomic structure of a) FeMo cluster and b) P-cluster with atom and residue names taken from the 3U7Q crystal structure.

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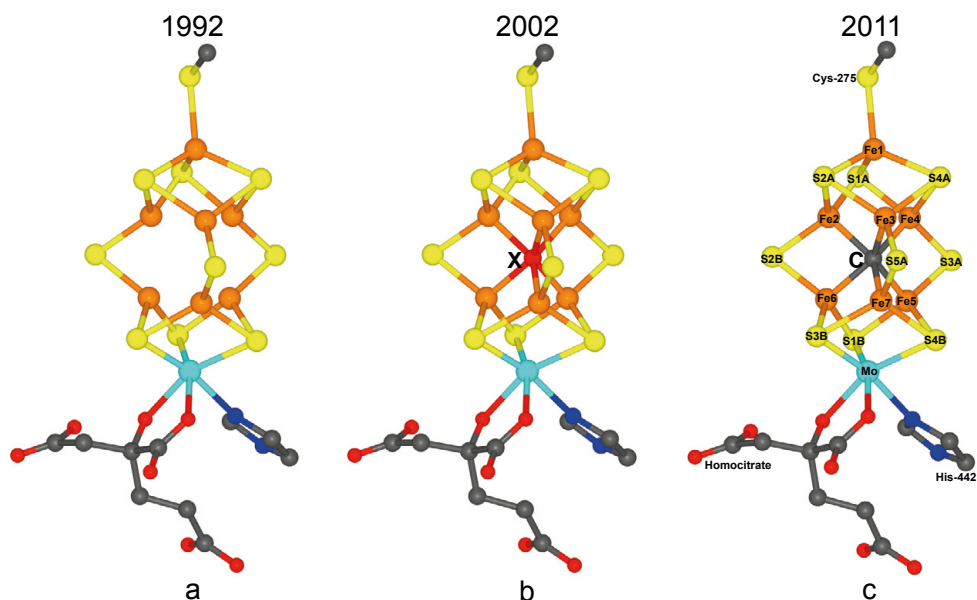
Figure 3.4: The structure of the FeMo cluster based on data from a) 1992 (PDB: 1N2C), b) 2002 (PDB: 1M1N), and c) 2011(PDB: 3U7Q).

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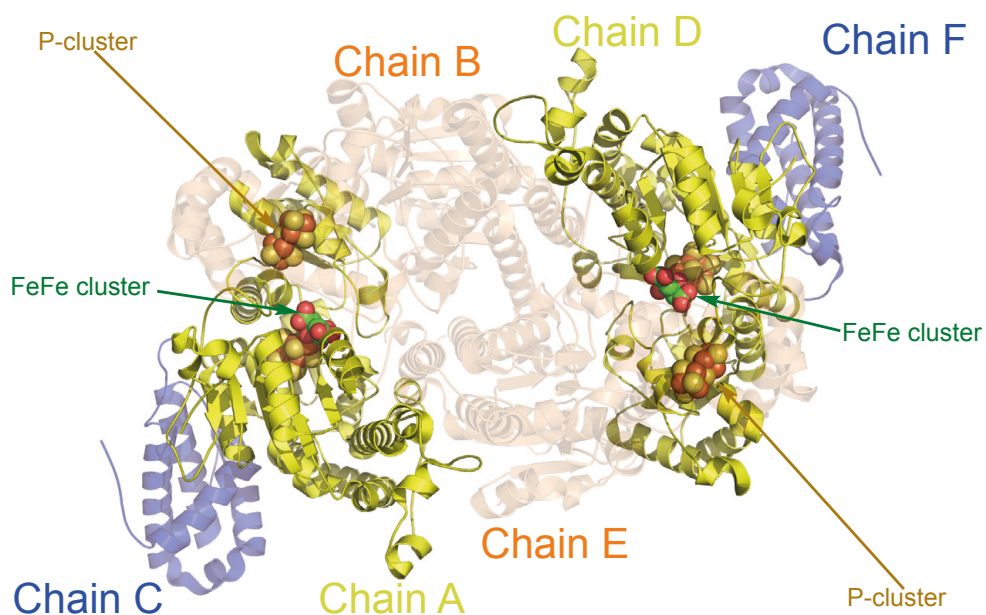
Figure 3.5: Structure of Fe-nitrogenase. Chains A and D are the D subunits and are shown in yellow. Chains B and E are the K subunits and are shown in transparent orange. Chains C and F are the G subunits and are shown in transparent blue (PDB 8BOQ).

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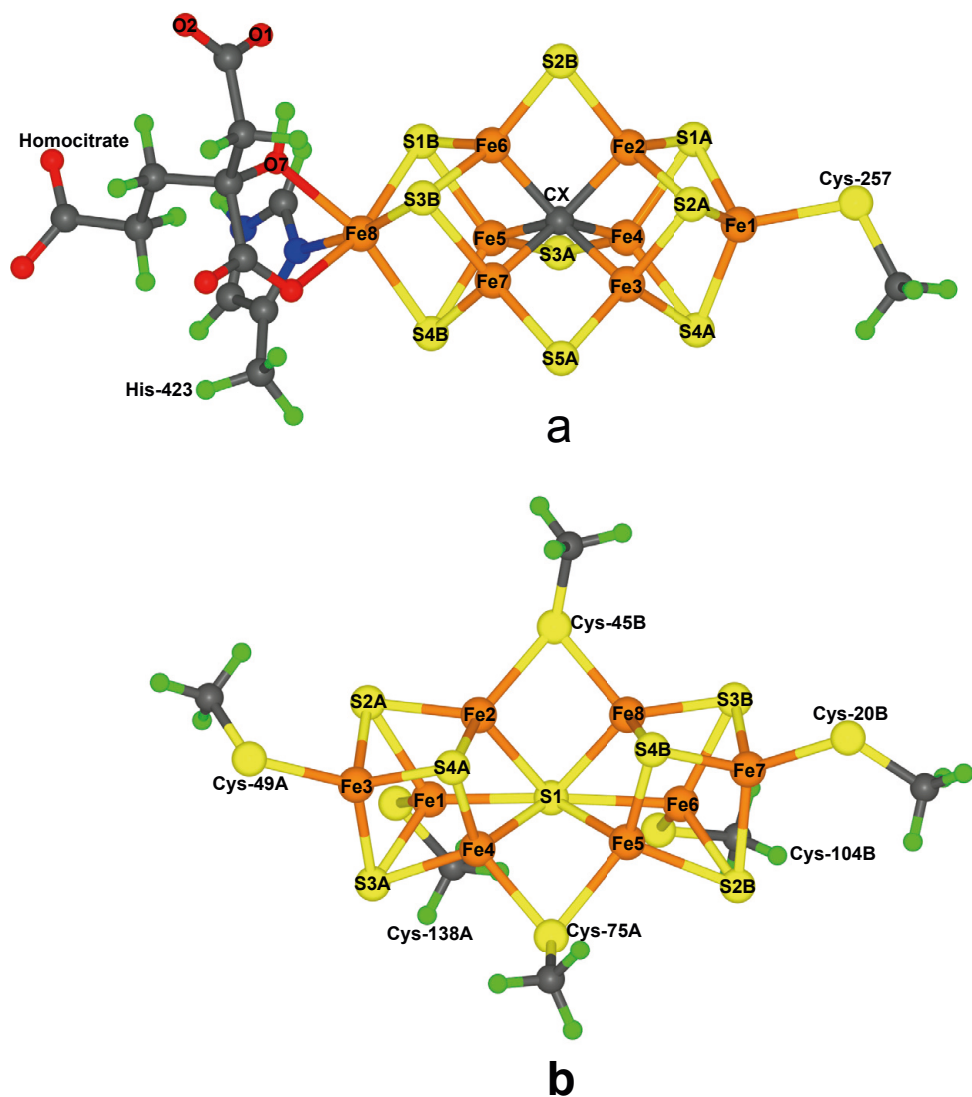
Figure 3.6: The atomic structures of the a) FeFe cluster and b) P-cluster in Fe-nitrogenase. The figure shows atom and residue names from the 8BOQ crystal structure.

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:

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \qquad (3.1)$$
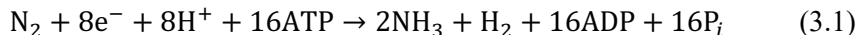
Nitrogenase has been studied extensively using spectroscopic, biochemical, and kinetic methods. [1,4,7,14,17,20,93,97] The reaction is commonly described by the Lowe–Thorneley cycle (Figure 3.7) for dinitrogen reduction, developed in the 1970s and 1980s. It describes the kinetics of transformations among 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.
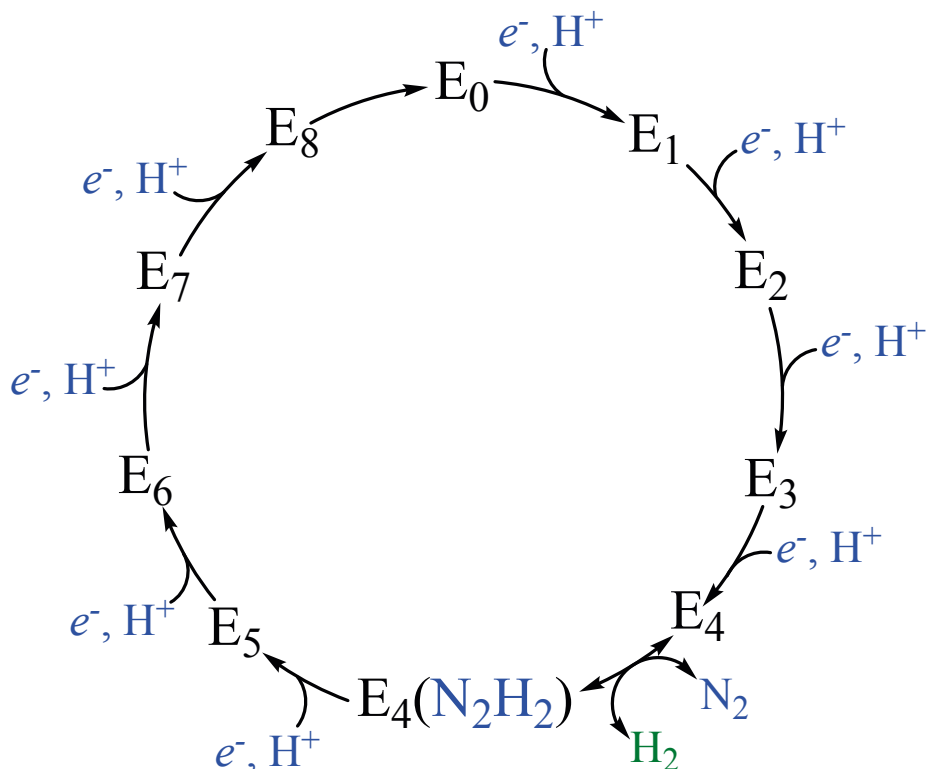
Figure 3.7: Simplified kinetic Lowe–Thorneley scheme.

3.2.1 E₀–E₄ 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.
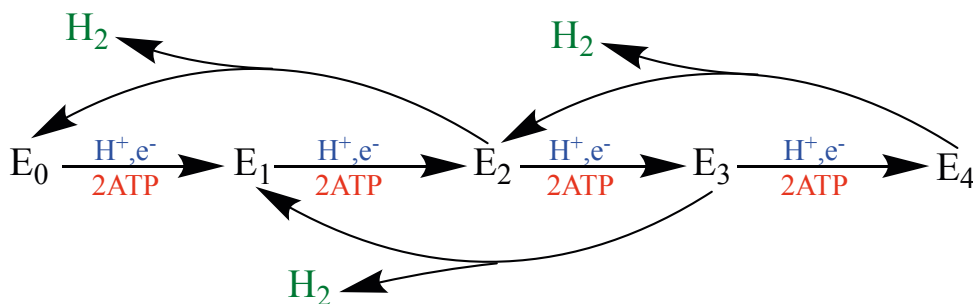
Figure 3.8: Early stages of the Lowe-Thorneley kinetic model for N_2 reduction.

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]

$$E_4(4H) \xrightleftharpoons[re/oa]{N_2 \quad H_2} E_4(N_2H_2)$$
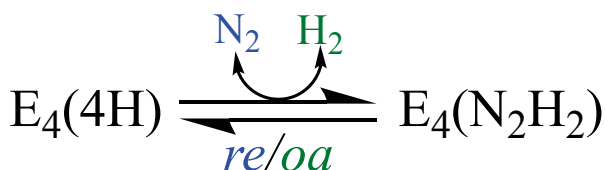
Figure 3.9: Scheme of *re/oa* mechanism.

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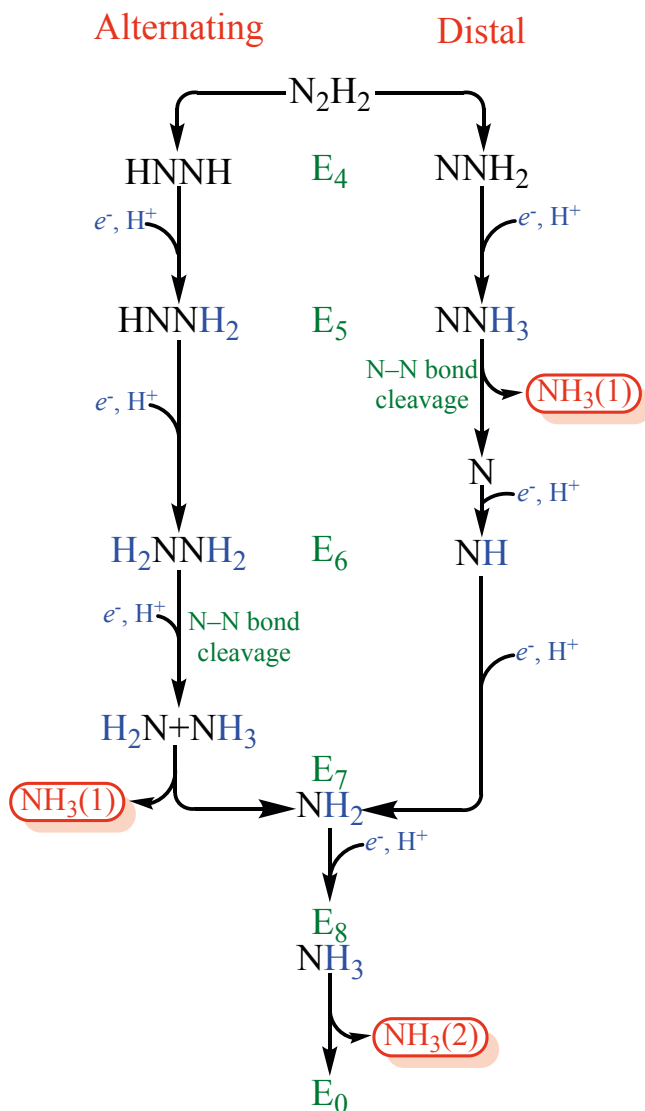
Figure 3.10: The *alternating* and *distal* mechanisms of N_2 reduction for the later parts of the Lowe–Thorneley cycle of nitrogenase.

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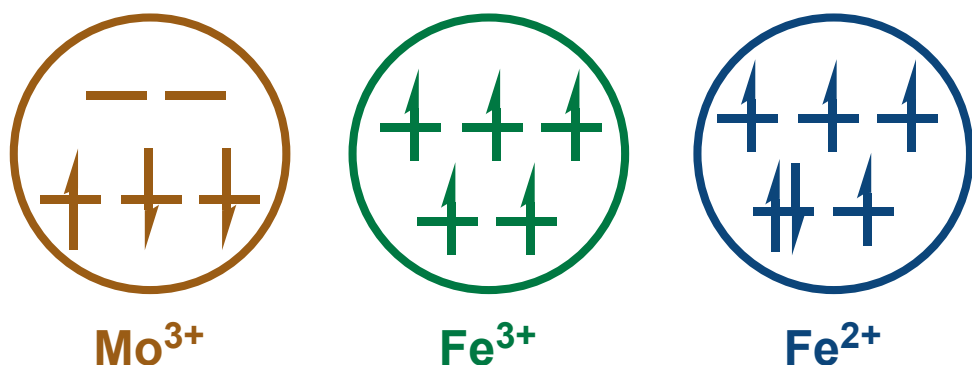
Figure 3.11: The electronic configurations for Mo^{3+}, Fe^{3+}, and Fe^{2+} in the FeMo cluster.

In early studies of nitrogenase, three different oxidation states were considered: Mo(IV)Fe(III)Fe(II)$_6$,[132] Mo(IV)Fe(III)$_3$Fe(II)$_4$,[107] and Mo(IV)Fe(III)$_5$Fe(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)$_4$Fe(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)$_4$Fe(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]
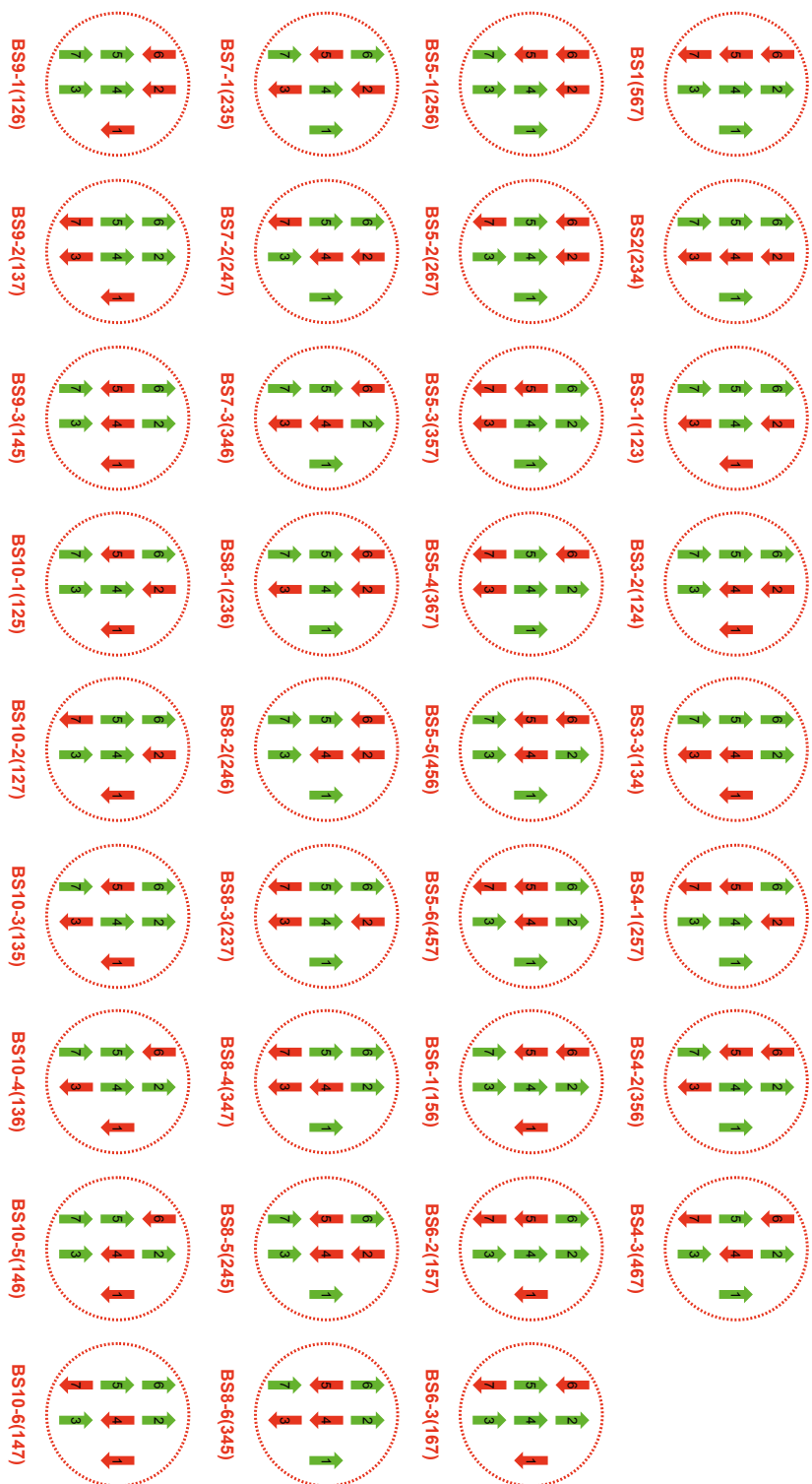
Figure 3.12: The 35 possible BS states of the FeMo cluster in nitrogenase.

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.

BS1358 BS1278 BS1247 BS1234

BS1367 BS1345 BS1248 BS1235

BS1368 BS1346 BS1256 BS1236

BS1378 BS1347 BS1257 BS1237

BS1456 BS1348 BS1258 BS1238

BS1457 BS1356 BS1267 BS1245

BS1458 BS1357 BS1268 BS1246
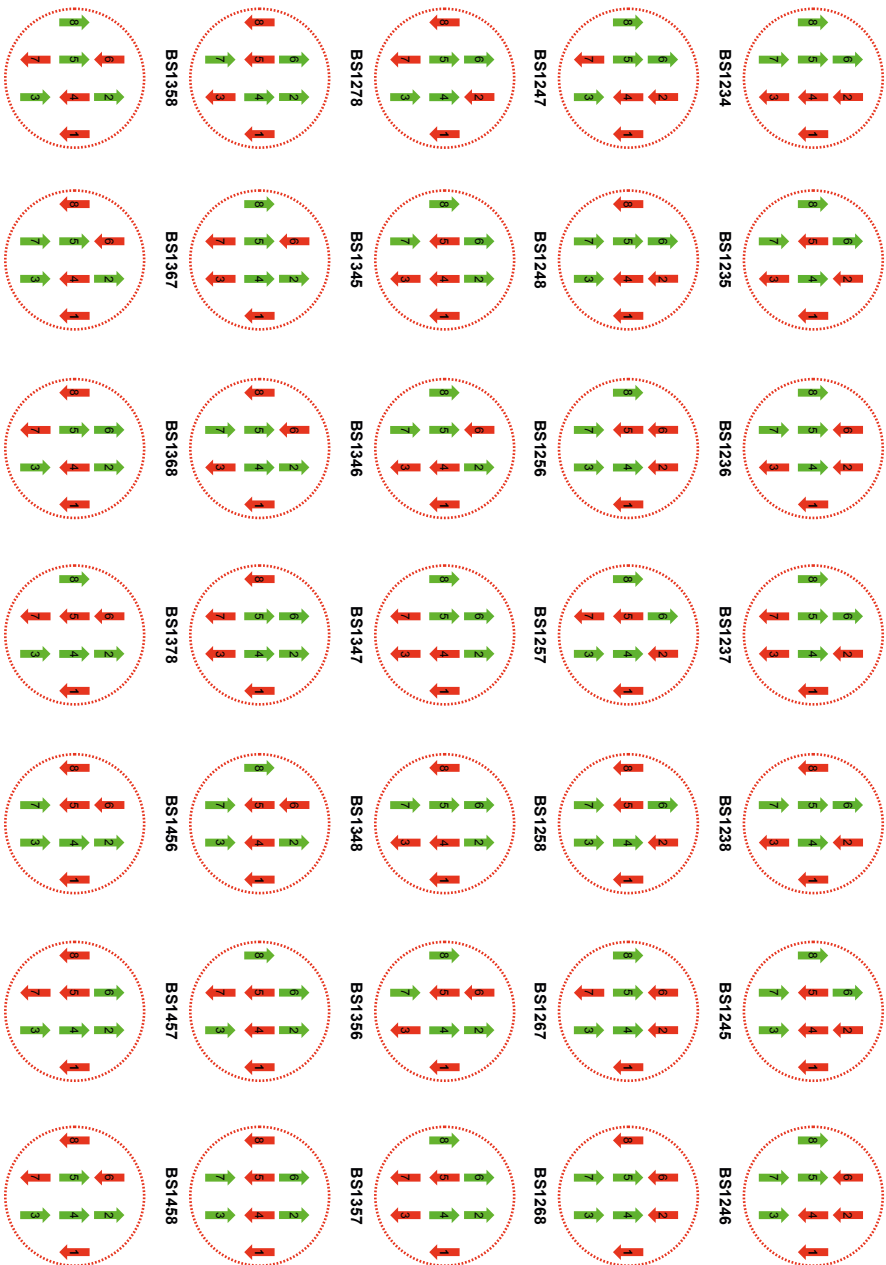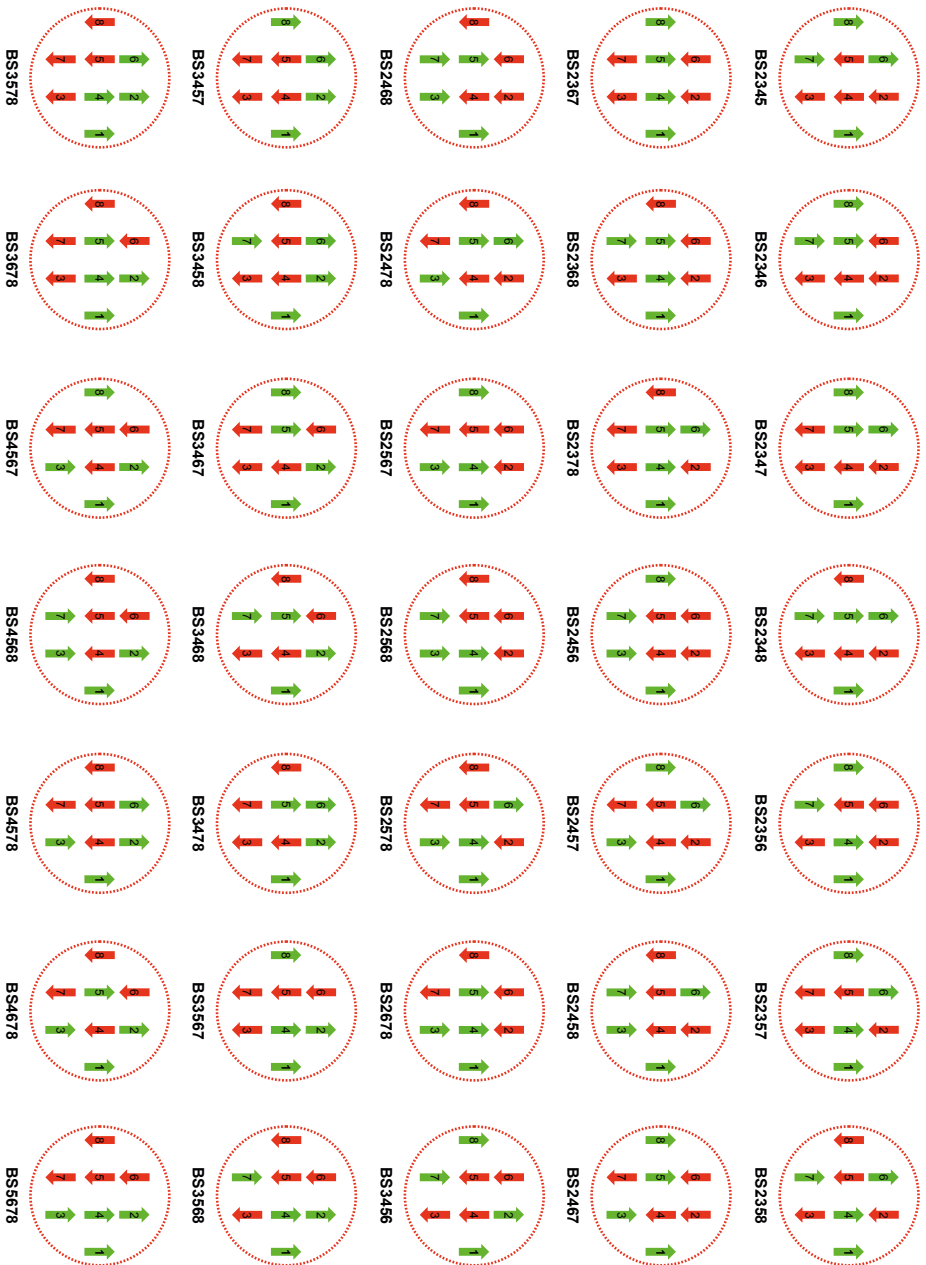
Figure 3.13: The 70 possible BS states of the FeFe cluster in Fe-onlu nitrogenase.

34

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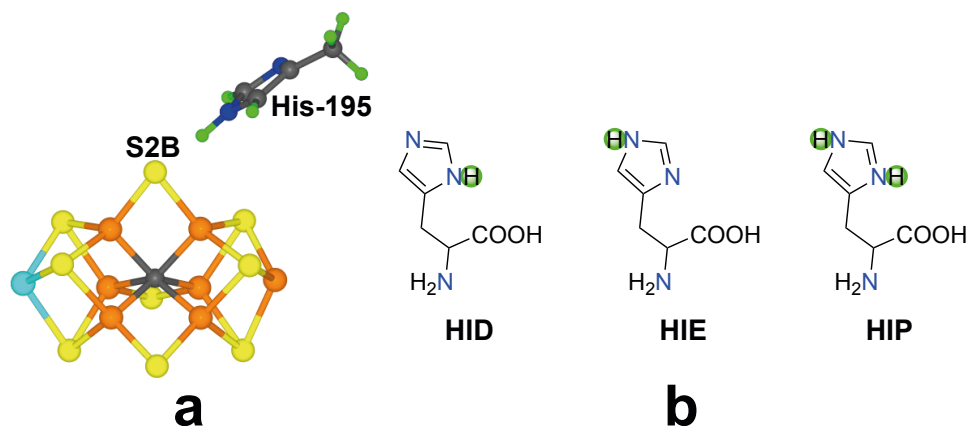
Figure 3.14: (a) The relative position of His195. (b) The three possible protonation states of histidine.

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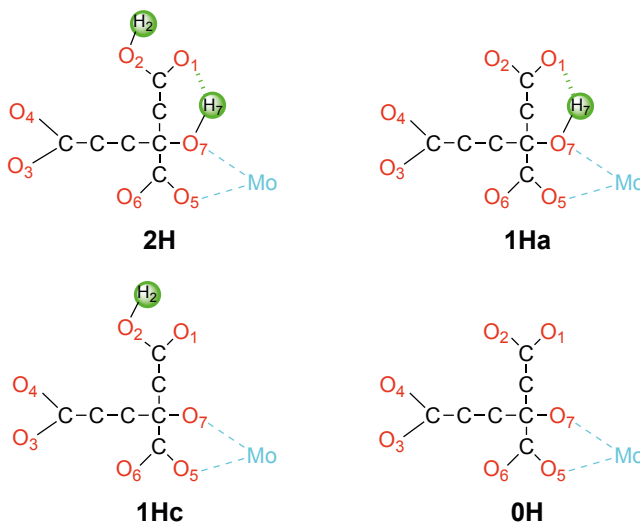
Figure 3.15: Four considered protonation states of homocitrate, 2H, 1Ha, 1Hc, and 0H. Atom numbers are also shown. Nonpolar H atoms are omitted. The charge of homocitrate is −2, −3, −3, and −4, respectively, in these four protonation states.

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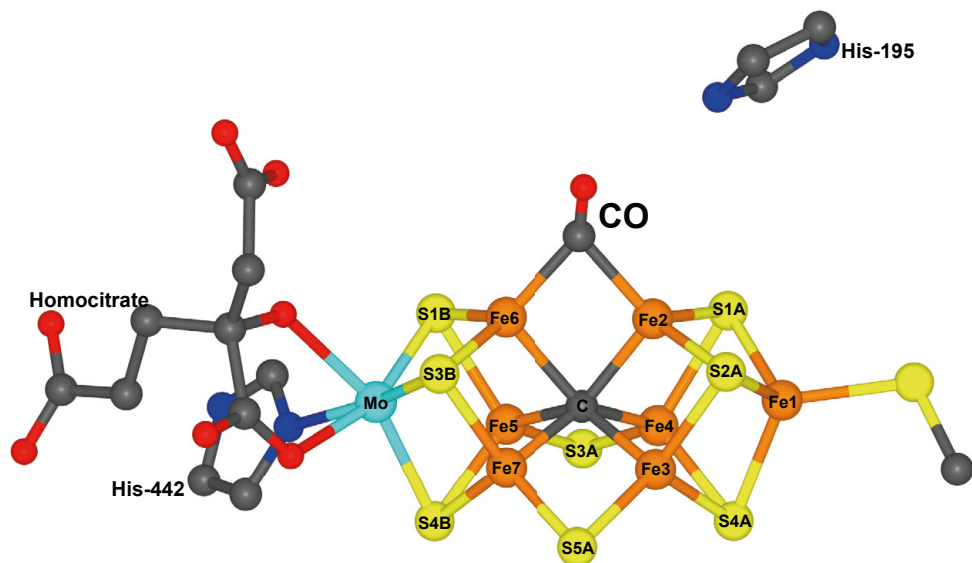
Figure 3.16: The active sites of CO-inhibited Mo-nitrogenase (PDB: 4TKV)

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]
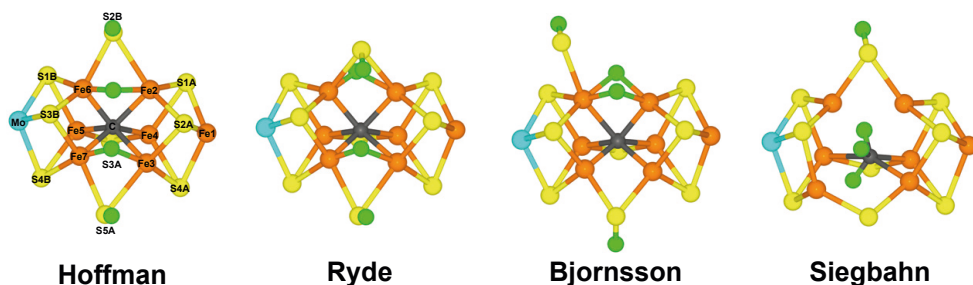
Figure 4.1: Previously suggested E$_4$ structures.

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^2SCAN 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 $3d$ 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.
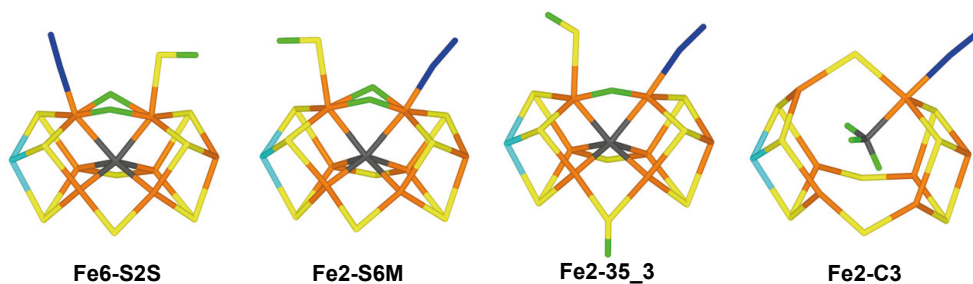
Figure 4.2: The best E_3 structures with N_2 bound, Fe6-S2S, Fe2-S6S, Fe2-3523 and Fe2-C3. The first three structures were optimised with TPSS, whereas Fe2-C3 was optimised with B3LYP.

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^2SCAN functionals, while B3LYP favors a triply deprotonated carbide ion. For E_4, the S6S structure is most stable with TPSS, TPSSh, and r^2SCAN, 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^2SCAN. TPSS generally favors N_2 binding to Fe6, while the other functionals prefer binding to Fe2.
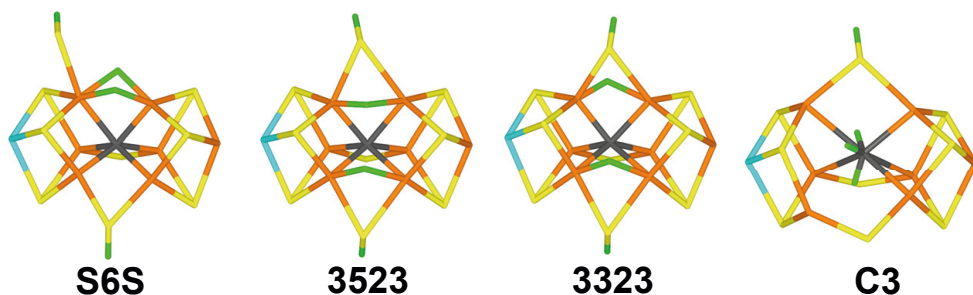
Figure 4.3: The best E_4 structures without N_2 bound.

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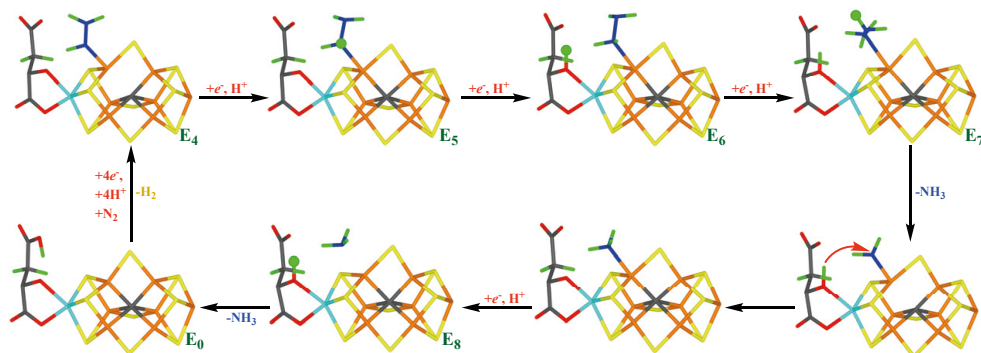
Figure 4.4: The suggested reaction mechanism for nitrogenase, assuming that S2B remains bound to the cluster.

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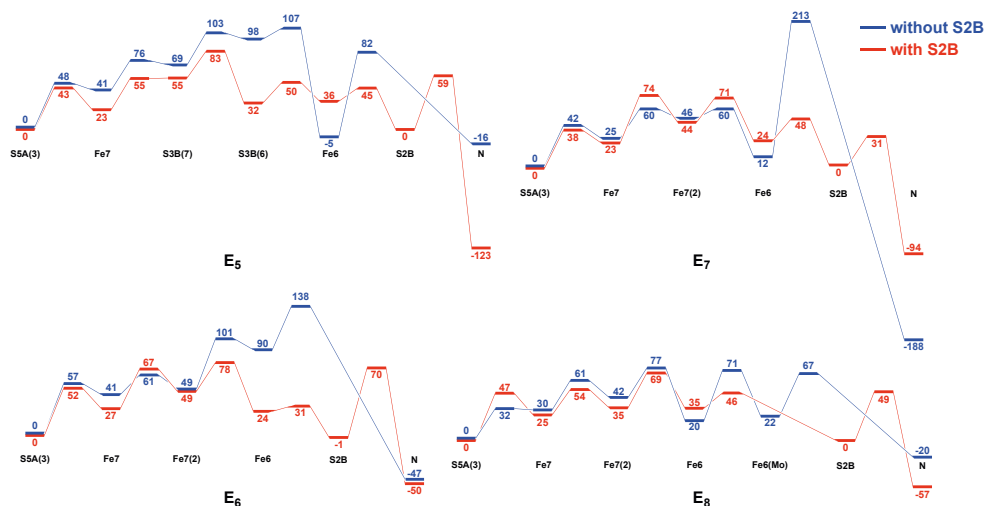
Figure 4.5: Relative energies for the proton-transfer reactions at the various E_n states with and without S2B bound.

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.
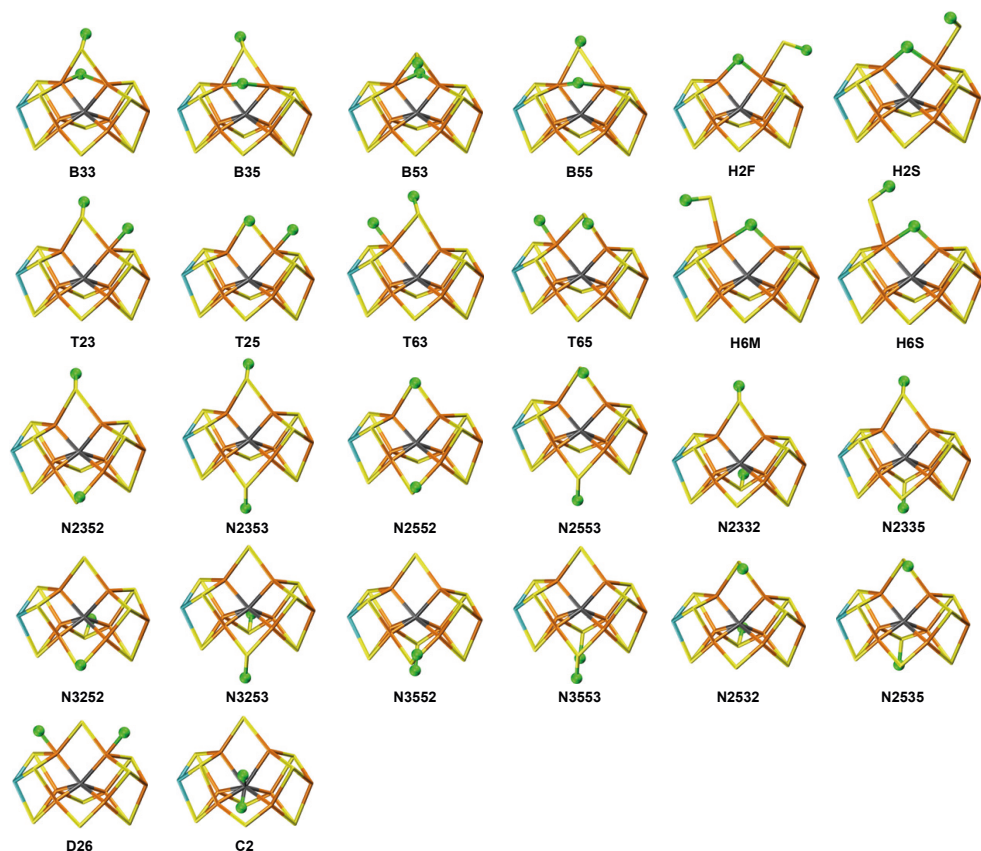
Figure 4.6: The 26 structures of the E_2 state investigated in **Paper V**.
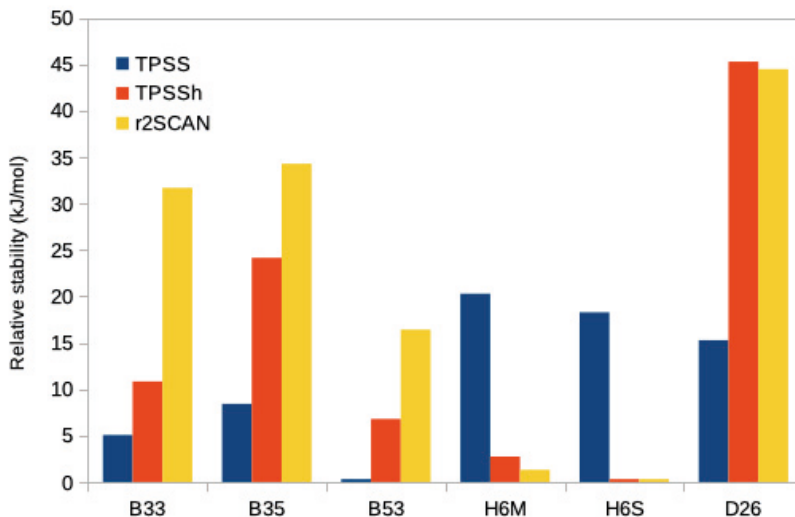
Figure 4.7: Relative stabilities of the best structures containing hydride ions for the TPSS, TPSSh, and r^{2}SCAN methods.

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^{2}SCAN, 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^2SCAN 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.
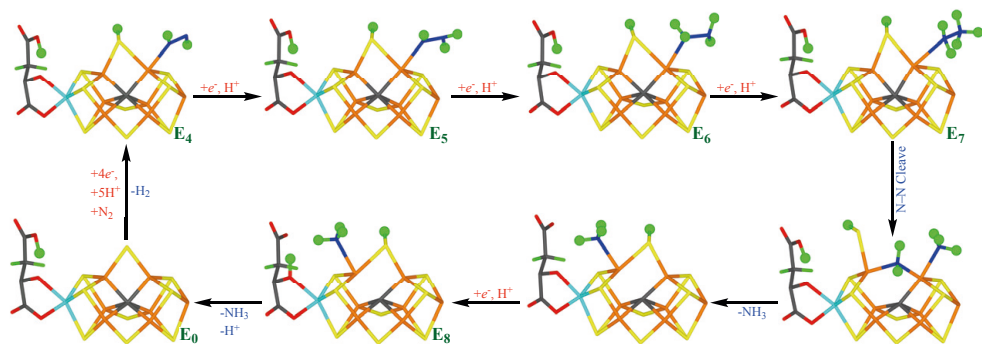
Figure 4.8: Suggested second part of the reaction mechanism for nitrogenase, assuming that the S2B ligand is protonated.

The two DFT functionals, r^2SCAN 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^2SCAN 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^o_{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^o_{calc}	ΔE^o_{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.
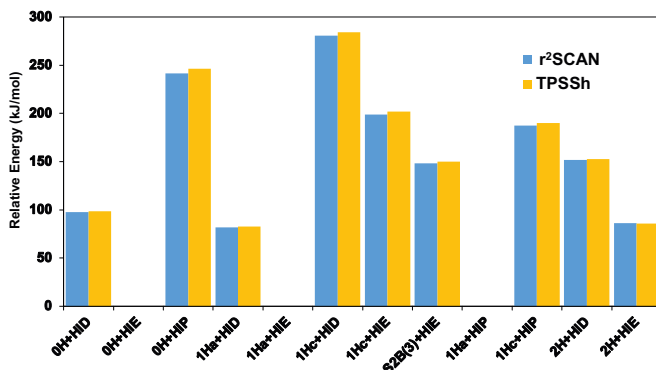
Figure 4.9: Relative energies (kJ/mol) of the various protonation states tested for Fe-nitrogenase in the E_0 state.

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^2SCAN, 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^2SCAN, 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]
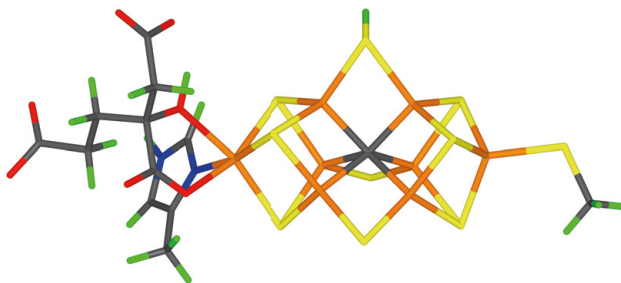
Figure 4.10: Best E_1 structure, protonated on the S2B atom.

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.

Paper I

PAPER

Check for updates

H$_2$ formation from the E$_2$–E$_4$ states of nitrogenase†

Hao Jiang [ID] and Ulf Ryde [ID] *

Nitrogenase is the only enzyme that can cleave the strong triple bond in N$_2$, making nitrogen available for biological lifeforms. The active site is a MoFe$_7$S$_9$C cluster (the FeMo cluster) that binds eight electrons and protons during one catalytic cycle, giving rise to eight intermediate states E$_0$–E$_7$. It is experimentally known that N$_2$ binds to the E$_4$ state and that H$_2$ is a compulsory byproduct of the reaction. However, formation of H$_2$ is also an unproductive side reaction that should be avoided, especially in the early steps of the reaction mechanism (E$_2$ and E$_3$). Here, we study the formation of H$_2$ for various structural interpretations of the E$_2$–E$_4$ states using combined quantum mechanical and molecular mechanical (QM/MM) calculations and four different density-functional theory methods. We find large differences in the predictions of the different methods. B3LYP strongly favours 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 E$_n$ and therefore need 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^{-1}, 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^{-1}).

Introduction

Nitrogenase (EC 1.18/19.6.1) is the only enzyme that can cleave the strong triple bond in N$_2$, making atmospheric nitrogen available to plant life.[1–4] Crystallographic studies have shown that the most active type of nitrogenase contains a MoFe$_7$S$_9$C(homocitrate) cofactor in the active site, called the FeMo cluster.[5–9]

Nitrogenase catalyses the chemical reaction[3,4]

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \tag{1}$$
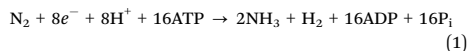
showing that eight electrons and protons are needed to convert N$_2$ to two molecules of ammonia. Consequently, the reaction is typically described by eight intermediates E$_0$–E$_7$, differing in the number of added electrons and protons.[10] It has been shown that the enzyme needs to be loaded by four electrons and protons (*i.e.* to E$_4$) before N$_2$ can bind.[3,4] It is currently

believed that the binding of N$_2$ is promoted by the dissociation of H$_2$, which is formed by reductive elimination of two hydride ions that bridge two Fe ions each.[11–15] This explains why H$_2$ is a compulsory byproduct in the reaction mechanism (*cf.* eqn (1)).

In spite of numerous spectroscopic, kinetic, biochemical and computational studies,[1–9,16,17] many details of the reaction mechanism of nitrogenase are still unknown.[4,17] An important reason for this is that different density-functional theory (DFT) methods give widely different predictions of the relative stability of various models of the active site of nitrogenase, *e.g.* different protonation states of the E$_n$ states.[17–20]

The structure of the resting E$_0$ state is well-known from accurate crystal structures, and combined quantum mechanical and molecular mechanical (QM/MM) calculations.[4,5,21–23] Moreover, quantum refinement has shown that this state does not contain any extra protons.[20] For the singly reduced and protonated E$_1$ state, DFT studies and EXAFS measurements indicate that one of the μ$_2$ bridging sulfide ions, S2B (atom names are shown in Fig. 1), is protonated.[20,24–26] However, spectroscopic studies of Fe-nitrogenase (in which the Mo ions is replaced by Fe), indicate that it instead should contain a dissociable hydride ions,[27] although this is not supported by recent QM/MM calculations.[28]

Department of Computational Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden. E-mail: Ulf.Ryde@compchem.lu.se; Fax: +46 2228648; Tel: +46 2224502

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3cp05181a
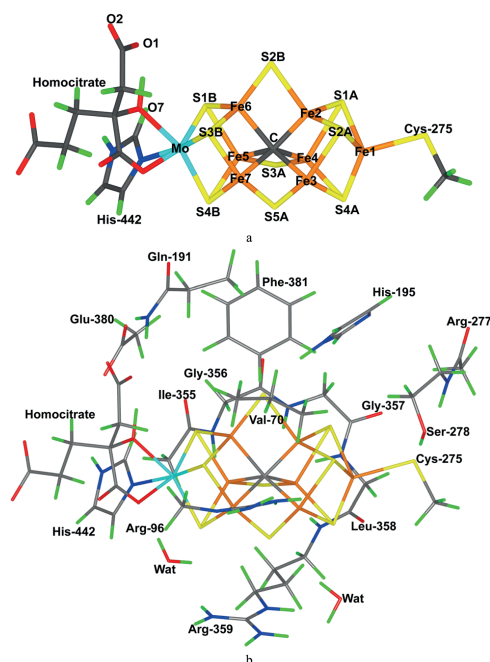
Fig. 1 (a) The FeMo cluster of nitrogenase, showing also the atom names. (b) The QM system used in the calculations, showing the names of the surrounding residues.

For the E_2 state, the results with different DFT methods start to differ.[20] Calculations with the pure meta generalised gradient approximation (mGGA) functional TPSS[29] suggest that the best structure involves a proton on S2B and a hydride ion bridging the Fe2 and Fe6 ions.[19] Other mGGA or hybrid functionals with a low amount of Hartree–Fock (HF) exchange (r^2SCAN[30] and TPSSh[31]) suggest a similar structure, but with the protonated S2B ion dissociated from one of the Fe ions.[32] A structure with both S2B and S5A protonated is also competitive with TPSSh.[33] The hybrid B3LYP functional (with 20% HF exchange)[34–36] suggests instead that the central carbide ion is doubly protonated. Likewise, for the E_3 state B3LYP suggests a structure with a triply protonated carbide ions, whereas the other functionals suggest a structure with S2B protonated and dissociated from one Fe ion, and two hydride ions both bridging between the Fe2 and Fe6 ions.[32] Other structures with a terminal hydride ion or a bound H_2 molecule have also been proposed.[37]

For the key E_4 intermediate, many different structures have been suggested.[17] ENDOR experiments show that E_4 should contain two bridging hydride ions[12,13,15] and based on this and computational studies, 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,

with all H atoms pointing towards the same face of the FeMo cluster.[38] We made a systematic search of structures with two bridging hydride ions and showed that a structure with the protons and the hydride ions at the same positions is most stable, but three of them were pointing towards to other faces of the FeMo cluster.[39] On the other hand, Bjornsson 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.[40] A thorough study of many of these structures indicated that with TPSS, the Hoffman and Bjornsson-type structures are nearly degenerate, whereas the latter are most stable with r^2SCAN and TPSSh.[32] A structure with a triply protonated carbide ion and the fourth proton on S2B is by far best by B3LYP (and competitive with TPSSh).[41,42] It has also been suggested that S2B may fully dissociate from the FeMo cluster,[43–46] inspired by crystallographic studies showing lability of this ligand.[8,47,48]

Recently, we studied the binding of N_2 to the E_0–E_4 states of nitrogenase and showed that only the TPSS[29] functional gave binding energies in accordance with experiments (i.e. a favourable binding to E_3 and E_4), but not to E_0–E_2.[32] The other tested functionals gave too weak binding to E_3 or E_4. Other studies have suggested stronger binding of N_2, but this depends on how the binding energy is defined and how the entropy loss of the N_2 ligand is treated.[40,49–51]

A related question is the formation and dissociation of H_2 from the FeMo cluster. This is a very important reaction. If H_2 forms from the E_2 and E_3 states of nitrogenase, the enzyme goes back to the E_0 and E_1 states, respectively, and two electrons and protons have been lost on an unproductive side reaction.[4,33] On the other hand, formation of H_2 is desirable for the E_4 state, but only if it is directly coupled to the binding of N_2.

Khadka and coworkers have studied H_2 evolution from the E_2 state with mediated bioelectrocatalysis and DFT calculations.[52] The experiments showed that the rate-limiting step for H_2 formation is hydride protonation. The calculations were performed with a minimal QM model of the active site in a continuum solvent. They were started from a structure with S2B protonated and a hydride ion bridging Fe2 and Fe6. A free-energy barrier of only 29 kJ mol^{-1} was obtained.

Thorhallsson and Bjornsson have studied the formation of H_2 from the E_2 state of nitrogenase by QM/MM calculations.[33] They found activation barriers of 86–95 kJ mol^{-1} from a structure with S2B dissociated from one of the Fe ions, depending on the protonation state of the nearby His-195 residue. Interestingly, the barrier was much lower (48 kJ mol^{-1}) when calculated with a small QM-cluster model, indicating that the surrounding protein counteracts this side reaction.

Here, we extend these investigations by studying the formation and dissociation of H_2 from the FeMo cluster in the E_2–E_4 states. The aim of this investigation is three-fold: The first is to study the H_2-formation reaction for many different structures: combining either two hydride ions or one proton and one hydride ion, and studying reactions for different positions in the cluster, e.g. both terminal and bridging hydride ions, located on different Fe ions or on the same Fe ion. The second is to study how easily different

This journal is © the Owner Societies 2024

Phys. Chem. Chem. Phys., 2024, **26**, 1364–1375 | **1365**

protonation states can be interconverted. This may answer the question whether the enzyme can avoid the formation of H_2 by posing the proton and hydride ions far from each other. The third is to compare the results of four different DFT functionals. We know that the enzyme needs to avoid the loss of H_2 from the E_2 and E_3 states, whereas H_2 formation is necessary for the E_4 state, but only in conjunction with the binding of N_2. If the H_2 formation is too facile for the E_2 and E_3 states, it is an indication of problems with the DFT method used or that the proper structure is still not found.

Methods

The protein

The calculations were based on the 1.0 Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[5] The setup of the protein is identical to that of our previous studies.[18,32,39,53] The entire heterotetramer was included in the calculations and the quantum mechanical (QM) calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule rather close to the active site (~ 11 Å) in the A subunit. The two P clusters and the FeMo cluster in subunit A were modelled by MM in the fully reduced and resting states, respectively, using a QM charge model.[53] The protonation states of all residues were the same as before,[53] and the homocitrate ligand was modelled in the singly protonated state with a proton shared between the hydroxyl group (O7 that coordinates to Mo) and the O1 carboxylate atom.[23,53] The protein was solvated in a sphere with a radius of 65 Å around the geometrical centre of the protein. Cl^- and Na^+ ions were added to an ionic strength of 0.2 M.[54] The final system contained 133 915 atoms. For the protein, we used the Amber ff14SB force field[55] and water molecules were described by the TIP3P model.[56] The metal sites were treated by a non-bonded model[57] and charges were obtained with the restrained electrostatic potential method.[58]

The FeMo cluster was modelled by $MoFe_7S_9C(homocitrate)$-$(CH_3S)(imidazole)$, where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included in the QM model, *viz.* Arg-96, Gln-191 and His-195 (sidechains), Ser-278 and Arg-359 (both backbone and sidechain, including the CA and C and O atoms from Arg-277), Gly-356, Gly-357 and Leu-358 (backbone, including the CA and C and O atoms from Ile-355), as well as two water molecules. Finally, the sidechain of Glu-380 was included because it forms hydrogen bonds to Gln191 and His-442, as well as the sidechains of Val-70 and Phe-381 because they are close to S2B, Fe2 and Fe6, *i.e.* the expected reactive site. The QM system contained 191–193 atoms depending on the E_n state and is shown in Fig. 1. The net charge of QM region was -4 e.

QM calculations

All QM calculations were performed with the Turbomole software (versions 7.5 and 7.6).[59] All structures were studied with

the TPSS,[29] r^2SCAN,[30] TPSSh[31] and B3LYP[34–36] functionals, whereas reactions were primarily studied with the TPSS and TPSSh functionals. TPSS and r^2SCAN are mGGA functionals, whereas TPSSh and B3LYP are hybrid functionals with 10 and 20% HF exchange, respectively. TPSS has been used in most of our previous studies of nitrogenase[20,39,53,60] and gives the best N_2 binding energies.[32] r^2SCAN and TPSSh have been shown to give accurate structures of nitrogenase models.[61] All calculations involved the def2-SV(P) basis set.[62] The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[63,64] Empirical dispersion corrections were included with the DFT-D4 approach,[65] as implemented in Turbomole.

In this investigation we study the E_0–E_4 states of the FeMo cluster. The resting E_0 state has the formal $Mo^{III}Fe_3^{II}Fe_4^{III}$ oxidation state[21,23,66] and is a quartet state according to experiments.[4] The other four states were obtained by successively adding one electron and one proton to the previous state. Several positions of the added protons were tested, based on previous investigations,[32] as will be discussed below. E_2 was studied in the quartet spin state and E_4 in the doublet state, in agreement with experiments[1–4,67–69] For E_1 and E_3, no experimental data are available and we assumed the quintet and triplet states, respectively (previous studies have shown that different spin states are close in energy).[19,20]

The electronic structure of all QM calculations was obtained with the broken-symmetry (BS) approach:[70] Each of the seven Fe ions was modelled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways.[71] The various BS states were obtained either by swapping the coordinates of the Fe ions[72] or with the fragment approach by Szilagyi and Winslow.[73] The BS states are named by listing the numbers of the three Fe ions with minority spin, *e.g.* BS-235. The selection of the BS states was based on our previous experience with the similar systems.[19,20,32,39,71,74]

For the H_2 dissociation energies (ΔE_{H_2}), H_2 was optimised in a conductor-like screening model (COSMO)[75,76] continuum solvent with a dielectric constant of 80. These calculations employed the default optimised COSMO atomic radius for H (1.30 Å) and a water solvent radius of 1.3 Å.[77] The COSMO solvation energy of H_2 is small, 1–3 kJ mol^{-1}, making this energy correction insignificant.

QM/MM calculations

QM/MM calculations were performed with the ComQum software.[78,79] In this approach, the protein and solvent are split into three subsystems: system 1 (the QM region) was relaxed by QM methods. System 2 was kept fixed at the original coordinates (equilibrated crystal structure), to avoid the risk that different calculations end up in different local minima.

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner. When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM system was capped with hydrogen atoms, the positions of which are

linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[78,80] All atoms were included in the point-charge model, except the CL atoms.[81] ComQum employs a subtractive scheme with van der Waals link-atom corrections.[82] No cut-off is used for the QM and QM–MM interactions. The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J mol^{-1} (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u. Approximate transition states for the formation of H_2 were obtained by performing relaxed scans of H–H distances.

Result and discussion

We have studied the formation of H_2 in nitrogenase. This is not possible until the E_2 state when two protons have accumulated on the cluster. Therefore, we have studied the E_2–E_4 states. The results at each reduction level are discussed in separate sections. We also study proton transfers within the FeMo cluster, connecting the various protonated structures.

H_2 formation in the E_2 state

We first studied the formation and dissociation of H_2 from the E_2 state. As discussed in the Introduction, different DFT functionals give different predictions of what is the most stable structure (protonation) of the E_2 state. Based on previous investigations,[19,32] we have studied H_2 formation and proton transfers within seven structural candidates for the E_2 state. These and some additional low-energy structural candidates are shown in Fig. 2. Relative energies calculated with four DFT functionals are shown in Table 1.

With the TPSS functional, the most stable structures have one proton bound to S2B (which bridges Fe2 and Fe6, *cf.* Fig. 1) and one hydride ion also bridging Fe2 and Fe6. There are four structural isomers of this structure depending on the direction of the two H atoms. The best one has the hydride ion on the same side of S2B as S3B (called Fe2/6(3)) and the proton on S2B also points towards S3A (called S2B(3)). This is structure B33 in Fig. 2 ("B" because both S2B and the hydride ion bridge Fe2 and Fe6). The other three isomers are called B35, B53 and B55, depending on whether the proton on S2B points towards S3A or

S5A (first number) and whether the hydride ion is on the same side of S2B as S3A or S5A (second number). They are 4–33 kJ mol^{-1} less stable than the B33 structure with TPSS.

For the B33 structure, the proton and the hydride ion are both on the same side of the cluster, with a distance of 2.32–2.33 Å in the optimised structures, and therefore the formation of H_2 from these two atoms is quite straight forward. The transition state is late, at a H–H distance of 1.0–1.1 Å, and the barrier is 86–91 kJ mol^{-1} (calculated with TPSS or TPSSh). For the B55 structure, the hydride ion and the proton are also on the same side of the cluster, at a distance of 2.21–2.23 Å. However, for this structure, the activation energy for H_2 formation is lower, 51–55 kJ mol^{-1}.

For the other two structures (B35 and B53), the proton and the hydride ion are on different sides of the cluster, with initial distances of 3.4–3.5 Å. Therefore, the reaction has to be performed in two steps. For example, B35 could first be isomerised to B33 by a rotation of the proton on S2B to the other side and then the transfer of the proton to the Fe2/6 hydride ion (which we have already studied). The rotation has a barrier of 55–62 kJ mol^{-1}.

With TPSSh and r^{2}SCAN, the best structure still has a proton on S2B and a hydride ion bridging Fe2/6, but S2B has dissociated from Fe2, but not from Fe6. This structure is called H6 in Fig. 2, indicating that S2B is half-dissociated, still binding to Fe6. The corresponding structure with S2B dissociated from Fe6 instead (H2 in Fig. 2) is 71–97 kJ mol^{-1} higher in energy. With TPSS, B33 is 13 kJ mol^{-1} more stable than H6, whereas the opposite is true with r^{2}SCAN and TPSSh by 41 and 14 kJ mol^{-1}, respectively.

For the H6 structure, the two H atoms are also on the same side of the cluster, with a distance of 2.43–2.48 Å. However, the activation barrier for H_2 formation is quite high, 79 kJ mol^{-1} with TPSS and 104 kJ mol^{-1} with TPSSh. The transition state has a H–H distance of 1.0–1.1 Å. For this structure, we also studied how easily the proton on the half-dissociated S2B may rotate 360° around the Fe6–S2B axis. We found two local minima. The lowest is that shown in Fig. 2, with the proton pointing between the hydride ion and S1B. In the second, which is 2–3 kJ mol^{-1} less stable, it instead points towards S3B. The two minima are separated by barriers of only 12–17 kJ mol^{-1}. Thus, the rotation of this proton is essentially free and it is enough to study the most stable structure.

Moreover, we studied the formation of H6 from B33 by cleavage of the S2B–Fe2 bond. This cleavage turned out to be quite facile, with a barrier of 35–39 kJ mol^{-1}. Similar reactions could also be used twice to move hydride ion in the B-type structures.

With B3LYP, the most stable structure instead has a doubly protonated central carbide ion. This structure is called C2 (two protons on the carbide ion). It is 88 kJ mol^{-1} more stable than the H6 state with B3LYP, whereas the opposite is true by 66–155 kJ mol^{-1} with the other functionals. We tried to form H_2 in the C2 structure using the B3LYP functional. Interestingly, even if the two protons on the carbide ion are quite close (1.71 Å), no reaction could be obtained; instead the energy increased monotonically to over 200 kJ mol^{-1} when the H–H

Fig. 2 The eight considered E_2 structures, showing the structure obtained with TPSS (B3LYP for C2). The labels are explained in the text. H–H distances in Å are marked in the figures.

Table 1   Structures and reactions within the E_2 state, listing the structure (Struct), the positions of the two added H atoms (H1 and H2), the BS state, the H–H distance (Å), the relative energy (ΔE), the H_2 dissociation energy (ΔE_{H_2}) relative the E_0 state and H_2 in a COSMO solvent, and the activation energy for H_2 formation ($\Delta E^{\ddagger}$; all energies in kJ mol^{-1})

				TPSS				r^{2}SCAN		TPSSh				B3LYP	
Struct.	H1	H2	BS	H–H	ΔE	ΔE_{H_2}	$\Delta E^{\ddagger}$	ΔE	ΔE_{H_2}	H–H	ΔE	ΔE_{H_2}	$\Delta E^{\ddagger}$	ΔE	ΔE_{H_2}
B55	S2B(5)	Fe2/6(5)	235	2.23	33	−133	51	63	−166	2.21	42	−146	59	171	−158
B53	S2B(5)	Fe2/6(3)	235	3.41	4	−104		46	−150	3.45	21	−125	62	152	−139
B33	S2B(3)	Fe2/6(3)	235	2.32	0	−100	86	41	−145	2.33	14	−119	91	142^a	−130
B35	S2B(3)	Fe2/6(5)	235	3.47	17	−117		17	−120	3.55	9	−114		133	−120
T53	S2B(3)	Fe5	235	5.82	19	−119	160	25	−129	5.82	16	−120		115	−102
H6	S2B(H6)	Fe2/6	235	2.43	13	−113	79	0	−104	2.48	0	−104	104	88^a	−75
H2	S2B(H2)	Fe2/6	235	3.66	75	−175		80	−183	3.78	71	−175		186	−173
C2	C2367	C3457	345	1.87	155^b	−255		113	−217	1.77	35	−139		0^c	13
C1	S2B(3)	C2456	346	3.90										32	−18
E_0 + H_2 from B55			235	0.75	−57	−43		−60	−44	0.75	−65	−40		−32	−43
E_0 + H_2 from B33			235	0.75	−44	−56		−40	−64	0.76	−48	−57		−9	−66
E_0 + H_2 from H6S			235	0.76	−105	5		−109	5	0.76	−109	5		−82	6

^{a} BS-346. ^{b} BS-235. ^{c} $\Delta E^{\ddagger}$ > 220 kJ mol^{-1}.

distance was decreased. We also tried to move the Fe2/6 hydride ion in the H6S structure to the central carbide ion. However, the barrier for such a movement was quite high, 106 kJ mol^{-1} with B3LYP. The product, which has protons on the carbide ion and on S2B (which rebound to Fe2 during the reaction) is 57 kJ mol^{-1} more stable than the reactant and therefore only 32 kJ mol^{-1} less stable than the C2 structure (*i.e.* the second-best structure with B3LYP; called C1 in Table 1).

Finally, we considered also H_2 formation from a structure with a hydride ion on Fe5 and a proton on S2B(3) (called T53, highlighting the terminal hydride ion). This structure is 16–25 kJ mol^{-1} less stable than the best one for TPSS, r^{2}SCAN and TPSSh (but 115 kJ mol^{-1} with B3LYP). The distance between the two H atoms is 5.82 Å. Therefore, the hydride ion on Fe5, first has to move from the *exo* position (*trans* to C), to an *endo* position, where it almost bridges to Fe4. The barrier for this is quite low, only 32 kJ mol^{-1}, and the intermediate is only 13 kJ mol^{-1} less stable than the starting T53 structure. However, the reaction from this intermediate with a H–H distance of 3.6 Å has a barrier of 160 kJ mol^{-1}, making the reaction prohibitive.

In all product structures of these reactions, the formed H_2 molecule has dissociated from the FeMo cluster. The structure of the cluster in these product states is similar because S2B binds back to Fe2 for the H6 structure, forming a normal E_0 structure. However, H_2 resides in different positions in the second coordination sphere. The most stable structure is the one starting from H6. In this structure, H_2 is 3.8 Å from Fe2 and it interacts weakly with His-195, Arg-277 and Ser-278 (*cf.* Fig. 3). This structure is 105–109 kJ mol^{-1} more stable than the best E_2 structure with the TPSS, r^{2}SCAN and TPSSh functionals, but it is 6 kJ mol^{-1} less stable than the C2 structure with B3LYP. This shows that formation of H_2 is strongly downhill with the former three functionals.

Another way to quantify this is to calculate the energy difference between a certain structure and the E_0 state plus H_2 in water solution (optimised with the COSMO continuum-solvation model). This H_2 dissociation energy (ΔE_{H_2}) is given
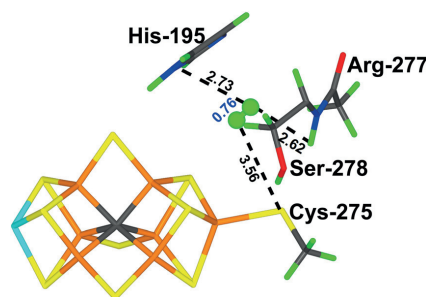
Fig. 3   The best E_2 structure with H_2 in the second coordination sphere. The H–H bond length and distances to the nearest residues are marked (in Å).

for all structures in Table 1. It can be seen that formation and dissociation of H_2 are strongly favourable for all structures considered, by 100–255 kJ mol^{-1} for TPSS, r^{2}SCAN and TPSSh. For B3LYP, the energies are lower, and for the most stable C2 state, the H_2 dissociation energy is actually unfavourable by 13 kJ mol^{-1}. Yet, H_2 dissociation would be further enhanced by translational and rotational entropy of the released H_2 molecule (typically estimated to 17–45 kJ mol^{-1}).[40,51,83,84] For the three structures with H_2 already formed and located in the second coordination sphere, ΔE_{H_2} is smaller and actually unfavourable by 5–6 kJ mol^{-1} for the best structure with all four functionals (but again this will be reversed by the entropy term).

H_2 dissociation from the E_3 state

Next, we studied H_2 formation from the E_3 state. As was discussed in the Introduction, the best E_3 structure with TPSS, r^{2}SCAN and TPSSh has S2B protonated and dissociated from Fe2, but still binding to Fe6, and two hydride ions both bridging Fe2 and Fe6 (called structure S6, because the two hydride ions bind to the same pair of Fe ions and S2B binds to

Fe6; *cf.* Fig. 4). A structure with the proton on S2B pointing to another direction (towards S3B rather than S1B) is only 2–9 kJ mol^{-1} less stable and since the results in the previous section indicated that the proton on the half-dissociated S2B may rotate quite freely around the Fe6–S2B axis, this structure is not further discussed. A structure with the proton and the hydride ions at the same positions, but with S2B binding to Fe2 instead (S2) is 52–85 kJ mol^{-1} less stable than S6.

On the other hand, with B3LYP, a structure with a triply protonated carbide ion, is 195 kJ mol^{-1} better (C3 in Fig. 4). With TPSS and r^{2}SCAN, C3 is 141–189 kJ mol^{-1} less stable than S6, but with TPSSh, the difference is only 4 kJ mol^{-1}. We have studied several other E$_3$ structures, of two different types. One is related to Hoffman's suggestion for the E$_4$ state (*i.e.* with protons on S2B and S5A and hydride ions bridging Fe2/6 and Fe3/7), but missing one of the H atoms for E$_3$.[38] They are named according to the direction of the four H atoms (towards, S3A, S2B or S5A), given by a number in the order S2B–Fe2/6–Fe3/7–S5A and with an underscore marking the missing H atom, *e.g.* 35_2, (*cf.* Fig. 4). The second type has a proton on S2B(3) (which bridges Fe2/6), a terminal hydride on Fe5 and either another terminal hydride on Fe4 (called 345) or a hydride ion bridging Fe2/6, pointing either towards S3A or S5A (called 335 or 355, respectively; *cf.* Fig. 4). The best structures of these two types are 21–46 and 44–58 kJ mol^{-1} less stable than the S6 structure (35_3 and 355 best), respectively, with TPSS, r^{2}SCAN and TPSSh, but 178 kJ mol^{-1} worse than C3 with B3LYP (35_2 best).

For all functionals, formation of H$_2$ is strongly exothermic, by 59–107 kJ mol^{-1} for the best structures (*i.e.* leading to H$_2$ and E$_1$ with S2B protonated; least for TPSS and most for TPSSh, ΔE_{H_2} in Table 2). The only exception is B3LYP, for which the C3 structure is 49 kJ mol^{-1} more stable than E$_1$ + H$_2$, *i.e.* an energy that is larger than the entropy gain from H$_2$ dissociation.
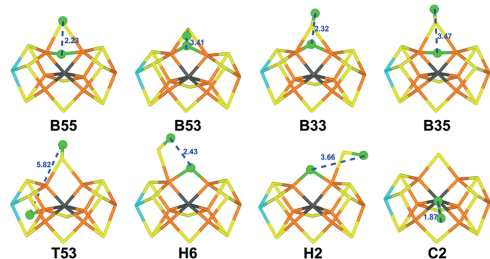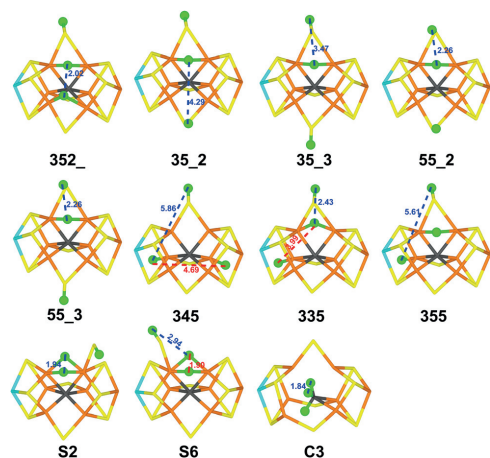

Fig. 4 The eleven considered E$_3$ structures, showing the structure obtained with TPSS (B3LYP for C3). The labels are explained in the text.

We studied H$_2$ formation for seven different structures. Interestingly, the activation barrier for the reactions depends mainly on what types of H atoms are involved. Connecting two Fe-bound hydride ions give small barriers. The two hydride ions in the S6 structure (at a H–H distance of 1.90 Å) can be connected with an activation energy of only 57 (TPSS) or 29 (TPSSh) kJ mol^{-1}. The reason for the large difference between the two functionals is that the reaction is much more exothermic with TPSSh than with TPSS (42 and 5 kJ mol^{-1}), because the product state contains S2B dissociated from Fe2 (and H$_2$ bound to Fe2). The transition state is later for TPSS (the H–H distances are 1.1 and 1.3 Å, respectively). Likewise, the barrier to connect the two Fe2/6 and Fe3/7 bridging hydride ions in the 352_ structure, at an initial distance of 2.02 Å, is only 34–39 kJ mol^{-1}.

More complicated reactions involving two hydride ions also give relatively low barriers. The Fe2/6 and Fe5 hydride ions in the 335 structure are initially at a distance of 3.91–3.99 Å. However, they can form H$_2$ by first moving the Fe2/6 hydride ion also to Fe5 (in the *endo* position) (*i.e.* almost bridging to Fe4) and then connect the two hydride ions. The first step has a barrier of 46–55 kJ mol^{-1}, whereas the second step is facile with a barrier of only 30–31 kJ mol^{-1}. Likewise, the two terminal hydride ions on Fe4 and Fe5 in the 345 structure, initially 5.9 Å apart, can be connected *via* an intermediate with the Fe5 hydride ion moved to S5A and a net barrier of 52 kJ mol^{-1} with TPSS and only 29 kJ mol^{-1} for TPSSh (in both cases for the first step).

On the other hand, reactions involving a proton and a hydride ion have appreciably higher activation energies. The proton on S2B and the hydride ion bridging Fe2/6 in the 55_2 structure are initially at a distance of 2.26–2.27 Å. They can form H$_2$ *via* an activation barrier of 75 (TPSS) to 115 kJ mol^{-1} (TPSSh). Likewise, the same two H atoms in the 335 structure (on the other side of the cluster), with an initial distance of 2.43–2.47 Å, can be connected with a barrier of 101–110 kJ mol^{-1}. The same applies to the proton on S2B and the closest Fe2/6 hydride ion in the S6 structure (initially 2.92–2.94 Å apart). They can be connected passing a barrier of 118 (TPSS) or 86 kJ mol^{-1} (TPSSh).

However, we failed to find any reaction of the two closes protons on the central carbide in the C3 structure with both B3LYP and TPSSh. Although the two protons are only 1.76 Å apart at the start, the reaction was monotonously uphill to over 200 kJ mol^{-1}.

We also studied some proton-transfer reactions within the cluster. For example, the Fe hydride ion bridging Fe3/7 could be moved to S5A in the 352_ structure (forming the 35_2 structure) with a barrier of 53 (TPSS) or 26 kJ mol^{-1} (TPSSh). Again, the reason for the lower barrier with TPSSh is that the reaction is more exothermic with that functional.

Among the product structures, the one started from the 345 structure was most stable, because it resulted in a E$_1$ structure with a proton on S2B(3) and H$_2$ in the second coordination sphere. It was 61–114 kJ mol^{-1} more stable than the best E$_3$ state with the TPSS, r^{2}SCAN and TPSSh functionals, but 42 kJ mol^{-1} less stable than the C3 structure with B3LYP. Three structures had H$_2$ coordinated to the cluster (either to Fe2 or Fe7), but these

This journal is © the Owner Societies 2024

Phys. Chem. Chem. Phys., 2024, **26**, 1364–1375 | **1369**

Table 2 Structures and reactions within the E_3 state, listing the structure (Struct), the positions of the three added H atoms (H1, H2 and H3), the BS state, the H–H distance (Å), the relative energy (ΔE), the H_2 dissociation energy (ΔE_{H_2}) relative the E_1 state and H_2 in a COSMO solvent, and the activation energy for H_2 formation ($\Delta E^\ddagger$; all energies in kJ mol^{-1}). For the 335 and S6 structures, two reactions were studied, connecting either a hydride ion and a proton (HP) or two hydride ions (HH; the two H atoms involved in the reactions are marked in bold face). The lower part of the table shows the products after the H_2-formation reactions, labelled after the starting E_3 structure and reaction

					TPSS				r²SCAN		TPSSh				B3LYP	
Struct.	H1	H2	H3	BS	H–H	ΔE	ΔE_{H_2}	$\Delta E^\ddagger$	ΔE	ΔE_{H_2}	H–H	ΔE	ΔE_{H_2}	$\Delta E^\ddagger$	ΔE	ΔE_{H_2}
352_	S2B(3)	**Fe2/6(5)**	**Fe3/7(2)**	147	2.02	57	−116	39	77	−169	2.03	62	−170	34	265	−216
35_2	S2B(3)	**Fe2/6(5)**	**S5A(2)**	147	4.29	55	−115	136	56	−148	4.26	32	−139	124	178	−129
35_3	S2B(3)	Fe2/6(5)	S5A(3)	147	3.47	46	−105		45	−137	3.50	21	−128		187	−138
55_2	**S2B(5)**	**Fe2/6(5)**	S5A(2)	147	2.26	74	−133	75	77	−168	2.27	51	−158	115	212	−163
55_3	S2B(5)	Fe2/6(5)	S5A(3)	147	2.26	63	−122		65	−156	2.26	40	−147		206	−157
345	**S2B(3)**	**Fe4**	Fe5	147	5.86	68	−127	52	58	−150	5.91	58	−165	29	248	−199
335 HP	**S2B(3)**	**Fe2/6(3)**	Fe5	147	2.43	61	−120	101	69	−161	2.47	56	−163	110	258	−209
HH		**Fe2/6(3)**	**Fe5**	147	3.99			55			3.91			46		
355	S2B(3)	Fe2/6(5)	Fe5	147	5.61	55	−115		58	−150	5.65	44	−152		242	−193
S2	S2B(H2)	Fe2/6(3)	Fe2/6(5)	147	1.94	59	−118		67	−158	1.93	52	−159		263	−214
S6 HP	**S2B(H6)**	**Fe2/6(3)**	**Fe2/6(5)**	147	1.90	0	−72	118	0	−91	1.90	0	−107	86	195	−146
HH	**S2B(H6)**	**Fe2/6(3)**			2.94		−59	57			2.92			29		
C3	C2367	C2456	C3457	147	1.84	189	−248		141	−233	1.80	4^a	−111		0	49^b
Product states = E_1 + H_2																
352_	S2B(3)	H_2 on Fe7		147	0.85	77	−136	139	104	−195	0.84	74	−181	187	233	−184
55_2	S5A(2)	H_2 dissociated		147	0.76	31	−90	92	12	−103	0.76	−1	−107	113	151	−102
345	S2B(3)	H_2 dissociated		147	0.77	−61	2	0	−109	18	0.76	−114	6	0	42	7
335 HP	Fe5	H_2 dissociated		147	0.76	−25	−34	37	−55	−36	0.76	−46	−61	68	134	−85
335 HH	S2B(3)	H_2 dissociated		147	0.76	−49	−11	13	−97	5	0.76	−101	−6	13	56	−7
S6 HP	Fe2,6(5)	H_2 on Fe2^c		147	0.79	49	−108	110	28	−119	0.78	34	−142	148	222	−173
S6 HH	S2B(H6)	H_2 on Fe2^d		147	0.76	9	−68	70	−38	−54	0.77	−42	−65	71	115	−67

^{a} BS-136. ^{b} Barrier >250 kJ mol^{-1}. ^{c} S2B dissociated from Fe2. ^{d} Protonated S2B dissociated from Fe2.

structures were at least 70–73 kJ mol^{-1} less stable than the best structure with H_2 dissociated. It should be noted, however, that we have not made any systematic investigation of H_2-bound (or H_2-dissociated) structures.

H_2 dissociation from the E_4 state

Finally, we studied H_2 formation for the E_4 state. We studied 24 reactions, starting from 14 different structures, described in Table 3 and in Fig. 5. Most of the structures are of two different types. The first have two protons on S2B and S5A and two hydride ions bridging Fe2/6 and Fe3/7, as suggested by Hoffman and coworkers.[38] As for the E_3 structures, they are named according to the direction of the four H atoms (towards S3A, S2B or S5A), given by a number in the order S2B–Fe2/6–Fe3/7–S5A, e.g. 5522 (cf. Fig. 5). We also studied two structures with S2B dissociated from Fe2 (3323H6 and 3523H6) The second group is related to the S2 and S6 structures for E_3. They also have protons on S2B and S5A, and two hydride ions, but both hydride ions bridge Fe2 and Fe6, and S2B has dissociated from either Fe2 (S6) or Fe6 (S2). Both structures have two local minima depending on the direction of the proton on S2B (cf. Table 3). In addition, we studied two differing structures: 3H has a proton on S2B and three hydride ions on Fe5, Fe6 and bridging Fe2/6. C3 has three protons on the central carbide and one on S2B.

The relative stabilities of the various E_4 states were discussed in a previous study,[32] showing that S61 is most stable with TPSS, TPSSh and r²SCAN, but C3 with B3LYP (which is 233 kJ mol^{-1} more stable than S61 with this functional, whereas C3 is 192,

146 and 1 kJ mol^{-1} less stable than S61 with TPSS, TPSSh and r²SCAN, respectively). However, with TPSS, several Hoffman-type structures are competitive, in particular 3323, which is only 3 kJ mol^{-1} less stable than S61 (27 and 11 kJ mol^{-1} less stable with r²SCAN and TPSSh).

With the TPSS, TPSSh and r²SCAN functionals, all studied E_4 structures are less stable than the best E_2 structure and H_2 in a COSMO solvent (i.e. ΔE_{H_2} in Table 3) by at least 72, 103 and 95 kJ mol^{-1}, respectively. On the other hand, the C3 structure is 3 kJ mol^{-1} more stable than the dissociated state with B3LYP. Thus, we can again conclude that formation and dissociation of H_2 from the FeMo cluster is highly thermodynamically favourable and should be kinetically controlled. Therefore, we performed a thorough investigation of the formation of H_2 from all pairs of H atoms on the same face of the FeMo cluster in 14 different starting structures.

The barriers of all 24 studied reactions are collected in Table 3. The reactions can be divided into five groups. The smallest barrier, 15–25 kJ mol^{-1}, is found for the reaction of two hydride ions bound to the same Fe6 ion, one terminal and the other bridging to Fe2 in the 3H structure. This gives an initial H–H distance of 2.30–2.40 Å. Likewise, reactions from the S61 and S64 structures, where the two hydride ions both bridge Fe2 and Fe6 also give quite small barriers of 34–47 kJ mol^{-1}. In the starting structure, the two hydride ions are only 1.74–1.87 Å apart.

Five reactions have barriers of 59–72 kJ mol^{-1} with TPSS and 48–81 kJ mol^{-1} with TPSSh. They all involve reactions between a proton on S2B and hydride ion bridging Fe2 and Fe6. They have initial H–H distances of 2.24–2.37 Å, irrespectively of

Table 3 Structures and reactions within the E$_4$ state, listing the structure (Struct), the positions of the four added H atoms (H1, H2, H3 and H4), the BS state, the H–H distance (Å), the relative energy (ΔE), the H$_2$ dissociation energy (ΔE$_{H_2}$) relative the E$_0$ state and H$_2$ in a COSMO solvent, and the activation energy for H$_2$ formation (ΔE‡; all energies in kJ mol^{-1}). The two H atoms involved in the reaction are marked in bold face. The lower part of the table shows the products after the H$_2$-formation reactions, labelled after the starting E$_3$ structure and reaction

					TPSS					r²SCAN			TPSSh					B3LYP		
Structure	H1	H2	H3	H4	BS	H–H	ΔE	ΔE$_{H_2}$	ΔE‡	BS	ΔE	ΔE$_{H_2}$	BS	H–H	ΔE	ΔE$_{H_2}$	ΔE‡	BS	ΔE	ΔE$_{H_2}$
3H	S2B(3)	**Fe2/6(3)**	**Fe5**	Fe6	14	4.02	40	−112	75	14	125	−228	147	3.90	112	−207	72	147	333	−330
		Fe2/6(3)		**Fe6**		2.40			25					2.30			15			
	S2B(3)	**Fe2/6(3)**				2.39			59					2.39			57			
3322	**S2B(3)**	Fe2/6(3)	Fe3/7(2)	S5A2(2)	14	2.36	15	−87	70	147	52	−156	147	2.57	39	−134	80	147	292	−289
		Fe3/7(2)		**S5A(2)**		2.55			101					2.48			105			
3323	**S2B(3)**	Fe2/6(3)	Fe3/7(2)	S5A(3)	14	2.37	3	−75	72	135	27	−131	135	2.45	11	−107	81	135	233	−230
3323H6	**S2B(H6S)**	Fe2/6(3)	Fe3/7(2)	S5A(3)	147	2.37	39	−111	76	147	44	−147	147	2.36	49	−145	88	147	313	−311
	Fe2/6(3)	**Fe3/7(2)**				3.24			92					3.09			56			
3522	S2B(3)	**Fe2/6(5)**	**Fe3/7(2)**	S5A(2)	14	2.04	24	−96	75	14	76	−179	14	2.09	48	−143	53	147	266	−264
		Fe3/7(2)		**S5A(2)**		2.48			99					2.51			74			
3523	S2B(3)	**Fe2/6(5)**	**Fe3/7(2)**	S5A(3)	14	2.08	14	−86	80	147	25	−129	135	2.26	18	−113	43	346	192	−189
3523H6	**S2B(H6S)**	**Fe2/6(5)**	Fe3/7(2)	S5A(3)	147	2.36	39	−111	77	147	44	−147	147	2.35	49	−145	88	147	293	−291
		Fe2/6(5)	**Fe3/7(2)**			3.16			91					3.09			56			
5322	**S2B(5)**	Fe2/6(3)	**Fe3/7(2)**	**S5A2(2)**	14	2.53	22	−93	98	147	59	−162	147	2.56	48	−143	102	147	305	−303
	S2B(5)			**Fe3/7(2)**		2.96			81					2.97			72			
5323	**S2B(5)**	Fe2/6(3)	Fe3/7(2)	S5A(3)	14	2.77	11	−83	78	147	45	−148	147	3.09	37	−132	67	147	296	−293
5522	**S2B(5)**	**Fe2/6(5)**	Fe3/7(2)	S5A(2)	14	2.24	41	−112	61	147	59	−162	147	2.32	51	−147	49	147	305	−303
		Fe2/6(5)	**Fe3/7(2)**			2.09			85					2.17			60			
		Fe3/7(2)		**S5A(2)**		2.51			102					2.51			77			
5523	**S2B(5)**	**Fe2/6(5)**	Fe3/7(2)	S5A(3)	14	2.25	30	−101	63	147	43	−146	147	2.32	40	−136	48	147	296	−294
		Fe2/6(5)	**Fe3/7(2)**			2.13			84					2.19			46			
S2	S2B(H2)	Fe2/6(3)	Fe2/6(5)	S5A(3)	147	1.88	54	−126		147	79	−182	147	1.86	56	−152		147	293	−290
S64	S2B(H64)	**Fe2/6(3)**	**Fe2/6(5)**	S5A(3)	147	1.85	1	−72	47	147	29	−132	147	1.87	18	−113	38	147	257	−254
S61	S2B(H61)	**Fe2/6(3)**	**Fe2/6(5)**	S5A(3)	147	1.82	0	−72	42	346	0	−103	157	1.74	0	−95	34	347	229	−226
C3	S2B(3)	C2367	C2456	C3457	234	1.83	192	−263					234	1.83				234	0	
C1	S2B(H6)	Fe2/6	C2367	S5A(3)						346	146	−249	234	1.83	1	−97		346	158	−156

Product states

3H Fe2/6–5	S2B(3)	Fe5			146	0.76	−34	−38		147	−56	−48	147	0.76	−56	−40		147	113	−111
3H Fe2/6–6	Fe5	Fe6	H$_2$ on Fe6		14	0.87	80	−152		147	131	−234	147	0.85	132	−227		147	423	−420
3322	Fe3/7	S5A(2)			147	0.76	−46	−26			−31	−73		0.76	−32	−64			228	−225
3322	S2B(3)	Fe2/6(3)			147	0.76	−28	−43			−44	−59		0.76	−54	−41			183	−181
3323	Fe3/7	S5A(3)			147	0.76	−60	−11			−47	−57		0.76	−44	−51			217	−214
3522	S2B(3)	S5A(2)			147	0.76	21	−93			11	−114		0.76	−11	−84			151	−148
3522	S2B(3)	Fe2/6(5)			147	0.75	−6	−66			−28	−75		0.76	−34	−61			179	−177
3523	S2B(3)	S5A(3)			146	0.75	5	−77			−9	−94		0.76	−24	−71			142	−139
5322	S5A(2)	Fe2/6(3)			235	0.76	41	−112			58	−162		0.75	48	−143			267	−265
5322	S2B(5)	Fe2/6(3)			147	0.76	−16	−56			−30	−73		0.76	−41	−54			198	−196
5323	S5A(3)	Fe2/6(3)			146	0.76	47	−119		147	56	−159	147	0.76	36	−132		147	246	−243
5522	Fe3/7	S5A(2)			147	0.76	1	−73			35	−138		0.75	30	−125			298	−295
5522	S2B(5)	S5A(2)			147	0.76	41	−112			29	−132		0.77	8	−103			187	−185
5522	S2B(5)	Fe2/6(5)			147	0.75	7	−79			−12	−91		0.75	−29	−66			195	−192
5523	Fe3/7	S5A(3)			147	0.76	−10	−61			20	−123		0.75	18	−113			288	−285
5523	S2B(5)	S5A(3)			146	0.76	21	−93			8	−111		0.76	−6	−89		147	131	−129
3323H6	Fe3/7(2)	S5A(3)			147	0.75	−34	−38			−44	−59		0.76	−41	−55			220	−217
3323H6	S2B(5)	S5A(3)			147	0.76	30	−101			17	−121		0.77	−4	−91			178	−176
3523	Fe3/7(2)	S5A(3)			147	0.75	−67	−5			−44	−59		0.76	−41	−55			255	−252
3523H6	S2B(5)	S5A(3)			147	0.76	39	−110			17	−121		0.76	−2	−94			172	−170
S64	S2B(H6M)	S5A(3)	H$_2$ on Fe2		147	0.94	40	−112			12	−116		0.80	12	−107		346	109	−106
S61	S2B(H6S)	S5A(3)	H$_2$ on Fe2		147	0.83	26	−98			9	−113		0.79	−5	−91			84	−82

whether the two H atoms are on the same side as S5A or S3A. The analogous reaction involving the Fe2/6 hydride and a proton on a half-dissociated S2B give a slightly higher barrier of 76–77 kJ mol^{-1} and 88 kJ mol^{-1}, respectively. Two reactions also involve the proton on S2B but the hydride ion bridging Fe3 and Fe7. They have slightly larger initial H–H distances, 2.77–3.09 Å and they give slightly larger barriers of 78–81 kJ mol^{-1} with TPSS, but lower barriers with TPSSh, 67–72 kJ mol^{-1}.

There are four reactions that involve the proton on S5A and the hydride ion bridging Fe3/7. They have initial H–H distances

of 2.48–2.55 Å, but the barriers are appreciably larger, 98–105 kJ mol^{-1} (except two barriers with TPSSh, 74–77 kJ mol^{-1}). Thus, the barriers depend more on the involved H atoms than on the initial H–H distances.

Finally, seven reactions involve two hydride ions on different Fe ions. Four of them involve hydrides on the Fe2/6 and Fe3/7 pairs, having initial H–H distances of 2.04–2.13 Å. They give barriers of 75–85 kJ mol^{-1} with TPSS and 43–60 kJ mol^{-1} with TPSSh. Two involve hydrides in the same positions, but with S2B dissociated from Fe2. This gives longer initial distances
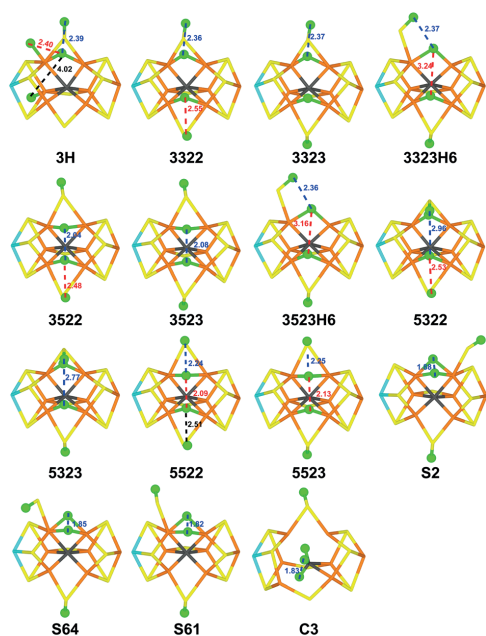
Fig. 5 The 15 considered E_4 structures, showing the structures obtained with TPSS (B3LYP for C3). The labels are explained in the text.

(3.16–3.24 Å) and higher barriers with TPSS (91–92 kJ mol^{-1}), but not with TPSSh (56 kJ mol^{-1}). The seventh reaction involves hydride ions on Fe2/6 and Fe5 in the 3H complex. In this case, the initial H–H distance is much larger, 4.02 Å. This gives a more complicated reaction, in which the Fe2/6 hydride ion first moves to Fe5, *via* a Fe5/6 bridging position (with a barrier of 72–75 kJ mol^{-1}) and then the two hydride ions on the Fe5 ion (with a H–H distance of 2.20 Å) connect with a barrier of 49–58 kJ mol^{-1} relative to the starting structure.

We also tried to connect the closest protons on the central carbide in the C3 structure. However, as for the corresponding E_2 and E_3 states, this was not be possible – the energy rose monotonically to >170 kJ mol^{-1}. On the other hand, it was possible to move the Fe3/7 hydride ion in 3523H6 structure to the central carbide ion. The barrier was only 47 kJ mol^{-1} when studied with B3LYP and the product (with protons also on S2B(H6), Fe2/6 and S5A(5)) was 33 kJ mol^{-1} more stable than the reactant (and therefore the second-best B3LYP structure), although it is still 158 kJ mol^{-1} less stable than the C3 structure. The corresponding reaction with TPSSh was not possible, mainly owing to the low stability of the product, but also to the fact that it was started from the 3523 structure with S2B still bridging and therefore the hydride on Fe2/6 was partly in the way for the movement of the Fe3/7 hydride ion.

In addition, we tried to convert the S61 structure to the 3323 structure. This requires two steps. First, one of the Fe2/6

hydrides needs to move to Fe3/7. This can be done with a barrier of 54–69 kJ mol^{-1}. Then, S2B should bind back to Fe2, which can be done with a minimal barrier of 5–7 kJ mol^{-1}.

As a result of the study of the H_2-formation reactions, we obtained 24 different product structures. Most of them have H_2 in the second coordination sphere of the FeMo cluster. The best structure with TPSS has a proton on S5A and a hydride bridging Fe3 and Fe7. With B3LYP, the best structure has two protons on S2B and S5A. With r^{2}SCAN and TPSSh, a structure with a proton on S2B and a terminal hydride ion on Fe5 is best. This is quite different from what was observed for the E_2 state without H_2, indicating that the second-sphere H_2 molecule has a significant influence on the relative energies and reflecting that we have not performed any systematic investigation of this type of complexes. This is also confirmed by the fact that none of the structures is more stable than the fully dissociated $E_2 + H_2$ state (ΔE_{H_2} is -5 to -111 kJ mol^{-1} for the best structures with the four DFT methods). In four of the product structures, H_2 coordinates side-on to Fe2 or Fe6. These structures have appreciably larger H–H bond lengths than the other structures (0.80–0.94 Å, compared to 0.75–0.76 Å). The best structure has H_2 bound to Fe2, S2B protonated and bound only to Fe6, and S5A protonated. However, this structure is 24–93 kJ mol^{-1} less stable than be best structure with H_2 in the second sphere.

Conclusions

We have studied the formation of H_2 in the E_2–E_4 states of Mo-nitrogenase. Even if there are quite some variations between the individual structures and the DFT functionals employed, there are some general trends. The lowest barrier for H_2 formation is obtained when combining two hydride ions, one terminal on Fe6 and one bridging Fe2/6, 15–25 kJ mol^{-1}. Combining two hydride ions both bridging Fe2/6 also give low barriers, 29–57 kJ mol^{-1}. Combining hydride ions bridging Fe2/6 and Fe3/7, give higher barriers with TPSS, 75–85 kJ mol^{-1}, except in the E_3 state (39 kJ mol^{-1}). With TPSSh, those barriers are lower, 34–60 kJ mol^{-1}. The reaction between a proton and a hydride ion gives similar barriers. With TPSS, combining a proton on S2B and a hydride ion bridging Fe2/6 typically gives lower barriers than combining two hydride ions on Fe2/6 and Fe3/7 or a proton on S5A and a hydride ion bridging Fe3/7. With TPSSh, barriers for two hydride ions are lower and there are smaller differences between the two proton–hydride reactions. Instead, there is some difference between whether the S2B proton and the Fe2/6 hydride ion are on the same side as S5A (lower barrier) or as S3A (higher barrier). For both functionals, it seems that the proton–hydride reactions give a lower barrier as the number of added electrons and protons increase (*i.e.* $E_2 > E_3 > E_4$). In general, TPSSh gives lower barriers than TPSS (by 9 kJ mol^{-1} on average), but for proton–hydride reactions, the opposite is often true.

We have also studied some other reactions, moving around the H atoms within the FeMo cluster. In agreement with previous studies of proton transfers in the FeMo cluster during the later (E_4–E_8) steps of the reaction mechanism,[60] we find

that individual barriers are rather small, 5–69 kJ mol^{-1}. The calculations show that a proton on S2B can rotate rather freely, either moving from the S2B(3) to the S2B(5) side (55–62 kJ mol^{-1}) when S2B is bridging Fe2/6 or rotating around the Fe6–S2B (12–17 kJ mol^{-1}) axis when it is dissociated from Fe6. The dissociation of S2B from Fe2 also has a low barrier (35–39 kJ mol^{-1}). In fact, it is also possible to move a Fe3/7 hydride ion to S5A (26–53 kJ mol^{-1}) or to move a bridging hydride ion from Fe2/6 to Fe3/7 (54–69 kJ mol^{-1}). A hydride ion can move to the central carbide with B3LYP (47 kJ mol^{-1}), but only in the E$_4$ state and if the starting structure is appropriate. This shows that the various isomers of the E$_n$ states can interconvert rather freely. Therefore, it is unlikely that the FeMo cluster may avoid unwanted side reactions by placing the H atoms on different faces of the FeMo cluster. Likewise, it seems impossible to avoid the most stable structures for each E$_n$ state, *i.e.* to stabilise metastable states by kinetic pathways and barriers.

Finally, we discuss the implications of the current calculations on the reaction mechanism of nitrogenase. From a functional point of view, it is clear that the enzyme should avoid the formation of H$_2$ before N$_2$ binds, *i.e.* at least for the E$_2$ and E$_3$ states. For the E$_2$ state, B33 is the most stable structure with TPSS. Formation of H$_2$ from this state has an activation barrier of 86–91 kJ mol^{-1}. This corresponds to a rate of 0.0008–0.0072 s^{-1} (3–26 h^{-1}; using classical transition-state theory with a pre-exponential factor of 6.2×10^{12} s^{-1}), which is smaller than the turnover rate of the enzyme 1–5 s^{-1}.[2,85] Thus, the protein should be able to avoid this side reaction if the electron flow is proper, but the rate constant is somewhat smaller than what has been estimated from kinetic measurements, 0.2 s^{-1},[85] corresponding to 77 kJ mol^{-1}. For the B55 structure, the activation energy for H$_2$ formation is lower, 51–55 kJ mol^{-1}. However, this structure is 33–42 kJ mol^{-1} less stable than the best state with TPSS or TPSSh, giving net barriers of 84–97 kJ mol^{-1}, corresponding to rates of 0.2–65 h^{-1}, still much lower than the turnover rate. For the H6 structure, which is the most stable structure with TPSSh and r^{2}SCAN, the activation barrier for H$_2$ formation is quite high, 79–104 kJ mol^{-1}. This corresponds to rates of 0.13–5 $\times 10^{-6}$ s^{-1} (0.02–460 h^{-1}), which are less than the turnover rate of the enzyme. For the best structure with B3LYP, C2, no H$_2$ formation could be obtained. Thus, we can conclude that H$_2$ formation does not seem to be any serious problem at the E$_2$ level of the protein.

Our calculated activation barrier for the H6 structure of E$_2$ is similar to that reported by Thorhallsson and Bjornsson, 86 kJ mol^{-1}, with TPSSh.[33] On the other hand, our barrier for the B55 structure (51 kJ mol^{-1}) is somewhat larger than that reported by Khadka and coworkers, 29 kJ mol^{-1}.[52] This might be caused by the difference in DFT methods (TPSS and BP86) but pure GGA functionals often give similar results for nitrogenase models.[18] It is more likely that the reason for the difference is that Khadka and coworkers used QM-cluster calculations, whereas we have performed QM/MM calculations. In fact, Thorhallsson and Bjornsson reported much lower activation energies with cluster models than QM/MM,[33] indicating that the surrounding enzyme disfavours the H$_2$ formation.

For the E$_3$ state, the results are more problematic. The S6 structure, which is preferred with the TPSS, r^{2}SCAN and TPSSh functionals, has two hydride ions both bridging Fe2/6, and the barrier of H$_2$ formation from these two hydrides is small 29–57 kJ mol^{-1}, corresponding to 7×10^2 to 5×10^7 s^{-1}, much faster than the turnover of the enzyme. We see no way for the enzyme to avoid this problematic side reaction with the current results. Either the TPSS, r^{2}SCAN and TPSSh methods give gravely unreliable results or we have not yet found the proper lowest-energy structure or BS state for E$_3$. Alternatively, the B3LYP results are trusted, which indicates that C3 structure is most stable. In our calculations, no H$_2$ formation is observed from this structure. On the other hand, B3LYP gives very weak N$_2$ binding in the E$_4$,[32] which has led Siegbahn to suggest that four additional reductions are needed before N$_2$ may bind,[46,83] something that does not find any experimental support.[13]

For the E$_4$ state, the situation is similar: The best structure with TPSS, r^{2}SCAN and TPSSh is S61, with two hydride ions both bridging Fe2/6. These can be connected with a barrier of only 34–47 kJ mol^{-1}, corresponding to rates of 5×10^4–7×10^6 s^{-1}, *i.e.* much faster than the turnover rate of the enzyme. However, with TPSS, the Hoffman-type structures are competitive in stability, in particular 3323, which is only 3 kJ mol^{-1} less stable than S61. From this structure, H$_2$ formation has a higher barrier of 72–81 kJ mol^{-1}, *i.e.* similar to the turnover rate of the enzyme. There are other Hoffman-type of complexes with lower barriers (down to 61 kJ mol^{-1} with TPSS and 46 kJ mol^{-1} with TPSSh), but they are too unstable compared to the best state so that the net barrier (counted from the best structure) becomes prohibitively high. Finally, the C3 structure is best with B3LYP and competitive with TPSSh, but no H$_2$ formation could be obtained from this structure. In conclusion, H$_2$ formation from the best E$_4$ structures is possible. In this case, it is unclear whether this reaction should take place before or after binding of N$_2$, *i.e.* if it should be avoided or not. The results are in accordance with a recent steady-state kinetic model of nitrogenase, indicating that the rate of H$_2$ formation is much higher from the E$_4$ state than from E$_2$.[85] In future studies, we will study the binding of N$_2$ to E$_4$ structures with H$_2$ and whether the binding of N$_2$ can be enhanced by the concerted formation or dissociation of H$_2$.

Author contributions

HJ performed all calculations and contributed to the analysis and the writing of the manuscript. UR directed the research, analysed the results and wrote the manuscript.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

References

1 B. K. Burgess and D. J. Lowe, *Chem. Rev.*, 1996, **96**, 2983.
2 B. Schmid *et al.*, *Handbook of Metalloproteins*, John Wiley & Sons, Ltd, 2006, p. 1025.
3 B. M. Hoffman, *et al.*, *Chem. Rev.*, 2014, **114**, 4041.
4 L. C. Seefeldt, *et al.*, *Chem. Rev.*, 2020, **120**, 5082.
5 T. Spatzal, *et al.*, *Science*, 2011, **334**, 940.
6 J. Kim and D. C. Rees, *Science*, 1992, **257**, 1677.
7 O. Einsle, *et al.*, *Science*, 2002, **297**, 1696.
8 T. Spatzal, *et al.*, *Science*, 2014, **345**, 1620.
9 O. Einsle, *J. Biol. Inorg. Chem.*, 2014, **19**, 737.
10 R. N. F. Thorneley, and D. J. Lowe, in *Molybdenum Enzymes*, ed. T. G. Spiro, Wiley, New York, 1985, p. 221.
11 Z. Y. Yang, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 16327.
12 D. Lukoyanov, *et al.*, *J. Am. Chem. Soc.*, 2016, **138**, 10674.
13 V. Hoeke, *et al.*, *J. Am. Chem. Soc.*, 2019, **141**, 11984.
14 H. Yang, *et al.*, *Inorg. Chem.*, 2018, **57**, 12323.
15 R. Y. Igarashi, *et al.*, *J. Am. Chem. Soc.*, 2005, **127**, 6231.
16 C. Van Stappen, *et al.*, *Chem. Rev.*, 2020, **120**, 5005.
17 I. Dance, *ChemBioChem*, 2020, **21**, 1671.
18 L. Cao and U. Ryde, *Phys. Chem. Chem. Phys.*, 2019, **21**, 2480.
19 H. Jiang, O. K. G. Svensson and U. Ryde, *Inorg. Chem.*, 2022, **61**, 18067.
20 L. Cao, O. Caldararu and U. Ryde, *J. Chem. Theory Comput.*, 2018, **14**, 6653.
21 T. Spatzal, *et al.*, *Nat. Commun.*, 2016, **7**, 10902.
22 B. Benediktsson and R. Bjornsson, *Inorg. Chem.*, 2017, **56**, 13417.
23 R. Bjornsson, F. Neese and S. DeBeer, *Inorg. Chem.*, 2017, **56**, 1470.
24 S. J. Yoo, *et al.*, *J. Am. Chem. Soc.*, 2000, **122**, 4926.
25 C. Van Stappen, *et al.*, *Inorg. Chem.*, 2019, **58**, 12365.
26 C. Van Stappen, *et al.*, *Chem. Sci.*, 2019, **10**, 9807.
27 D. A. Lukoyanov, *et al.*, *Inorg. Chem.*, 2022, **61**, 5459.
28 H. Jiang, K. J. M. Lundgren and U. Ryde, *Inorg. Chem.*, 2023, **62**(48), 19433–19445.
29 J. Tao, *et al.*, *Phys. Rev. Lett.*, 2003, **91**, 146401.
30 J. W. Furness, *et al.*, *J. Phys. Chem. Lett.*, 2020, **11**, 8208.
31 V. N. Staroverov, *et al.*, *J. Chem. Phys.*, 2003, **119**, 12129.
32 H. Jiang and U. Ryde, *Dalton Trans.*, 2023, **52**, 9104.
33 A. T. Thorhallsson and R. Bjornsson, *Chem. – Eur. J.*, 2021, **27**, 16788.
34 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.
35 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.
36 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.
37 Y. Pang and R. Bjornsson, *Phys. Chem. Chem. Phys.*, 2023, **25**, 21020.
38 S. Raugei, L. C. Seefeldt and B. M. Hoffman, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 10521.
39 L. Cao and U. Ryde, *J. Chem. Theory Comput.*, 2020, **16**, 1936.
40 A. T. Thorhallsson, B. Benediktsson and R. Bjornsson, *Chem. Sci.*, 2019, **10**, 11110.
41 P. E. M. Siegbahn, *J. Am. Chem. Soc.*, 2016, **138**, 10485.
42 P. E. M. Siegbahn, *J. Comput. Chem.*, 2018, **39**, 743.
43 J. B. Varley, *et al.*, *Phys. Chem. Chem. Phys.*, 2015, **17**, 29541.
44 M. Rohde, *et al.*, *Biochemistry*, 2018, **57**, 5497.
45 W.-L. Li, *et al.*, *Chem. Catal.*, 2023, **3**(7), 100662.
46 P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.*, 2023, **25**, 23602.
47 D. Sippel, *et al.*, *Science*, 2018, **359**, 1484.
48 W. Kang, *et al.*, *Science*, 2020, **368**, 1381.
49 I. Dance, *J. Am. Chem. Soc.*, 2007, **129**, 1076.
50 I. Dance, *Z. Anorg. Allg. Chem.*, 2015, **641**, 91.
51 I. Dance, *Dalton Trans.*, 2021, **50**, 18212.
52 N. Khadka, *et al.*, *J. Am. Chem. Soc.*, 2017, **139**, 13518.
53 L. Cao, O. Caldararu and U. Ryde, *J. Phys. Chem. B*, 2017, **121**, 8242.
54 B. M. Barney, *et al.*, *Biochemistry*, 2007, **46**, 6784.
55 J. A. Maier, *et al.*, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
56 W. L. Jorgensen, *et al.*, *J. Chem. Phys.*, 1983, **79**, 926.
57 L. Hu and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 2452.
58 C. I. Bayly, *et al.*, *J. Phys. Chem.*, 1993, **97**, 10269.
59 F. Furche, *et al.*, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 91.
60 H. Jiang, *et al.*, *Angew. Chem., Int. Ed.*, 2022, **61**, e202208544.
61 B. Benediktsson and R. Bjornsson, *J. Chem. Theory Comput.*, 2022, **18**, 1437.
62 A. Schäfer, H. Horn and R. Ahlrichs, *J. Chem. Phys.*, 1992, **97**, 2571.
63 K. Eichkorn, *et al.*, *Chem. Phys. Lett.*, 1995, **240**, 283.
64 K. Eichkorn, *et al.*, *Theor. Chem. Acc.*, 1997, **97**, 119.
65 E. Caldeweyher, *et al.*, *J. Chem. Phys.*, 2019, **150**, 154122.
66 R. Bjornsson, *et al.*, *Chem. Sci.*, 2014, **5**, 3096.
67 D. J. Lowe, R. R. Eady and R. N. F. Thorneley, *Biochem. J.*, 1978, **173**, 277.
68 D. Lukoyanov, *et al.*, *Inorg. Chem.*, 2014, **53**, 3688.
69 D. A. Lukoyanov, *et al.*, *Inorg. Chem.*, 2018, **57**, 6847.
70 T. Lovell, *et al.*, *J. Am. Chem. Soc.*, 2001, **123**, 12392.
71 L. Cao and U. Ryde, *Int. J. Quantum Chem.*, 2018, **118**, e25627 (16 pages).
72 C. Greco, *et al.*, *Int. J. Quantum Chem.*, 2011, **111**, 3949.
73 R. K. Szilagyi and M. A. Winslow, *J. Comput. Chem.*, 2006, **27**, 1385.
74 L. Cao and U. Ryde, *J. Biol. Inorg. Chem.*, 2020, **25**, 521.
75 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799.
76 A. Schäfer, *et al.*, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2187.
77 A. Klamt, *et al.*, *J. Chem. Phys. A*, 1998, **102**, 5074.
78 U. Ryde, *J. Comput.-Aided Mol. Des.*, 1996, **10**, 153.

79 U. Ryde and M. H. M. Olsson, *Int. J. Quantum Chem.*, 2001, **81**, 335.

80 N. Reuter, *et al.*, *J. Phys. Chem. A*, 2000, **104**, 1720.

81 L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput.*, 2011, 7, 761.

82 L. Cao and U. Ryde, *Front. Chem.*, 2018, **6**, 89.

83 P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.*, 2019, **21**, 15747.

84 W.-J. Wei and P. E. M. Siegbahn, *Chem. – Eur. J.*, 2022, **28**, e202103745.

85 D. F. Harris, A. Badalyan and L. C. Seefeldt, *Biochemistry*, 2022, **61**, 2131.

Paper II

Check for updates

N2 binding to the E0–E4 states of nitrogenase†

Hao Jiang [ID] and Ulf Ryde [ID] *

Nitrogenase is the only enzyme that can convert N_2 into NH_3. The reaction requires the addition of eight electrons and protons to the enzyme and the mechanism is normally described by nine states, E_0–E_8, differing in the number of added electrons. Experimentally, it is known that three or four electrons need to be added before the enzyme can bind N_2. We have used combined quantum mechanical and molecular mechanics methods to study the binding of N_2 to the E_0–E_4 states of nitrogenase, using four different density functional theory (DFT) methods. We test many different structures for the E_2–E_4 states and study binding both to the Fe2 and Fe6 ions of the active-site FeMo cluster. Unfortunately, the results depend quite strongly on the DFT methods. The TPSS method gives the strongest bonding and prefers N_2 binding to Fe6. It is the only method that reproduces the experimental observation of unfavourable binding to the E_0–E_2 states and favourable binding to E_3 and E_4. The other three methods give weaker binding, preferably to Fe2. B3LYP strongly favours structures with the central carbide ion triply protonated. The other three methods suggest that states with the S2B ligand dissociated from either Fe2 or Fe6 are competitive for the E_2–E_4 states. Moreover, such structures with two hydride ions both bridging Fe2 and Fe6 are the best models for E_4 and also for the N_2-bound E_3 and E_4 states. However, for E_4, other structures are often close in energy, *e.g.* structures with one of the hydride ions bridging instead Fe3 and Fe7. Finally, we find no support for the suggestion that reductive elimination of H_2 from the two bridging hydride ions in the E_4 state would enhance the binding of N_2.

Introduction

Nitrogenase (EC 1.18/19.6.1) is the only enzyme that can cleave the triple bond in molecular N_2,[1–4] thereby making nitrogen available for biological lifeforms. X-ray crystallographic studies have shown that nitrogenase contains a complicated MoFe$_7$S$_9$C (homocitrate) cluster in the active site, called the FeMo cluster (Fig. 1).[5–9] Alternative nitrogenases exist, in which the Mo ion is replaced by V or Fe, but they have lower activities.[10,11] The nitrogenase reaction is demanding, requiring eight electrons and 16 ATP molecules for each N_2 molecule processed:[3,4]

$$N_2 + 8e^- + 8H^+ + 16\ ATP \rightarrow 2\ NH_3 + H_2 + 16\ ADP + 16\ P_i \quad (1)$$
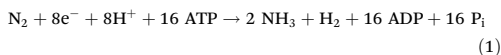
Nitrogenase has been extensively studied by spectroscopic, biochemical and kinetic methods.[1–9,12] The reaction is normally described by the Lowe–Thorneley cycle,[13] which involves nine intermediates, E_0–E_8, differing in the number of added electrons and protons. The E_0 resting state has been thoroughly characterised by crystallography, spectroscopic and compu-

tational studies.[4,7,14–16] The E_1 state has been studied by X-ray absorption and Mössbauer spectroscopy[17,18] and most likely contains a proton on the S2B μ_2 bridging sulfide ion (see Fig. 1b for atom names).[19] The E_2 state is known to involve two conformers, of which at least one contains an iron-bound hydride ion.[20–23] The E_4 state has been characterised by EPR and ENDOR spectroscopy, and has been shown to contain two hydride ions that bridge between two Fe ions of the FeMo cluster.[24–26] It has been shown that N_2 binds to the E_3 and E_4 states, but not the E_0–E_2 states.[1,25,27–30] In connection with the binding of N_2, H_2 is released by reductive elimination, *i.e.* by the formation of H_2 from two hydride ions.[25,26,31,32] Then, N_2 is successively reduced and protonated to two molecules of NH_3. Mutational studies have indicated that the Fe2–Fe3–Fe6–Fe7 face of the FeMo cluster is the primary site for N_2 reduction and that Fe2 or Fe6 are the most likely binding sites of N_2.[33,34]

Nitrogenase has also been extensively studied by computational methods, using density functional theory (DFT).[35] However, such studies are complicated by the fact that there are very many possibilities for the binding of up to four protons to the cluster, that different DFT method give widely different relative energies of the various protonation states and that the electronic structure is complicated (there are 35 possible broken-symmetry states).[35–38]

Several DFT studies have been devoted to the binding of N_2 to the FeMo cluster in different E_n states.[35] Early investigations

Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden. E-mail: Ulf.Ryde@teokem.lu.se; Fax: +46-46 2228648; Tel: +46-46 2224502
†Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3dt00648d
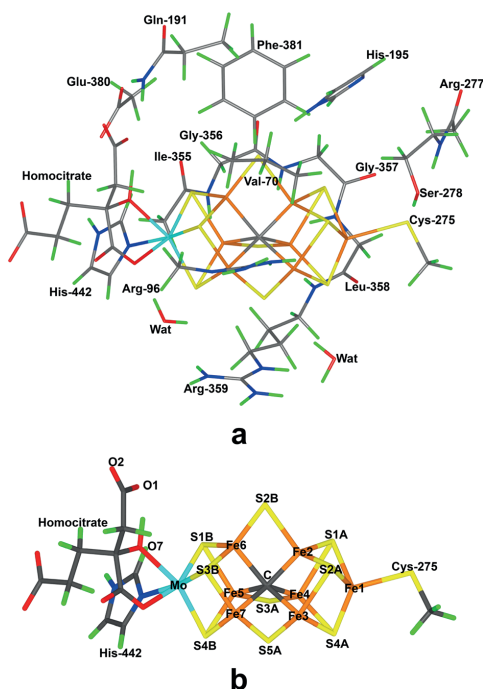
Fig. 1 Structure of the FeMo cluster in the E_0 state. (a) Illustrates the QM system used in all calculations, as well as the names of the nearby residues; (b) shows the FeMo cluster with atom names indicated. H, C, N, O, S, Fe and Mo atoms are shown in green, grey, blue, red, yellow, orange and cyan, respectively. All figures show the same orientation and colouring scheme.

suffered from incomplete knowledge of the composition of the cluster, its net charge and the sequence of proton and N_2 binding.[39–42] However, also later studies have led to disparate suggestions. Blöchl, Kästner and coworkers suggested that N_2 binds to Fe7 after dissociation of S5A from this ion (atom names are shown in Fig. 1).[43,44]

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 Fe6,[45–47] and crystallographic studies have indicated that S2B may sometimes be replaced by other ligands,[8,48,49] indicating that sulfide lability may be mechanistically relevant.[50,51] In particular, Bjornsson and coworkers showed that N_2 may bind to Fe2 or Fe6 in E_4 with favourable binding energies of 56 or 43 kJ mol^{-1}, respectively.[52] They did not find any binding of N_2 to the E_0 and E_1 states, and a less favourable binding (29 kJ mol^{-1}) to the E_2 state. Recently, they have published a more thorough study, reporting unfavourable N_2-binding energies to E_0, E_1 and E_2 (by 69, 41 and 8 kJ mol^{-1}), but a slightly favour-

able binding energy to E_4, 17 kJ mol^{-1}.[53] They emphasize the importance of two doubly occupied 3d orbitals on the Fe ion binding N_2, which can donate electron density into the N_2 π^* orbitals.

Hoffman and coworkers have suggested that reductive elimination of H_2 from the E_4 state of nitrogenase is necessary for the binding of N_2.[54,55] Based on ENDOR experiments and DFT calculations, they suggested a structure of the E_4 state with two protons on S2B and S5A (both remaining bound to both Fe ions) and two hydride ions bridging Fe2/6 and Fe3/7, all located on the same face of the FeMo cluster. In such a structure, H_2 may form from the two hydride ions and N_2 can bind to this state with the concurrent release of H_2. DFT calculations suggested a favourable binding free energy of 13 kJ mol^{-1} and metadynamics simulations indicated that formation of H_2 is endergonic by 20 kJ mol^{-1}, with a barrier of 49 kJ mol^{-1} from the E_4 ground state.

Dance has presented several studies of N_2 binding to nitrogenase.[56–58] He showed that side-on binding of N_2 is less favourable than end-on binding and that N_2 bridging between two Fe ions is unfavourable. The early studies suggested preferable binding to Fe6. However, recently he suggested that first a promotional, but unreactive, N_2 binds to Fe2 in the exo-position and then a second reactive N_2 binds in the endo-position of Fe6.[59] He reported favourable binding energies of up to 38 kJ mol^{-1}. The binding is somewhat enhanced in structures with a bound H_2 molecule (up to 59 kJ mol^{-1}). In all structures, S2B remains bound to both Fe2 and Fe6.

On the other hand, Siegbahn argued that N_2 binding to the E_0–E_4 states is endergonic.[60,61] Therefore, he suggested that nitrogenase needs to be reduced by four additional electrons before it can bind N_2, i.e. that the E_0–E_3 states are outside the catalytic cycle and the E_4 state becomes the E_0 state in his catalytic cycle. Thereby, the cluster reaches a state with two Fe(ɪ) ions and five Fe(ɪɪ), which enhances binding of N_2. In his first study, N_2 was suggested to bind bridging between Fe4 and Fe6 in a reaction that is endergonic by 13 kJ mol^{-1}, but in a later study, he suggested that S2B dissociates from the cluster and N_2 binds to Fe4 with a slightly less endergonic free energy of 10 kJ mol^{-1}. The binding energy strongly depends on the amount of Hartree–Fock exchange in the functional.

Apparently, there is no consensus in how N_2 binds to the FeMo cluster and this is partly caused by the disagreement regarding the structure of the E_4 state and the large differences in the structures and energies obtained with different DFT functionals. Therefore, we here study the binding of N_2 to nitrogenase with four different DFT methods. We study the binding of N_2 to the five E_0–E_4 states and see how well different DFT functionals reproduce the experimental observation the N_2 binds only to the E_3 and E_4 states.[1,25,28–30] For the E_0–E_2 states, there is reasonable consensus regarding the preferred protonation states.[18,19,35,37,62,63] For the E_3 and E_4 states, we enhance previous studies of the preferred protonation state,[35,37,38,52,64–67] in particular with structures where S2B has dissociated from either Fe2 or Fe6.

Methods

The protein

The calculations are based on the 1.0 Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[7] The setup of the protein is identical to that of our previous studies.[38,66,68,69] The entire heterotetramer was considered in the calculations and the quantum mechanical (QM) calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule from the solvent rather close to the active site (~11 Å) in the A subunit. The two P clusters and the FeMo cluster in subunit A were modelled by MM in the fully reduced and resting states, respectively, using a QM charge model.[68] The protonation states of all residues were the same as before,[68] and the homocitrate ligand was modelled in the singly protonated state with a proton shared between the hydroxyl group (O7 that coordinates to Mo) and the O1 carboxylate atom.[16,68] The protein was solvated in a sphere with a radius of 65 Å around the geometrical centre of the protein. Cl^- and Na^+ ions were added to an ionic strength of 0.2 M.[70] The final system contained 133 915 atoms. For the protein, we used the Amber ff14SB force field[71] and water molecules were described by the TIP3P model.[72] The metal sites[68,73] were treated by a non-bonded model[74] and charges were obtained with the restrained electrostatic potential method.[75]

QM calculations

All QM calculations were performed with the Turbomole software (versions 7.5 and 7.6).[76] All structures were studied with the TPSS,[77] r^2SCAN,[78] TPSSh[79] and B3LYP[80–82] functionals. The former two are meta generalised gradient approximation (GGA) functionals, whereas the other two are hybrid functionals with 10 and 20% Hartree–Fock exchange, respectively. r^2SCAN and TPSSh have been shown to give very accurate structures of nitrogenase models.[83] We employed the def2-SV(P) basis set.[84] Previous studies have shown that increasing the basis set to def2-TZVPD changes the relative energies by up to 11–20 kJ mol^{-1}.[37,66,69,73] Test calculations for the best structures in this study, shown in Table S1 in the ESI,† confirm that this is also the case for the current structures (mean signed and unsigned changes of 2 and 6 kJ mol^{-1}), except in a few structures, in which the electronic structure changes extensively (Table S2†). However, for N_2-binding energies, the larger basis set gives more unfavourable binding energies by 9–21 kJ mol^{-1} (average 15 kJ mol^{-1}; *cf.* Table S3†), probably reflecting that the binding with the smaller basis set is enhanced by the basis-set superposition error. The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[85,86] Empirical dispersion corrections were included with the DFT-D4 approach,[87] as implemented in Turbomole.

The FeMo cluster was modelled by $MoFe_7S_9C$(homocitrate) (CH_3S)(imidazole), where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included in the QM

model, *viz.* Arg-96, Gln-191 and His-195 (sidechains), Ser-278 (both sidechain and backbone, including some atoms from Arg-359), Gly-356, Gly-357 and Leu-358 (backbone), as well as two water molecules. Finally, the sidechain of Glu-380 was included because it forms hydrogen bonds to Gln191 and His-442, as well as the sidechains of Val-70 and Phe-381 because they are close to S2B, Fe2 and Fe6, *i.e.* the expected binding site of N_2. The QM system involved 189–195 atoms (depending on the E_n state and whether N_2 was included or not) and is shown in Fig. 1a. The net charge of QM region was always $-4e$. His-195 was always neutral and protonated on the NE2 atom, because this state has been found to be most stable in our previous studies.[37,68,88]

In this investigation we study the E_0–E_4 states of the FeMo cluster with or without N_2. The resting E_0 state has the formal $Mo^{III}Fe_3^{II}Fe_4^{III}$ oxidation state[14,16,89] and is a quartet state according to EPR experiments.[1–4] The other four states were obtained by successively adding one electron and one proton to the previous state. Several positions of the added protons were tested, based on previous investigations,[18,19,35,37,38,52,62–67] as will be discussed below. E_0 and E_2 were studied in the quartet spin state and E_4 in the doublet state, in agreement with experiments.[1–4,22,23,90] For E_1 and E_3, no experimental data are available and we assumed the quintet and triplet states (previous studies have shown that different spin states are close in energy).[37,63]

The electronic structure of all QM calculations was obtained with the broken-symmetry (BS) approach:[36] each of the seven Fe ions was modelled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways.[69] The various BS states were obtained either by swapping the coordinates of the Fe ions[91] or with the fragment approach by Szilagyi and Winslow.[92] The various BS states are named by listing the number in the Noodleman nomenclature (BS1–10),[36] followed by the numbers of the three Fe ions with minority spin. The selection of the BS states was based on our previous experience with the similar systems.[37,63,66] For E_0–E_2, we tested mainly the BS7-235 state. The E_3 structures were studied mainly in the BS10-147 state. For E_4, an initial investigation was performed in the BS10-147 and BS-14 states. For the best four (without N_2) or six (with N_2) structures, we tested eight additional BS states (BS7-235, BS7-247, BS7-346, BS2-234, BS6-157, BS8-347, BS10-127 and BS10-135) with all four functionals.

We study the binding of N_2 to nitrogenase. We will discuss three types of structures, *viz.* without any N_2 molecule (E_n), with N_2 bound directly to either Fe2 or Fe6 (denoted E_n–N_2), *i.e.* in the first coordination sphere with a Fe–N distance of typically 1.8–2.0 Å, but occasionally longer (especially with B3LYP), up to 2.5 Å (but with the N_2 molecule directed towards the Fe ion), or with N_2 unbound, but residing in the second coordination sphere of the Fe ion (denoted E_n + N_2). Naturally, the latter structures are less well defined, but stable structures were typically found with a Fe–N distance of ~3.7 Å from Fe6 and ~2.8 Å from Fe2. The N_2 molecule is no longer directed towards the Fe ion but forms weak dispersive interactions with

the surrounding residues. If no local minimum was found for either E_n–N_2 or E_n + N_2 (i.e. if the geometry optimisation converged to the other type), we obtained a structure with the Fe–N distance restrained to 1.9, 3.7 or 2.8 Å, for E_n–N, or second-sphere bonding to Fe6 and Fe2, respectively. Based on previous mutational and computational studies,[1–4,56–58] we study only binding to Fe2 or Fe6, and only end-on binding in the exo position (i.e. trans to the central carbide ion).

There are several ways to calculate N_2 binding energies. We use three different definitions in this article. First, we define the N_2 binding energy, ΔE_{N_2}, as the QM/MM energy difference between the N_2-bound structure and the best (i.e. the structure with the lowest QM/MM energy) optimised structure without N_2 at this E_n level (denoted E_n(best)), and a free N_2 molecule optimised in a conductor-like screening model (COSMO)[93,94] continuum solvent with a dielectric constant of 80, the default optimised COSMO atomic radius for N (1.83 Å) and a water solvent radius of 1.3 Å:[95]

$$\Delta E_{N_2} = E^{QM/MM}(E_n-N_2) - E^{QM/MM}(E_n(\text{best})) - E^{COSMO}(N_2)$$
(2)

It should be noted that the COSMO solvation free energy of N_2 is only 2 kJ mol^{-1}, so it does not matter much whether it is calculated in vacuum or in the continuum solvent. This seems to be the definition used by Siegbahn[60,61] and by Bjornsson and coworkers in their latest study.[53]

Second, we define the direct N_2 binding energy (ΔE_{db}) as the difference in energy between the same type of structure (i.e. the same E_n and protonation state; denoted E_n + N_2(same)) with N_2 in the second or first coordination sphere:

$$\Delta E_{db} = E^{QM/MM}(E_n-N_2) - E^{QM/MM}(E_n + N_2(\text{same}))$$
(3)

This is the definition used by Dance.[59] We have also followed such binding reactions by starting from the second-sphere structure and adding a restraint on the Fe–N distance, which is successively decreased to a typical bonding distance (~1.9 Å) and finally removing the restraints. The resulting potential-energy surfaces also give approximate activation energies for the binding, which are reported.

A third way to define binding energies, intermediate between the other two, is to use the same type of complex without N_2 bound (i.e. the same E_n and protonation state; denoted E_n(same)) and free N_2 as the reference:

$$\Delta E_{bn} = E^{QM/MM}(E_n-N_2) - E^{QM/MM}(E_n(\text{same})) - E^{COSMO}(N_2)$$
(4)

This seems to be the definition used by Bjornsson and coworkers in their first study[52] and called single-step N_2 binding energy in their second study.[53] In all three cases, a negative binding energy indicates a favourable binding.

QM/MM calculations

QM/MM calculations were performed with the ComQum software.[96,97] In this approach, the protein and solvent are split into two subsystems: system 1 (the QM region) was

relaxed by QM methods, whereas system 2 contained the remaining part of the protein and the solvent, and it was kept fixed at the original coordinates (equilibrated crystal structure, to avoid the risk that different calculations end up in different local minima).

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: the QM system was capped with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[96,98] All atoms were included in the point-charge model, except the CL atoms.[99] ComQum employs a subtractive scheme with van der Waals link-atom corrections.[100] No cutoff is used for any of the QM or MM interactions. The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J mol^{-1} (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u. Approximate transition states for the binding N_2 were obtained by first optimising free N_2 at a distance of 2.5–4 Å from Fe2 or Fe6 and then performing a relaxed scan of Fe–N distances until a bound state was found.

Results and discussion

We have studied the binding of N_2 to the FeMo cluster in nitrogenase. We will discuss the results for different E_n states in separate sections.

N_2 binding to the E_0 and E_1 states

We first studied the binding of N_2 to the resting E_0 state of nitrogenase (using BS7-235;[69] shown in Fig. 1b). As expected, no N_2-bound state was found with any of the four DFT methods. Bonded structures could be obtained by restraining the Fe–N distance of 1.90 Å. However, the binding energy for such restrained structures is unfavourable, more for the binding to Fe2 than to Fe6, e.g. ΔE_{N_2} = 44 and 34 kJ mol^{-1} for TPSS (cf. Table 1). With the other three functionals, the energies are slightly more unfavourable, 44–61 kJ mol^{-1} for Fe6 and 49–69 kJ mol^{-1} for Fe2, with the trend B3LYP < TPSSh < r^{2}SCAN.

For the E_1 state, we added the proton to S2B in agreement with previous QM/MM[35,37] and experimental studies.[18,19] We assumed that the proton points towards S3A and that the FeMo cluster remains in the BS7-235 state.[35,37,69]

In this case, a state with N_2 bound end-on to both Fe2 and Fe6 could be found with TPSS (Fig. 2). They have both a Fe–N distance of 1.92 Å (cf. Table 1). However, the ΔE_{N_2} binding energies are still unfavourable, by 26 and 33 kJ mol^{-1} for Fe6 and Fe2, respectively. A state with N_2 in the second coordination sphere of Fe6 (with a Fe6–N distance of 3.68 Å, Fig. 2)

Table 1 Structures of the E_0 and E_1 states with N_2 in the second coordination sphere ($+N_2$) or bound ($-N_2$) to either Fe2 or Fe6. For each structure and each of the four DFT methods, the Fe–N bond length (in Å; bold face indicates a restrained distance), the relative energy (ΔE in kJ mol^{-1}, within the same column and section) and the ΔE_{N_2} binding energy (eqn (2) in kJ mol^{-1}). Fe–N$_{TS}$ and ΔE_{TS} are the bond length and activation energy for the transition state for the binding of N_2. All structures were studied in the BS7-235 state. All E_1 structures were protonated on S2B, with the proton directed towards S3A

| E_n | Struct. | TPSS | | | | | r²SCAN | | | TPSSh | | | B3LYP | | |
		Fe–N	ΔE	ΔE_{N_2}	Fe–N$_{TS}$	ΔE_{TS}	Fe–N	ΔE	ΔE_{N_2}	Fe–N	ΔE	ΔE_{N_2}	Fe–N	ΔE	ΔE_{N_2}
E_0	Fe2 + N$_2$	2.79	27.6	28.6				39.4	35.8		31.6	33.3		37.7	35.1
	Fe6 + N$_2$	3.31	0.0	1.0				0.0	−3.6		0.0	1.6		0.0	−2.5
	Fe2–N$_2$	**1.90**	42.6	43.5			**1.90**	69.0	65.4	**1.90**	60.0	61.6	**1.90**	49.5	46.9
	Fe6–N$_2$	**1.90**	33.4	34.3			**1.90**	60.9	57.3	**1.90**	56.0	57.7	**1.90**	44.0	41.5
E_1	Fe2 + N$_2$	2.84	31.0	32.8	2.4	39.7	2.86	39.1	37.4	2.81	32.0	35.4	2.83	39.0	39.1
	Fe6 + N$_2$	3.68	0.0	1.9	2.0	32.4	3.63	0.0	−1.7	3.64	0.0	3.4	3.63	0.0	0.1
	Fe2–N$_2$	1.92	31.1	32.9			1.97	42.9	41.2	2.01	46.4	49.9	**1.90**	80.8	80.9
	Fe6–N$_2$	1.92	24.1	26.0			**1.90**	60.3	58.6	**1.90**	58.6	62.1	**1.90**	73.7	73.9
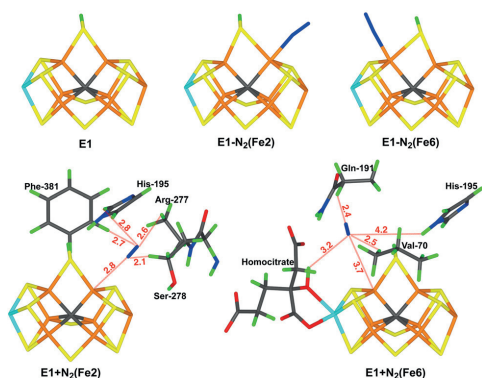
Fig. 2 The best E_1 structures, obtained with TPSS: E_1 without N_2, Fe2–N_2 and Fe6–N_2 with N_2 coordinating to Fe, as well as Fe2 + N_2 and Fe6 + N_2 with N_2 in the second coordination sphere, showing also nearby residues.

has ΔE_{N_2} = 2 kJ mol^{-1}, showing that we could also have used such a structure as the reference state. In this structure N_2 resides in a cavity between homocitrate, His-195, Gln-191 and Val-70, forming weak interactions with in particular the latter two residues (N···H distances of 2.4–4.2 Å). The activation energy for the binding of N_2 to Fe6 from this second-sphere structure is only 32 kJ mol^{-1}.

Around Fe2, a second-sphere structure with a Fe2–N distance of 2.84 Å can be found, but it is 31 kJ mol^{-1} less favourable (ΔE_{N_2} = 33 kJ mol^{-1}; Fig. 2). On the other hand, binding from this structure to Fe2 has an activation energy of only 9 kJ mol^{-1}. In this structure, N_2 interacts weakly with Ser-278 (2.1 Å), Arg-277 (2.6 Å), Phe-381 (2.7 Å) and His-195 (2.8 Å).

With the other three functionals, no structure with N_2 bound to Fe6 was found. However, structures with N_2 bound to Fe2 were found for TPSSh and r²SCAN, but not with B3LYP. They have Fe2–N distances of 1.97 and 2.01 Å, and ΔE_{N_2}

unfavourable binding energies of 41 and 50 kJ mol^{-1} for r²SCAN and TPSSh, respectively. Restraining the Fe–N distance to 1.90 Å, we obtain a ΔE_{N_2} binding energy for B3LYP of 81 kJ mol^{-1}. For Fe6, such restrained structures give ΔE_{N_2} = 59–74 kJ mol^{-1}. For B3LYP, binding to Fe6 is stronger, whereas for the other two functionals binding to Fe2 is stronger. Thus, we can conclude that all four functionals suggest that N_2 binding to the E_0 and E_1 states is unfavourable, in agreement with experiments.[1,25,28–30] For E_1, the bond strengths are in the order TPSS > r²SCAN > TPSSh > B3LYP, showing a decreasing trend with respect to the amount of HF exchange in the functional (10% for TPSSh and 20% with B3LYP).

N_2 binding to the E_2 state

Next, we considered the binding of N_2 to the E_2 state. We first studied ten structures for the unligated E_2 state with the H atoms on S2B, S5A, Fe5, the central carbide or bridging Fe2 and Fe6 (denoted Fe2/6). In some structures, the protonated S2B group had dissociated from either Fe2 or Fe6. The structures are described in Table 2 and are shown in Fig. 3. Most of them were included also in our previous study[63] and we use the naming convention from that study (explained in detail in Table 2): Structures starting with a "B" have a hydride ion _bridging_ Fe2 and Fe6, and S2B is protonated and also _bridging_ Fe2 and Fe6. The two numbers indicate the direction of the proton on S2B and the hydride ion (in this order), _viz._ pointing towards S3A(3) or towards S5A(5). Structures starting with "H" has the proton and the hydride in the same positions, but S2B has dissociated from either Fe2 or Fe6, but not the other Fe ion (it is _half-dissociated_). The number indicates which Fe ion it still binds to, and the final letter indicates whether the proton on S2B points towards Fe, S or Mo. Structures starting with "N" have _no_ bridging hydride ion, but instead protons on S2B and S5A. The numbers indicate the direction of the two protons in this order (for that on S5A, either towards S2B or S5A). The T53 structure has a _terminal_ hydride ion on Fe5 and a proton on S2B, directed towards S3A. The C2 structure had a doubly protonated central _carbide_ ion.

Table 2 The ten structures studied for the E_2 state without N_2. The names are the same as in our previous study of this state.[63] ΔE is the relative energy for each DFT method (kJ mol^{-1}). The H1 and H2 columns describe which atom is protonated and the direction of the proton. S2B (3) or S2B(5) means that S2B is protonated with the proton directed towards the S3A or S5A atoms. Fe2/6(5) means that the H atom bridges Fe2 and Fe6 on the same side of S2B as S5A. C2367 and C3457 means that the central carbide is protonated with the proton pointing to the Fe2–Fe3–Fe6–Fe7 or Fe3–Fe4–Fe5–Fe7 face. S2B(H6S) means that S2B is protonated and is dissociated from Fe2, but remains bound to Fe6, with the proton directed towards S1B. Likewise, S2B(H2F) means that S2B is protonated and is dissociated from Fe6, but remains bound to Fe2, with the proton directed towards Fe1. The structures were studied in the BS7-235 state, unless otherwise stated

Structure	H1	H2	ΔE (kJ mol^{-1})			
			TPSS	r^{2}SCAN	TPSSh	B3LYP
B33	S2B(3)	Fe2/6(3)	0.0	41.0	14.5	55.0
B35	S2B(3)	Fe2/6(5)	17.0	16.7	9.5	44.5
B53	S2B(5)	Fe2/6(3)	4.4	46.0	20.6	63.9
B55	S2B(5)	Fe2/6(5)	32.7	62.6	42.3	82.8
H6S	S2B(H6S)	Fe2/6	13.3	0.0	0.0	0.0^a
H2F	S2B(H2F)	Fe2/6	75.3	79.6	70.8	97.4
N33	S2B(3)	S5A(3)	26.4	54.6	27.0	27.0
N52	S2B(5)	S5A(2)	45.7	77.1	45.8	44.5
T53	S2B(3)	Fe5	19.0	24.9	16.0	26.8
C2^b	C2367	C3457	157.7	113.1	35.1	−88.5

^{a} Studied in the BS7-346 state. ^{b} Studied in the BS8-345 state.
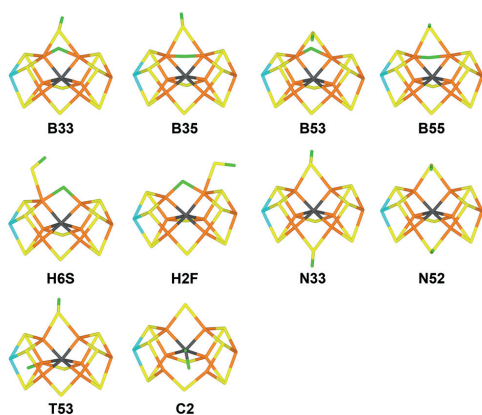
Fig. 3 The ten E_2 structures without N_2 bound. The positions of the added H atoms are described in Table 2 and the labels are explained in the text.

The relative stabilities of these structures are also shown in Table 2. It can be seen that with the TPSS functional, the B33 structure (with H atoms on S2B and bridging Fe2/6; Fig. 3) is most stable, 13 kJ mol^{-1} better than the half-dissociated H6S structure (with H atoms in the same positions, but with S2B dissociated from Fe2; Fig. 3). With the other three functionals,

the situation is opposite by 14–55 kJ mol^{-1} (in both cases, some other structures are intermediate in energy). However, with B3LYP, the structure with the central carbide ion doubly protonated (C2; Fig. 3) is 88 kJ mol^{-1} more stable, whereas this structure is disfavoured by 35–158 kJ mol^{-1} with the other functionals. Such structures are strongly distorted.

Next, we studied the binding of N_2 to the best of these structures. The results are collected in Table 3. With the TPSS functional, N_2 was found to bind to both Fe2 and Fe6 for all types of structures. The most favourable structure has N_2 bound to Fe6 and S2B bound to Fe2 but dissociated from Fe6 (Fig. 4). Such a structure naturally forms from the H2F structure, but it arose also from the B33 and B35 structures, because S2B automatically dissociated from Fe6 during the binding of N_2 (the three structures are isoenergetic within 0.5 kJ mol^{-1} and we describe the best, Fe6–B35, in the following). It has a Fe6–N bond length of 1.81 Å and a favourable ΔE_{N_2} of −11 kJ mol^{-1}. The corresponding structure with N_2 bound to Fe2 and S2B bound only to Fe6 is 7 kJ mol^{-1} less stable (H6S; again it arose also from B33 or B35 by spontaneous S2B dissociation from Fe2; Fig. 4). The Fe2–N bond length is 1.85 Å and ΔE_{N_2} is −5 kJ mol^{-1}. Structures with S2B still bridging both Fe2 and Fe6 are 52–67 kJ mol^{-1} less stable when N_2 binds to Fe6 and 64–78 kJ mol^{-1} less stable when N_2 binds to Fe2, in both cases following the order N33 < T53 < N52. Structures with N_2 binding in the second sphere can be found for all structures, except H6S and H2F, but they are 30–78 kJ mol^{-1} less stable than the best bound N_2-bound structure. The activation energy for N_2 binding is rather small for all structures, 5–46 kJ mol^{-1}, and barrierless in three cases.

With the other three DFT functionals, N_2 does not bind to Fe6 for the N33 and N52 structures, and it does not bind to Fe2 for the T53 structure (very weakly for B3LYP with a Fe2–N distance of 2.42 Å). In fact, most B3LYP structures with N_2 binding to Fe2 have very long Fe2–N distances, 2.28–2.53 Å (1.98–2.15 Å, 2.01–2.02 Å and 1.85–1.95 Å for the corresponding structures with TPSSh, r^{2}SCAN and TPSS, respectively). As with TPSS, S2B dissociates from the Fe ion binding N_2 for the B33 and B35 structures, giving structures virtually identical to those started from the half-dissociated H6S or H2F structures. The same happens for N_2 binding to Fe6 in the T53 structure with all three functionals and for N_2 binding to Fe2 in the N33 and N52 structures with r^{2}SCAN and TPSSh. With
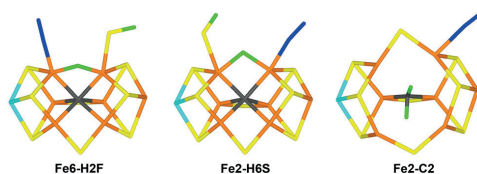
Fig. 4 The best E_2 structures with N_2 bound, Fe6-H2F, Fe2-H6S and Fe2–C2. The first two were optimised with TPSS, whereas Fe2–C2 was optimised with B3LYP.

Table 3 Structures studied for the E_2 state with N_2. The structures are the same as in Table 2 and the entries are the same as in Table 1. ΔE_{db} is the direct binding energy relative to the same type of structure with N_2 in the second sphere (upper part of the table; eqn (3)). All Fe—N bond lengths are in Å (a distance in bold face indicates a restrained bond) and all energies are in kJ mol^{-1}. All structures were studied in the BS7–235 state, if not otherwise stated

| Structure		TPSS						r2SCAN				TPSSh				B3LYP			
		Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	FeN$_{TS}$	ΔE_{TS}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}
Fe2 + N2	B33	2.86	55.0	43.6				2.85	55.6	70.1		2.88	58.0	66.7		2.87	66.1	193.5	
	B35	2.80	63.8	52.4				2.04	6.9	21.5		2.08	12.6	21.4		2.46	9.7	137.0	
	T53	2.89	62.7	51.3				2.91	53.7	68.3		2.87	48.8	57.6		2.81	65.7	193.0	
	N33	2.81	69.5	58.1				2.79	77.6	92.2		2.70	50.1	58.9		2.64	18.6	145.9	
	N52	2.89	78.2	66.7				2.86	89.5	104.0		2.80	62.7	71.5		2.80	31.1	158.4	
	H6S	**2.80**	48.3	36.9				**2.80**	13.5	28.1		**2.80**	19.3	28.1		**2.80**	4.5	131.8	
	C2	2.75^a	192.1	180.7				2.70^a	128.1	142.7		2.70^a	55.1	63.9		2.86	-95.7	31.7	
Fe6 + N2	B33	3.40	29.8	18.4				3.54	25.7	40.3		3.42	31.7	40.5		3.50	33.5	160.8	
	B35	3.56	34.1	22.7				3.52	4.9	19.5		3.52	8.7	17.5		3.51	12.6	139.9	
	T53	3.66	34.0	22.6				3.61	8.9	23.4		3.66	11.9	20.6		3.60	-10.4	116.9	
	N33	3.60	42.2	30.8				3.58	40.7	55.2		3.56	21.8	30.5		3.50	-16.9	110.4	
	N52	3.24	56.1	44.7				3.35	60.0	74.6		3.25	40.9	49.6		3.36	5.2	132.5	
	H2F	**3.70**	76.7	65.3				**3.70**	46.8	61.4		**3.70**	49.6	58.4		**3.70**	44.2	171.5	
	C2	3.20^a	169.9	158.4				3.32^a	93.2	107.7		3.27^a	26.7	35.5		3.22	-135.8	-8.5	
Fe2–N2	B33^b	1.85	7.0	-4.4	-48.0	2.40	4.9	2.01	2.1	16.7	-53.5	2.00	2.1	10.9	-55.9	2.52	1.6	128.9	-8.1
	B35^b	1.85	6.9	-4.5	-56.9	No barrier		2.01	0.0	14.6	-6.9	1.98	0.0	8.8	-12.6	2.53	0.0	127.3	-9.7
	T53^c	1.91	66.6	55.1	3.9	2.30	12.2	1.90	75.4	90.0	21.7	1.90	86.3	95.1	37.5	2.42	67.7	195.0	2.0
	N33	1.93	63.9	52.5	-5.6	2.40	4.6	2.01^b	77.9	92.5	0.3	2.15^b	47.6	56.4	-2.5	2.28	16.1	143.4	-2.6
	N52	1.95	78.5	67.1	0.3	2.30	7.3	2.02^b	89.6	104.1	0.1	2.14^b	64.2	73.0	1.5	2.30	31.3	158.6	0.2
	H6S	1.85	6.7	-4.7	-41.6	No barrier		2.01	2.1	16.6	-11.5	2.00	1.2	10.0	-18.1	2.01	1.8	129.1	-2.7
	C2	2.29^a	184.8	173.4	-7.3			2.54^a	126.2	140.8	-1.9	2.35^a	56.5	65.3	1.4	2.11	-105.3	22.0	9.6
Fe6–N2	B33^b	1.80	0.4	-11.0	-29.4	2.30	11.0	1.83	32.6	47.2	6.9	1.83	34.0	42.8	2.4	1.93	74.3	201.6	40.8
	B35^b	1.81	0.0	-11.4	-34.1	2.30	46.4	1.87	20.6	35.2	15.7	1.86	21.9	30.7	13.2	1.98	51.2	178.5	38.6
	T53	1.83	56.9	45.5	22.9			1.86^b	121.3	135.9	112.5	1.86^b	102.2	111.0	90.4	1.89^b	151.3	278.7	161.8
	N33	1.88	51.6	40.2	9.4	2.20	16.4	1.90	75.1	89.7	34.4	1.90	57.1	65.9	38.3	1.90	46.8	174.2	61.2
	N52	1.95	66.9	55.5	10.8	2.30	14.0	1.90	92.8	107.4	32.8	1.90	72.9	81.7	32.1	1.90	59.2	186.5	53.9
	H2F	1.81	0.5	-10.9	-76.2	No barrier		1.87	19.9	34.5	-26.9	1.85	19.3	28.1	-30.4	1.96	51.0	178.3	6.8
	C2	1.81	175.5	164.1	37.3			2.08^a	97.5	112.1	4.4	2.13	46.2	55.0	-0.6	1.81^a	6.2	133.5	136.4

^{a} Studied in the BS8–345 state. ^{b} S2B dissociates spontaneously from the Fe ion not binding N_2. ^{c} Studied in the BS10–147 state.

the latter two functionals, the most favourable structure is Fe2-H6S (which arises also from B33 and N35). This structure is 20 and 19 kJ mol^{-1} more stable than the corresponding structure with N$_2$ bound to Fe6 (Fe6-H2F/B33/B35; which was most stable with TPSS), respectively. ΔE_{N_2} of the best structure is unfavourable by 9–15 kJ mol^{-1}. Second-sphere N$_2$ binding is found to both Fe2 and Fe6 for all structures except H6S and H2F. The best is Fe6 + B35, but the corresponding Fe2 structure is only 2–4 kJ mol^{-1} less stable. They are 5–9 kJ mol^{-1} less stable the best first-sphere N$_2$-bound structure.

With B3LYP, the structure with N$_2$ binding to Fe2 in the C2 structure is by far most stable, 105 kJ mol^{-1} more stable than the *e2e* structure (Fe2 or Fe6) structure (Fig. 4). However, the ΔE_{N_2} binding energy is unfavourable by 22 kJ mol^{-1}.

It is notable that for several structures, the direct N$_2$ binding energy ΔE_{db} is favourable for all four methods, by up to 76, 53, 56 and 10 kJ mol^{-1} for TPSS, r^{2}SCAN, TPSSh and B3LYP, respectively. However, this mainly reflects problems with the definition of ΔE_{db}. For r^{2}SCAN and TPSSh, the strongly favourable ΔE_{db} energies comes from the B33 structure, which reorganises to a H6S structure when N$_2$ binds (by dissociation of S2B from one of the Fe ions). If we instead use the (restrained) H6S structure as the reference, ΔE_{db} becomes much less favourable, −11 and −18 kJ mol^{-1}. For TPSS, favourable ΔE_{db} are also obtained for structures in which S2B is already half-dissociated, for which no minimum with N$_2$ in the second coordination sphere is found. The large difference between ΔE_{db} and ΔE_{N_2} is caused by the unfavourable energies of all structures with N$_2$ in the second coordination sphere. In reality, the binding takes place to the most stable E$_2$ structure without N$_2$ and forms the most stable E$_2$–N$_2$ structure (unless kinetic barriers are large). Therefore, ΔE_{N_2} of the best E$_2$–N$_2$ structure should be the most relevant binding energy and we will not discuss ΔE_{db} for the other E$_n$ states (it is still included in the tables).

N$_2$ binding to the E$_3$ state

Next, we turn to the E$_3$ state. This state is less thoroughly studied than the other states.[37] However, the results from the E$_2$ and E$_4$ states give some clues of possible protonation states also for E$_3$.[18,19,35,37,38,52,54,62–67] We have optimised 16 different structures without N$_2$. The protonation states and the nomenclature are described in Table 4 and the structures are shown in Fig. 5. Two types of structures were studied. One is based on the suggestions by Hoffman and coworkers[54] that E$_4$ has protons on S2B and S5A, as well as two hydride ions bridging Fe2/6 and Fe3/7. Each H atom can attain two conformations, *e.g.* directed towards S3A or S5A for the one on S2B (but the hydride ion on Fe3/7 was always on the S2B side of S5A). This gives eight possibilities for E$_4$ and we studied six variants of these with either the H atom on Fe3/7 or S5A is deleted for E$_3$. They are denoted in the same way as for E$_4$ below, *i.e.* with four numbers showing the conformation of the H atoms on S2B, Fe2/6, Fe3/7 and S5A in this order, using underscore to indicate a vacancy, *e.g.* 332_. Two variants of the 35_3 structure were also studied with S2B dissociated from either Fe2 or Fe6, but still binding to Fe6 (H6S) or to Fe2 (H2F; the final letter reflect the direction of the proton on S2B, towards S or Fe). The second type of structures is based on the suggestion by Bjornsson and coworkers that two hydride ions may both bridge Fe2/6, especially if S2B is protonated and has dissociated from either Fe2 or Fe6.[62] The names of these structures start with a "S" (the hydride ions bind the *same* pair of Fe ions). Four such structures were studied, depending on the direction of the proton on S2B (S2F, S2S, S6M and S6S), indicating that the proton points towards Fe, S or Mo). In addition, a structure with the central carbide ion triply protonated (C3) and three structures with hydride ions terminally bound on one or two Fe ions were studied (345, 335 and 355; explained in Table 4).[37]

Table 4 The 16 structures studied for the E$_3$ state without N$_2$. ΔE is the relative energy for each DFT method (kJ mol^{-1}). The H1, H2 and H3 columns describe which atom is protonated and the direction of the proton in the same way as was described in the legend of Table 2. The structures were studied in the BS10-147 state, unless otherwise stated

Structure	H1	H2	H3	ΔE (kJ mol^{-1})			
				TPSS	r^{2}SCAN	TPSSh	B3LYP
352_	S2B(3)	Fe2/6(5)	Fe3/7(2)	43.5	77.5	62.4	86.7
35_2	S2B(3)	Fe2/6(5)	S5A(2)	42.1	56.1	31.7	0.0
35_3	S2B(3)	Fe2/6(5)	S5A(3)	32.2	45.1	20.9	9.3
H2F	S2B(H2F)	Fe2/6	S5A(3)	101.1	109.3	81.8	70.4
H6S	S2B(H6S)	Fe2/6	S5A(3)	57.0	64.8	50.8	31.0
552_	S2B(5)	Fe2/6(5)	Fe3/7(2)	58.8	94.3	79.4	104.0
55_2	S2B(5)	Fe2/6(5)	S5A(2)	60.5	76.5	51.1	34.2
55_3	S2B(5)	Fe2/6(5)	S5A(3)	49.9	64.7	39.9	28.0
345	S2B(3)	Fe4	Fe5	54.1	58.1	58.0	70.1
335	S2B(3)	Fe2/6(3)	Fe5	47.7	69.3	55.6	80.2
355	S2B(3)	Fe2/6(5)	Fe5	42.1	58.4	44.3	64.1
S2F	S2B(H2F)	Fe2/6	Fe2/6	45.3	66.9	51.8	84.7
S2S	S2B(H2S)	Fe2/6	Fe2/6	49.7	69.6	56.7	88.9
S6M	S2B(H6M)	Fe2/6	Fe2/6	1.3^a	1.8	9.1	28.1
S6S	S2B(H6S)	Fe2/6	Fe2/6	0.0	0.0	0.0	16.7
C3	C2367	C2456	C3457	175.8	141.1	3.6^b	−177.9

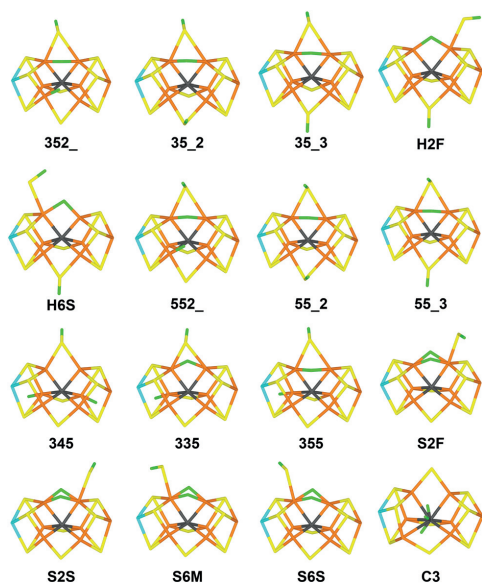^{a} Studied in the BS-14 state. ^{b} Studied in the BS10-136 state.

Fig. 5 The 16 E_3 structures without N_2 bound. The positions of the added H atoms are described in Table 4 and the nomenclature is explained in the text.

The relative stabilities of these structures are shown in Table 4. It can be seen that with TPSS, r^2SCAN and TPSSh, the S6S structure is most stable, *i.e.* with two Fe2/6 hydride ions and S2B protonated and binding only to Fe6 (Fig. 5). The S6M structure with the proton on S2B pointing in a different direction is close in energy (1–9 kJ mol^{-1}). The third-best state is 35_3 (H atoms on S2B(3), Fe2/6(5) and S5A(3); Fig. 5), which is 32, 45 and 21 kJ mol^{-1} less stable with the three functionals, respectively. With B3LYP, instead the C3 structure is best (with a triply protonated carbide ion; Fig. 5), 178 and 195 kJ mol^{-1} more stable than the 35_2 and S6S structures. The C3 structure is only 4 kJ mol^{-1} less stable than S6S with TPSSh, whereas it is 176 and 141 kJ mol^{-1} less stable with TPSS and r^2SCAN, confirming the previous observation[38] that the stability of structures with the central carbide protonated depends strongly on the amount of HF exchange in the functional. Interestingly, no functional indicates that the half-dissociated variants of 35_3 (H2F and H6M) are more stable than non-dissociated variant.

For the most stable and interesting structures, we then studied binding of N_2 (results in Table 5). All tested structures gave stable N_2-bound states at both Fe ions, except Fe6-35_2. With r^2SCAN, no 345 structure with N_2 bound to Fe2 or Fe6 was found. For the 355, 35_2, 35_3 and 352_ structures with N_2 binding to Fe2, S2B dissociated from Fe2. With TPSS, the most stable structure was S2S with N_2 bound to Fe6 (Fig. 6). It gave a Fe–N bond length of 1.80 Å. ΔE_{N_2} is favourable by 46 kJ mol^{-1}. The other three structures with two hydrides bridging

Fe2/6 (S2F, S6M and S6S) are rather close in energy (14–23 kJ mol^{-1} less stable), whereas the other structures are appreciably worse (at least 66 kJ mol^{-1} less stable than Fe6-S2S).

With r^2SCAN and TPSSh, instead the two S6M and S6S structures with N_2 bound to Fe2 (Fig. 6) are most stable and degenerate within 1 kJ mol^{-1}. They have Fe–N bond lengths of 1.84 Å, but unfavourable ΔE_{N_2} = 9 or 3 kJ mol^{-1}, respectively. The Fe6-S2S structure is 21–30 kJ mol^{-1} less stable and the Fe2-35_3 structure (with S2B dissociated from Fe2; Fig. 6) is 28 (TPSSh) or 48 (r^2SCAN) kJ mol^{-1} worse.

With B3LYP, the situation is similar: Fe2-35_3 and Fe6-S2S are 9 and 27 kJ mol^{-1} less stable than Fe2-S6S, the latter with a Fe–N bond length of 1.88 Å. However, the C3 structures are by far the most stable, by 180 and 164 kJ mol^{-1} for N_2 bound to Fe2 and Fe6, respectively (Fig. 6). For the latter two, ΔE_{N_2} is unfavourable by 28–44 kJ mol^{-1}.

Structures with N_2 in the second coordination sphere of the Fe6 ion were found for most structures, but not for the C3 structure or for any of the structures with two Fe2/6 hydride ions. With N_2 in the second coordination sphere of Fe2, only three structures were found, 345, 355 and 35_2. These structures are much less stable than the N_2-bound structures with TPSS, r^2SCAN and TPSSh (by at least 76, 39 and 26 kJ mol^{-1}). However, with B3LYP, they are slightly more stable than the corresponding N_2 bound structures (by up to 15 kJ mol^{-1}), but they are still much less stable than the N_2-bound C3 structures (for which no second-sphere structures are found; by 165 kJ mol^{-1}). Activation barriers for the binding of N_2 from the second coordination sphere are low, 4–35 kJ mol^{-1}, and many of the reactions are barrierless (*cf.* Table 5).

N_2 binding to the E_4 state

Finally, we studied also the E_4 state using 20 different structures, described in Table 6 and shown in Fig. 7. The naming of the structures follows the same philosophy as for the E_3 structures and in analogy we investigated mainly two types of structures. The first is structures with protons on S2B and S5A and two hydride ions bridging Fe2/6 and Fe3/7, as suggested by Hoffman and coworkers.[54] We studied eight such structures with the H atoms pointing in different directions,[66] as is described in Table 6 and Fig. 7. The 5522 structure is the one advocated by Hoffman and coworkers,[54] whereas the 3323 structure was lowest in energy in our previous study.[66] We also studied four variants of these with the protonated S2B group dissociated from either Fe2 or Fe6 (H6 or H2). Second, we studied six structures with two protons still on S2B and S5A, but both hydride ions bridging Fe2/6 and with S2B dissociated from either Fe2 or Fe6 (S2F, S2S, S6M, S6S, S6M2 and S6S2, named the same way as for the E_3 structures, as is described in Table 6 and shown in Fig. 7; the two structures with a final "2" have the proton on S5A pointing towards S2B rather than towards S3A). Finally, we studied one structure with a proton on S2B and hydride ions on Fe5, Fe6 and bridging Fe2/6 (called 3H)[37] and one structure with a proton on S2B and the central carbide ion triply protonated (C3).[37]

Dalton Transactions

Table 5 Structures studied for the E_4 state with N_2. The structures are the same as in Table 4 and the entries are the same as in Table 3. All Fe–N bond lengths are in Å (a distance in bold face indicates a restrained bond) and all energies in kJ mol^{-1}. The structures were studied in the BS10-147 state, unless otherwise stated.

		TPSS						r2SCAN				TPSSh				B3LYP			
Group		Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	FeN$_{TS}$	ΔE_{TS}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}
F2 + N2	345	2.91[a]	112.9	67.4				2.75	84.0	93.2		2.69	75.8	78.6		2.66	58.5	266.3	
	335	**2.80**	119.1	73.5				**2.80**	92.5	101.6		**2.80**	72.8	75.5		**2.80**	89.2	297.0	
	355	2.84[a]	119.5	74.0				2.54	83.5	92.7		2.63	65.7	68.4		2.66	57.7	265.5	
	35_2	2.75	123.9	78.3				2.02	62.3	71.5		2.70	67.0	69.8		2.61	23.5	231.3	
	35_3	**2.80**	110.9	65.4				**2.80**	64.0	73.1		**2.80**	54.0	56.8		**2.80**	16.6	224.4	
	352_	**2.80**	114.3	68.8				**2.80**	94.8	104.0		**2.80**	85.5	88.3		**2.80**	98.6	306.4	
	S6M	**2.80**	58.6	13.1				**2.80**	31.0	40.2		**2.80**	18.6	21.4		**2.80**	−7.3	200.5	
	S6S	**2.80**[b]	58.2	12.6				**2.80**	47.5	56.7		**2.80**	9.6	12.3		**2.80**	−17.9	189.9	
	C3	**2.80**	250.8	205.3				**2.80**[c]	156.7	165.9		**2.80**[c]	34.8	37.6		**2.80**[c]	−180.2	27.6	
Fe6 + N2	345	3.67	103.6	58.1				3.71	50.0	59.2		3.72	60.3	63.0		3.72	44.8	252.6	
	335	3.65	75.7	30.1				3.78	71.8	81.0		3.76	56.8	59.6		3.79	77.4	285.2	
	355	3.70	97.5	52.0				3.65	64.4	73.6		3.74	68.1	70.9		3.72	74.4	282.2	
	35_2	3.69	92.6	47.0				3.70	50.3	59.5		3.77	36.8	39.6		3.83	−6.9	200.9	
	35_3	3.72	82.9	37.4				3.69	39.2	48.4		3.74	25.7	28.4		3.74	−14.6	193.2	
	352_	3.62	96.9	51.4				3.58	73.0	82.2		3.61	69.2	72.0		3.62	65.5	273.3	
	S2F	**2.80**[b]	63.2	17.6				**2.80**	21.6	30.8		**2.80**	11.6	14.3		**2.80**	25.7	233.5	
	S2S	**2.80**[b]	74.1	28.6				**2.80**[d]	98.6	107.8		**2.80**[d]	72.8	75.5		**2.80**	148.2	355.9	
	C3	**3.70**	337.1	291.6				**3.70**	235.6	244.8		**3.70**	131.2	134.0		**3.70**	−104.1	103.7	
Fe2–N2	345	1.91	116.5	70.9	3.5	No barrier		2.02	87.4	96.5	3.4	1.91	145.3	148.1	69.5	1.89	135.5	343.3	77.0
	335	1.94	99.3	53.8	−19.8	No barrier		2.03	76.4	85.5	−7.1	2.09	68.5	71.3	−4.2	2.34	67.2	275.0	−22.0
	355	1.93	91.8	46.3	−27.7	2.4	4.5	1.93	61.8	71.0	−0.5	2.08	63.6	66.4	−2.0	2.31	80.9	288.7	23.2
	35_2	1.89	86.7	41.2	−37.1	No barrier		1.93	48.2	57.3	−15.8	1.97	52.7	55.4	−14.4	2.13	23.1	230.8	−0.4
	35_3	1.90	78.4	32.9	−32.5	No barrier		1.88	79.4	88.5	−15.5	1.97	28.0	30.7	−26.1	2.14	9.1	216.9	−7.5
	352_	1.85	73.9	28.4	−40.4	No barrier		1.84	0.4	9.6	−30.6	1.88	69.4	72.2	−16.1	1.90	88.9	296.7	−9.7
	S6M	1.85[b]	22.9	−22.6	−35.7	No barrier		1.84	0.0	9.2	−47.5	1.84	0.0	2.8	−18.6	1.90	1.1	208.9	8.4
	S6S	1.83[b]	15.8	−29.8	−42.4							1.84	1.1	3.9	−8.4	1.88	0.0	207.8	17.9
	C3	1.78[c]	227.4	181.9	−16.9			1.76[c]	154.1	163.3	1.3	1.88	56.8	59.5	36.2	2.28	−179.7	28.1	0.5
Fe6–N2	345	1.81	130.0	84.5	26.4	2.1	34.7	1.80	71.9	81.0	0.1	1.87	120.5	123.3	60.2	1.90	120.3	328.1	0.5
	335	1.81	66.1	20.6	−9.5	2.3	19.1	1.80	76.2	85.4	11.8	1.83	51.9	54.7	−4.9	1.86	81.4	289.2	4.0
	355	1.81	87.2	41.7	−11.9	2.3	12.0	1.90	96.3	105.4	0.0	1.82	59.5	62.3	−8.6	1.85	88.3	296.1	13.9
	35_2	1.90[b]	100.5	54.9	7.9			1.85	94.0	103.1	54.8	1.90[b]	67.9	70.7	31.1	1.90[b]	121.4	329.1	128.3
	35_3	1.85	89.4	43.9	6.5	2.3	18.4	1.83	142.5	151.7	69.5	1.87	69.0	71.8	43.4	1.93	95.5	303.3	110.1
	352_	1.83	120.2	74.7	23.3	2.3	27.4	1.86	36.8	46.0	15.2	1.84	117.0	119.7	47.8	1.90	173.3	381.1	107.8
	S2F	1.79[b]	14.4	−31.2	−48.8	No barrier		1.79[b]	29.8	39.0	−68.8	1.80[b]	31.6	34.4	20.0	2.41	27.2	235.0	1.5
	S2S	1.80[b]	0.0	−45.5	−74.1	No barrier		1.96[c]	101.6	110.8	−134.1	1.81[b]	20.9	23.7	−51.8	2.31	27.2	235.0	−121.0
	C3	1.73[e]	186.0	140.4	−4.8							1.76	52.4	55.2	−78.8	1.80	−163.6	44.2	−44.8

[a] Studied in the BS7-235 state. [b] Studied in the BS-14 state. [c] Studied in the BS7-136 state. [d] Studied in the BS10-136 state. [e] Studied in the BS8-14 state.
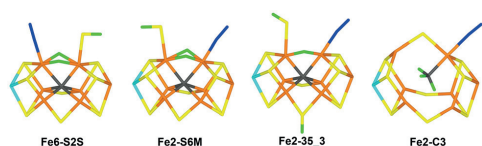
Fig. 6 The best E_3 structures with N_2 bound, Fe6-S2S, Fe2-S6M, Fe2-35_3 and Fe2–C3. The first three structures were optimised with TPSS, whereas Fe2–C3 was optimised with B3LYP.

The relative stability of the various E_4 states without any N_2 are shown in Table 6. The S6S structure (with two hydride ions bridging Fe2/6 and the protonated S2B dissociated from Fe2; Fig. 7) is most stable with TPSS, r^2SCAN and TPSSh. However, with TPSS, eight structures are within 15 kJ mol^{-1}. Three of these, S6M, 3323 and 3523 (Fig. 7), are within 25–29 and 11–18 kJ mol^{-1} of S6S for r^2SCAN and TPSSh, respectively. Changing the proton on S5A in the S6S structure so that it points towards S2B instead of S3A decreases the stability by 9–13 kJ mol^{-1}. Likewise, dissociating S2B from either Fe2 or Fe6 in the 3323 or 3523 structures also makes the structures less stable. With B3LYP, instead the structure with a triply protonated carbide ion (C3; Fig. 7) is by far the best, 192 and 229 kJ mol^{-1} more stable than 3523 and S6S, respectively. With TPSSh, the C3 structure is only 1 kJ mol^{-1} less stable than S6S, whereas it is 146–192 kJ mol^{-1} less stable than the best structure for the other two functionals. Thus, the E_4 state remains a challenge for computational methods in that several

structures are close in energy and the preferred structure depends strongly on the DFT functional.

For the best structures, we studied the binding of N_2. The results are shown in Table 7. N_2-bound structures were found for all structures studied and both Fe ions, except Fe6-3323H2 and Fe6-5522 with B3LYP. For all four functionals, S2B dissociates spontaneously from Fe2 when N_2 binds to this ion for all structures with the Fe2/6 hydride on the S3A side (with B3LYP also the 3322 and 3522 structures). In principle, such half-dissociation of S2B should remove the dependency on the conformations of the H atoms on S2B and Fe2/6. However, in practice there are still distinct local minima for both H atoms (for example, the hydride bridging Fe2/6 can still bend towards S3A or towards S5A, although S2B is no longer in between the two conformations), but the barriers between them are most likely appreciably lower. S2B does not dissociate spontaneously when N_2 binds to Fe6, but half-dissociated structures with N_2 bound are typically lower in energy for the cases we have tested.

With TPSS, the S2S structure with N_2 bound to Fe6 is most stable (Fig. 8). It gives a Fe6–N distance of 1.79 Å and a favourable ΔE_{N_2} binding energy of -51 kJ mol^{-1}. The S6S and 3523 structures with N_2 binding to Fe2 are 9 and 32 kJ mol^{-1} less stable, respectively (Fig. 8). With r^2SCAN, instead Fe2-S6S is most stable, with Fe–N = 1.79 Å and $\Delta E_{N_2} = -44$ kJ mol^{-1}. The Fe6-S2S structure is 11 kJ mol^{-1} less stable. With TPSSh, the same structures are also among the best ones and degenerate within 0.2 kJ mol^{-1}. However, the Fe6–C3 structure (Fig. 8) is actually 1 kJ mol^{-1} more stable. It has Fe–N bond length of 1.83 Å and $\Delta E_{N_2} = -27$ kJ mol^{-1}. With B3LYP, the Fe2–C3 struc-

Table 6 The 20 structures studied for the E_4 state without N_2. ΔE is the relative energy for each DFT method (kJ mol^{-1}). The H1, H2, H3 and H4 columns describe positions and directions of the four H atoms. The nomenclature is the same as in Table 2. The structures were studied in the BS10-147 state, unless otherwise stated.

					ΔE			
	H1	H2	H3	H4	TPSS	r2SCAN	TPSSh	B3LYP
3H	S2B(3)	Fe2/6(3)	Fe5	Fe6	40.3^a	125.0	112.0	141.2
3322	S2B(3)	Fe3/7(2)	Fe3/7(2)	S5A(3)	15.1	52.5	38.6	100.5
3323	S2B(3)	Fe2/6(3)	Fe3/7(2)	S5A(3)	3.2^a	27.4^b	11.3^b	41.5^b
3323H2	S2B(H2F)	Fe2/6(3)	Fe3/7(2)	S5A(3)	118.0^a	79.3	102.6	202.9
3323H6	S2B(H6S)	Fe2/6(3)	Fe3/7(2)	S5A(3)	39.0	43.5	49.2	121.6
3522	S2B(3)	Fe2/6(5)	Fe3/7(2)	S5A(2)	24.0	76.2^a	48.0^a	74.7
3523	S2B(3)	Fe2/6(5)	Fe3/7(2)	S5A(3)	14.0^a	25.4	17.9^b	0.0^c
3523H2	S2B(H2F)	Fe2/6(5)	Fe3/7(2)	S5A(3)	114.9			203.0
3523H6	S2B(H6S)	Fe2/6(5)	Fe3/7(2)	S5A(3)	39.3	43.7	49.2	101.8
5322	S2B(5)	Fe2/6(3)	Fe3/7(2)	S5A(2)	21.7^a	59.0	48.1	113.9
5323	S2B(5)	Fe2/6(3)	Fe3/7(2)	S5A(3)	11.1^a	45.0	37.0	104.1
5522	S2B(5)	Fe2/6(5)	Fe3/7(2)	S5A(2)	40.6^a	59.0	51.3	113.8
5523	S2B(5)	Fe2/6(5)	Fe3/7(2)	S5A(3)	29.7^a	43.2	40.4	104.7
S2F	S2B(H2F)	Fe2/6(5)	Fe2/6(5)	S5A(3)	54.1	78.7	56.2	101.3
S2S	S2B(H2S)	Fe2/6(3)	Fe2/6(5)	S5A(3)	54.8	76.4	56.4	108.2
S6M	S2B(H6M)	Fe2/6(3)	Fe2/6(5)	S5A(3)	0.6	28.5	18.1	65.5
S6S	S2B(H6S)	Fe2/6(3)	Fe2/6(5)	S5A(3)	0.0	0.0^c	0.0^d	37.0^e
S6M2	S2B(H6M)	Fe2/6(3)	Fe2/6(5)	S5A(2)	10.1	39.8	25.5	76.4
S6S2	S2B(H6S)	Fe2/6(3)	Fe2/6(5)	S5A(2)	9.4	76.4	29.6	64.5
C3	S2B(3)	C2367	C2456	C3457	191.7^f	145.9^c	1.3^f	-191.5^f

^{a} Studied in the BS-14 state. ^{b} Studied in the BS10-135 state. ^{c} Studied in the BS7-346 state. ^{d} Studied in the BS6-157 state. ^{e} Studied in the BS8-347 state. ^{f} Studied in the BS2-234 state.
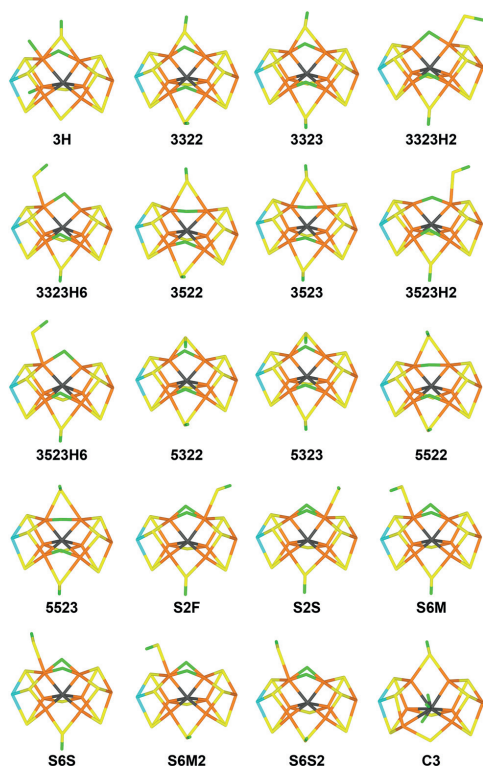
Fig. 7 The 20 E_4 structures without N_2 bound. The positions of the added H atoms are described in Table 6 and the nomenclature is explained in the text.

ture is the most stable, 46 kJ mol⁻¹ better than Fe6–C3 and 225 kJ mol⁻¹ better than Fe2-S6S. It has a Fe–N bond length of 2.18 Å and the ΔE_{N_2} binding energy is −13 kJ mol⁻¹.

With TPSS, we could find structures with N_2 residing in the second coordination sphere of Fe2 or Fe6 for all the Hoffman-type and 3H structures, but not for the structures with S2B half-dissociated or the C3 structures. With the other functionals, the same applied for N_2 binding in the second sphere of Fe6, but for Fe2, second-sphere structures were found only for 3H, 5322 and 5323 (also 3322 and 3323 with r²SCAN and TPSSh). The Fe6-3323 structure is the best for TPSS and TPSSh, whereas Fe6-3523 is best with r²SCAN, and Fe6-3522 is best for B3LYP. In all cases, the structures with N_2 in the second sphere are appreciably less stable than those with N_2 binding in the first sphere, by at least 59, 50, 51 and 274 kJ mol⁻¹ for TPSS, r²SCAN, TPSSh and B3LYP, respectively. Therefore, the barriers for N_2 binding to Fe2 are all low, below 14 kJ mol⁻¹ and often barrierless. For the binding to Fe6, the barriers are higher, 15–39 kJ mol⁻¹, but 79 kJ mol⁻¹ for the 3H structure.

Discussion

The prime results in this investigation are the binding energies. Unfortunately, they depend on the DFT functional and on how it is defined. We argue that it is most reasonable to use the QM/MM energy of the best structure of the same E_n state and free N_2 in a water-like continuum solvent as the reference (ΔE_{N_2}). For most functionals, this gives binding energies of the most stable N_2-bound structures that become increasingly negative (favourable) when going from E_0 to E_4 (employing restraints to obtain a bound state if no such state is found): 34, 26, −11, −46 and −51 kJ mol⁻¹ for TPSS, 57, 41, 15, 9 and −44 kJ mol⁻¹ for r²SCAN, 58, 50, 9, 3 and −27 kJ mol⁻¹ for TPSSh, and 41, 74, 22, 28 and −6 kJ mol⁻¹ for B3LYP (somewhat less regular). As mentioned in the Method section, it is likely that the def2-SV(P) basis set give ΔE_{N_2} binding energies that are ∼14 kJ mol⁻¹ too negative (too favourable, cf. Table S3†), so we will in the following discussion add 14 kJ mol⁻¹ to the ΔE_{N_2} energies.

As mentioned in the Introduction, Bjornsson and co-workers report more favourable binding energies, based on TPSSh calculations, −29 kJ mol⁻¹ for E_2 (Fe6-H₂S) and −43 or −56 kJ mol⁻¹ for E_4 (Fe2–H6 or Fe6–H2).[52] The reason for this is mainly that he uses another definition for the binding energy (ΔE_{bn} in eqn (4)). With this definition, we obtain ΔE_{bn} = −43 kJ mol⁻¹ for E_2-Fe6-H₂S and −82 kJ mol⁻¹ for E_4-Fe6-H₂S (because the H₂S is not the best structure neither for E_2 nor for E_4). Likewise, Dance obtains more favourable binding energies[59] than we do because he uses the ΔE_{db} definition in eqn (3). Moreover, all his E_4 structures involve a bound H₂ molecule, i.e. a type of structures not included in our investigation.

It should be noted that binding energies discussed so far are pure (electronic) energies. To compare with experimental results, we need to use free energies, i.e. to add enthalpic and entropic corrections. In particular, N_2 loses translational and rotational entropy upon binding. Unfortunately, there is no consensus in the size of this entropic penalty. Bjornsson and coworkers, as well as Siegbahn, use DFT frequency calculations to estimate an entropic penalty of 41–45 kJ mol⁻¹.[52,60,61] On the other hand, Dance argues that the relevant dissociated state is N_2 at a diffusible position inside the protein, where it has already lost most of its translational and rotational entropy. Therefore, he suggests a much smaller entropic penalty of ∼17 kJ mol⁻¹.[58]

With the entropy correction of Bjornsson and Siegbahn, together with the basis-set correction, none of the DFT methods give favourable N_2 binding for any of the E_n states, although for E_4 with TPSS, ΔE_{N_2} is only slightly positive (∼6 kJ mol⁻¹). With Dance's entropy penalty (and the basis-set correction), TPSS suggests that N_2 cannot bind to E_0–E_2, but that it binds to E_3 and E_4, in accordance with the experimental data.[1,25,28–30] B3LYP and TPSSh still give no favourable N_2 binding to any E_n state (although that to E_4 for TPSSh is only slightly positive, by 4 kJ mol⁻¹). For r²SCAN, ΔE_{N_2} is favourable to E_4 (by 13 kJ mol⁻¹) and unfavourable to the other states.

Table 7 Structures studied for the E$_4$ state with N$_2$. The structures are the same as in Table 4 and the entries are the same as in Table 3. All Fe–N bond lengths are in Å (a distance in bold face indicates a restrained bond) and all energies in kJ mol^{-1}. The structures were studied in the BS-14 state, unless otherwise stated.

		TPSS						r2SCAN				TPSSh				B31LYP			
		Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	FeN$_{TS}$	ΔE_{TS}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}
Fe2 + N$_2$	3H[a]	2.97	87.3	36.5				3.12	119.0	100.5		3.06	124.0	97.2		3.08	174.9	393.9	
	3322	2.85	95.9	45.1				2.90	102.7	84.1		2.80	95.1	68.2		2.80	83.7	302.7	
	3323	2.77	86.8	36.0				2.79	117.3	98.7		2.71	95.0	68.1					
	3323H6	2.80	115.1	64.2				2.80	93.1	74.5		2.80	85.2	58.4					
	3522	2.68	112.0	61.1															
	3523	1.87	59.8	8.9															
	5523H6	2.80	102.2	51.4				2.80	80.3	61.8		2.80	85.7	58.9		2.80	15.3	234.4	
	5322	2.82	93.5	42.7				2.80	104.8	86.3		2.80	90.4	63.6		2.80	91.3	310.3	
	5323	2.79	59.8	34.4				3.02	104.0	85.5		2.88	96.5	69.6		2.85	108.0	327.0	
	5522	2.70	127.6	76.8				2.96	91.6	73.0		2.88	84.3	57.5		2.78	85.6	304.7	
	5523	2.71	120.6	69.8															
	S6M	2.80	70.0	19.1				2.80	68.7	50.1		2.80	60.6	33.8		2.80	−8.3	210.7	
	S6S	2.80	75.7	24.9				2.80	47.0	28.5		2.80	26.7	−0.1		2.80	−6.6	212.4	
	S6M2	2.80	78.9	28.0				2.80	108.9	90.3		2.80	74.5	47.6		2.80	99.4	318.5	
	C3[a]	2.80	322.7	271.8				2.80	198.8	180.2		2.80	28.5	1.7		2.80	−157.2	61.9	
Fe6 + N$_2$	3H[a]	3.79	72.1	21.2				3.81	91.7	73.2		3.85	100.5	73.7		3.77	147.6	366.6	
	3322	3.69	70.2	19.4				3.72	67.5	48.9		3.72	70.0	43.1		3.80	54.7	273.7	
	3323	3.72	59.0	8.1				3.70	51.9	33.4		3.71	51.2	24.3		3.74	67.0	286.1	
	3523	3.57	82.2	31.3				3.56	65.5	47.0		3.60	68.1	41.3		3.58	48.3	267.3	
	3522	3.57	72.7	21.8				3.57	50.0	31.5		3.60	58.2	31.4		3.60	69.6	288.7	
	5323	3.16	68.5	17.6				3.32	102.0	83.4		3.24	76.8	50.0		3.50	64.0	283.0	
	5322	3.19	60.5	9.6				3.33	63.0	44.4		3.24	65.3	38.4		3.75	59.0	278.0	
	5523	3.56	98.7	47.9				3.55	115.0	96.5		3.58	103.3	76.5		3.53	91.8	310.8	
	5522	3.54	87.6	36.7				3.52	63.5	45.0		3.55	75.5	48.7		3.50	51.5	270.5	
	S2F	2.80	80.6	29.7				2.80	64.2	45.7		2.80	60.1	33.3		2.80	14.6	233.6	
	S2S	2.80	80.7	29.9				2.80	68.5	50.0		2.80	65.1	38.3		2.80	62.3	281.3	
	C3[a]	3.70	246.2	195.3				3.70	142.5	123.9		3.70	129.9	3.1		3.70	−212.2	6.8	
Fe2–N$_2$	3H[a]	1.83	76.9	26.0	−10.4	2.5	14.0	1.86	134.5	116.0	15.5	1.84	130.0	103.2	6.0	1.87	182.2	401.2	7.3
	3322	1.89	67.2	16.3	−28.7	2.6	0.1	1.90	105.0	86.5	−22.7	1.91	81.6	54.7	−22.8	1.97	80.4	299.4	−60.6
	3323	1.89	56.6	5.7	−30.2	2.6	0.8	1.90	92.0	73.5	−25.2	1.90	81.3	54.5	−13.6	1.96	75.3	294.3	−55.1
	3323H6	1.84	57.6	6.8	−57.4			1.86	66.6	48.0	−26.5	1.86	48.3	21.5	−36.9	1.87	53.2	272.2	−30.6
	3522	1.84	49.1	−1.8	−62.9	No barrier		1.96	62.8	44.3		1.87	53.8	27.0	−24.5	2.12	46.8	265.8	
	3523	1.83	32.0	−18.8	−73.7	No barrier		1.98	47.4	28.8	−33.0	1.89	29.6	2.8	−56.1	2.29	22.9	241.9	7.6
	3523H6	1.84	40.4	−10.5	−61.9	2.6	1.4	1.86	57.7	39.1	−70.5	1.87	42.7	15.9	−47.7	1.90	58.5	277.5	−32.8
	5322	1.87	70.8	19.9	−22.7	2.5	2.4	1.87	107.8	89.3	−9.6	1.89	75.6	48.7	−29.8	1.93	109.5	328.5	−44.5
	5323	1.88	60.2	9.3	−25.1	2.5	1.7	1.87	93.5	75.0	−23.9	1.89	62.1	35.3	−33.0	1.93	97.3	316.4	−46.0
	5522	1.83	66.3	15.5	−61.3	2.5	0.1	1.85	84.9	66.4	−68.5	1.86	68.4	41.5	−2.4	1.88	62.5	281.6	−105.1
	5523	1.83	57.1	6.2	−63.6	No barrier		1.85	72.6	54.1	−72.5	1.86	56.4	29.6	−3.4	1.88	53.0	272.0	−106.9
	S6M	1.84	34.2	−16.7	−35.8	No barrier		1.84	28.5	9.9	−40.2	2.19	54.8	27.9	−5.9	2.36	54.1	273.2	62.5
	S6S	1.83	9.4	−41.5	−66.3	No barrier		1.79	0.0	−18.5	−47.0	1.84	1.2	−25.6	−25.5	1.87	0.0	219.0	6.6
	S6M2	1.83	44.9	−6.0	−33.9	No barrier		1.84	41.5	23.0	−67.3	1.85	41.1	14.3	−33.4	1.88	35.6	254.6	−63.8
	C3[a]	1.88	202.3	151.4	−120.4			1.77	67.7	49.2	−131.1	1.81	21.8	−5.0		2.18	−178.9	40.1	−21.7
Fe6–N$_2$	3H[a]	1.80	123.6	72.7	51.5	2.2	79.2	1.79	143.8	125.2	52.0	1.82	149.6	122.7	49.0	1.86	190.9	409.9	43.3
	3322	1.82	72.0	21.1	1.8	2.3	29.4	1.83	112.5	93.9	17.5	1.85	99.6	72.7	20.9	1.87	143.8	362.8	20.6
	3323	1.80	52.6	1.7	−6.4	2.3	31.1	1.79	73.6	55.1	25.6	1.93	48.3	54.5	33.0	1.89	98.1	317.1	58.5
	3323H2	1.82	56.2	5.3				1.85	96.0	77.5		1.88	75.4	48.6					
	3522	1.82	101.4	50.5	19.2	2.2	39.0	1.81	138.2	119.6	39.0	1.84	118.1	91.3	33.0	1.94	158.5	377.5	25.3
	3523	1.82	91.8	41.0	19.2	2.2	38.1	1.81	126.3	107.8	76.3	1.84	108.2	81.3	49.9	1.84	150.6	369.6	26.6

Table 7 *(Contd.)*

	TPSS						r2SCAN				TPSSh				B3LYP			
	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	FeN_{TS}	ΔE_{TS}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}	Fe–N	ΔE	ΔE_{N_2}	ΔE_{db}
3523H2	1.79	58.3	7.5	0.0			1.85[g]	65.8	47.3	−0.8	1.81[b]	52.0	25.2	0.1	2.28[b]	96.5	315.6	−66.7
5322	1.84	71.8	21.0	3.3	2.3	21.1	1.81[b]	101.2	82.7	−12.7	1.99	86.2	59.4	−0.2	3.25[b]	63.8	282.8	−66.6
5323	1.84	63.5	12.6	3.0	2.3	21.0	2.05[b]	79.2	60.7	15.7	1.99	76.1	49.3	13.8	3.25[b]	52.6	271.7	
5522	1.83	94.2	43.3	−4.5	2.2	26.5	1.92	130.7	112.2		1.90	117.1	90.3					
5523	1.83	83.2	32.3	−4.4	2.2	25.8	2.03[b]	90.0	71.4	−46.5	2.11[b]	89.9	63.1	−46.5	1.91	146.4	365.4	6.7
S2F	1.79	15.7	−35.2	−64.9	2.4	14.9	1.81[b]	41.9	23.4	−31.1	1.81[b]	34.9	8.0	−25.2	1.86[b]	67.0	286.0	27.9
S2S	1.79[c]	0.0	−50.9	−80.7	No barrier		1.79[d]	10.7	−7.8	−57.8	1.80[d]	1.4	−25.4	−63.7	1.84[e]	41.3	260.3	−21.0
C3	1.74[a]	201.7	163.1	5.6			1.84[f]	109.5	90.9	−33.0	1.83	0.0	−26.8	−29.9	2.18[a]	−225.3	−6.2	−13.0

a Studied in the BS2-234 state. *b* Studied in the BS10-147 state. *c* Studied in the BS2-234 state. *d* Studied in the BS10-135 state. *e* Studied in the BS7-346 state. *f* Studied in the BS8-347 state. *g* Studied in the BS6-157 state. *h* Studied in the BS10-146 state.

Fe6-S2S   Fe2-S6S   Fe2-3523   Fe2-C3
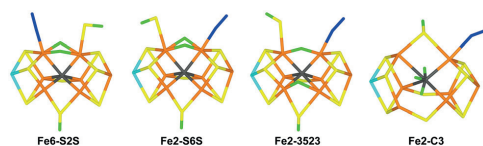
Fig. 8 The best E_4 structures with N_2 bound, Fe6-S2S, Fe2-S6S, Fe2-3523 and Fe2–C3. The first three structures were optimised with TPSS, whereas Fe2–C3 was optimised with B3LYP.

These results can be interpreted in several ways. One interpretation is that TPSS is the only DFT method that gives reasonable binding energies. This is the most direct interpretation of the results, but it is weakened by the fact that TPSSh and r²SCAN give better geometries of the FeMo cluster[83] and that TPSS typically does not give the most accurate energies, neither for main-group molecules,[101] nor for nitrogenase-type reactions[102,103] (but for hydrogenase models, TPSS has been shown to give more accurate results than B3LYP[104,105]).

As already discussed, there are different definitions of the binding energy. ΔE_{N_2} has the advantage of using well-defined states and the most stable structures for both the bound and unbound states. However, it is sensitive to that we really find the best possible structures, as well as spin and BS state of the two structures, which is a formidable task. The direct binding energy, ΔE_{db} in eqn (3) is typically more favourable (negative) than ΔE_{N_2}, mainly because it is not based on the most stable structure for the unbound state. Unfortunately, owing to the non-polar nature of N_2, the unbound complexes are weak and therefore rather poorly defined and quite often the binding is barrierless, so that no unbound structure of the same type is found. Then ΔE_{db} is undefined or has to be based on a structure with a restrained Fe–N distance. Our calculations are based on the philosophy that the relevant states involved in the reaction mechanism are those with the lowest energies (less stable state should have minor populations under normal conditions, unless they are strongly kinetically favoured). However, if we consider the best bound state and use restrained structures when no unbound state is found, we get ΔE_{db} energies of 33, 24, −34, −74 and −81 kJ mol⁻¹ for TPSS, 61, 4, −7, −48 and −47 kJ mol⁻¹ for r²SCAN, 56, 14, −13, −19 and −30 kJ mol⁻¹ for TPSSh and 44, 74, 10, 0 and −13 kJ mol⁻¹ for B3LYP for the E_0–E_4 states. Thus, with the basis-set and entropy corrections, no state is bound for B3LYP and TPSSh, E_3 and E_4 are bound with TPSS, whereas whether E_2 is bound with TPSS, E_3 and E_4 are bound with r²SCAN depends on the size of the entropy correction (*i.e.* bound with Dance's penalty, but not with that of Bjornsson or Siegbahn).

A third interpretation is that N_2 actually does not bind directly to E_4 (or E_3). In fact, it has been suggested that formation of H_2 by reductive elimination is a necessary requisite for the binding of N_2.[54,55] In particular, H_2 should be formed by a reaction between the two hydride ions in E_4 and it has been suggested that thereby a reactive state of E_2 is formed

with two protonated sulfide ions, *i.e.* a state in which the FeMo cluster is formally two steps more reduced than the most stable E_2 state with one proton and one hydride ion (if H on sulfide is considered as a proton and H on Fe as a hydride ion, the two states would formally be $Fe_5^{II}Fe_2^{III}H_2^+$ and $Fe_3^{II}Fe_4^{III}H^+H^-$). However, Dance has argued that this is not supported by QM calculations, showing only minor differences between H atoms on S or Fe[35,106] and calculations of redox potentials support the latter view.[107] The results in Table 2 shows that the best structure with two protons on sulfide ions (N33) is 11–55 kJ mol^{-1} less stable than the most stable E_2 structures with the four DFT methods. Moreover, neither N33 nor the N25 structure shows any enhanced N_2 binding energies (Table 3), irrespectively if using the ΔE_{N_2} (40–67 kJ mol^{-1}, compared to −11 kJ mol^{-1} for Fe6–B33, B35 and H2F with TPSS) or the ΔE_{db} definition (−6 to +11 kJ mol^{-1}, compared to −29 to −76 kJ mol^{-1} Fe6–B33, B35 and H2F with TPSS; qualitatively similar results are obtained also with the other functionals). In a future study, we will study the dissociation of H_2 from the various E_n states, whether N_2 may bind concomitantly with the dissociation of H_2 from E_4 or if a bound H_2 molecule may enhance the binding of N_2.

A fourth possible interpretation of the poor binding with hybrid DFT functionals has been given by Siegbahn, who has suggested that the FeMo cluster needs to be reduced by four more electrons before N_2 may bind favourably.[60,61] This is an interesting suggestion, but the need of such additional reduction steps is not supported by experimental data.[3,4]

Our results are quite similar to those found in other studies. For example, the recent study by Pang and Bjornsson reports N_2 binding energies (calculated with the r^2SCAN functional) of 69, 41, 8 and −17 kJ mol^{-1} to the E_0, E_1, E_2 and E_4 states, respectively,[53] which are similar to our results 57, 41, 15 and −44 kJ mol^{-1}. Our suggestions of the most stable states with or without N_2 bound also reasonably agree, but we find a smaller energy difference (still with the r^2SCAN functional) between the best H6S or S6S structures for the E_2 and E_4 states without N_2 and alternative structures B35 (17 kJ mol^{-1}) or 3523/3323 (25–27 compared to 70–73 kJ mol^{-1}) and also a preference for N_2 binding to Fe2 rather than Fe6. The differences are understandable considering the differences in details of the calculations and the many possible BS states and conformations for the added protons.

is much more stable (and this state is only 4 kJ mol^{-1} less stable than the best by TPSSh). For the E_4 state, the situation is slightly more complicated. Because several structures have comparable energies: With the TPSS, TPSSh and r^2SCAN functionals, the S6S structure is most stable, but the 3323 and 3523 structures are within 3–27 kJ mol^{-1}. The C3 structure is best for B3LYP and degenerate with S6S for TPSSh.

• Binding of N_2 is observed for the E_2–E_4 states. It binds end-on in the exo position (*i.e.* trans to the central carbide) to either Fe2 or Fe6. Typical Fe–N bond lengths are 1.80–1.98 Å.

• Half-dissociation of S2B enhances the binding of N_2, especially to Fe2. As observed before for the E_2 state,[63] the preference for half-dissociation is lower with the TPSS functional, than with the other functionals.

• TPSS consistently favours binding of N_2 to Fe6, whereas the other three functionals mostly prefer binding to Fe2.

• The binding free energy depends on the DFT functional, the entropy correction and on how the binding is defined. With the large entropy correction of Bjornsson and Siegbahn (41–45 kJ mol^{-1}), no functional gives favourable binding to any E_n state. Using the QM/MM energy of the best structure of the same E_n state and free N_2 in a water-like continuum solvent as the reference (ΔE_{N_2}) and Dance's lower entropy penalty (17 kJ mol^{-1}), TPSS gives favourable binding to the E_3 and E_4 states and r^2SCAN only to E_4. B3LYP and TPSSh still give no favourable N_2 binding to any E_n state.

Thus, our results show that computational results for the N_2 binding to the FeMo cluster strongly depends on the DFT method employed, with hybrid functionals giving a weaker binding, favouring binding to Fe2 and protonation of the central carbide. However, it is likely that structures with the S2B ligand dissociated from either Fe2 or Fe6, as well as structures with two hydride ions both bridging Fe2 and Fe6 are involved in the reaction mechanism. On the other hand, we find no support to the suggestion that reductive elimination of the two hydride ions in E_4 would enhance the binding of N_2. Clearly, further studies of the dissociation of H_2 from the FeMo cluster and how it affects the binding of N_2 are needed.

Conflicts of interest

There are no conflicts to declare.

Conclusions

In this work, we have studied the binding of N_2 to the E_0–E_4 states of Mo-nitrogenase with four different DFT methods. This has given a number of interesting and useful results.

• We provide further information about the stability of various structures of the E_3 and E_4 states. We show that S6M and S6S states (with two hydride ions both bridging Fe2/6 and with a protonated S2B ligand dissociated from Fe2) are the best models for the E_3 state with the TPSS, TPSSh and r^2SCAN functionals, but with B3LYP a triply deprotonated carbide ion

Acknowledgements

References

1 B. K. Burgess and D. J. Lowe, *Chem. Rev.*, 1996, **96**, 2983–3012.

2 B. Schmid, H.-J. Chiu, V. Ramakrishnan, J. B. Howard and D. C. Rees, *Handbook of Metalloproteins*, John Wiley & Sons, Ltd, 2006, pp. 1025–1036, DOI: **10.1002/0470028637.met174**.

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

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

5 J. Kim and D. C. Rees, *Science*, 1992, **257**, 1677–1682.

6 O. Einsle, F. A. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard and D. C. Rees, *Science*, 2002, **297**, 1696–1696.

7 T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees and O. Einsle, *Science*, 2011, **334**, 940–940.

8 T. Spatzal, K. A. Perez, O. Einsle, J. B. Howard and D. C. Rees, *Science*, 2014, **345**, 1620–1623.

9 O. Einsle, *J. Biol. Inorg. Chem.*, 2014, **19**, 737–745.

10 R. R. Eady, *Chem. Rev.*, 1996, **96**, 3013–3030.

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

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

13 R. N. F. Thorneley and D. J. Lowe, Molybdenum Enyzymes, ed. T. G. Spiro, Wiley, New York, 1985, pp. 221–284.

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

15 B. Benediktsson and R. Bjornsson, *Inorg. Chem.*, 2017, **56**, 13417–13429.

16 R. Bjornsson, F. Neese and S. DeBeer, *Inorg. Chem.*, 2017, **56**, 1470–1477.

17 S. J. Yoo, H. C. Angove, V. Papaefthymiou and B. K. Burgess, *J. Am. Chem. Soc.*, 2000, **122**, 4926–4936.

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

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

20 K. Fisher, W. E. Newton and D. J. Lowe, *Biochemistry*, 2001, **40**, 3333–3339.

21 D. Lukoyanov, B. M. Barney, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 1451–1455.

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

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

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

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

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

27 D. J. Lowe and R. N. Thorneley, *Biochem. J.*, 1984, **224**, 877–886.

28 R. N. F. Thorneley and D. J. Lowe, *Biochem. J.*, 1984, **224**, 887–894.

29 B. D. Howes, K. Fisher and D. J. Lowe, *Biochem. J.*, 1994, **297**, 261–264.

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

31 Z. Y. Yang, N. Khadka, D. Lukoyanov, B. M. Hoffman, D. R. Dean and L. C. Seefeldt, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 16327–16332.

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

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

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

35 I. Dance, *ChemBioChem*, 2020, **21**, 1671–1709.

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

37 L. Cao, O. Caldararu and U. Ryde, *J. Chem. Theory Comput.*, 2018, **14**, 6653–6678.

38 L. Cao and U. Ryde, *Phys. Chem. Chem. Phys.*, 2019, **21**, 2480–2488.

39 I. Dance, *Chem. Commun.*, 1997, 165–166, DOI: **10.1039/A607136H**.

40 T. H. Rod and J. K. Nørskov, *J. Am. Chem. Soc.*, 2000, **122**, 12751–12763.

41 Z. Cao, Z. Zhou, H. Wan, Q. Zhang and W. Thiel, *Inorg. Chem.*, 2003, **42**, 6986–6988.

42 L. C. Seefeldt, I. G. Dance and D. R. Dean, *Biochemistry*, 2004, **43**, 1401–1409.

43 J. Schimpl, H. M. Petrilli and P. E. Blöchl, *J. Am. Chem. Soc.*, 2003, **125**, 15772–15778.

44 P. P. Hallmen and J. Kästner, *Z. Anorg. Allg. Chem.*, 2015, **641**, 118–122.

45 M. L. McKee, *J. Comput. Chem.*, 2007, **28**, 1342–1356.

46 I. Dance, *Dalton Trans.*, 2019, **48**, 1251–1262.

47 I. Dance, *Dalton Trans.*, 2022, **51**, 15538–15554.

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

49 W. Kang, C. C. Lee, A. J. Jasniewski, M. W. Ribbe and Y. Hu, *Science*, 2020, **368**, 1381–1385.

50 K. L. Skubi and P. L. Holland, *Biochemistry*, 2018, **57**, 3540–3541.

51 T. M. Buscagan and D. C. Rees, *Joule*, 2019, **3**, 2662–2678.

52 A. T. Thorhallsson, B. Benediktsson and R. Bjornsson, *Chem. Sci.*, 2019, **10**, 11110–11124.

53 Y. Pang and R. Bjornsson, *Inorg. Chem.*, 2023, **62**, 5357–5375.

54 S. Raugei, L. C. Seefeldt and B. M. Hoffman, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 10521–10530.

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

56 I. Dance, *J. Am. Chem. Soc.*, 2007, **129**, 1076–1088.

57 I. Dance, *Z. Anorg. Allg. Chem.*, 2015, **641**, 91–99.

58 I. Dance, *Dalton Trans.*, 2021, **50**, 18212–18237.

59 I. Dance, *Dalton Trans.*, 2022, **51**, 12717–12728.

60 P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.*, 2019, **21**, 15747–15759.

61 W.-J. Wei and P. E. M. Siegbahn, *Chem. – Eur. J.*, 2022, **28**, e202103745–e202103745.

62 A. T. Thorhallsson and R. Bjornsson, *Chem. – Eur. J.*, 2021, **27**, 16788–16800.

63 H. Jiang, O. K. G. Svensson and U. Ryde, *Inorg. Chem.*, 2022, **61**, 18067–18076.

64 M. Rohde, D. Sippel, C. Trncik, S. L. A. Andrade and O. Einsle, *Biochemistry*, 2018, **57**, 5497–5504.

65 P. E. M. Siegbahn, *J. Comput. Chem.*, 2018, **39**, 743–747.

66 L. Cao and U. Ryde, *J. Chem. Theory Comput.*, 2020, **16**, 1936–1952.

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

68 L. Cao, O. Caldararu and U. Ryde, *J. Phys. Chem. B*, 2017, **121**, 8242–8262.

69 L. Cao and U. Ryde, *Int. J. Quantum Chem.*, 2018, **118**, e25627 (25616 pages).

70 B. M. Barney, J. McClead, D. Lukoyanov, M. Laryukhin, T.-C. Yang, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, *Biochemistry*, 2007, **46**, 6784–6794.

71 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

72 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

73 L. Cao and U. Ryde, *J. Biol. Inorg. Chem.*, 2020, **25**, 521–540.

74 L. Hu and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 2452–2463.

75 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

76 F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, *Wiley Interdisc. Rev.: Comput. Mol. Sci.*, 2014, **4**, 91–100.

77 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401–146401.

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

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

80 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

81 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.

82 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1372.

83 B. Benediktsson and R. Bjornsson, *J. Chem. Theory Comput.*, 2022, **18**, 1437–1457.

84 A. Schäfer, H. Horn and R. Ahlrichs, *J. Chem. Phys.*, 1992, **97**, 2571–2577.

85 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283–289.

86 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Acc.*, 1997, **97**, 119–124.

87 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2019, **150**, 154122–154122.

88 L. Cao and U. Ryde, *J. Catal.*, 2020, **391**, 247–259.

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

90 D. J. Lowe, R. R. Eady and R. N. F. Thorneley, *Biochem. J.*, 1978, **173**, 277–290.

91 C. Greco, P. Fantucci, U. Ryde and L. D. Gioia, *Int. J. Quantum Chem.*, 2011, **111**, 3949–3960.

92 R. K. Szilagyi and M. A. Winslow, *J. Comput. Chem.*, 2006, **27**, 1385–1397.

93 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 799–805.

94 A. Schäfer, A. Klamt, D. Sattel, J. C. W. Lohrenz and F. Eckert, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2187–2193.

95 A. Klamt, V. Jonas, T. Bürger and J. C. W. Lohrenz, *J. Phys. Chem. A*, 1998, **102**, 5074–5085.

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

97 U. Ryde and M. H. M. Olsson, *Int. J. Quantum Chem.*, 2001, **81**, 335–347.

98 N. Reuter, A. Dejaegere, B. Maigret and M. Karplus, *J. Phys. Chem. A*, 2000, **104**, 1720–1735.

99 L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 761–777.

100 L. Cao and U. Ryde, *Front. Chem.*, 2018, **6**, 89–89.

101 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.

102 I. Dance, *Mol. Simul.*, 2018, **44**, 568–581.

103 M. Torbjörnsson and U. Ryde, *Electron. Struct.*, 2021, **3**, 34005–34005.

104 M. G. Delcey, K. Pierloot, Q. M. Phung, S. Vancoillie, R. Lindh and U. Ryde, *Phys. Chem. Chem. Phys.*, 2014, **16**, 7927–7938.

105 G. Dong, Q. M. Phung, S. D. Hallaert, K. Pierloot and U. Ryde, *Phys. Chem. Chem. Phys.*, 2017, **19**, 10590–10601.

106 I. Dance, *Dalton Trans.*, 2015, **44**, 9027–9037.

107 H. Jiang, O. K. G. Svensson and U. Ryde, *Molecules*, 2023, **28**, 65.

Paper III

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Check for updates

Chemistry
Europe
European Chemical
Societies Publishing

www.chemeurj.org

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

Hao Jiang[a] and Ulf Ryde*[a]

Abstract: We have used combined quantum mechanical and molecular mechanical (QM/MM) calculations to study the reaction mechanism of nitrogenase, assuming that none of the sulfide ligands dissociates. To avoid the problem that there is no consensus regarding the structure and protonation of the E_4 state, we start from a state where N_2 is bound to the cluster and is protonated to N_2H_2, after dissociation of H_2. We show that the reaction follows an alternating mechanism with HNNH (possibly protonated to $HNNH_2$) and H_2NNH_2 as intermediates and the two NH_3 products dissociate at the E_7 and E_8 levels. For all intermediates, coordination to

Fe6 is preferred, but for the E_4 and E_8 intermediates, binding to Fe2 is competitive. For the E_4, E_5 and E_7 intermediates we find that the substrate may abstract a proton from the hydroxy group of the homocitrate ligand of the FeMo cluster, thereby forming $HNNH_2$, H_2NNH_2 and NH_3 intermediates. This may explain why homocitrate is a mandatory component of nitrogenase. All steps in the suggested reaction mechanism are thermodynamically favourable compared to protonation of the nearby His-195 group and in all cases, protonation of the NE2 atom of the latter group is preferred.

Introduction

Nitrogen is an essential element of all lifeforms, being a component of all amino acids and nucleic acids. Although the atmosphere of Earth contains 78% of N_2, nitrogen is still a limiting element for plant growth and a prominent component of fertilizers. The reason for this is the strong triple bond in N_2, which makes it chemically inert.[1,2] Industrially, N_2 is converted to ammonia through the Haber–Bosch process, which requires high temperature and pressure.[2] Only a single group of enzymes can cleave the N–N bond in N_2, the nitrogenases (EC 1.18/19.6.1), which work at ambient temperature and pressure.[1,3,4]

Crystallographic studies have shown that the most active type of nitrogenase contains a $MoFe_7S_9C$(homocitrate) cluster (the FeMo cluster) in the active site, connected to the protein by a histidine and a cysteine residue at the opposite ends of the cluster (Figure 1).[5–9] There also exist alternative nitrogenases with the Mo ion replaced with either vanadium or iron, which have lower activities towards N_2.[10]

The nitrogenases catalyse the reaction

$$N_2 + 8\,e^- + 8\,H^+ + 16\,ATP \rightarrow 2\,NH_3$$
$$+\,H_2 + 16\,ADP + 16\,P_i \tag{1}$$
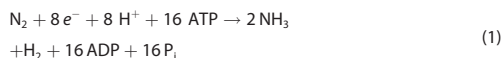
The mechanism is normally discussed in terms of nine intermediates E_0–E_8, differing in the number of added electrons and protons, according to the Lowe–Thorneley scheme.[11] Thorough biochemical, kinetic and spectroscopic studies have indicated that the resting E_0 state needs to be reduced to the E_4 state before N_2 may bind.[1,3,4,12–17] It has also been suggested that H_2 formation through reductive elimination is a prerequisite for the binding of N_2, explaining why H_2 is a compulsory byproduct in the reaction. It is normally assumed that N_2 is directly reduced and protonated to N_2H_2 upon binding to the enzyme.[1,18]

It has long been debated whether the nitrogenases follow a sequential or alternating reaction mechanism. In the sequential mechanism, the first three protons bind to the same N atom of N_2, which then dissociates as NH_3 from the E_5 intermediate, before the second N atom starts to be protonated. This mechanism was originally suggested by Chatt and has gained support from inorganic model complexes.[19–23] In the alternating mechanism, the protons are instead added alternatively to the two N atoms, so that HNNH and H_2NNH_2 (hydrazine) are intermediates and the first NH_3 product does not dissociate until the E_7 state. It is supported by the fact that nitrogenase can use hydrazine as a substrate and that hydrazine is released upon acid or base hydrolysis of the enzyme during turnover.[1,3,24,25] Moreover, it has been shown that N_2, N_2H_2, CH_3NH_2 and N_2H_4 all react via a common intermediate.[1,26] The nitrogenases have been thoroughly studied also by computational methods.[1,13,35–42,27–34] Unfortunately, these studies have given very diverging and disparate suggestions. In fact,

[a] H. Jiang, U. Ryde
Department of Theoretical Chemistry
Lund University
Chemical Centre, P. O. Box 124, 221 00 Lund (Sweden)
E-mail: Ulf.Ryde@teokem.lu.se

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202103933

Part of the Chemistry Europe joint Special Collection on Quantum Bioinorganic Chemistry.

Chem. Eur. J. 2022, 28, e202103933 (1 of 18)

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
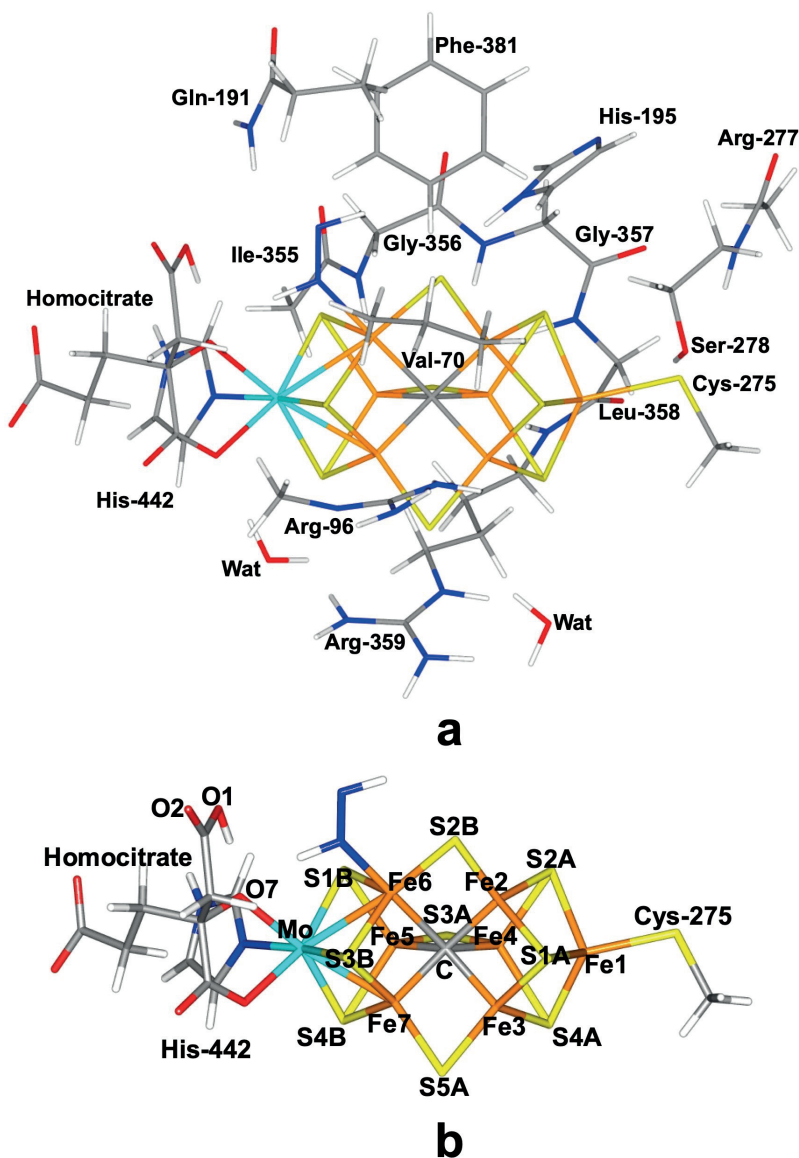European Chemical
Societies Publishing

Figure 1. Structure of the FeMo cluster (with *trans*-HNNH bound to Fe6), illustrating also the QM system used in all calculations, as well as the names of the nearby residues (a). (b) shows only the FeMo cluster with atom names indicated.

there is not even any consensus about the structure of the key E_4 intermediate. Important reasons for this are that different density-functional theory (DFT) methods give very different predictions of the relative stability of various intermediates, with differences of 600 kJ/mol[43] and that there are very many

possibilities for the structures and electronic states of the intermediates.[44,45]

Hoffman and coworkers have suggested a structure of the E_4 intermediate with two hydride ions bridging the Fe2 and Fe6 ions, as well as the Fe3 and Fe7 ions, and with two protons on

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

the S2B and the S5A sulfides, all positioned on the same face of the FeMo cluster[15,46] (the name of the various Fe and sulfide ions are shown in Figure 1b). They have shown that this structure is lower in energy than a few other structures and that it may bind N_2 after reductive elimination of the two hydride ions, leaving the cluster in a doubly reduced state.[15,41,47,48]

On the other hand, Siegbahn has suggested that the FeMo cluster needs to be reduced by four electrons from the resting state before the true E_0 state is reached, which involves a triply protonated central carbide and a strongly distorted cluster.[49] Then, this state is reduced by another four electrons to reach the E_4 state, from which H_2 dissociates and N_2 binds, bridging two Fe ions. It is successively protonated in a manner that is a mixture of the alternating and sequential mechanism, involving NNH_2, $HNNH_2$, H_2NNH_2, but also $HNNH_3$. The first NH_3 dissociates at the E_7 level.

Dance has presented a mechanism in which E_4 contains two terminal hydride ions on Fe2 and Fe6, and two protons on S2B and S3B. N_2 then binds side-on to Fe6, without any dissociation of H_2 and is alternatively protonated to H_2NNH_2, at which level the N–N bond is cleaved, forming two NH_2 fragments on Fe2 and Fe6.[27,50,51]

On the other hand, Nørskov and coworkers have suggested a mechanism in which the E_0 state is doubly protonated and a sulfide ligand dissociates from the cluster during the reaction mechanism.[37] This forms a binding site, where N_2 binds in an end-on fashion, bridging two Fe ions and it is then sequentially protonated on the outer N atom. The dissociation of the sulfide ion was inspired by several crystallographic studies of both Mo and V nitrogenase, showing that the S2B group can be replaced by several other ligands, for example CO, OH⁻ and Se.[8,52–56]

Recently, we studied a similar mechanism, involving dissociation of S2B.[57] We used a larger and more realistic model system, which was studied with the combined quantum mechanical and molecular mechanics (QM/MM) approach. Our study indicated that the conversion of N_2H_2 to two NH_3 molecules is thermodynamically favourable, but it follows a mainly alternating pathway (although the first intermediate involved a bridging NNH_2 group, which is normally connected to a sequential mechanism).

Naturally, such studies do not prove that the nitrogenase mechanism actually involves a dissociated S2B group. To that end, it must be shown that the replacement of S2B by N_2 is energetically favourable, which has been questioned by Dance.[58] Moreover, it should be shown that a reaction without replacement of S2B is not possible or at least is less favourable. Here, we make an investigation of the latter reactions, that is, the formation of ammonia from bound N_2 for Mo nitrogenase without dissociation of S2B. We show that also such a reaction is possible and thermodynamic favourable, following an alternating mechanism.

Methods

The protein

The calculations were based on the 1.0-Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[7] The setup of the protein is identical to that of our previous studies.[43,59–61] The entire heterotetramer was considered in the calculations, because the various subunits are entangled without any natural way to separate them. The quantum mechanical (QM) calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule from the solvent rather close to the active site (~ 11 Å) in the A subunit. The two P-clusters and the FeMo cluster in subunit A were modelled by MM in the fully reduced and resting states, respectively, using a QM charge model.[59]

The protonation states of all residues were the same as before:[59] All Arg, Lys, Asp and Glu residues were assumed to be charged, except Glu-153, 440 and 231D (a letter "D" after the residue number indicates that it belongs to that subunit; if no letter is given, it belongs to subunit C; subunits A and B are identical to the C and D residues). Cys residues coordinating to Fe ions were assumed to be deprotonated. His-274, 451, 297D, 359D and 519D were assumed to be protonated on the ND1 atom, His-31, 196, 285, 383, 90D, 185D, 363D and 457D were presumed to be protonated on both the ND1 and NE2 atoms (and therefore positively charged), whereas the remaining 14 His residues were modelled with a proton on the NE2 atom. The homocitrate ligand was modelled in the singly protonated state with a proton shared between the hydroxy group (which coordinates to Mo) and the O1 carboxylate atom. This protonation state was found to be the most stable one in an extensive QM/MM, molecular dynamics and quantum-refinement study[59] and this protonation state is also supported by another QM/MM study.[62]

The protein was solvated in a sphere with a radius of 65 Å around the geometrical centre of the protein. 160 Cl⁻ and 182 Na⁺ ions were added at random positions (but not inside the protein[59]) to neutralise the protein and give an ionic strength of 0.2 M.[63] The final system contained 133 915 atoms. The added protons, counter ions and water molecules were optimised by a simulated annealing calculation (up to 370 K), followed by a minimisation, keeping the other atoms fixed at the crystal-structure positions.[59]

All MM calculations were performed with the Amber software.[64] For the protein, we used the Amber ff14SB force field[65] and water molecules were described by the TIP3P model.[66] For the metal sites, the MM parameters were the same as in our previous investigation.[59] The metal sites[45,59] were treated by a non-bonded model[67] and charges were obtained with the restrained electrostatic potential method, obtained at the TPSS/def2-SV(P) level of theory[68,69] and sampled with the Merz–Kollman scheme.[70]

The FeMo cluster was modelled by $MoFe_7S_9C$-(homocitrate)(CH_3S) (imidazole), where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included in the

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

QM model, viz. Arg-96, Gln-191 and His-195 (sidechains), Ser-278 and Arg-359 (both backbone and sidechain, including the Cα and C and O atoms from Arg-277), Gly-356, Gly-357 and Leu-358 (backbone, including the Cα and C and O atoms from Ile-355), as well as two water molecules. Finally, Phe-381 and Val-70 were also included because they are close to the putative N_2 binding site and therefore may affect the binding of the substrate. The QM system involved 183–190 atoms in total (depending on the number of added protons and N atoms) and is shown in Figure 1a. The net charge of QM region was −3.

QM calculations

All QM calculations were performed with the Turbomole software (version 7.5).[71] All structures were studied with both the TPSS[68] and B3LYP[72–74] functionals with def2-SV(P) basis set.[69] The most stable states were examined also with the larger def2-TZVPD basis set. The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[75,76] Empirical dispersion corrections were included with the DFT-D4 approach,[77] as implemented in Turbomole. All minima were fully optimized without any restraints. Transition states (for N–N cleavage and NH_3 dissociation) were determined as the highest point on the potential energy surface along the reaction coordinates, which were scanned with a step of 0.1 Å near the transition states.

Experiments have shown that the ground spin state of E_4 is a doublet,[1,78] and we used this state for E_4 models. For the other E_n states, we used mainly the doublet or triplet states, but for the most interesting structures, we checked which of the two or three lowest spin states has the most favourable energy at the TPSS and B3LYP/def2-SV(P) levels of theory.

The electronic structure of all QM calculations was obtained with the broken-symmetry (BS) approach:[79] Each of the seven Fe ions were modelled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways $\left(\frac{7!}{3!\,4!}\right)$.[60] The various BS states were obtained either by swapping the coordinates of the Fe ions[80] or with the fragment approach by Szilagyi and Winslow.[81] The various BS states are named by listing the number in the Noodleman nomenclature (BS1—10),[79] followed by the numbers of the three Fe ions with minority spin (however, in the tables, only the latter three numbers are given). Most structures were studied in the BS10-147 state, i.e. with β spin on Fe1, Fe4 and Fe7, because it was found to be lowest in both this and in our previous study.[57] However, sometimes the calculations converged to other states (especially BS7-235). For twelve of the most stable structures, the relative stabilities of all 35 states were examined (with structures fully optimised for each BS state). Moreover, for all structures within 20 kJ/mol of the most stable structure at each E_n level, the BS7-235 state was also studied.

As have been discussed before,[43,60] TPSS/def2-SV(P) calculations give geometries that reproduce the crystal structure of the resting state of nitrogenase excellently with average and

maximum deviations of 0.05 and 0.09 Å for the metal–metal distances, and 0.02 and 0.06 Å for metal—ligand distances, and a root-mean-squared-deviation (RMSD) of 0.06 Å for the metals and the first-sphere ligands. This is similar to the results obtained with the TPSSh[62] approach and appreciably better than with the B3LYP/def2-SV(P) method, which gives average and maximum deviations of 0.08 and 0.12 Å for the metal–metal and 0.04 and 0.11 Å metal–ligand distances, respectively and a RMSD of 0.08 Å. Therefore, we discuss primarily the TPSS/def2-SV(P) results.

QM/MM calculations

QM/MM calculations were performed with the ComQum software.[82,83] In this approach, the protein and solvent are split into three subsystems: System 1 (the QM region) was relaxed by QM methods. System 2 contained all residues and water molecules with at least one atom within 6 Å of any atom in system 1 and it was optionally relaxed by MM. It included residues 49, 59–74, 92, 95–98, 189–199, 226–231, 234, 235, 253–255, 273–282, 300, 353–355, 358–364, 377–383, 385, 386, 401 422–427, 438, 440–444, 450 and 451 from subunit C and residues 93, 97, 98, 101 and 105 from subunit D, in total 94 residues and 39 water molecules). Finally, system 3 contained the remaining part of the protein and the solvent, and it was kept fixed at the original coordinates (equilibrated crystal structure to avoid the risk that different calculations end up in different local minima). The total system was spherical and non-periodic with 133 915 atoms. Most calculations were performed without relaxing system 2, but for the most interesting structures, calculations with relaxed surroundings were also performed. The effect of the relaxed surroundings are described in the Supporting Information.

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM system was capped with hydrogen atoms (hydrogen link atoms, HL), the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[82,84] All atoms were included in the point-charge model, except the CL atoms.[85]

The total QM/MM energy in ComQum is calculated as[82,83]

$$E_{QM/MM} = E_{QM1+ptch23}^{HL} + E_{MM123,\,q_1=0}^{CL} - E_{MM1,\,q_1=0}^{HL} \qquad (2)$$

in which $E_{QM1+ptch23}^{HL}$ is the QM energy of the QM system truncated by HL atoms and embedded in the set of point charges modelling systems 2 and 3 (but excluding the self-energy of the point charges). $E_{MM1,\,q_1=0}^{HL}$ is the MM energy of the QM system, still truncated by HL atoms, but without any electrostatic interactions. Finally, $E_{MM123,\,q_1=0}^{CL}$ is the classical energy of all atoms in the system with CL atoms and with the

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

**Chemistry
Europe**
European Chemical
Societies Publishing

charges of the QM region set to zero (to avoid double-counting of the electrostatic interactions). Thus, ComQum employs a subtractive scheme with van der Waals link-atom corrections.[86] No cut-off is used for any of the interactions in the three energy terms in Equation (3).

The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J/mol (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u.

QM/MM calculations give comparable energies only if they contain exactly the same number of electrons and atoms of each element in both the QM and MM systems. Therefore, we compare only structures within the same E_n level. On the other hand, it means that we can study proton transfers within the QM system, for example from the homocitrate ligand or from His-195 to the substrate. For each transition from E_n to E_{n+1}, an electron and a proton is added to the QM system, and we compare the energies of structures with this proton in different positions.

Result and Discussion

In this investigation, we study the later part of the reaction mechanism of nitrogenase, assuming that the S2B ligand does not dissociate. We describe in separate sections states at different oxidation levels, from E_4 to E_8.

N$_2$-bound E$_4$ structures

We start with the N_2-bound E_4 state. As in our previous study of the reaction mechanism with a dissociated S2B ligand,[57] we avoid the problem there is no consensus regarding the protonation of the E_4 state[27,38,43,59,61,87–89] by starting from a state where N_2 has already bound to the cluster and is protonated to N_2H_2. The immediate protonation of the substrate upon binding is normally assumed,[1,18] although it has not been experimentally observed. Mutations and other studies have shown that the substrate most likely bind either the Fe2 or Fe6 ions of the FeMo cluster[1,90,91] and this is also supported by a systematic scan of all possible N_2H_2 binding positions.[45] Therefore, we look for the best structure with N_2H_2 bound to either Fe2 or Fe6, or

to both. In the latter case, the substrate may bind on two different sides of the bridging S2B ligand and we call these structures Fe2/6(3) and Fe2/6(5), depending on whether it is on the same side as S3A or S5A. Likewise, when N_2H_2 binds only to Fe2, the non-bonding N atom point either to the S3A or S5A sides of the cluster, which will be called Fe2(3) or Fe2(5) (and similar for binding to Fe6). We have considered three isomers of N_2H_2, viz. NNH$_2$, cis-HNNH or trans-HNNH (the latter two are abbreviated cHNNH or tHNNH in the following). Moreover, we have tested three protonation states of His-195: with protons on either ND1 (HID) or NE2 (HIE) or on both (HIP; but this adds an extra proton to the system and was therefore studied for the next higher E_n level). The results are collected in Table 1 and are shown Figure S1.

It can be seen that in nearly all structures, the HIE protonation state was more favourable than HID protonation by 21–132 kJ/mol. The only exception was the Fe2/6(3)-cHNNH state with B3LYP. All structures were studied in the BS10-147 state, but sometimes it shifted to the BS7-235 state (the spin on Fe6 is often small and may change sign; the latter state was also studied for all low-energy structures). However, BS10-147 was always 9–37 kJ/mol lower in energy than BS7-235 when both states were found, except for the Fe6-cHNNH(3) state, for which BS7-235 was 1 kJ/mol more stable at both the TPSS and B3LYP levels.

The most favourable structure has trans-HNNH end-on bound to Fe6 (Figure 2a). The Fe6-N distance is 1.91 Å and the N–N bond length is 1.26 Å, which is slightly longer than in isolated trans-HNNH, optimised with the same level of theory, 1.25 Å. The Mulliken spin populations are (in absolute terms) 3.2–2.7 e on the seven Fe ions, except Fe6, which has only 1.6 e (cf. Table S1 in the Supporting Information). Mo has a population of -0.3 e. This structure is stabilised by a hydrogen bond from the HN group bound to Fe6 to the alcohol O7 atom of homocitrate (which coordinates to Mo), with a H···O distance of 1.95 Å (cf. Figure 2a). The other N atom of the substrate receives a hydrogen bond from the HE1 atom of Gln-191 (the HE1···N distance is 2.17 Å) and the other H atom of the substrate is directed towards S2B, with a H···S distance of 2.29 Å, but the N-H···S angle is only 124°. S2B receives another hydrogen bond from the HE2 atom of His-195 (2.13 Å, with a more ideal geometry). The corresponding structure in the quartet state is 30 kJ/mol less stable with the TPSS functional, but 7 kJ/mol
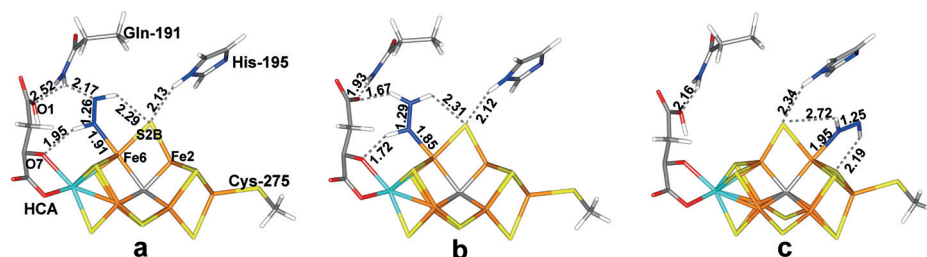

Figure 2. The best E$_4$ structures: (a) Fe6-tHNNH, (b) Fe6-HNNH$_2$ and (c) Fe2-tHNNH(3), all with the HIE state of His-195.

Chem. Eur. J. **2022**, *28*, e202103933 (5 of 18)

© 2022 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

Table 1. Energies (in kJ/mol), N–N and Fe–N distances (in Å) of the various structures of the E_4 states. All states were studied in the $S = \frac{1}{2}$ state unless otherwise stated. His is the protonation state of His-195. The BS state is identified by the Fe ions with minority spin (e.g. 147 = BS10-147). Up to four different energies are listed: TPSS-D4/def2-SV(P) (TP), B3LYP-D4/def2-SV(P) (B3), TPSS-D4/def2-TZVPD (TZ), all based on TPSS-D4/def2-SV(P) structures obtained with fixed surroundings, as well as TPSS-D4/def2-SV(P) with relaxed surroundings (Rlx). When multiple Fe–N distances are given, Fe2 comes before Fe6.

Structure	His	BS	TP	B3	TZ	Rlx	N-N	Fe-N
Fe2-cHNNH(3)	HID	147	139	126			1.26	1.92
	HIE	147	35	21	55	28	1.26	1.92
Fe2-cHNNH(5)	HIE	147	91	75			1.25	1.92
		235	124	97			1.25	1.98
Fe6-cHNNH(3)	HID	147	182	186			1.26	1.91
	HIE	147	55	54			1.25	1.91
		235	54	52			1.26	1.86
Fe6-cHNNH(5)	HID	147	133	125			1.25	1.95
	HIE	147	48	37	58	78	1.26	1.92
Fe2/6-cHNNH(3)	HID	147	227	222			1.31	1.88,1.90
	HIE	147	206	239			1.38	1.97,1.90
Fe2/6-cHNNH(5)	HID	147	220	229			1.30	1.92,1.88
	HIE	147	161	199			1.35	1.98,1.89
Fe2/6-cHNNH[a]	HIE	147	204	237			1.36	1.91,1.88,2.05
Fe2/6-cHNNH[b]	HIE	147	152	139			1.36	1.94,1.93,1.95
Fe2-tHNNH(3)	HID	147	108	94			1.26	1.94
	HIE	147	20	5	22	34	1.25	1.95
		235	51	30			1.25	2.01
Fe2-tHNNH(5)	HID	147	176	94			1.25	1.95
	HIE	147	61	48			1.25	1.93
Fe6-tHNNH	HID	147	86	86			1.26	1.91
	HIE	147	0	0	0	0	1.26	1.91
		235	10	10			1.26	1.87
	$S = 3/2$	147	30	−7			1.26	1.94
Fe2-NNH$_2$(3)	HID	147	194	192			1.27	1.83
	HIE	147	66	42			1.26	1.89
Fe2-NNH$_2$(5)	HID	147	222	211			1.25	1.91
	HIE	147	99	82			1.25	1.89
		235	129	119			1.25	1.91
Fe2/6-NNH$_2$(3)[c]	HID	235	244	275			1.32	1.83,1.77
	HIE	235	113	166			1.38	1.82,1.79
Fe2/6-NNH$_2$(5)[c]	HID	147	255	217			1.30	1.85,1.80
	HIE	147	108	195			1.36	1.80,1.77
Fe6-HNNH$_2$	HID	147	87	75			1.26	1.85
	HIE	147	4	-6	4	3	1.29	1.85
		235	18	3			1.29	1.84

[a] S2B has dissociated from Fe6; One atom of N$_2$ bridges Fe2/6, the other binds to Fe6. [b] S2B has dissociated from Fe2; One atom of N$_2$ bridges Fe2/6, the other binds to Fe2. [c] S2B has dissociated from Fe6.

more stable with the B3LYP. A full investigation of all 35 BS states, collected in Table S2, shows that the BS10-147 state is indeed the most stable BS state, 10–67 kJ/mol more stable than the other BS states (BS7-235 second lowest).

A structure with NNH$_2$ bound end-on to Fe6 is only 4 kJ/mol less stable with TPSS (also with the larger def2-TZVPD basis set and 3 kJ/mol with relaxed surroundings) and it is actually 6 kJ/mol more stable with B3LYP. In this structure, the substrate has abstracted the alcohol proton from homocitrate, giving HNNH$_2$ (Figure 2b). It has a Fe6-N distance of 1.85 Å and a N–N bond length of 1.29 Å (1.21 Å in neutral NNH$_2$ and 1.23 Å in HNNH$_2^+$ optimised with the same method). It is stabilised by hydrogen bonds from the Fe6-bound NH group to the alcohol O7 atom of homocitrate (H···O7 = 1.72 Å), from the NH$_2$ to one of the carboxylic groups of homocitrate (H···O1 = 1.67 Å) and from the other atom of NH$_2$ group to S2B (H···S2B = 2.31 Å; Figure 2b). The S2B atom also receives a hydrogen bond from His-195 (HE2···S2B = 2.12 Å) and the O1 atom of homocitrate receives another hydrogen bond from Gln-191 (HE1···O1 = 1.93 Å). This

structure has also a low spin population on Fe6 (1.8 e). Again, a full investigation of all 35 BS states showed that BS10-147 is the most favourable BS (Table S2), although it is only 4 kJ/mol more stable than BS10-135 (14 kJ/mol more stable than BS7-235).

A structure with *trans*-HNNH binding to Fe2, directed to the S3A side (Figure 2c), is 20 kJ/mol less stable than Fe6-tHNNH structure (22 kJ/mol with the larger basis set, 34 kJ/mol with relaxed surroundings and 5 kJ/mol with B3LYP). The Fe2-N distance is 1.95 Å and the N–N bond length is 1.25 Å. The two H atoms of the substrate point towards S2B (H···S2B = 2.72 Å) and S1A (H···S1A = 2.19 Å), but the N–H···S angles are far from straight (83° and 131°, respectively). The Fe spin population on Fe2 (2.1 e) is only slightly less than on the other Fe ions (3.2–2.4 e in absolute terms). A full investigation of all 35 BS (Table S2) shows that BS10-147 indeed is the most stable state. The structure with the substrate directed to the S5A side is 41 kJ/mol less stable.

Other structures studied are appreciably less stable. Structures with *cis*-HNNH bound side-on to Fe2 and Fe6 are 161–

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

**Chemistry
Europe**
European Chemical
Societies Publishing

239 kJ/mol less stable than Fe6-tHNNH, structures with *cis*-HNNH bound end-on to either Fe2 or Fe6 are 35–182 kJ/mol less stable and structures with NNH_2 bound to Fe2 are 66–255 kJ/mol less stable.

We also considered structures with S2B dissociated from either Fe2 or Fe6 (but still bound to the other ion), because such structures have been suggested to be competitive by other authors.[35,42,48,58] However, with our methods, such structures were always high in energy, by 108–255 kJ/mol for structures with NNH_2 bridging Fe2 and Fe6 with the terminal N atom and by 152–204 kJ/mol for structures with *cis*-HNNH bridging Fe2 and Fe6.

The present results are somewhat different from those obtained in our previous study of the binding of N_2H_2 to the FeMo cluster.[45] In that study, we found that the structure with *trans*-HNNH bound to Fe2 was 10 kJ/mol more stable than the Fe6-tHNNH state and 9 kJ/mol more stable than the Fe6-HNNH$_2$ state (19 and 3 kJ/mol with the larger def2-TZVPD basis set). The difference is most likely connected to the larger QM system used in the present study (models of Val-70, Gln-191 and Phe-381, all situated around the binding site, were not included in the previous study). The present results should be more reliable.

The results also differ from those obtained with a dissociated S2B ligand (and a rotated conformation of Gln-191),[57] for which a structure with NNH_2 bridging Fe2 and Fe6 was found to be most favourable. Such structures were at least 108 kJ/mol less stable than Fe6-tHNNH in this study and led to half-dissociation of S2B. Clearly, the active site with a bridging S2B group is so crowded that it disfavours structures with N_2H_2 simultaneously bridging Fe2 and Fe6.

Bjornsson and coworkers studied structures with *trans*-HNNH bound to the FeMo cluster in the E_4 state.[42] They also found that binding to Fe6 was more favourable than to Fe2, in agreement with our results, but the difference was larger, 69 kJ/mol. In their models, S5A was protonated and one of the protons on HNNH was abstracted from homocitrate.

E$_5$ structures

Next, we added an electron and a proton to the FeMo cluster (i.e. to the QM system) to obtain structures at the E_5 level. They were studied in the triplet state with BS10-147. The structures are described in Table 2 and the best are shown in Figures 3 and S2. As for the E_4 structures, HIE structures were always more stable than the corresponding HID structures, by 16–173 kJ/mol.

The most favourable state has H_2NNH_2 (hydrazine) bound to Fe6, where the extra proton is abstracted from the hydroxy group of homocitrate (Figure 3a). The Fe6-N distance is 2.09 Å and the N–N bond length is 1.43 Å, which is the same as for isolated hydrazine, optimised at the same level of theory. The non-coordinating NH_2 group is directed to the S3A side of the cluster. The two H atoms of the Fe-bound NH_2 group forms hydrogen bonds to O7 of homocitrate (H···O7 = 1.94 Å) and S3B (H···S3B = 2.73 Å). The other two H atoms of hydrazine point towards O1 of homocitrate (H···O1 = 2.66 Å) and S2B (H···S2B = 2.34 Å). The spin density on Fe6 is 2.1 e (Table S3). The singlet state was 16 kJ/mol more stable than the triplet state with TPSS, but 38 kJ/mol less stable with B3LYP. The quintet was 4–33 kJ/mol less stable than the triplet. A full investigation of all BS states (Table S2) showed that BS7-235 is actually 9 kJ/mol lower in energy than BS10-147. In fact, nine different BS states were found within 11 kJ/mol of the lowest state. There are several structures with similar energies with slight variations in the hydrogen-bond lengths and the relative conformations of the two NH_2 groups. For example, a structure with the non-bonded NH_2 group directed towards S5A is only 1 kJ/mol less stable (6 kJ/mol more stable by B3LYP, but 38 kJ/mol less stable with relaxed surroundings).

Other structures are appreciably less stable. The second-best structure had HNNH$_3$ bound end-on to Fe6 with the NH group and with the NH$_3$ group directed toward S5A (Figure 3b; again with a proton abstracted from homocitrate). It is 54–68 kJ/mol less stable than the Fe6-H$_2$NNH$_2$ structure at the
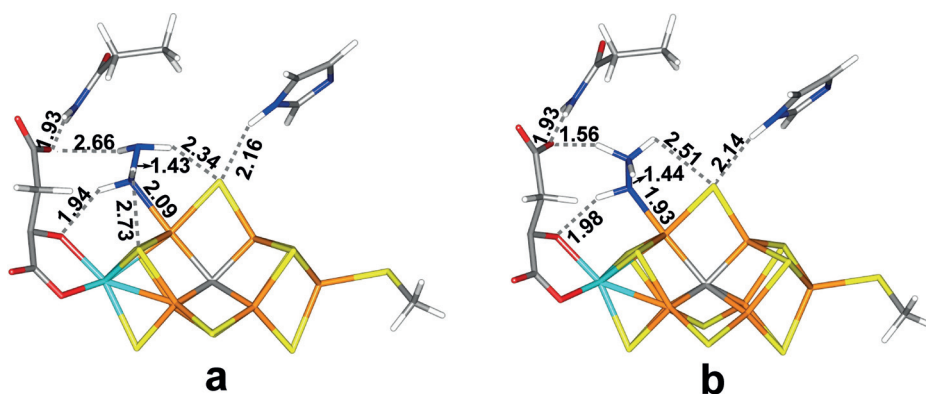

Figure 3. The best E$_5$ structures: (a) Fe6-H$_2$NNH$_2$(3) and (b) Fe6-HNNH$_3$(5) both with the HIE state of His-195.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing

Table 2. Energies (in kJ/mol), N–N and Fe–N distances (in Å) of the various structures of the E_5 states. All states were studied in the $S=1$ state unless otherwise stated. The entries are the same as in Table 1.

Structure	His	BS	TP	B3	TZ	Rlx	N-N	Fe-N
Fe2-HNNH$_2$(3)	HID	147	209	209			1.37	1.92
	HIE	147	80	76	94	54	1.36	1.94
Fe2-HNNH$_2$(5)	HID	147	272	186			1.34	1.95
	HIE	147	131	126			1.35	1.94
Fe6-HNNH$_2$	HID	147	205	152			1.38	1.39
	HIE	147	75	82	84	66	1.38	1.88
	$S=0$	147	86	123			1.39	1.59
	$S=2$	147	122	115			1.38	1.88
Fe2/6-HNNH$_2$(3)	HID	147	320	303			1.42	1.98,1.90
	HIE	147	247	250			1.45	1.95,2.00
Fe2/6-HNNH$_2$(5)	HID	147	305	324			1.44	2.08,1.90
	HIE	147	248	291			1.45	2.06,1.93
Fe2/6-H$_2$NNH(3)	HID	167	272	298			1.44	1.93,1.98
	HIE	147	223	217			1.45	2.02,1.98
Fe2/6-H$_2$NNH(5)	HID	147	337	366			1.43	1.90,2.04
	HIE	147	182	193			1.45	2.00,2.02
Fe6-H$_2$NNH$_2$(3)	HID	147	130	129			1.43	2.10
	HIE	147	0	0	0	0	1.43	2.09
		235	−9	−3			1.43	2.08
	$S=0$	147	−16	38			1.43	2.11
		235	15	37			1.43	2.06
	$S=2$	147	33	4			1.43	2.09
Fe6-H$_2$NNH$_2$(5)	HID	147	128	121			1.43	2.14
	HIE	147	1	−6	−1	38	1.43	2.13
		235	0	−1			1.43	2.15
Fe2-NNH$_3$	HID	147	344	334			1.45	1.86
	HIE	147	196	184			1.44	1.84
Fe2/6-NNH$_3$(5)	HID	147	407	471			1.46	1.92,1.91
	HIE	147	301	367			1.47	1.93,1.91
Fe6-HNNH$_3$(3)	HID	147	215	213			1.43	1.93
	HIE	147	80	75	87	72	1.43	1.92
Fe6-HNNH$_3$(5)	HID	147	151	148			1.44	1.93
	HIE	147	67	65	68	54	1.44	1.93
Fe2-cHNNH(5)	HIP	147	333	344			1.25	1.91
Fe6-cHNNH(3)	HIP	147	300	290			1.27	1.91
Fe6-cHNNH(5)	HIP	147	279	273			1.27	1.88
Fe2-tHNNH(3)	HIP	147	259	76			1.28	1.99
Fe2-tHNNH(5)	HIP	147	295	322			1.25	1.93
Fe6-tHNNH	HIP	147	239	236			1.25	1.89
Fe2-NNH$_2$(3)	HIP	147	323	353			1.32	1.90
Fe2-NNH$_2$(5)	HIP	147	353	363			1.26	1.88
Fe2/6-NNH$_2$(3)	HIP	147	379	265			1.42	1.82,1.89
Fe2/6-NNH$_2$(5)	HIP	147	430	417			1.30	1.85,1.81
Fe6-HNNH$_2$	HIP	147	233	170			1.30	1.87

various levels of theory. A similar structure with the NH$_3$ group directed towards S3A is 10–18 kJ/mol less stable. A structure with HNNH$_2$ bound end-on to Fe6 with the NH group (i.e. without the proton transfer from homocitrate) is 66–84 kJ/mol less stable than the Fe6-H$_2$NNH$_2$ structure. A structure with HNNH$_2$ bound to Fe2 with the NH$_2$ group directed towards S3A is 54–94 kJ/mol less stable than the best structure. The structure with the NH$_2$ group pointing in the opposite direction is 50 kJ/mol less stable. Structures with NNH$_3$ bound to Fe2 are 184–196 kJ/mol less stable than the best Fe6-HNNH$_3$ structure.

We also studied a number of structures with HNNH$_2$ or NNH$_3$ bridging Fe2 and Fe6 on either side of S2B, but all of them were high in energy (182–471 kJ/mol less favourable than Fe6-HNNH$_3$). Moreover, we studied structures with a proton on His-195, instead of on the substrate, giving the HIP state. These structures were at least 233–430 kJ/mol higher than Fe6-H$_2$NNH$_2$-HIE, with Fe6-HNNH$_2$ lowest, 6 kJ/mol lower than Fe6-

tHNNH. This may illustrate a possible path for the transfer of protons to the substrate and it is apparently strongly downhill.

For the best Fe6-HNNH$_3$ structure, we tested to cleave the N–N bond. However, this reaction turned out to prohibitive with an activation barrier of 119 kJ/mol and a reaction energy of 78 kJ/mol to a product with NH bound to Fe6 and NH$_3$ dissociated from the cluster, but forming hydrogen bonds to O1 of homocitrate, S2B and S3B. This indicates that nitrogenase does not follow a sequential reaction mechanism. On the other hand, the N–N bond can be cleaved in Fe2-NNH$_3$, with a barrier of 49 kJ/mol, but the reactant and the Fe2-N product are 196 and 149 kJ/mol less stable than the Fe6-H$_2$NNH$_2$ structure, showing that they are not expected to form during the reaction mechanism.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing

E₆ structures

Adding a proton and an electron to the previous structures gives intermediates at the E$_6$ level. These were studied primarily in the doublet BS10-147 state. The results are collected in

Table 3 and the best structures are shown in Figures 4 and S3. As for the E$_4$ and E$_5$ structures, HIE protonation was found also to be 1–138 kJ/mol more favourable than HID, except for Fe2/6(3)-H$_2$NNH$_2$ with TPSS (2 kJ/mol), where His-195 accepts a hydrogen bond from H$_2$NNH$_2$ in the HID state.

Table 3. Energies (in kJ/mol), N–N and Fe–N distances (in Å) of the various structures of the E$_6$ states. All states were studied in the $S=1/2$ state unless otherwise stated. The entries are the same as in Table 1.

Structure	His	BS	TP	B3	TZ	Rlx	N-N	Fe-N
Fe2-H$_2$NNH$_2$(3)	HID	147	163	163			1.43	2.10
	HIE	147	33	29	31	28	1.43	2.12
		157	17				1.43	2.11
Fe2-H$_2$NNH$_2$(5)	HID	147	163	161			1.42	2.10
	HIE	147	60	55			1.42	2.06
Fe6-H$_2$NNH$_2$(3)	HID	147	83	84			1.45	2.10
	HIE	147	0	0	0	0	1.45	2.10
		235	51	9			1.45	2.05
[a]		147	43	−38			1.43	2.27
	$S=3/2$	147	16	−53			1.45	2.12
Fe6-H$_2$NNH$_2$(5)	HID	147	93	87			1.45	2.15
	HIE	147	2	−33	19	9	1.45	2.15
		235	17	−18			1.45	2.09
Fe2/6-H$_2$NNH$_2$(3)	HID	235	245	213			1.46	2.10,1.99
	HIE	235	247				1.46	2.15,1.98
Fe2/6-H$_2$NNH$_2$(5)	HID	147	226	76			1.47	2.12,2.02
	HIE	147	201				1.45	2.36,2.06
Fe2-HNNH$_3$(3)	HID	147	232	177			1.44	2.09
	HIE	147	109	82			1.45	2.07
Fe2-HNNH$_3$(5)	HID	147	282	246			1.44	2.13
	HIE	147	150	108			1.44	2.09
Fe6-HNNH$_3$(3)	HID	147	243	197			1.44	2.02
	HIE	147	105	108			1.44	2.02
Fe6-HNNH$_3$(5)	HID	147	200	204			1.44	2.11
	HIE	147	119	122			1.44	2.11
Fe2/6-NH(3) + NH$_3$	HID	147	220	179				1.81,1.88
	HIE	147	117	194				1.87,1.90
Fe2/6-NH(5) + NH$_3$	HID	147	309	286				1.91,1.87
	HIE	147	265	285				1.92,1.87
Fe6-H$_2$NNH$_3$(5)	HID	147	110	116			1.44	2.11
	HIE	147	28	31	29	43	1.44	2.11
Fe2-HNNH$_2$(5)	HIP	147	342	345			1.37	1.96
Fe2/6-HNNH$_2$(3)	HIP	147	478	547			1.45	1.95,2.01
Fe2/6-HNNH$_2$(5)	HIP	147	445	501			1.46	1.97,1.99
Fe6-HNNH$_2$	HIP	147	285	249			1.42	1.89
Fe6-H$_2$NNH$_2$(3)	HIP	147	156	167			1.44	2.09
Fe6-H$_2$NNH$_2$(5)	HIP	147	164	169			1.45	2.15
Fe2-NNH$_3$	HIP	147	409	409			1.45	2.05
Fe2/6-NNH$_3$(5)	HIP	147	502	537			1.47	1.91,1.91
		235	469	533			1.46	1.91,1.92
Fe6-HNNH$_3$(5)	HIP	147	254	267			1.45	1.99

[a] A structure with a H–N–N–H torsion of 98–99°.
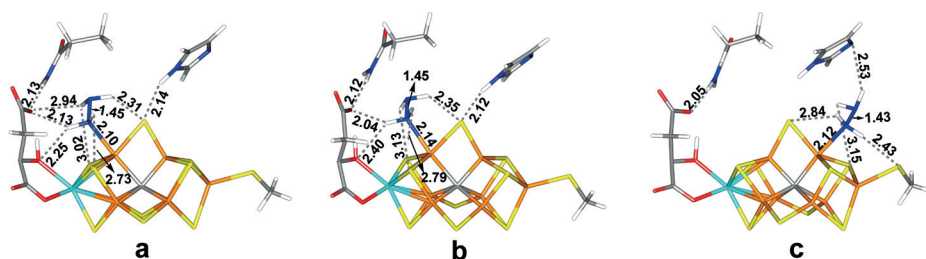
Figure 4. The best E$_6$ structures: (a) Fe6-H$_2$NNH$_2$(3), (b) Fe6-H$_2$NNH$_2$(5) and (c) Fe2-H$_2$NNH$_2$(3), all with the HIE state of His-195.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

The best structure at the TPSS level has H_2NNH_2 bound end-on to Fe6, with the non-coordinating NH_2 group pointing towards S3A (thus, the new proton in the E_6 state is added to homocitrate and not the substrate; Figure 4a). It has a Fe6-N distance of 2.10 Å and a N–N bond length of 1.45 Å, which is slightly longer than for the E_5 structures and the same molecule optimised in vacuum, 1.43 Å. The Fe ions have spin populations of 2.2–3.2 e, but 2.0 e on Fe2 and 1.4 e on Fe6 (Table S4). The spin on Mo is minor and slightly positive, 0.1 e. It is stabilised by a hydrogen bond from the Fe6-bound NH_2 group to the hydroxy O7 atom of homocitrate (H⋯O7 = 2.25 Å) and by a hydrogen bond from the other NH_2 group to S2B (H⋯S2B = 2.31 Å). The other H atom of the latter NH_2 group points in the direction of S1B, but the distance is long, 3.02 Å. It is also 2.94 Å from O1 of homocitrate. The fourth H atom of the substrate is 2.73 Å from S3B, but the geometry is far from ideal. It is also close to a methyl group of Val-70 (1.83 Å H–H distance). The quartet state was 16 kJ/mol less stable at the TPSS level, but 53 kJ/mol more stable with B3LYP. An investigation of all BS states (Table S2) showed that BS10-147 is most stable, 6 kJ/mol better than BS10-135. With B3LYP, a structure in which the two NH_2 groups are twisted with respect to each other (H–N–N–H torsion of 99°), as in the structure for the free hydrazine, is 38 kJ/mol more stable than the structure in Figure 4a, but at the TPSS level, the other structure is 43 kJ/mol more stable.

The corresponding structure with the non-coordinating NH_2 group pointing towards S5A (Figure 4b) is only 2–19 kJ/mol less stable with TPSS but 33 kJ/mol more stable B3LYP. It is stabilised by a hydrogen bond to the O1 atom of homocitrate (H⋯O1 = 2.04 Å), whereas the other H atoms interact with S2B, S3B and S1B as for the other conformation (H⋯S distances of 2.35, 3.13 and 2.79 Å). With B3LYP, this structure is further stabilised by 62 kJ/mol when studied in the BS7-235 state, whereas with TPSS the BS10-147 state is 48 kJ/mol more stable.

The corresponding structure with hydrazine bound end-on to Fe2 (Figure 4c) is 17–33 kJ/mol less stable. The two H atoms of the NH_2 group bound to Fe2 form hydrogen bonds to SG of Cys275 (H⋯SG = 2.43 Å) and S2B (H⋯S2B = 2.84 Å). The other two H atoms interact with ND1 of His-195 and S1A (H⋯ND1 = 2.53 and H⋯S1A = 3.15 Å), but with poor geometries. A full investigation of all BS states (Table S2) showed that this structure is most stable in the BS6-157 state, which is 16 kJ/mol more stable than the BS10-147 state. End-on bound $HNNH_3$ structures are 105–282 kJ/mol less stable, whereas side-on (Fe2/6) structures are 117–309 kJ/mol less stable and those with $HNNH_3$ dissociate to NH and NH_3.

We have also studied the same structures as for E_5, but with HIP and an extra electron. They were all high in energy, 156–547 kJ/mol less stable than the best Fe6-H_2NNH_2 structure, showing that proton transfer from His-195 is strongly favourable. In several of this type of structures, the substrate automatically abstracts a proton from His-195, forming HID states instead (which are less stable than the HIE structures, as has already been discussed).

E$_7$ structures

After adding yet another proton and electron, we reach the E_7-level intermediates. They were studied in the triplet BS10-147 state. The structures are listed in Table 4 and the most stable structures are shown in Figures 5 and S4. As usual, the HID structures were less stable than the corresponding HIE structures by 3–176 kJ/mol.

Only a few structures were obtained with H_2NNH_3 bound to the FeMo cluster. They had a N–N distance of 1.42–1.43 Å. The most stable one had H_2NNH_3 bound end-on to Fe6, with the NH_3 directed towards the S5A side (Figure 5a), but binding to Fe2 was only 4 kJ/mol less favourable with TPSS (30 kJ/mol by B3LYP). An investigation of all BS states (Table S2) showed that this complex is most stable in the BS2-234 state, which is actually 36 kJ/mol more stable than the BS10-147 state. The N–N bond can readily be cleaved in this structure with an activation barrier of only 32 kJ/mol and an exothermic reaction energy of −154 kJ/mol. The product has NH_2 bound end-on to Fe6 and NH_3 dissociated, but hydrogen bonded to the cluster (Figure 5b). The two H atoms of NH_2 form hydrogen bonds to O1 of homocitrate (H⋯O1 = 2.47 Å) and S2B (H⋯S2B = 2.53 Å). The dissociated NH_3 molecule forms hydrogen bonds to NH_2 (H⋯N = 2.06 Å) and O2 of homocitrate (H⋯O2 = 2.23 Å), whereas the third H atom does not form any favourable interaction, but instead is quite close to a methyl group of Val-70 (2.03 Å H–H distance). A structure with H_2NNH_3 dissociated from the Fe ions, but still hydrogen-bonded to the cluster is 83 kJ/mol more stable than the bound Fe6-H_2NNH_3 structure (Figure 5c). When it is dissociated, it is appreciably harder to cleave the N–N bond – the calculated barrier is 91 kJ/mol.

Structures with a cleaved N–N bond and both NH_2 and NH_3 coordinated to the cluster are up to 84 kJ/mol more stable than the bound Fe6-H_2NNH_3 structure, but 21 kJ/mol less stable than the structure with NH_3 dissociated. The best one has NH_2 bridging Fe2 and Fe6 on the side facing S5A, whereas NH_3 binds to Fe6 (Figure 5d). The two Fe–NH_2 distances are 1.93 and 1.95 Å, whereas the Fe6—NH_3 distance is 2.27 Å. S2B has moved considerably, but it still binds to Fe2 and Fe6, and it receives a hydrogen bond from His-195 (HE2⋯S2B = 2.36 Å). NH_3 forms a hydrogen bond to O1 of homocitrate (H⋯O1 = 2.00 Å), whereas the second H atom points towards S2B (H⋯S2B = 2.65 Å). The third H atom does not form any favourable interactions. The two H atoms of NH_2 point towards S3A and S2B (H⋯S3A = 3.21 and H⋯S2B = 2.63 Å). The spin populations on Fe2 and Fe6 are relatively low, 2.1 and 2.2 e, respectively, but that on Fe7 is even lower, 1.5 e (Table S5). NH_3 may dissociate from this structure, but the activation barrier is rather high, 78 kJ/mol.

There are several other structures with comparable energies (cf. Table 4), for example with NH_2 bridging Fe2 and Fe6 on either side of S2B and with NH_3 either on Fe2 or Fe6. The relative energies sometimes differ rather much between TPSS and B3LYP.

As for the other E_n states, we tested also structures with HIP protonated, but all these were at least 183 kJ/mol less stable.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

Table 4. Energies (in kJ/mol), N–N and Fe–N distances (in Å) of the various structures of the E_7 states. All states were studied in the $S = 1$ state unless otherwise stated. The entries are the same as in Table 1. Fe–N distances of NH_2 precede those of NH_3.

Structure	His	BS	TP	B3	TZ	Rlx	N-N	Fe-N
Fe2-H$_2$NNH$_3$(3)	HID	147	277	297			1.42	2.14
	HIE	147	137	162			1.43	2.05
Fe2-H$_2$NNH$_3$(5)	HID	147	248	256			1.43	2.17
	HIE	147	143	159			1.43	2.14
Fe6-H$_2$NNH$_3$(3)	HID	147	193	216			1.43	2.28
Fe6-H$_2$NNH$_3$(5)	HID	147	217	237			1.43	2.26
	HIE	147	133	−12	28	132	1.43	2.38
		234	97				1.42	2.43
H$_2$NNH$_3$ dissociated	HIE	147	50				1.44	
Fe2-NH$_2$(3)-Fe6-NH$_3$	HID	147	94	109				1.84,2.06
	HIE	147	61	68				1.88,2.08
Fe6-NH$_2$(3)-Fe2-NH$_3$	HID	147	76	147				1.98,1.91
	HIE	147	73	57				2.06,1.92
Fe6-NH$_2$(5)-Fe2-NH$_3$	HID	147	85	154				2.01,1.91
	HIE	147	88	126				2.03,1.93
Fe2/6-NH$_2$(3)-Fe2-NH$_3$	HID	147	104	166				1.93,1.95,2.01
	HIE	147	0	110				1.93,1.95,2.00
Fe2/6-NH$_2$(3)-Fe6-NH$_3$	HID	147	61	86				1.94,1.97,2.09
	HIE	147	36	51				1.96,2.02,2.27
Fe2/6-NH$_2$(5)-Fe2-NH$_3$	HID	147	107	188				1.95,1.94,2.05
	HIE	147	26	38				1.93,1.96,2.05
Fe2/6-NH$_2$(5)-Fe6-NH$_3$	HID	147	83	105				1.93,1.97,2.24
	HIE	147	0	0	0	0		1.93,1.95,2.27
		14	−50					1.92,1.94,2.09
	$S=0$	147	−31	37				1.92,1.95,2.16
	$S=2$	147	20	97				1.92,1.95,2.21
Fe2/6-NH$_3$(3)-Fe2-NH$_2$	HID	147	110	85				2.57,2.03,1.89
	HIE	147	83	75				2.70,2.03,1.88
Fe2/6-NH$_3$(3)-Fe6-NH$_2$	HIE	147	132	145				2.45,2.13,2.02
Fe2/6-NH$_3$(5)-Fe2-NH$_2$	HID	147	135	169				2.04,2.95,2.00
	HIE	147	43	−7				2.02,2.95,1.98
Fe2/6-NH$_3$(5)-Fe6-NH$_2$	HID	147	198	231				2.86,2.04,2.01
	HIE	147	136	200				2.64,2.01,2.02
Fe6-NH$_2$ + NH$_3$	HIE	147	−21					1.89
Fe2-H$_2$NNH$_2$(3)	HIP	147	210	224			1.45	2.16
Fe2-H$_2$NNH$_2$(5)	HIP	147	240	252			1.42	2.08
Fe6-H$_2$NNH$_2$(3)	HIP	147	183	202			1.45	2.09
Fe6-H$_2$NNH$_2$(5)	HIP	147	191	206			1.46	2.15
Fe2/6-H$_2$NNH$_2$(3)	HIP	235	424	426			1.46	2.12,1.99
Fe2/6-H$_2$NNH$_2$(5)	HIP	147	374	205			1.48	2.15,2.03
Fe2-HNNH$_3$(5)	HIP	147	368	229			1.43	2.14
Fe6-HNNH$_3$(3)	HIP	147	292	313			1.44	2.02
Fe6-HNNH$_3$(5)	HIP	147	307	326			1.44	2.10
Fe6-H$_2$NNH$_3$(5)	HIP	147	217	239			1.44	2.10

E_5–E_8 structures with only one N atom

We studied also structures with only a single N atom, i.e. after N–N bond cleavage and dissociation of a NH_3 product. These were studied at four levels of oxidation and protonation (E_5–E_8), even if the results in the previous subsections indicate that only the E_7 and E_8 states are involved in the reaction mechanism. The results are collected in Table 5 and the best structures are shown in Figures 6, 7 and S5.

The best E_5 structure has N bound end-on to Fe6 with a Fe–N distance of 1.60 Å (Figure 6a). The N atom receives a hydrogen bond from HE1 of Gln-191 (2.55 Å, but this hydrogen also forms a hydrogen bond to O1 of homocitrate with a H···O1 distance of 2.40 Å). We tested also the singlet and quartet states for this structure. The latter was 38 kJ/mol less stable at the TPSS level (33 kJ/mol with B3LYP). However, the singlet was 39 kJ/mol more stable with TPSS, but 11 kJ/mol less stable with B3LYP. A structure with NH bound to Fe6 (with the proton abstracted from homocitrate) is only 2 kJ/mol less stable (38 kJ/mol by B3LYP), but it is 14 kJ/mol more stable with the larger basis set and 9 kJ/mol more stable if the surroundings are relaxed.

The corresponding structure with N bound end-on to Fe2 (Figure 6b) is also 2 kJ/mol less stable (35 kJ/mol with B3LYP), but 1 kJ/mol more stable with the larger basis set, 41 kJ/mol more stable if the surroundings are relaxed. It has an even shorter Fe2–N bond length of 1.54 Å. The N atom does not receive any polar hydrogen bond, but it is 2.00 Å from a HB atom of Ser-278. The corresponding structures with N bridging Fe2 and Fe6 are 30 kJ/mol (on the S3A side) and 57 kJ/mol (S5A side) less stable. In both cases, S2B moves to a position where it interacts with more Fe ions than Fe2 and Fe6. Moreover, N receives the hydrogen bond from His-195 (instead of S2B; 2.44 and 2.50 Å, respectively). The corresponding HID structures are 78–126 kJ/mol less stable. Interestingly, for the three most stable structures BS7-235 was found to be 5–37 kJ/mol more

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing
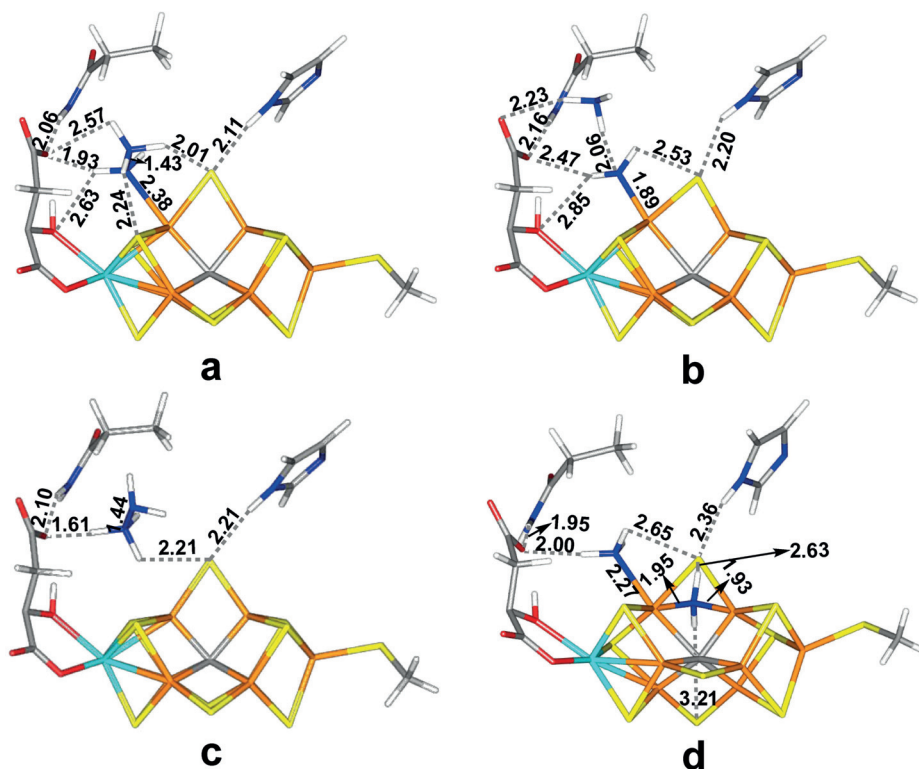
Figure 5. The best E_7 structures: (a) Fe6-H$_2$NNH$_3$(5), (b) Fe6-NH$_2$+NH$_3$, (c) H$_2$NNH$_3$ dissociated and (d) Fe2/6-NH$_2$(5)-Fe6-NH$_3$, all with the HIE state of His-195.

stable than the BS10-147 state. However, the Fe6-N structure was still the best structure.

The most stable E_6 structure has NH$_2$ bound end-on to Fe6 with the extra proton on the substrate abstracted from homocitrate (Figure 6c). It has a Fe6–N bond length of 1.85 Å and the H atoms of NH$_2$ point toward O1 and O7 of homocitrate (H···O1 = 2.01 Å, H···O7 = 2.17 Å) and S2B (H···S2B = 2.59 Å). The spin population on Fe6 (2.0 e) is slightly lower than that on Fe2 (2.1 e; Table S6). This structure is 126 kJ/mol more stable (63–90 kJ/mol with B3LYP, the larger basis set or relaxed surroundings) than a structure with NH binding end-on to Fe6 (Fe6–N bond length of 1.73 Å). The corresponding HID structures are 51–221 kJ/mol higher in energy. Structures with HIP are 305–373 kJ/mol higher in energy, showing that proton transfer from His-195 is favourable.

The most stable E_7 structure has NH$_3$ bound end-on to Fe6 (with a proton abstracted from homocitrate). The Fe6–N bond length is 2.06 Å (Figure 6d). One of the three H atoms forms hydrogen bonds to both O1 and O7 of homocitrate (H···O1 = 2.01 and H···O7 = 2.03 Å). The other two are rather close to S2B (H···S2B = 2.76 and 2.86 Å), but also to S1B and S3B (H···S1B =

2.96 Å and H···S3B = 3.09 Å). The spin population of Fe6 is 2.1 e (Table S6). NH$_3$ cannot dissociate from this structure (the energy keeps rising by more than 120 kJ/mol when the Fe6–N bond is elongated). A full investigation of all BS states (Table S2) showed that BS6-156 is actually lowest, but the BS10-147 state is only 8 kJ/mol higher in energy and there are five BS states within 10 kJ/mol of BS6-156.

This Fe6-NH$_3$ structure is 99–113 kJ/mol more stable than a structure with NH$_2$ bound to Fe6 (and a proton on homocitrate). The structure with NH$_2$ bound to Fe2 is 95–119 kJ/mol less stable. Structures with NH$_2$ bridging Fe2 and Fe6 are 191–350 kJ/mol higher in energy than the best structure. Structures with HID are ~90 kJ/mol less stable than the corresponding HIE structures. Structures with HIP are 281–418 kJ/mol less stable, again indicating that proton transfer from His-195 is strongly favourable.

Finally, we studied also E_8 states. Again, the most stable structure has NH$_3$ bound to Fe6 with a Fe–N distance of 2.13 Å (Figure 7a). NH$_3$ forms a hydrogen bond to O1 of homocitrate with a H–O distance of 1.93 Å. One of the other H atoms is 2.68 Å from S2B, but the last atom does not form any

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

Table 5. Energies (in kJ/mol), N–N and Fe–N distances (in Å) of the various structures of the E_5-E_8 states with a single N atom. All states were studied in the $S=1$ (E_5 or E_7) or $S=1/2$ (E_6 or E_8) state unless otherwise stated. The entries are the same as in Table 1.

State	Structure	His	BS	TP	B3	TZ	Rlx	Fe-N
E_5	Fe2-N	HID	147	124	157			1.53
		HIE	147	2	35	−1	−41	1.54
			235	−5	87			1.52
	Fe6-N	HID	147	126	126			1.60
		HIE	147	0	0	0	0	1.60
			235	−37	−13			1.52
		$S=0$	147	−39	11			1.52
		$S=2$	147	38	33			1.60
	Fe2/6-N(3)	HID	147	155	309			1.69,1.79
			235	134	256			1.71,1.72
		HIE	147	57	23			1.69,1.81
	Fe2/6-N(5)	HID	43	111	302			
		HIE	127	30	199			
			14	64	−64			
	Fe6-NH	HID	147	117	16			1.77
		HIE	147	2	38	−14	−9	1.72
			235	−17	28			1.62
E_6	Fe2-NH	HID	147	226	199			1.85
		HIE	147	132	118	139	69	1.83
	Fe6-NH	HID	147	204	196			1.89
		HIE	147	126	90	131	69	1.79
		$S=3/2$	147	154	133			1.77
	Fe2/6-NH(3)	HID	147	239	328			1.80,1.87
		HIE	147	129	107	149	79	1.84,1.86
		$S=3/2$	147	124	167			1.86,1.87
	Fe2/6-NH(5)	HID	147	250	302			1.85,1.90
		HIE	147	138	188	175	116	1.85,1.89
	Fe6-NH$_2$	HID	147	86	113			1.86
		HIE	147	0	0	0	0	1.85
			235	26	−48			1.84
	Fe6-N	HIP	147	305	273			1.53
	Fe2/6-N(3)	HIP	147	373	288			1.72,1.85
	Fe6-NH	HIP	147	319	296			1.78
E_7	Fe2-NH$_2$	HID	147	207	201			1.89
		HIE	147	117	107	119	95	1.87
	Fe6-NH$_2$	HID	147	200	195			1.86
		HIE	147	112	109	113	99	1.85
		$S=0$	147	118	150			1.87
		$S=2$	147	149	101			1.85
	Fe2/6-NH$_2$(3)	HID	147	242	213			1.97,1.96
		HIE	147	224	240			1.98,2.03
	Fe2/6-NH$_2$(5)	HID	147	281	350			1.94,1.95
		HIE	147	191	263			1.92,1.94
	Fe6-NH$_3$	HID	147	89	85			2.06
		HIE	147	0	0	0	0	2.06
			156	−8				2.09
		$S=0$	147	−13	47			2.08
		$S=2$	147	33	−37			2.06
	Fe6-NH	HIP	147	418	394			1.82
	Fe6-NH$_2$	HIP	147	283	281			
E_8	Fe2-NH$_3$	HID	147	126	128			2.06
		HIE	147	11	8	14	−28	2.06
			235	35	−2			2.13
	Fe6-NH$_3$	HID	147	88	81			2.12
		HIE	147	0	0	0	0	2.13
			235	17	12			2.07
		$S=3/2$	147	23	−35			2.20
	Fe2/6-NH$_3$(3)	HID	147	326	336			2.09,2.51
		HIE	147	324	301			2.11,2.34
	Fe2/6-NH$_3$(5)	HID	147	273				2.10,3.14
	NH$_4$ dissociated	HID	147	88	74			3.04
		HIE	147	−2	6	1	−32	3.03
			235	17	−20			3.09
	Fe6-NH$_2$	HIP	147	292	228			1.88
	Fe6-NH$_3$	HIP	147	152	146			2.06

favourable interactions. The Fe spin populations are 2.2–3.2 e, but 2.0 e on Fe6 (Table S6). The structure was studied in the doublet state and the corresponding quartet state is 23 kJ/mol less stable with TPSS, but 35 kJ/mol more stable with B3LYP. A

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing
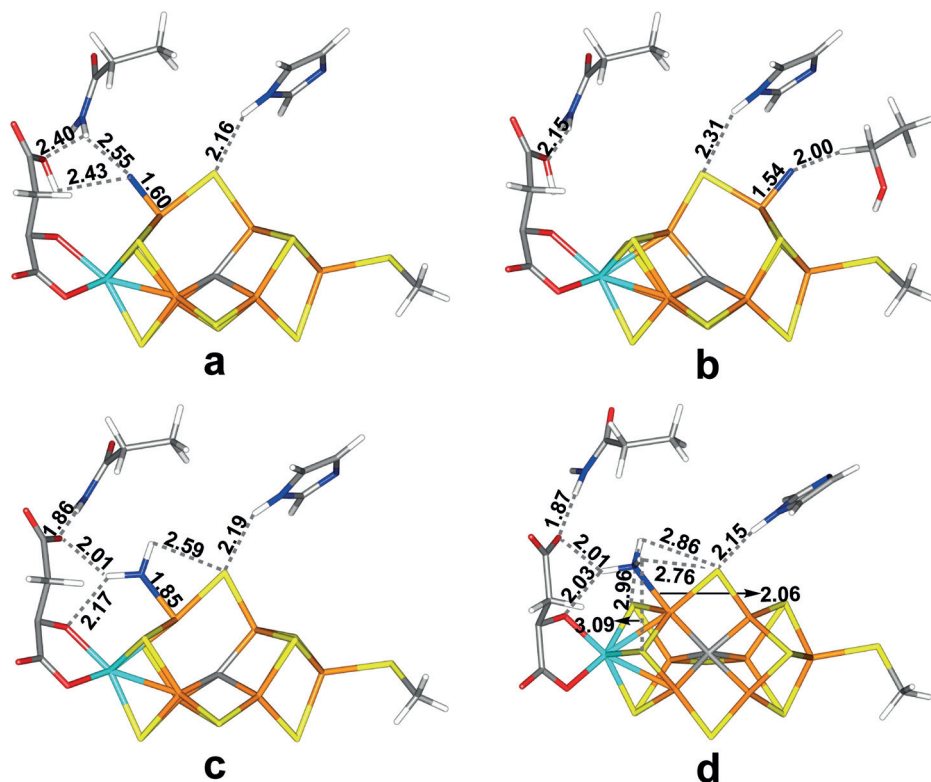
Figure 6. The best structures with only one N atom: (a) E_5-Fe6-N, (b) E_5-Fe2-N, (c) E_6-Fe6-NH$_2$, (d) E_7-Fe6-NH$_3$, all with the HIE state of His-195.
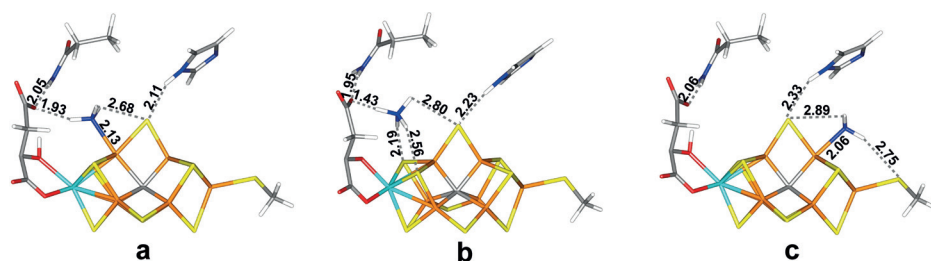
Figure 7. The E_8 structures: (a) Fe6-NH$_3$, (b) NH$_4$ dissociated and (c) Fe2-NH$_3$, all with the HIE state of His-195.

full investigation of all possible BS states (Table S2) showed that the BS10-147 state is best, but there are four other BS states within 10 kJ/mol. NH$_3$ can dissociate from this structure with a barrier of 42 kJ/mol. The dissociation energy of NH$_3$ (compared to the quartet BS7-235 E_0 state and NH$_3$ in a water-like continuum solvent with a dielectric constant of 80) is 16 kJ/mol (−20 kJ/mol with B3LYP), which can easily be overcome by the

gain in translational and rotational entropy of the released NH$_3$ ligand, ~ 60 kJ/mol).[92,93]

A structure, in which NH$_3$ has abstracted the proton from O7 of homocitrate, forming NH$_4^+$, is actually 2 kJ/mol more stable than the Fe6-NH$_3$ structure. NH$_4^+$ has dissociated from Fe6 and the four H atoms form hydrogen bonds to O1 of

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing

homocitrate (H···O1 = 1.43 Å), S2B (H···S2B = 2.80 Å), S1B (H···S1B = 2.19 Å) and S3B (H···S3B = 2.56 Å; Figure 7b).

A structure with NH_3 bound to Fe2 is 8–11 kJ/mol less stable than the Fe6-NH_3 structure (but 28 kJ/mol more stable with relaxed surroundings). It has a slightly shorter Fe—N bond (2.06 Å; cf. Figure 7c). Two of the H atoms of NH_3 points approximately towards SG of Cys-275 and S2B (H···SG = 2.75 Å and H···S2B = 2.89 Å). A full investigation of all BS states showed that BS10-147 indeed is the most stable state, but only 2 kJ/mol more stable than the BS6-157 state. We also tried to find structures with NH_3 bridging Fe2 and Fe6, but they were at least 326 kJ/mol less stable. Structures with HID were 81–120 kJ/mol less stable than the corresponding HIE structures. Structures with HIP were 152–292 kJ/mol less stable, showing that proton transfer from His-195 is strongly favourable.

Conclusions

In this investigation, we have studied possible reaction intermediates of nitrogenase, assuming that the S2B remains bound to the FeMo cluster. To avoid the problem that the structure of the E_4 intermediate is not known and that different DFT functionals give very different relative stabilities of various protonation states,[43,44] we started our study after H_2 has dissociated and N_2 has bound to the cluster and has become doubly protonated to N_2H_2,[1,18] so that no protons remain bound to the cluster. Based on the accumulated experimental evidence,[1,90,91] as well a systematic study of the binding of N_2H_2 to the FeMo cluster,[45] and in agreement with most previous computational studies,[27] we have assumed that N_2H_2 binds either to Fe2 or Fe6.

Our study has led to the following conclusions:

• For the E_4 state, Fe6-tHNNH, Fe2-tHNNH and Fe6-HNNH$_2$ structures are all competitive (within 5 kJ/mol with at least one of the four levels of theory included in Table 1).

• For the E_5 state, Fe6-H_2NNH_2 is lowest in energy. Fe6-NNH_3 is 54–68 kJ/mol higher and the N—N bond in cannot be cleaved.

• For the E_6 state, Fe6-H_2NNH_2 structure is lowest in energy, 28–43 kJ/mol lower than Fe6-H_2NNH_3. Cleavage of N—N in the latter has a barrier of 95 kJ/mol.

• The N—N bond in the H_2NNH_3 E_7 complexes can easily be cleaved, the reaction is exothermic and NH_3 moves spontaneously into the second coordination sphere of the cluster, whereas NH_2 binds with similar affinities to both Fe2 and Fe6. However, the most stable structure is obtained if NH_2 abstracts the hydroxy proton from homocitrate, forming NH_3 bound to Fe6, which cannot dissociate at this level of reduction.

• In the E_8 state, NH_3 binds preferably to Fe6 (binding to Fe2 is 8–11 kJ/mol higher in energy). It can readily dissociate from the FeMo cluster.

Based on these results, we suggest the reaction mechanism in Figure 8. In this mechanism, the substrate binds to Fe6. In the E_5 state, it is protonated to H_2NNH_2, whereas in E_6, the proton is added to homocitrate, so that the ligand remains H_2NNH_2. In the E_7 state, the substrate is protonated to H_2NNH_3, in which the N—N bond is readily cleaved and NH_3 automatically dissociates. The resulting NH_2 group remains bound to Fe6 and is protonated to NH_3. In the E_8 state, NH_3 dissociates and the resting E_0 state is formed.

It has been much discussed whether the reaction mechanism of nitrogenase follows a sequential or alternating reaction mechanism.[1] Our results suggest that the enzyme follows an alternating mechanism, with HNNH and H_2NNH_2 as intermediates (although the former may be protonated by homocitrate to HNNH$_2$). Moreover, the N—N bond is cleaved in the E_7 state and that the NH_3 products dissociate at the E_7 and E_8 levels, in accordance with an alternating mechanism.

Between the various E_n levels in Figure 8, we have assumed that an electron and a proton are delivered to the FeMo cluster. The electrons are provided by the Fe protein, via the P-cluster.[1,4] The protons ultimately come from the solvent. Dance
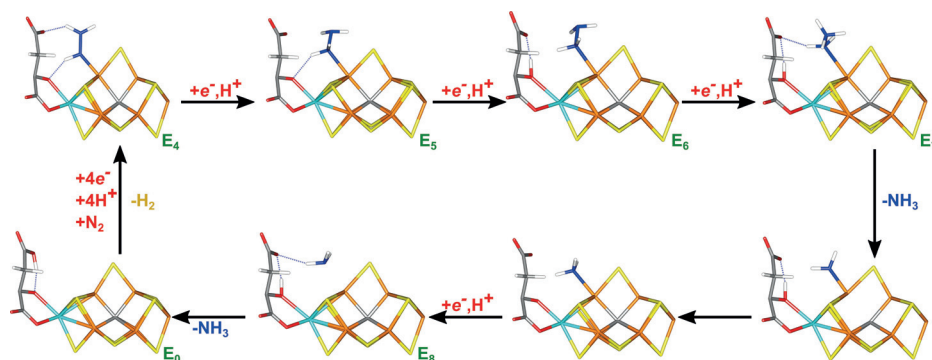
Figure 8. Suggested reaction mechanism for nitrogenase, assuming that S2B remains bound to the cluster.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

has investigated different proton paths from the solvent to the FeMo cluster. He suggested that protons are transferred from the surface via a chain of water molecules to a water molecule close to homocitrate and the S3B atom.[94–97] He has also studied how the proton may be transferred within the FeMo cluster, starting on the S3B atom and ultimately ending up close to the substrate-binding sites.[96] His-195 may also provide protons to the substrate, as has often been assumed.[38,49] In fact, our calculations show that such a transfer (via the HIP state of His-195) is always downhill. However, this gives the HID state of His-195, which is nearly always unfavourable. Moreover, Dance has shown that rotation of the sidechain of His-195 is unlikely in the protein.[94,98] Therefore, His-195 can probably not provide more than one proton to the substrate (at the most).

However, it should be remembered that QM studies of nitrogenase are extremely complicated. Our survey of the most stable structures at the various E_4–E_8 levels have indicated that several structures often have quite similar energies and that here are many conformations of the bound intermediates, depending on the direction of the non-coordinated N-group, the H–N–N–H dihedral angle and the formation of hydrogen bonds and other polar interactions. Moreover, sometimes other BS states than BS10-147 are most stable and a full BS investigation of every structure is currently too demanding. This makes it harder to settle the most stable structures and it cannot be excluded that we might have overlooked some low-energy structures.

For all states, we have optimised structures with both TPSS and B3LYP. In general, the two methods give relative energies that agree reasonably. However, in some cases, differences >50 kJ/mol are observed, without any significant differences in the geometry. This is often observed for the various spins states (B3LYP typically prefers higher spin states than TPSS, as expected). However occasionally, such differences also occur for structures with differences only in the hydrogen-bond pattern or coordination mode, making interpretation of the results harder.

For the most interesting states, we have also recalculated energies with the larger def2-TZVPD basis set. As in our previous studies,[43,59–61] this has typically only a minor effect on the energies, <20 kJ/mol. However, calculations with relaxed surroundings occasionally have larger effects, up to 26 kJ/mol, which may indicate that the structure need to relax more than is allowed in the rather large QM system or problems with local minima in the surroundings. However, a more detailed study of the relaxation of the surroundings (in the Supporting Information) indicates that the large energy differences are connected to major movements of water molecules and other groups far from the substrate-binding site, suggesting that it reflects more occurrences of multiple local minima, rather than important relaxation of the surroundings.

For many of the E_n states, we have observed that a transfer of the hydroxy proton of homocitrate to the N_2-intermediate is favourable. This indicates homocitrate may constitute a proton buffer, which can be used to stabilise certain intermediates of the reaction, especially H_2NNH_2 and NH_3. This provides an attractive explanation why homocitrate is mandatory for the

function of nitrogenase.[88,99] Bjornsson and coworkers have also suggested that the hydroxy proton of homocitrate is employed to form trans-HNNH from N_2 in the E_4 intermediate, although the reaction was uphill.[42]

For all E_n levels, binding of the intermediates to Fe6 seems to be preferred and therefore we have suggested such binding in the mechanism in Figure 8. This site has more hydrogen-bond possibilities (besides sulfide ions), involving His-195, Gln-191 and the homocitrate ligand. It can also employ the suggested proton buffer of homocitrate and it is also closer to the end of the suggested proton-transfer path, involving a chain of water molecules, ending close to homocitrate and S3B.[94–97] However, for the E_4 and E_8 states, binding of the N_2-intermediates to Fe2 are competitive. Moreover, it has been suggested that a likely N_2-binding channel ends at Fe2.[27]

We have previously studied the reaction mechanism of nitrogenase, assuming that S2B dissociates from the FeMo cluster, opening up for an obvious binding site of the substrate.[57] This gave rise to a mainly alternating reaction mechanism, in which the substrate and the intermediates bound in a bridging mode (with one or both N atoms) between Fe2 and Fe6. In the present study, bridging intermediates were always found to be much less stable than end-on intermediates, except for E_7 structures with both NH_2 and NH_3. The reason for this is that the FeMo cluster is too crowded if both the substrate and S2B bridges Fe2 and Fe6. In fact, it is often observed that S2B moves to other positions or reacts with the substrate or other sulfide ions in such high-energy structures with a bridging substrate. Thus, it seems clear that bridging substrate structures are unlikely when S2B remains bound.

In conclusion, this study shows that the second part of the nitrogenase reaction (after binding of N_2) is possible also if the S2B ligand has not dissociated. Such a reaction follows an alternating mechanism with the substrate and intermediates binding to Fe6. It has also pointed out an important role for the homocitrate ligand as a proton buffer. In future studies, we study the binding of N_2 and dissociation of H_2 to the E_4 state of the cluster.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry
Europe
European Chemical
Societies Publishing

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

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

[2] B. E. Smith, *Science* **2002**, *297*, 1654–1655.

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

[4] B. Schmid, H.-J. Chiu, V. Ramakrishnan, J. B. Howard, D. C. Rees, in *Handb. Met.*, John Wiley & Sons, Ltd, **2006**, pp. 1025–1036.

[5] J. Kim, D. C. Rees, *Science* **1992**, *257*, 1677–1682.

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

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

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

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

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

[11] R. N. F. Thorneley, D. J. Lowe, in *Molybdenum Enzym.* (Ed.: T. G. Spiro), Wiley, New York, **1985**, pp. 221–284.

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

[13] 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.

[14] 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.

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

[16] L. E. Roth, F. A. Tezcan, *Methods Mol. Biol.* **2011**, *766*, 147–164.

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

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

[19] J. Chatt, *Annu. Proc. Phytochem. Soc. Eur.* **1980**, *18*, 1–18.

[20] J. Chatt, J. R. Dilworth, R. L. Richards, *Chem. Rev.* **1978**, *78*, 589–625.

[21] D. V. Yandulov, R. R. Schrock, *Science* **2003**, *301*, 76–78.

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

[23] R. R. Schrock, *Angew. Chem. Int. Ed.* **2008**, *47*, 5512–5522.

[24] R. N. F. Thornely, R. R. Eady, D. J. Lowe, *Nature* **1978**, *272*, 557–558.

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

[26] D. Lukoyanov, S. A. Dikanov, Z.-Y. Yang, B. M. Barney, R. I. Samoilova, K. V. Narasimhulu, D. R. Dean, L. C. Seefeldt, B. M. Hoffman, *J. Am. Chem. Soc.* **2011**, *133*, 11655–11664.

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

[28] F. Tuczek, in *RSC Met. Ser. 7* (Eds.: R. Hille, C. Schulzke, M. L. Kirk), Royal Society Of Chemistry, Cambridge, **2017**, pp. 223–274.

[29] I. Dance, *J. Biol. Inorg. Chem.* **1996**, *1*, 581–586.

[30] K. K. Stavrev, M. C. Zerner, *Int. J. Quantum Chem.* **1998**, *70*, 1159–1168.

[31] P. E. M. Siegbahn, J. Westerberg, M. Svensson, R. H. Crabtree, *J. Phys. Chem. B* **1998**, *102*, 1615–1623.

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

[33] H. Xie, R. Wu, Z. Zhou, Z. Cao, *J. Phys. Chem. B* **2008**, *112*, 11435–11439.

[34] J. Kästner, P. E. Blöchl, *J. Am. Chem. Soc.* **2007**, *129*, 2998–3006.

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

[36] I. Dance, *Zeitschr. Anorg. Allg. Chem.* **2015**, *641*, 91–99.

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

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

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

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

[41] S. Raugei, L. C. Seefeldt, B. M. Hoffman, *Proc. Nat. Acad. Sci.* **2018**, *115*, 10521–10530.

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

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

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

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

[46] 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.

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

[48] 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.

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

[50] I. Dance, *Dalton Trans.* **2008**, *2*, 5977–5991.

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

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

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

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

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

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

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

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

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

[60] L. Cao, U. Ryde, *Int. J. Quantum Chem.* **2018**, *118*, e25627 (16 pages).

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

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

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

[64] D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. J. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, **2018**.

[65] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[66] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.

[67] L. Hu, U. Ryde, *J. Chem. Theory Comput.* **2011**, *7*, 2452–2463.

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

[69] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571–2577.

[70] B. H. Besler, K. M. Merz, P. A. Kollman, *J. Comput. Chem.* **1990**, *11*, 431–439.

[71] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, F. Weigend, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 91–100.

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

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

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

[75] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, *Chem. Phys. Lett.* **1995**, *240*, 283–289.

[76] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.* **1997**, *97*, 119–124.

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

[78] 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.

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

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

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

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

[83] U. Ryde, M. H. M. Olsson, *Int. J. Quantum Chem.* **2001**, *81*, 335–347.

[84] N. Reuter, A. Dejaegere, B. Maigret, M. Karplus, *J. Phys. Chem. A* **2000**, *104*, 1720–1735.

[85] L. Hu, P. Söderhjelm, U. Ryde, *J. Chem. Theory Comput.* **2011**, *7*, 761–777.

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

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

[88] 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.

Chemistry–A European Journal

Research Article
doi.org/10.1002/chem.202103933

Chemistry Europe
European Chemical
Societies Publishing

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

[90] L. C. Seefeldt, I. G. Dance, D. R. Dean, *Biochemistry* **2004**, *43*, 1401–1409.

[91] S. J. George, B. M. Barney, D. Mitra, R. Y. Igarashi, Y. Guo, D. R. Dean, S. P. Cramer, L. C. Seefeldt, *J. Inorg. Biochem.* **2012**, *112*, 85–92.

[92] L. Watson, O. Eisenstein, *J. Chem. Educ.* **2002**, *79*, 1269.

[93] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Ltd, Chichester, **2017**.

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

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

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

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

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

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

Paper IV

Nitrogenase

Proton Transfer Pathways in Nitrogenase with and without Dissociated S2B

Hao Jiang, Oskar K. G. Svensson, Lili Cao, and Ulf Ryde*

Abstract: Nitrogenase is the only enzyme that can convert N_2 to NH_3. Crystallographic structures have indicated that one of the sulfide ligands of the active-site FeMo cluster, S2B, can be replaced by an inhibitor, like CO and OH^-, and it has been suggested that it may be displaced also during the normal reaction. We have investigated possible proton transfer pathways within the FeMo cluster during the conversion of N_2H_2 to two molecules of NH_3, assuming that the protons enter the cluster at the S3B, S4B or S5A sulfide ions and are then transferred to the substrate. We use combined quantum mechanical and molecular mechanical (QM/MM) calculations with the TPSS and B3LYP functionals. The calculations indicate that the barriers for these reactions are reasonable if the S2B ligand remains bound to the cluster, but they become prohibitively high if S2B has dissociated. This suggests that it is unlikely that S2B reversibly dissociates during the normal reaction cycle.

Introduction

Nitrogenase (EC 1.18/19.6.1) is the only group of enzymes that can cleave the strong triple bond in N_2, making nitrogen available for biological lifeforms. It is found in some groups of bacteria and archaebacteria, but many higher plants live in symbiosis with such organisms. Crystal structures of the most active form of nitrogenase show that the active site consists of a $MoFe_7S_9C$(homocitrate) cluster (the FeMo cluster, shown in Figure 1).[1–5] It is connected to the protein by a cysteine ligand to one of the Fe ions and a histidine ligand to the Mo ion. There also exist alternative nitrogenases with the Mo ion replaced with either vanadium or iron, which have lower activities towards N_2.[6]

Nitrogenase catalyses the reaction:

$$N_2 + 8e^- + 8H^+ + 16\,ATP \rightarrow$$
$$2\,NH_3 + H_2 + 16\,ADP + 16\,P_i \tag{1}$$
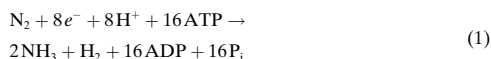
Thus, it requires eight electrons and protons, and consumes 16 molecules of ATP. H_2 is a mandatory by-product. The reaction is often described by the Lowe–Thorneley scheme,[7] which contains nine states E_0 to E_8, differing in the number of added electrons and protons. Thorough biochemical, kinetic, structural and spectroscopic studies have suggested that the enzyme typically needs to be loaded with four electrons and protons before the N_2

[*] H. Jiang, O. K. G. Svensson, Dr. L. Cao, Prof. U. Ryde
Theoretical Chemistry, Lund University, Chemical Centre
P. O. Box 124, 22100 Lund (Sweden)
E-mail: Ulf.Ryde@teokem.lu.sel

substrate can bind, through reductive elimination of H_2.[8–16] The substrate probably binds to the Fe2 or Fe6 ions (atom names are shown in Figure 1).[12,17]

Several crystallographic studies have shown that one of the sulfide ligands, S2B, which bridges the Fe2 and Fe6 ions, can be replaced by inhibitors, like CO and OH^-.[4,18] The process is reversible and a putative storage site for the dissociated SH^- ion has been identified. Therefore, it has been suggested that S2B may dissociate also during the normal reaction mechanism, opening up an obvious binding site for the substrate.[19,20] In fact, a recent crystal structure was suggested to show such replacement of S2B (and also the S3A and S5A sulfide ions) by N_2,[21] although the interpretation has been disputed.[22,23] Nitrogenase has also been extensively investigated by computational methods,[24] but so far no consensus has been reached regarding the reaction mechanism or even of the structure and protonation of the crucial E_4 state.[24–31]

Between each E_n level, one electron and one proton are delivered to the FeMo cluster. The electrons are provided by the Fe protein via the P-cluster,[11,12] and they can move
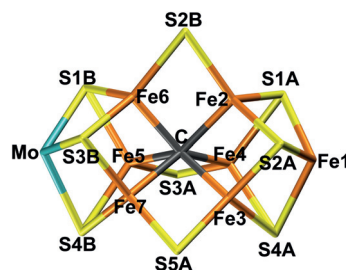
Figure 1. The FeMo cluster with atom names indicated. All structures in the figures use the same perspective.

freely within the FeMo cluster. However, movements of the protons are more restricted. They should ultimately come from the bulk solvent. His-195 forms a hydrogen bond to the S2B sulfide ion and it has therefore often been assumed to provide protons to the substrate.[27,32] However, to donate more than one proton to the substrate, the sidechain of His-195 must rotate and calculations by Dance have indicated that such rotations are unlikely in the protein.[33,34]

Therefore, he investigated alternative proton pathways from the solvent to the FeMo cluster and found that there is a conserved chain of water molecules from the surface that ends with a water molecule close to the S3B atom of the FeMo cluster.[33,35–37] He suggested that protons are delivered to the FeMo cluster on the S3B atom and that they are then transferred to the substrate via various Fe and sulfide ions in the cluster. He identified six local minima for the binding of a proton on S3B and showed that the proton may move between these with barriers of 10–60 kJ mol^{-1}.[33,38] Based on this finding, he also studied possible proton transfers within the FeMo cluster, starting from the S3B atom and ending up at the substrate-binding site at Fe6 and Fe2.[33,36,38] However, he studied mainly the first four steps of the reaction mechanism (E$_0$ to E$_4$) and partly with an outdated model of the FeMo cluster.

Recently, we have studied putative reaction mechanisms of nitrogenase, starting from bound and protonated N$_2$H$_2$ and going to two NH$_3$ molecules, with either S2B bound to or dissociated from the cluster.[20,39] In both cases, reasonable pathways could be identified, following mainly alternating mechanisms (i.e. the two N atoms are protonated alternatively and the products do not dissociate before the E$_6$ intermediate). Therefore, the calculations could not discriminate between the two scenarios. Moreover, between each E$_n$ state, protons were simply added to every possible site on the substrates, assuming that they can freely move around in the FeMo cluster. Therefore, the studies mainly determined the thermodynamically most stable protonation state and binding conformation of the substrate at each E$_n$ level.

In this study, we go one step further and study proton-transfer reactions within the FeMo cluster, from the sulfide ions closest to the end of the water chain and to the substrate. In this case, we observe a significant effect of the S2B ligand.

Results and Discussion

We have studied possible proton transfers within the FeMo cluster for the E$_5$–E$_8$ states both when S2B has dissociated and when it remains bound to the cluster. Following the investigations by Dance,[33,36,38] we assume that the protons are delivered to the cluster via the proton wire ending close to S3B.[33,35–37] However, looking at the most accurate crystal structure of nitrogenase (3U7Q[3]), the last water molecule in the chain (HOH-519) is close to three sulfide ions, S3B (4.05 Å), S5A (4.03 Å) and S4B (3.77 Å). Therefore, we considered all three ions as possible entry sites for the proton.

First, we studied possible proton-transfer paths at the E$_5$ level and with S2B dissociated. We started from the best E$_4$ structure in our previous investigation,[20] i.e. with NNH$_2$ bridging Fe2 and Fe6. To this structure, we added an electron and a proton that initially was placed on either S5A, S3B or S4B. The results are shown in Figure 2. It turned out that protonation of S5A is 97 and 69 kJ mol^{-1} more favourable than protonation of S4B and S3B. The proton on each sulfide ion can point in several different directions that represent different local minima. For example, the proton on S5A can point either towards S2B or S3A (the two conformations are called S5A(2) or S5A(3) in Figure 2. The latter is 16 kJ mol^{-1} more stable.

Then, we tried to find a proton-transfer pathway from the three sulfide ions to the NNH$_2$ ligand. The proton can first be transferred from S5A to Fe7 with a barrier of 48 kJ mol^{-1}. During this reaction the most stable broken-symmetry (BS) state changes from BS10-147 to BS7-235. Similar changes are found throughout the reaction mechanism and in the figures only the energy of the lowest BS state is indicated. A proton on S4B can easily be transferred to Fe7 with a barrier of only 3 kJ mol^{-1}. Next, the proton on Fe7 changes its conformation to point towards Fe2 with a barrier of 68 kJ mol^{-1}. Then, it is transferred to Fe6 (11 kJ mol^{-1} barrier) and to Fe2 (11 kJ mol^{-1} barrier), before it can be transferred to the substrate with a barrier of 40 kJ mol^{-1}. This results in a HNNH$_2$ state with only the NH group bridging Fe2 and Fe6. It is 16 kJ mol^{-1} more stable than the starting S5A structure, but it needs to be reorganised by a barrier of 30 kJ mol^{-1} before it reaches the most stable HNNH$_2$ structure with also the NH$_2$ group coordinating to Fe6, 55 kJ mol^{-1} more stable than the initial state with the proton on S5A.

Unfortunately, even if none of the individual barriers is prohibitively high, the early part of proton-transfer path is uphill and the highest net barrier (calculated from S5A(3)) is actually 114 kJ mol^{-1}. This is too high, compared to the net reaction rate of nitrogenase, ≈ 5 s^{-1},[12,26] corresponding to a barrier of ≈ 70 kJ mol^{-1}. Owing to this prohibitively large barrier, we looked for alternative paths. Several of these are included in Figure 2 and paths with net barriers of 107 or 109 kJ mol^{-1} can be found, leading to an alternative HNNH$_2$ product with a slightly different orientation, 4 kJ mol^{-1} less stable than the one of the other path. We also tested different variants of the method and the model (e.g. adding an extra proton on S5A). The results of these calculations are described in the Supporting Information. For the E$_5$ intermediates, it turned out that more reasonable barriers could be obtained for calculations with the B3LYP functional, giving a maximum barrier of 80 kJ mol^{-1}.

Next, we added an electron and a proton to the HNNH$_2$ structure to study proton transfers for the E$_6$ state. In this case, protonation of S5A was 67 and 78 kJ mol^{-1} more favourable than protonation of S3B and S4B. As can be seen in Figure 3, the proton can be transferred from S5A to Fe7 and, after some changes in the conformation, to Fe6 and finally to the ligand, forming H$_2$NNH$_2$ (hydrazine). The individual barriers are 13–55 kJ mol^{-1}. However, because all intermediate states are higher in energy than the initial
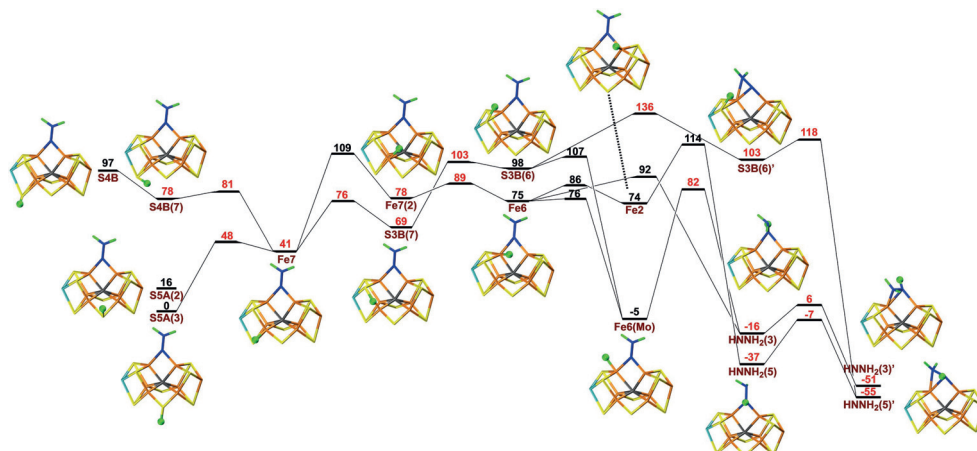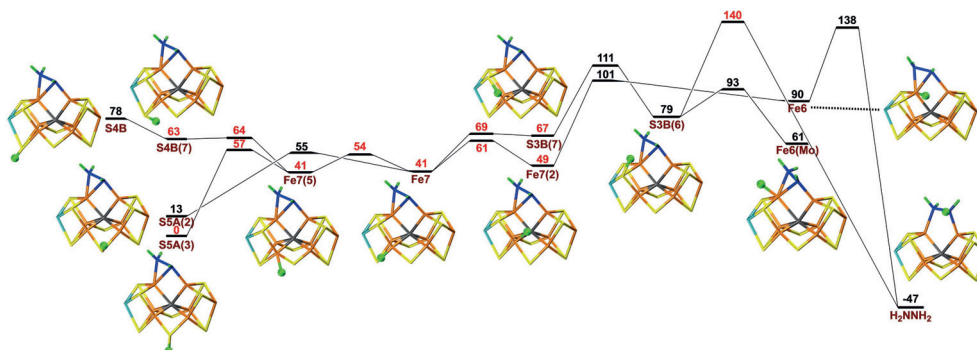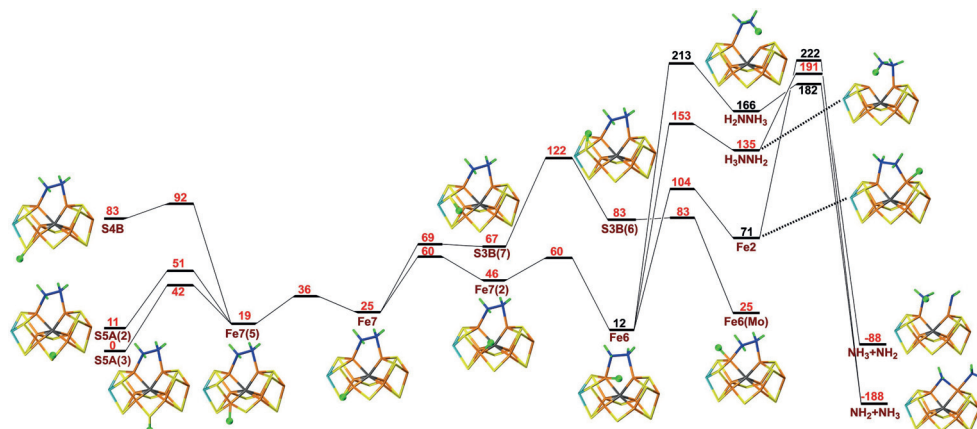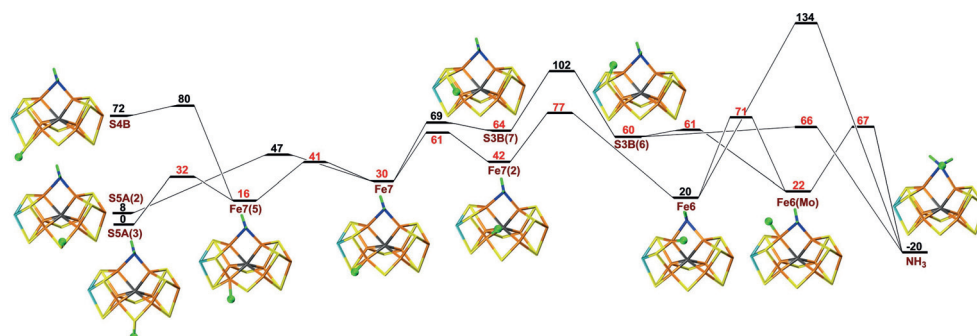
Figure 2. Reaction and activation energies, as well as structures for proton transfers in the E_5 state without S2B bound to the cluster, from S5A, S4B or S3B to the NNH$_2$ intermediate bound to Fe6 and Fe2, leading to a HNNH$_2$ product. Energies in black and red indicate that the BS10-147 and BS7-235 states are most stable, respectively. S4B(7) differs from S4B in that the proton points towards Fe7, rather than towards Mo. Fe7(2) differs from Fe7 in that the proton points towards Fe2, rather than towards S4B. In S3B(7) and S3B(6), the proton points towards Fe7 or Fe6, respectively. Fe6(Mo) differs from Fe6 in that the proton points towards Mo, rather than towards Fe2 (sometimes called the exo and endo positions, respectively).[33,35–37] HNNH$_2$(3) differs from HNNH$_2$(5) in that the added proton points towards S5A rather than S3A. A final prime in the name of the structures (e.g. S3B(6)′) indicates that the NNH$_2$ substrate has changed its conformation so that the NH$_2$ group binds to Fe6 (while the other N atom still bridges Fe2 and Fe6).

Figure 3. Reaction and activation energies, as well as structures for proton transfers in the E_6 state without S2B bound to the cluster. Energies in black and red indicate that BS10-147 and BS7-235 states are most stable, respectively. Fe7(5) differs from Fe7 in that the proton points towards S5A.

protonation of S5A, the net barrier is 138 kJ mol^{-1}, which is prohibitively large. An alternative path involving protonation of S3B has a slightly larger net barrier of 140 kJ mol^{-1}. Unfortunately, the barrier becomes even higher with B3LYP, 162 kJ mol^{-1}.

Next, we added an electron and a proton to the H$_2$NNH$_2$ structure, to study possible proton transfers for the E_7 state. As for the E_5 and E_6 states, protonation of S5A was 67 and 83 kJ mol^{-1} more stable than the S3B and S4B states.

Figure 4 shows that the proton on S5A can be transferred to Fe7 and, after some reorientations, to Fe6 and finally to H$_2$NNH$_2$. However, the barriers for latter step is prohibitive, 153 or 213 kJ mol^{-1} relative to S5A(3) and the resulting H$_2$NNH$_3$ states (bound to Fe2 or Fe6) are 135 or 166 kJ mol^{-1} less stable than the starting S5A(3) state. However, if the N–N bond is cleaved, a state with NH$_2$ bridging between Fe2 and Fe6 and with NH$_3$ coordinated to Fe6 is 188 kJ mol^{-1} more stable than the best starting state.
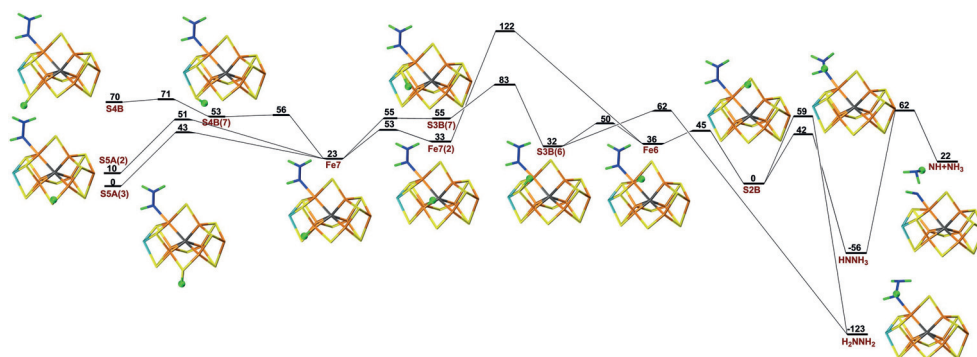
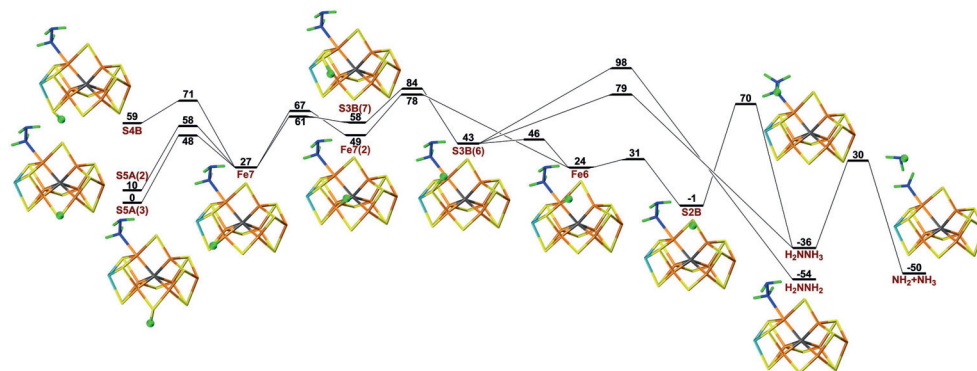**Figure 4.** Reaction and activation energies, as well as structures for proton transfers in the E_7 state without S2B bound to the cluster. Energies in black and red indicate that BS10-147 and BS7-235 states are most stable, respectively.

Yet, to reach that state, barriers of 182–222 kJ mol^{-1} need to be passed. The B3LYP method also gives high barriers, at least 215 kJ mol^{-1}.

Finally, we studied proton transfers within the E_8 state, starting from NH$_2$ bridging Fe2 and Fe6 (i.e. after dissociation of one NH$_3$ molecule). As usual, protonation of S5A(3) was 72 and 60 kJ mol^{-1} more favourable than protonation of S4B and S3B. The results in Figure 5 show that the proton on S5A can be transferred to Fe7 and, after some rotations, further to Fe6, before it reaches NH$_2$, forming the NH$_3$ product. The individual barriers are rather small, 3–51 kJ mol^{-1}, but most of the steps are upwards, leading to a net barrier of 77 kJ mol^{-1} for the transfer of the proton from Fe7 to Fe6. An alternative path involving S3B has a net barrier of 102 kJ mol^{-1}. At the B3LYP level, the net barriers are higher, at least 128 kJ mol^{-1}, because Fe-bound hydride ions are strongly disfavoured, compared to sulfur-bound protons.

We have also performed similar calculations for structures with S2B remaining bound to the FeMo cluster. In that case, the binding site of the substrate is less clear. In our previous studies, we found that binding to Fe2 or Fe6 is thermodynamically most favourable for the E_4 state,[20,40] in agreement with mutation studies.[12] For the E_5 state, binding of H$_2$NNH$_2$ to Fe6 is at least 54 kJ mol^{-1} more stable than any state binding to Fe2.[39] When binding to Fe6, the substrate can form hydrogen bonds to His-195, Gln-191 and the homocitrate ligand. Moreover, homocitrate may constitute a proton buffer, by donating a proton to the substrate, thereby stabilising certain protonation states, in particular H$_2$NNH$_2$ and NH$_3$. Consequently, we considered only states with the substrate bound to Fe6.

Figure 5. Reaction and activation energies, as well as structures for proton transfers in the E_8 state without S2B bound to the cluster. Energies in black and red indicate that BS10-147 and BS7-235 states are most stable, respectively.

First, we studied the E_5 state. We started from the Fe6-HNNH$_2$ structure, which was our most favourable E_4 structure[39] and aimed for the Fe6-H$_2$NNH$_2$ structure, which was the most stable E_5 structure. In both cases, the substrate has abstracted a proton from homocitrate. The suggested proton-transfer path is shown in Figure 6. As usual, protonation of S5A is more favourable than protonation of S4B or S3B by 70 and 32 kJ mol^{-1}, respectively.

The proton on S5A or S4B can be transferred to Fe7 (with barriers of 2 or 43 kJ mol^{-1}), then to S3B (35 kJ mol^{-1} barrier) and after a change in conformation (28 kJ mol^{-1} barrier), the proton can be transferred to the substrate (30 kJ mol^{-1} barrier). The resulting H$_2$NNH$_2$ intermediate is 123 kJ mol^{-1} more stable than the starting S5A(3) structure. Figure 6 also shows some alternative paths, involving Fe6 and other product states. The highest barrier along the entire proton-transfer reaction path is 83 kJ mol^{-1} (for the rotation of the proton at S3B). This is slightly higher than the turnover rate of the enzyme. However, considering the

approximations involved in the calculations (approximate transition states, no entropies, no tunnelling) and the known DFT-functional sensitivity of energetic result for nitrogenase,[30] the result is acceptable.

Next, we studied the transfer of a proton in the E_6 state. Since the most stable intermediate of both the E_5 and E_6 states have H$_2$NNH$_2$ bound to Fe6, we expected the proton to end up on the alcohol oxygen of homocitrate. The results are shown in Figure 7. As usual, protonation of S5A was more favourable than protonation of S4B or S3B (by 59 and 43 kJ mol^{-1}, respectively). From S5A or S4B, the proton can be transferred to Fe7 (barriers of 12 or 52 kJ mol^{-1}), then to S3B (34 kJ mol^{-1}) and it many then change its conformation to point towards Fe6 (26 kJ mol^{-1} barrier).

From this conformation, there are several possible paths. First, it can be transferred directly to homocitrate, passing a barrier of 56 kJ mol^{-1} and giving the most stable product, which is 54 kJ mol^{-1} more stable than the starting S5A(3) state. However, the maximum net barrier (for the last

Figure 6. Reaction and activation energies, as well as structures for proton transfers in the E$_5$ state with S2B remaining bound to the cluster.

Figure 7. Reaction and activation energies, as well as structures for proton transfers in the E$_6$ state with S2B remaining bound to the cluster, from S5A, S4B or S3B to the H$_2$NNH$_2$ intermediate bound to Fe6, leading to H$_2$NNH$_2$ or NH$_2$ product. In the H$_2$NNH$_2$ structure, the extra proton has been transferred to homocitrate.

proton transfer) is prohibitively high, 98 kJ mol⁻¹. Second, the proton can be transferred to the substrate, giving H_2NNH_3 with the homocitrate still deprotonated. It has a lower barrier of 36 kJ mol⁻¹, but the product is 18 kJ mol⁻¹ less stable than the H_2NNH_2 isomer. Third, the proton may be transferred to Fe6 with a barrier of only 4 kJ mol⁻¹ and then to S2B with a barrier of 7 kJ mol⁻¹. Finally it can be transferred to the substrate (again forming H_2NNH_3) with a barrier of 71 kJ mol⁻¹. Thus, this third pathway has a slightly lower barrier than the second path, and both are dominated by the net barrier of forming S3B(6) from S3B(7), which is 84 kJ mol⁻¹ above S5A(3). However, Fe6 can also be reached via Fe7 with a slightly lower net barrier of 78 kJ mol⁻¹.

Interestingly, the N–N bond in H_2NNH_3 can be cleaved with a barrier of 65 kJ mol⁻¹, forming a product with NH_3 dissociated from the FeMo cluster and NH_2 coordinated to Fe6 (and homocitrate still deprotonated). This state is only 4 kJ mol⁻¹ less stable than the H_2NNH_2 isomer (with homocitrate protonated). According to the activation barriers, this is actually the preferred path for the E_6 state, since the barrier forming the H_2NNH_2 product is 14 kJ mol⁻¹ higher than that forming the $NH_3 + NH_2$ product. Thus, we can conclude that this reaction is possible with a maximum barrier of 78 kJ mol⁻¹.

Next, we considered proton-transfer reactions for the E_7 state. We started from two different structures, viz. with either H_2NNH_2 or NH_2 bound to Fe6. According to the previous paragraph, the latter should be the proper starting state, but since hydrazine (H_2NNH_2 is a known substrate of the enzyme, the former reaction should also be possible and was therefore also tested.

Possible proton-transfer reactions from the sulfide ions to H_2NNH_2 are shown in Figure 8. Protonation of S5A is 57 and 29 kJ mol⁻¹ more stable than protonation on S4B and S3B. The proton on S5A or S4B can first be transferred to Fe7 with barriers of 16 or 38 kJ mol⁻¹. The proton can then be transferred to S3B with a barrier of 60 kJ mol⁻¹. This proton can rotate to point towards Fe6 with a barrier of only 11 kJ mol⁻¹, leading to a stabilisation by 46 kJ mol⁻¹. Then,

the proton can be transferred successively to Fe6, S2B and finally to the substrate, passing barriers of 20, 24 and 31 kJ mol⁻¹, with reaction energies of −5, −24 and −19 kJ mol⁻¹. The product is the H_2NNH_3 intermediate, but it does not directly coordinate to the cluster. Therefore, the barrier for the cleavage of the N–N is rather high, 91 kJ mol⁻¹. The net barrier of the reaction (for the isomerisation of the S3B-protonated state) is 86 kJ mol⁻¹. However, there is also an alternative path, involving another conformation of the proton on Fe7 (Fe7(2), i.e. with the proton pointing towards Fe2) and with no proton transfer to S3B, giving a lower net barrier of 74 kJ mol⁻¹ (also shown in Figure 8).

Then, we studied the protonation of NH_2 to NH_3 in the E_7 state (in this case, the substrate has abstracted the hydroxyl proton from homocitrate). The results in Figure 9 show that protonation of S5A is 68 and 34 kJ mol⁻¹ more favourable than protonation of S4B and S3B. The proton can be transferred to Fe7 with a barrier of 28 kJ mol⁻¹. After a rotation of the proton (15 kJ mol⁻¹ barrier), it can be transferred to S3B with a barrier of 28 kJ mol⁻¹. After another rotation (with a barrier of 18 kJ mol⁻¹), the proton can be transferred to NH_2, passing a barrier of only 6 kJ mol⁻¹. This step is strongly exothermic (229 kJ mol⁻¹). The net barrier of the full reaction (compared to the state protonated on S5A) is 73 kJ mol⁻¹, for the rotation of the proton on S3B.

Finally, we studied the transfer of a proton from the sulfide ions to NH_3 in the E_8 state (again, with a proton abstracted from homocitrate). The results in Figure 10 show that it is 51 or 47 kJ mol⁻¹ more favourable to have the proton on S5A than on S4B or S3B. The proton can be transferred to Fe7 with a barrier of 2–47 kJ mol⁻¹. After a rotation (29 kJ mol⁻¹ barrier), it can be transferred to Fe6 (34 kJ mol⁻¹ barrier) and then to S2B (with a barrier of 11 kJ mol⁻¹) before it can move to NH_3 with a barrier of 49 kJ mol⁻¹. The product is NH_4^+, which dissociates from Fe6. The highest effective barrier of the reaction (relative to protonated S5A) is 69 kJ mol⁻¹ for the transfer of the proton
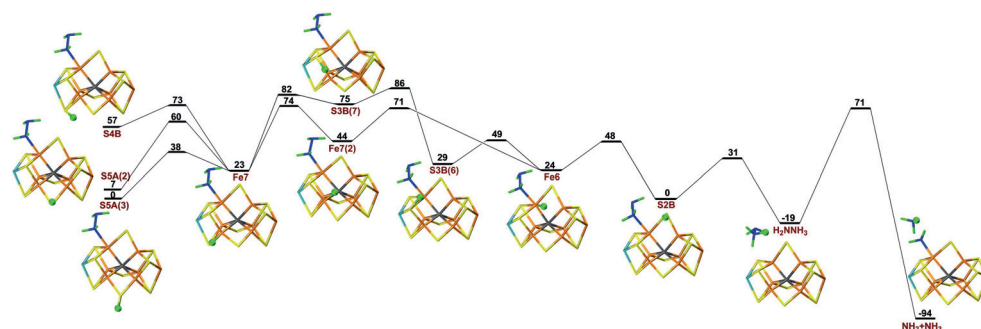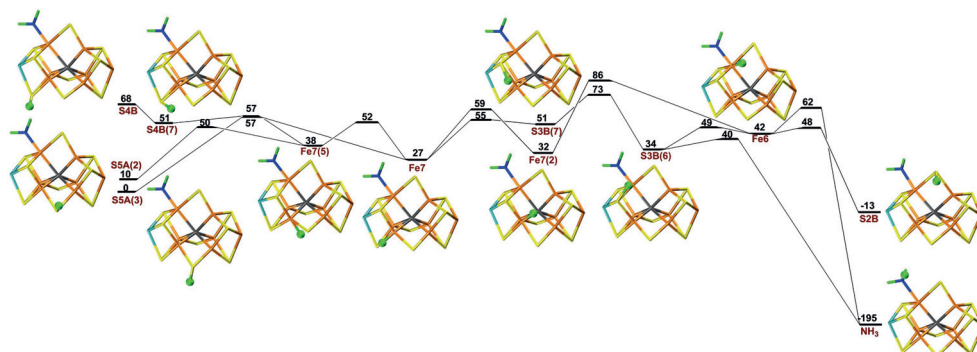

Figure 8. Reaction and activation energies, as well as structures for proton transfers in the E₇ state with S2B remaining bound to the cluster, from S5A, S4B or S3B to the H₂NNH₂ intermediate bound to Fe6, leading to a NH₂ product.
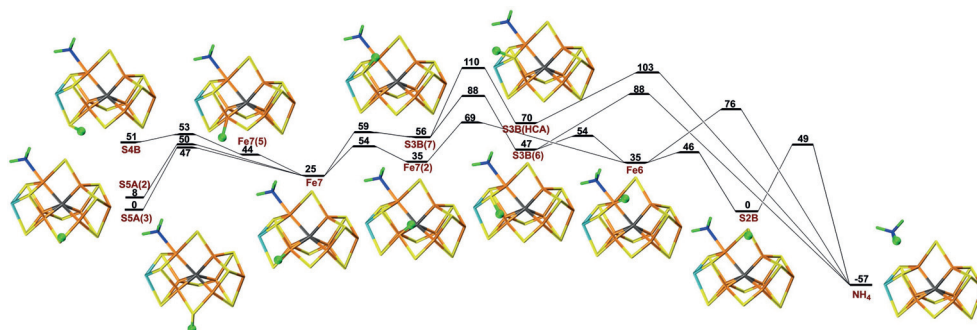
Figure 9. Reaction and activation energies, as well as structures for proton transfers in the E_7 state with S2B remaining bound to the cluster, from S5A, S4B or S3B to the NH_2 intermediate bound to Fe6, leading to a NH_3 product.

Figure 10. Reaction and activation energies, as well as structures for proton transfers in the E_8 state with S2B remaining bound to the cluster. In the S3B(HCA) structure, the extra proton points towards homocitrate.

from Fe7(2) to Fe6. There are several other possible paths with slightly higher barriers shown in Figure 10.

Conclusion

In this investigation, we have studied proton transfers within the FeMo cluster, assuming that the proton enters on either S3B, S4B or S5A, and is then transported to the substrate via the sulfide and Fe ions. Interestingly, we find that the net barriers for the proton transfers are in general higher when S2B has dissociated from the cluster than if S2B remains bound. In fact, in the former case, the maximum barriers are prohibitively large ($107–213$ kJ mol^{-1}) for the E_5–E_7 levels. When S2B remains bound, the maximum barriers are lower, $69–83$ kJ mol^{-1}. We have checked that the barriers cannot be lowered by relaxing the surrounding of the QM system or by using another DFT method. Thus, our results provide a strong argument against the dissociation of S2B.

For all E_n levels, protonation of S5A is always $29–98$ kJ mol^{-1} more favourable than protonation of S4B and S3B. States with Fe7 and Fe2 are also $16–74$ kJ mol^{-1} less stable. This shows that even if the proton initially is delivered to S3B, as Dance suggested, it would rapidly be transferred to S5A, which is thermodynamically more stable and the barriers for such a transfer is typically not higher than those involving a transfer towards the substrate.

However, a problem with the stable S5A protonation is that it becomes a thermodynamic sink for the reactions, increasing the effective barriers for the proton-transfer reactions. The individual barriers for the proton-transfer and proton-rotation reactions are typically $6–67$ kJ mol^{-1} for the seven paths in Figures 2–10. The highest individual barrier is typically observed for the first step of the path (i.e. moving the proton from S5A to Fe7) or the last step (moving the proton to the substrate). This corresponds to rates that are faster than the net reaction rate of nitrogenase. However, if the barriers are compared to the S5A state, the maximum effective barriers increase to $69–$
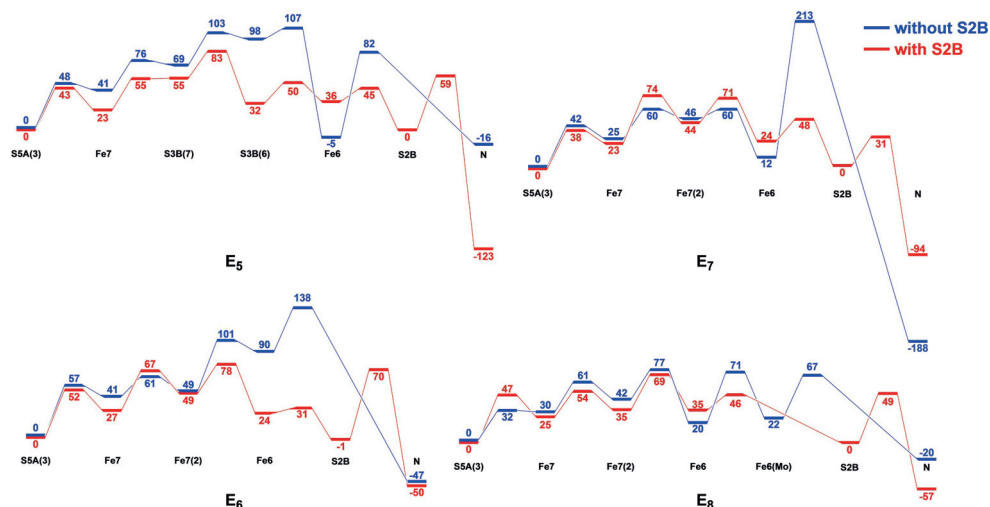
Figure 11. Relative energies for the proton-transfer reactions at the various E_n states with and without S2B bound. As shown in Figures 2–10, there are often multiple possible paths, but this figure shows only the most favourable ones (those with the lowest net barriers). Alternative paths, inspired by the results for the other E_n states were also tested, but sometimes failed owing to subtle differences in the structures.

83 kJ mol^{-1} (with S2B bound), which are higher than the experimental reaction rate. The reason for this may be that proton tunnelling has not been considered, that the transition states are approximate, that no entropy and thermal effects are considered and that TPSS may not give accurate results for such reactions. However, an alternative explanation may be that S5A actually is always protonated throughout the reaction mechanism of nitrogenase. We have tested such reactions for one state and found that the barriers are indeed reduced by 16 kJ mol^{-1} (discussed in the Supporting Information).

Dance also studied proton-transfer reactions within the FeMo cluster.[33, 35–37] However, he never studied transfers to or from S5A and therefore did not observe its high stability. Consequently, he underestimated the barriers. S3B is in the middle of our reaction mechanisms, 29–69 kJ mol^{-1} above the S5A(3) state. The barriers backwards to S5A are always lower than those forwards towards the substrate.

When S2B is bound, the energies of the various intermediates and transition state vary by only 4–20 kJ mol^{-1} between the E_5 to E_8 states (cf. Table S3), except for the last step. Two competing pathways are observed, either via S3B(7) and S3B(6) or via Fe7(2). Most pathways involve the transfer of the proton from S2B to the substrate.

When S2B has dissociated, the variation in the energies is larger and the barriers are higher. The same two pathways are observed but the proton needs to be transferred directly from Fe6 to the substrate (because S2B has dissociated). Figure 11 compares the energies with and without S2B. They are similar in the early steps but differ at the end. When S2B is present, it is normally used for the final transfer of the proton to the substrate. When it is not

present, the final proton transfer to the substrate is often problematic, in three cases leading to prohibitively large barriers. This explains why the proton transfers are significantly higher if S2B has dissociated than if S2B is still bound to the cluster.

We have tried to gain further understanding how the surrounding protein affects the proton transfers by dividing the total QM/MM energy into components from MM and from the point-charge model, indicating the importance of steric and electrostatic effects from the surrounding protein and solvent, outside the QM system. The MM energy correction is quite small (-11 to 15 kJ mol^{-1}), with a distribution that is only slightly biased to positive values (average 2 kJ mol^{-1}). It often increases slightly as the proton approaches the substrate.

On the other hand, the point-charge model has a quite large influence on the relative reaction energies, by -18 to 84 kJ mol^{-1}. It is positive for most intermediates and transition states (average 28 kJ mol^{-1}), indicating that protonation of S5A is more favoured by the electrostatics of the surrounding than the other protonation states.

Acknowledgements

© 2022 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

[1] J. Kim, D. C. Rees, *Science* **1992**, *257*, 1677–1682.

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

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

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

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

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

[7] R. N. F. Thorneley, D. J. Lowe in *Molybdenum Enzym* (Ed.: T. G. Spiro), Wiley, New York, **1985**, pp. 221–284.

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

[9] 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.

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

[11] B. Schmid, H.-J. Chiu, V. Ramakrishnan, J. B. Howard, D. C. Rees, *Handbook of Metalloproteins*, Wiley, Hoboken, **2006**, pp. 1025–1036.

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

[13] 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.

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

[15] L. E. Roth, F. A. Tezcan, *Methods Mol. Biol.* **2011**, *766*, 147–164.

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

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

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

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

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

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

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

[23] J. W. Peters, O. Einsle, D. R. Dean, S. DeBeer, B. M. Hoffman, P. L. Holland, L. C. Seefeldt, *Science* **2021**, *371*, eabe5481.

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

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

[26] 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.

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

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

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

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

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

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

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

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

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

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

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

[38] I. Dance, *Biochemistry* **2006**, *45*, 6328–6340.

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

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

Paper V

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

Hao Jiang, Oskar K. G. Svensson, and Ulf Ryde*

Cite This: _Inorg. Chem._ 2022, 61, 18067−18076

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Nitrogenase is the only enzyme that can cleave the triple bond in N_2, making nitrogen available for all lifeforms. Previous computational studies have given widely diverging results regarding the reaction mechanism of the enzyme. For example, some recent studies have suggested that one of the μ_2-bridging sulfide ligands (S2B) may dissociate from one of the Fe ions when protonated in the doubly reduced and protonated E_2 state, whereas other studies indicated that such half-dissociated states are unfavorable. We have examined how the relative energies of 26 structures of the E_2 state depend on details of combined quantum mechanical and molecular mechanical (QM/MM) calculations. We show that the selection of the broken-symmetry state, the basis set, relativistic effects, the size of the QM system, relaxation of the surroundings, and the conformations of the bound protons may affect the relative energies of the various structures by up to 12, 22, 9, 20, 37, and 33 kJ/mol, respectively. However, they do not change the preferred type of structures. On the other hand, the choice of the DFT functional strongly affects the preferences. The hybrid B3LYP functional strongly prefers doubly protonation of the central carbide ion, but such a structure is not consistent with experimental EPR data. Other functionals suggest structures with a hydride ion, in agreement with the experiments, and show that the ion bridges between Fe2 and Fe6. Moreover, there are two structures of the same type that are degenerate within 1−5 kJ/mol, in agreement with the observation of two EPR signals. However, the pure generalized gradient approximation (GGA) functional TPSS favors structures with a protonated S2B also bridging Fe2 and Fe6, whereas r^2SCAN (meta-GGA) and TPSSh (hybrid) prefer structures with S2B dissociated from Fe2 (but remaining bound to Fe6). The energy difference between the two types of structure is so small (7−18 kJ/mol) that both types need to be considered in future investigations of the mechanism of nitrogenase.
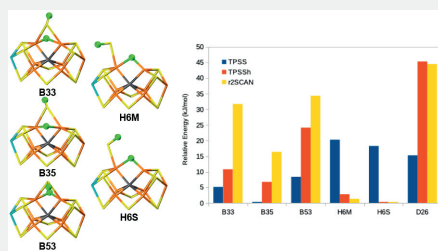
INTRODUCTION

Nitrogenases (EC 1.18/19.6.1) are the only group of enzymes that can cleave the inert triple bond in N_2, making atmospheric nitrogen available for all lifeforms.[1−3] Crystal structures have shown that the active site of Mo-nitrogenase is a complicated $MoFe_7S_9C$(homocitrate) cluster (the FeMo cluster), which is bound to the protein with one His and one Cys residue (Figure 1).[4−8] Alternative nitrogenases also exist with the Mo ion replaced with either vanadium or iron, but they have lower activities toward N_2.[9]

The nitrogenase reaction requires 16 ATP molecules to convert one N_2 molecule to two NH_3 molecules:[1−3]

$$N_2 + 8e^- + 8H^+ + 16ATP$$
$$\rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \quad (1)$$
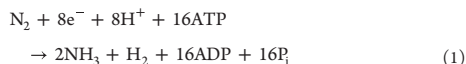
In addition, H_2 seems to be a mandatory byproduct. The reaction consumes eight electrons and protons. It is normally assumed that each reduction of the cluster leads to the uptake of one proton. Therefore, the reaction is normally described by eight intermediates, $E_0−E_7$, differing in the number of added electrons and protons.[10] Extensive biochemical, kinetic, and spectroscopic studies have indicated that the enzyme needs to be reduced four times (from E_0 to E_4) before N_2 can bind through reductive elimination of H_2.[1,2,11−17] Electron nuclear double resonance (ENDOR) experiments suggest that the E_4 state contains two hydride ions that bridge a pair of Fe ions each.[2,14,15]

Nitrogenase has also been extensively studied by computational methods.[18] Unfortunately, they have suggested very disparate reaction mechanisms. In fact, they do not even agree on the structure of the key E_4 state.[3,18−26] Structures have been suggested with the central carbide ion triply protonated or various combinations of protonated sulfide and iron ions (both bridging and terminal). Even for structures with two bridging hydride ions, the suggested models differ in what iron ions are
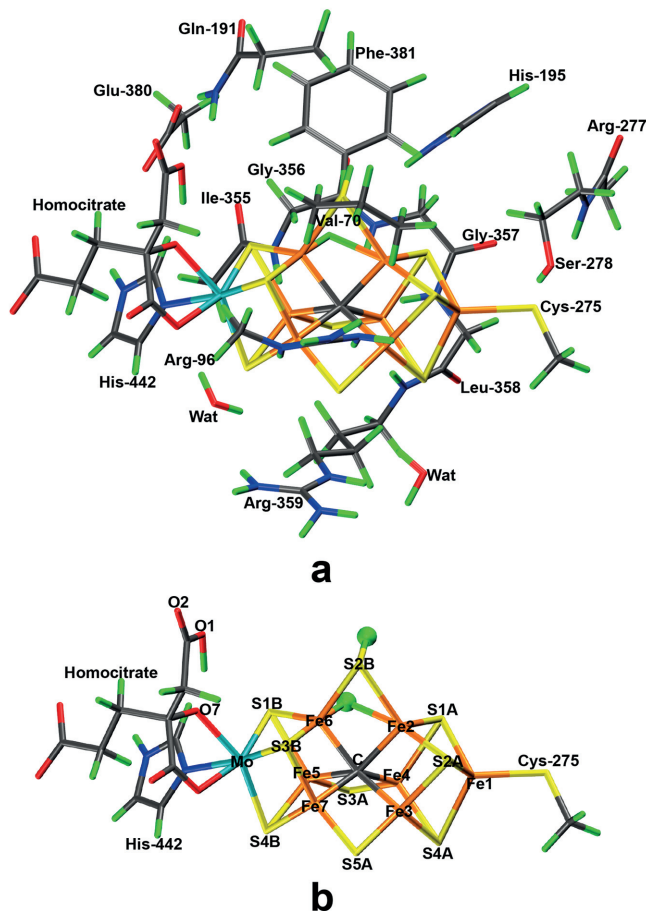
Figure 1. (a) Structure of the FeMo cluster (B33 structure), also illustrating the QM system used in the calculations, as well as nearby residues. (b) The FeMo cluster with atom names indicated.

involved (typically Fe2/Fe6, Fe3/Fe7, or Fe 4/Fe5, but possibly the same pair for both hydrides; atom names are indicated in Figure 1b), the position of the hydride ion relative to the μ_2-bridging sulfide ions, and whether the latter remains bound or not. An important reason for this discrepancy is that different DFT method give relative energies of different protonation states that can differ by over 600 kJ/mol, depending mainly on the amount of Hartree−Fock exchange in the method.[23]

A way to sort out these problems is to study simpler states with a lower number of possibilities. Most computational[18,27] and experimental[28,29] studies agree that the E_1 state involves the protonation of the S2B μ_2-bridging sulfide ligand of the FeMo cluster (but a recent experimental study of Fe-nitrogenase instead suggested a hydride ion[30]).

However, for the E_2 state, the predictions of different DFT methods start to diverge. Pure generalized gradient approx-

imation (GGA) functionals suggest that the second proton binds to a Fe ion.[27] Thereby, it formally forms a hydride ion and brings the oxidation state of the cluster back to that of the resting E_0 state, explaining why the same source of the electrons can be used for all E_n states.[31] Experimentally, the E_2 state has been studied by electron paramagnetic resonance (EPR) spectroscopy.[32−34] Two signals are observed, which have been interpreted as structures that both contain a hydride ion bound to one or two Fe ions. The two structures are nearly degenerate (within 4−8 kJ/mol) and isomerization between them involves some structural relaxation of the surroundings of the FeMo cluster.

Recently, several groups have suggested that the protonated S2B ligand may dissociate from one of its two Fe ions (Fe2 or Fe6).[26,35−37] In particular, Thorhallsson and Bjornsson (T&B) performed a study of the E_2 state of the FeMo cluster with the TPSSh functional.[38] They compared the relative energies of 18
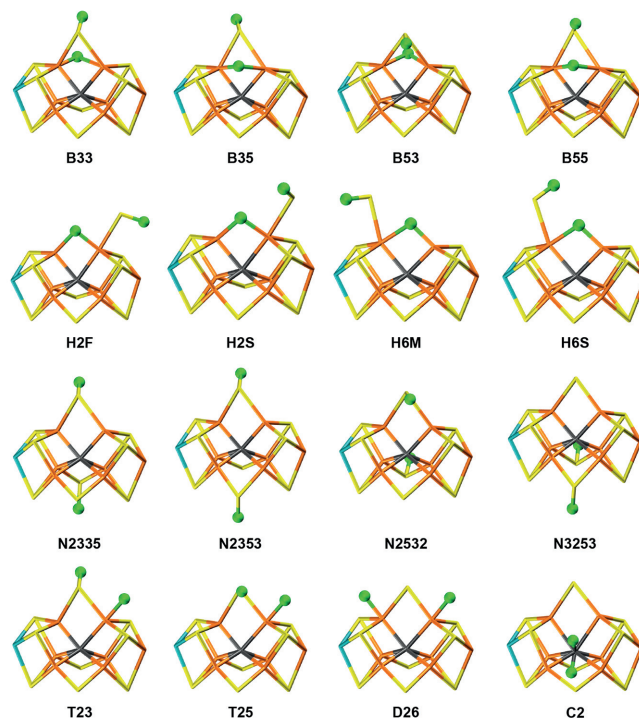
Figure 2. Sixteen of the 26 structures of the E$_2$ state investigated in this work (all 26 structures are shown in Figure S2 in the Supporting Information).

different states, involving protonation of either the three μ_2 bridging sulfide ions, the Fe ions, or the central carbide. With a QM/MM model, the most favorable structures had either the two protons on S2B and S5A, or a bridging hydride between Fe2 and Fe6 and a proton on S2B, which had dissociated from Fe2.

These results are quite different from what we obtained from a systematic study of ~40 different protonation states of E$_2$, all with a proton on S2B:[27] With the TPSS functional, states with a bridging hydride ion between Fe2 and Fe6 (with the protonated S2B ligand still binding to both Fe2 and Fe6) were most stable and the two states with the hydride ion on either side of S2B differed by only 2 kJ/mol. However, a state with a terminal hydride ion on Fe5 was only 3 kJ/mol less stable. On the other hand, states with the second proton on S5A (pointing either toward 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 they were not systematically explored.

Since half-dissociated S2B states have repeatedly been suggested to be involved in the reaction mechanism of nitrogenase,[26,35−37] it is important to sort out whether these discrepancies in the computational predictions are caused by the QM model used, the DFT method or by other details in the calculations. Therefore, we here perform a systematic study of 26 different E$_2$ structures at different levels of theory.

METHODS

The Protein. The calculations were based on the 1.0-Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[6] The setup of the protein is identical with that of our previous studies.[22−24,39] The entire heterotetramer was considered in the calculations, and the quantum mechanical (QM) calculations were concentrated on the FeMo clusters in the C subunit, because there is a buried imidazole molecule from the solvent rather close to the active site (~11 Å) in the A subunit. The two P clusters and the FeMo cluster in subunit A were modeled by MM in the fully reduced and resting states, respectively, using a QM charge model.[22] The protonation states of all residues were the same as before,[22] and the homocitrate ligand was modeled in the singly protonated state with a proton shared between the hydroxyl group (O7 that coordinates to Mo) and the O1 carboxylate atom[22,40] (all structures give two H−O distances of ~1.1 and ~1.4 Å; sometimes the proton is closer O1, sometimes O7, and when both structures can be found, they are typically degenerate within a few kJ/mol[22,27]). The protein was solvated in a sphere with a radius of 65 Å around the geometrical center of the protein. Cl$^-$ and Na$^+$ ions were added to an ionic strength of 0.2 M.[41] The final system contained 133 915 atoms. For the protein, we used the Amber ff14SB force field,[42] and water molecules were described by the TIP3P model.[43] The metal sites[22,44] were treated by a nonbonded model,[45] and charges were obtained with the restrained electrostatic potential method.[46]

The FeMo cluster was modeled by MoFe$_7$S$_9$C(homocitrate)-(CH$_3$S)(imidazole), where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included in the QM model, viz. Arg-96, Gln-

191 and His-195 (side chains), Ser-278 and Arg-359 (both backbone and side chain, including the CA and C and O atoms from Arg-277), Gly-356, Gly-357 and Leu-358 (backbone, including the CA and C and O atoms from Ile-355)), as well as two water molecules. Finally, the side chain of Glu-380 was included because it forms hydrogen bonds to Gln191 and His-442, as well as the side chains of Val-70 and Phe-381 because they are close to S2B, Fe2 and Fe6, i.e., the prime binding sites of the two added protons or hydride ions. The QM system involved 191 atoms and is shown in Figure 1a. The net charge of QM region was −4 e. In one set of calculations, we instead used the same QM model as T&B,[38] which contains 144 atoms and is shown in Figure S1 in the Supporting Information (the net charge is still −4 e).

QM Calculations. All QM calculations were performed with the Turbomole software (version 7.5).[47] All structures were studied with the TPSS,[48] TPSSh,[49] B3LYP,[50−52] and r²SCAN[53] functionals with the def2-SV(P) basis set.[54] In one set of calculations, we instead used the def2-TZVPD basis set.[55] The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[56,57] Empirical dispersion corrections were included with the DFT-D4 approach,[58] as implemented in Turbomole.

This investigation concentrates on the E_2 state of the FeMo cluster. Thus, we added two electrons and two protons to the resting E_0 state, which is at the formal $Mo^{III}Fe_3^{II}Fe_4^{III}$ oxidation state.[12,20,40,59] It was studied in the quartet spin state, in agreement with experiments.[32−34] Twenty-six different positions of the added protons were tested, as will be discussed below.

The electronic structure of all QM calculations was obtained with the broken-symmetry (BS) approach:[60] Each of the seven Fe ions was modeled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways.[39] The various BS states were obtained either by swapping the coordinates of the Fe ions[61] or with the fragment approach by Szilagyi and Winslow.[62] The various BS states are named by listing the number in the Noodleman nomenclature (BS1−10),[60] followed by the numbers of the three Fe ions with minority spin.

QM/MM Calculations. QM/MM calculations were performed with the ComQum software.[63,64] In this approach, the protein and solvent are split into three subsystems: System 1 (the QM region) was relaxed by QM methods. System 2 contained all residues and water molecules with at least one atom within 6 Å of any atom in system 1 and it was optionally relaxed by MM. It included residues 59, 61, 62, 65−74, 92, 95−98, 191−199, 226−231, 234, 235, 253−255, 273−282, 300, 353−355, 358−364, 377−383, 385, 386, 401 422−427, 438, 440−444, 450, and 450 from subunit C and residues 97, 98, 101, and 105 from subunit D, in total 1488 atoms from 87 residues and 35 water molecules). Finally, system 3 contained the remaining part of the protein and the solvent, and it was kept fixed at the original coordinates (equilibrated crystal structure, to avoid the risk that different calculations end up in different local minima).

In the QM calculations, system 1 was represented by a wave function, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarization of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM system was capped with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[63,65] All atoms were included in the point-charge model, except the CL atoms.[66] ComQum employs a subtractive scheme with van der Waals link-atom corrections.[67] No cutoff is used for the QM and QM−MM interactions. However, for the optional MM optimization of system 2, a 30-Å cutoff for the nonbonded interactions had to be used. The geometry optimizations were continued until the energy change between two iterations was less than 2.6 J/mol ($10^{−6}$ a.u.) and the maximum norm of the Cartesian gradients was below $10^{−3}$ a.u.

RESULTS AND DISCUSSION

Relative Energies of Different Protonation States at the TPSS − BS7−235 Level. In this investigation, we have examined 26 different possible structures for the E_2 state, differing in the positions of the two added protons. We examined six types of structures (illustrated in Figure 2):

1. With a hydride ion bridging the Fe2 and Fe6 ions and the second proton on S2B, which is also bridging Fe2 and Fe6. The proton on S2B can point in two directions, viz. toward either S3A or S5A, called B3 and B5. Likewise, the hydride ion can be on either side of S2B (viz. on the same side as either S3A or S5A), giving the second number in our structure code, e.g., B35. The four conformations are shown in Figure 2. T&B called this type of structures bH(2,6)-CBS(S2B) and considered only three of the four conformations (not B53).

2. With a hydride ion bridging the Fe2 and Fe6 ions and a protonated S2B that is binding only to either Fe2 or Fe6. Thus, S2B is half-dissociated and the structures are called H2 and H6, depending on which Fe ion S2B still binds to. A second letter indicates whether the proton points toward Fe1, Mo, or a sulfide ion (F, M, or S; e.g., H2F). The four structures are shown in Figure 2. T&B called these structures bH(2,6)-OBS(2) or bH(2,6)-OBS(6) for H2S and H6S and considered only three of the four conformations (not H2S).

3. With two terminal hydride ions, one on Fe2 and the other on Fe6 (called D26 and shown in Figure 2). The hydride ions always bind trans to the central carbide ion.

4. With a terminal hydride ion on either Fe2 (T2) or Fe6 (T6) and a protonated and bridging S2B. Again, the proton on S2B can point in two directions, viz. toward either S3A or S5A, giving the second number in our structure code, e.g., T23, cf. Figure 2.

5. With no hydride ions, but instead the two protons on either S2B, S3A, or S5A. Each proton can point in two directions, viz. toward the other two of these three μ_2 sulfide ions. The structures are named with a N (no hydride) followed by two pairs of numbers, the first indicating which sulfide is protonated and the second indicating the direction of that proton, e.g., N2532, indicating that S2B has a proton pointing toward S5A and S3A has the other proton pointing toward S2B (four examples are shown in Figure 2). T&B called these structures noH−CBS(S2B,S3B) and similar. They considered only four out of the 12 possible combinations and conformations.

6. With the two protons on the central carbide ion (C2). We considered only the state with the two protons directed toward the Fe2−Fe3−Fe6−Fe7 and Fe3−Fe4−Fe5−Fe7 faces of the cluster (shown in Figure 2), which was most favorable in our previous study.[27]

We have not studied states with S2B fully dissociated, because it is hard to accurately compare the energies of states with and without S2B, owing to uncertainties in the protonation state of S2B and the FeMo cluster, as well as the location of S2B after dissociation.

We first studied the 26 structures with TPSS-D4 and the BS-235 state. The results are shown in Table 1. It can be seen that the most stable structure is B33, i.e., with a bridging hydride between Fe2 and Fe6 on the side facing S3A and with the other proton on S2B (remaining bound to both Fe2 and Fe6),

Table 1. Relative Energies (kJ/mol) of the 26 Structures, Calculated with TPSS and the BS7-235 State[a]

Structure	TPSS	TPSSh[38]
B33	0.0	19.4
B35	12.2	20.1
B53	4.2	
B55	32.9	37.5
H2F	75.1	35.0
H2S	83.9	
H6M	15.2	0.3
H6S	13.2	1.4
D26	13.3	37.3
T23	17.4	25.0
T25	46.2	33.0
T63	25.7	28.2
T65	37.2	35.3
N2352	35.5	
N2353	25.1	0.0
N2552	46.2	
N2553	33.6	14.1
N2332	58.1	
N2335	53.5	
N2532	72.7	68.7
N2535	63.7	
N3252	79.6	
N3253	72.6	48.4
N3552	76.7	
N3553	70.5	
C2	138.8	

[a]The third column shows the corresponding results from the T&B article (relative QM/MM energies with His-195 protonated on NE2 from Table S2), obtained with TPSSh.[38]

also pointing toward S3A. The structure with the S2B proton pointing in the opposite direction (B53) is only 4 kJ/mol less stable. Structures with the hydride bridge on the other side of S2B are 12 kJ/mol less stable than B33 if the S2B proton points toward S3B (B35), whereas if the S2B proton points in the opposite direction (B55), the structure is 33 kJ/mol less stable.

We tested four structures with a half-dissociated protonated S2B ligand. The two structures with S2B dissociated from Fe6 are strongly unfavorable, 75−84 kJ/mol less stable than the B33 structure. However, the two structures with S2B dissociated from Fe2 are more stable, only 13−15 kJ/mol less stable than our best structure.

The structure with two terminal hydride ions on Fe2 and Fe6 (D26), as well as one structure with a terminal hydride ion on Fe2 (T23) are also competitive, 13 and 17 kJ/mol less stable than B33, respectively. However, the other structures with a terminal hydride on Fe2 or Fe6 are less stable, 26−46 kJ/mol worse than B33.

The 12 structures with both protons on the μ_2 bridging sulfide ions (S2B, S5A, or S3A) are unfavorable. The best is N2353 (with S2B and S5A protonated, both protons pointing toward S3A), 25 kJ/mol less stable than B33. Structures with S2B and S3A protonated are worse, and those with S3A and S5A are worst, 71−80 kJ/mol less stable than B33, indicating that the preference of protonation follow the order S2B > S5A > S3A.

Finally, we also tested the C2 structure with a doubly protonated carbide ion. However, with TPSS it is strongly unfavorable, 139 kJ/mol higher in energy than B33.

These results are quite similar to what was found in our previous study,[27] e.g., that the structure with a Fe2/6 bridging hydride and S2B protonated are most stable, better than structures with a terminal hydride or two protonated sulfide groups. However, the relative energies differ by up to 12 kJ/mol, owing to differences in the QM model and the BS state. Table 1 includes also the results from T&B.[38] It can be seen that the results diverge quite strongly from ours (by up to 40 kJ/mol). They reported three structures that are lowest in energy and nearly degenerate (within 1.4 kJ/mol), viz. one structure with S2B and S5A protonated (in our nomenclature N2353) and two structures with a bridging hydride and a protonated half-dissociated S2B (H6M and H6S). In our calculations these structures are 13−25 kJ/mol less stable than the best B33 structure. Likewise, T&B reported that the structures with both the hydride ion and S2B bridging Fe2 and Fe6 (they studied B33, B35, and B55) are 19−37 kJ/mol less stable than the N2354 structure. Naturally, such qualitative differences are alarming, considering that both studies use similar QM/MM methods. In the following sections, we examine possible reasons for this discrepancy.

Effect of the BS States. T&B tested four different BS states for their structures, viz. the three BS7 states (BS7−235, BS7−247, and BS7−346) and the BS10−147 state.[38] Still, energies for all four states are reported only for five structures, whereas only one state is reported for ten of the 24 structures studied. We decided to do a full BS investigation of all 35 BS states for one structure of each of the six types of structures mentioned above (and also for one of each of the three combinations of protonation of either S2B, S3A, and S5A, as well as for all four structures with half-dissociated S2B). For the other three structures of the same type, we investigated at least the BS7−235, BS7−247, BS7−346, and BS10−147 states and possibly additional states that were found to be low in energy for similar structures in the full investigation. In total 437 BS states were obtained with the TPSS-D4 functional for the 26 structures. The relative energies and Mulliken spin populations of the metal ions are listed in Table S1.

Table 2 shows the relative energies of the best BS state for the 26 different structures (second column). It can be seen that in most cases the change in relative energies is quite small, up to 14 kJ/mol (20 kJ/mol for C2). Table 2 also shows the optimum BS state (third column). For 20 of the structures, it is one of the three BS7 states (but BS7−235 in only five cases). BS6−157 is most stable for the two half-dissociated H2F and H2S structures, whereas the BS2−234, BS8−347, BS10−127, and BS10−147 states are most stable for one structure each (D26, T65, N2532, and C2, respectively). In general, there are several BS states within 7 kJ/mol of the best one.

Most importantly, it can be seen from Table 2 that the BS-state investigation does not solve the discrepancy between our results and those of T&B: The B-type structures with both the hydride ion and the protonated S2B bridging Fe2 and Fe6 are lowest in energy, with B35 best, 5, 8, and 23 kJ/mol below B33, B53, and B55. D26 (with two terminal hydrides) is 15 kJ/mol less stable than B35. The two structures with S2B dissociated from Fe2 (H6M and H6S) are 18−20 kJ/mol higher in energy than B35, whereas the two structures with S2B dissociated from Fe6 are appreciably less stable (74−84 kJ/mol). The best structure with one terminal hydride ion

Table 2. Relative Energies (kJ/mol) of the 26 Different Structures, Calculated with Different Methods[a]

Structure	TPSS						B3LYP	TPSSh		r²SCAN	
	SV		TZ	r.e.	T&B	Relax	SV	SV		SV	
B33	5.1	235	5.3	**0.0**	11.7	25.4	119.7	10.8	157	31.7	247
B35	**0.0**	247	1.2	5.0	**0.0**	**0.0**	117.8	6.8	147	16.4	235
B53	8.4	346	**0.0**	8.5	17.4	4.8	136.1	24.2	346	34.3	247
B55	22.6	247	20.2	29.0	23.2	24.1	138.0	26.7	147	39.5	235
H2F	74.4	157	72.2	71.9	80.8	56.4	196.9	71.2	235	78.6	346
H2S	83.5	157	78.8	82.8	83.8	66.6	198.7	78.4	135	82.3	346
H6M	20.3	235	23.8	14.5	20.3	44.0	82.8	2.7	235	1.3	235
H6S	18.3	235	20.2	12.7	19.3	43.4	80.7	**0.0**	235	**0.0**	235
D26	15.3	234	10.6	6.6	16.1	38.6	181.4	45.3	235	44.5	235
T23	22.5	235	11.3	13.9	23.9	24.9	123.8	31.3	346	42.9	157
T25	39.8	247	27.3	37.7	40.8	32.6	138.1	41.7	346	50.6	157
T63	29.1	247	38.2	25.6	34.7	33.9	130.6	30.2	235	48.4	235
T65	41.9	347	35.9	39.1	45.6	35.7	134.4	41.0	235	58.1	235
N2352	29.9	247	16.6	30.3	39.0	31.6	99.8	21.5	247	52.7	247
N2353	18.2	346	26.1	16.8	30.0	19.3	86.6	14.2	346	46.6	147
N2552	48.0	247	33.7	51.6	63.2	42.8	113.6	38.8	346	59.2	346
N2553	29.7	346	37.6	31.8	40.6	44.3	95.8	22.1	247	49.0	346
N2332	57.9	247	36.4	66.1	66.7	61.8	115.9	50.0	247	68.7	235
N2335	49.8	346	43.5	58.0	62.1	86.6	74.0	27.8	235	60.0	147
N2532	71.8	127	121.2	79.2	114.0	97.8	124.3	45.0	235	76.2	235
N2535	61.6	346	56.8	72.5	71.4	63.5	104.0	47.4	346	69.1	346
N3252	84.6	235	65.4	86.3	98.8	90.6	150.9	82.3	346	96.1	346
N3253	75.1	346	63.4	81.6	84.6	82.3	129.3	66.5	247	84.7	346
N3552	75.8	346	60.2	84.6	86.5	80.9	132.1	64.9	247	86.8	346
N3553	61.7	346	47.6	71.2	73.0	65.0	122.6	56.1	346	75.6	346
C2	143.8	147	153.6	155.7	164.3	146.4	**0.0** (-27.2)[b]	27.6	147	107.9	147

[a]Four different DFT methods were used, TPSS, B3LYP, TPSSh, and r²SCAN. All results were obtained with the def2-SV(P) basis set (SV), except those in the TZ column, which used the def2-TZVPD basis sets. In the r.e. column, relativistic effects are included. The "T&B" column shows the results with the smaller QM system used by T&B (still with TPSS). The "Relax" column shows the results obtained with the surrounding protein and water (within 6 Å of the QM system) allowed to relax by MM (also with TPSS). For TPSS, TPSSh, and r²SCAN, an investigation of the best BS state was performed and the best BS state is given after the energy, described by the three Fe ions with minority spin. The TZ, r.e., T&B, Relax, and B3LYP results were obtained for the same BS state as for TPSS. The most stable state for each method is marked in bold face. [b]For BS8−345.

(T23) is 23 kJ/mol less favorable than B35, and the best structure with no hydride ion (N2353) is 18 kJ/mol less favorable. The C2 structure is strongly disfavored. Thus, the BS states cannot explain the difference between our and the T&B results.

Table S1 also shows the TPSS spin populations on the metals. It can be seen that the largest Fe spin (in absolute terms) is 2.7−3.5 e (3.2 e on average). The average sorted spin populations decrease by ∼0.2 for the following four Fe ions, 2.9, 2.7, 2.6, and 2.4 e, still representing high-spin Fe. However, one or two of the Fe ions often have an appreciably lower spin population, as frequently observed for the FeMo cluster when hydride ions or other ligands coordinate to Fe:[18,27,44,68] 75% of the studied structures and BS states have one Fe ion with a spin population below 2 e and 37% have two such Fe ions (2% have three).

For the best BS state of each structure, Fe1 always has the largest spin population. The half-dissociated structures and C2 have no Fe ion with a low spin (the lowest one is 2.2−2.8 e). For the other structures, Fe6 has a low spin population (0.2− 1.8 e), except when S3A and S5A are both protonated (then instead Fe7 has a low spin population of 0.7−1.5 e). Sometimes, Fe7 (especially when S2B and S5A protonated), Fe5 (0.2 e for three structures with S2B and S3A protonated), or Fe2 (1.5−1.6 e for B53, T23, and T25) also have low spin

populations. The spin population on Mo is small and negative, −0.2 e on average for the best BS states.

Effect of the Basis Sets, Relativistic Effect, and the Model Size. Next, we studied how the basis sets affect the results. We calculated single-point energies for all 26 structures with the much larger def2-TZVPD basis set. The results in the fourth (TZ) column in Table 2 show that the basis set has a relatively small effect on the relative energies (as has also been observed before[27,69]); they change by up to 22 kJ/mol (−3 kJ/ mol on average). The effect is largest for the complexes with two protonated sulfide ions, whereas those with a bridging sulfide ion change by less than 8 kJ/mol. The B53, B35, and B33 structures are still most stable and nearly degenerate (with 5 kJ/mol). However, the D26 and T23 structures are only 11 kJ/mol less stable, whereas the best half-dissociated structures (H6M and H6S) are 20−24 kJ/mol higher in energy.

Likewise, relativistic effects (estimated from the mass− velocity and Darwin terms) change the results by up to 12 kJ/ mol (2 kJ/mol on average). The same three states (B33, B35, and B53) are still most stable, within 9 kJ/mol, but D26 is also within this range and the H6M and H6S half-dissociated structures are only 13−14 kJ/mol worse.

T&B used a smaller QM model than the one used in our study (144 atoms, compared to 191 atoms; cf. Figures 1 and S1). In particular, their model is missing the backbone of

residues 355–359, which forms four hydrogen bonds to S3A, and two water molecules, which both form hydrogen bonds to S5A. To investigate the effect of this smaller QM model, we optimized the 26 protonation states also with their 144-atom model (still with the TPSS-D4 method and with the best BS state for the larger QM model). This had a rather small effect on the relative energies of the structures, 1–20 kJ/mol. In particular, all structures with no hydride ions are destabilized by 9–15 kJ/mol kJ/mol. However, the ordering of the relative stability of the structures is not changed. B35 is most stable, 12–19 kJ/mol more stable than B33, D26, B53, and H6S. Thus, the size of the QM model cannot explain the discrepancy between our and the T&B results.

Relaxation of the Surrounding Protein. T&B allowed 1001 atoms surrounding the FeMo cluster to relax during the geometry optimization, whereas we by default keep all atoms outside the QM system fixed at the crystal structure (to avoid different states converge to different local minima of the surroundings). Even if most groups close to the S2B ligand are included in the QM system, it is possible that the surroundings may relax if S2B dissociates from one of the Fe ions, favoring such half-dissociated structures, especially as EPR experiments have indicated that some structural relaxation of the surroundings are involved in the conversion of the two observed E_2 states.[34] Therefore, we also run one set of calculations in which all residues with any atom within 6 Å of the QM system are relaxed by a MM optimization in every QM/MM geometry optimization step (1488 atoms). The results of these calculations are also included in Table 2 (column Relax).

It can be seen that this had a slightly larger effect on the relative energies, up to 37 kJ/mol. The two structures with S2B dissociated from Fe6 (H2F and H2S) are favored by 17–18 kJ/mol, but the two structures with S2B dissociated from Fe2 (H6M and H6S), as well as the D26 structure (with two hydride ion) are disfavored by 23–25 kJ/mol. Consequently, B35 is still the best structure, 5 kJ/mol better than B53, whereas N2353 is the third-best structure, 19 kJ/mol less stable than B35. The best half-dissociated structure, H6S, is 44 kJ/mol less stable than B35.

The movements of the surroundings are modest, with root-mean-squared movements of less 0.4 Å for the protein residues and somewhat larger for some water molecules (up to 0.7 Å). The movements are similar for all structures.

Effect of the DFT Method. Previous studies have shown alarming differences between relative energies of nitrogenase FeMo cluster isomers calculated with different DFT methods.[23,27] T&B also studied a few of their structures with both the TPSSh and the TPSS functionals[38] and found quite large differences for the relative energies calculated with the two methods (up to 59 kJ/mol). Therefore, we reoptimized all our 26 structures also with the r^{2}SCAN, TPSSh, and B3LYP functionals. The first is a meta-GGA functional that has been recommended for nitrogenase models,[70] whereas the other two are hybrid functionals with 10 and 20% Hartree–Fock exchange, respectively. T&B used mainly TPSSh in their study. The results are also included in Table 2.

It can be seen that the DFT functional indeed has a strong influence on what E_2 structure is preferred. B3LYP strongly prefers the C2 state, as was also previously observed.[27] B3LYP also disfavors all states with Fe-bound hydride ions.

TPSSh has a smaller effect on the relative energies, up to 30 kJ/mol (besides the C2 structure, which is stabilized by 116

kJ/mol). It also somewhat disfavors structures with hydride ions and favors the half-dissociated structures, especially those for which S2B has dissociated from Fe2. Consequently, H6S becomes the best structure, 3 kJ/mol more stable than H6M and 7 kJ/mol more stable than the nondissociated B35 structure. The best structure with two protonated sulfide ions is N2353, 14 kJ/mol less stable than H6M. C2 is strongly stabilized, but it is still 28 kJ/mol less stable than H6S.

These results are based on the best BS state according to a restricted scan of a few of the best BS for TPSS (at least the BS7–235, BS7–247, BS7–346, and BS10–147 states and typically a few more states; a full investigation was performed on H6S; relative energies and Mulliken metal spin populations are shown in Table S2). The best BS state agrees with that suggested by T&B for seven of the 15 overlapping structures and they involve either the BS7 states or BS10–147 for all except two structures (BS6–157 for B33 and HS2).

The Mulliken spin populations calculated by TPSSh are in general larger and more even than those obtained with TPSS. The average absolute values of the sorted spin populations are 3.6, 3.5, 3.4, 3.3, 3.2, 2.9, and 2.6 e, i.e., 0.4–1.2 e larger than for the corresponding TPSS spin populations. No structure has two Fe ions with a low (<2 e) spin population, but 15% of the structures have one Fe ion with a spin population less than 2 e. However, among the best BS states, only one structure (B33) has a Fe spin population of 2.0 e; for the other structures, the lowest Fe population is 2.4–3.3 e. Thus, with TPSSh, a low spin population typically indicates convergence to a suboptimal wave function or BS. Our TPSSh spin populations are typically ~0.2 e larger than those reported by T&B,[38] reflecting differences in the QM model, the basis sets, and other details of the calculations. The difference is never larger than 0.5 e. The Mo spin population for the best BS state is −0.2 to −0.7 e (average −0.4 e) with a single exception (B33 has a positive population of 0.5 e).

The results obtained with the r^{2}SCAN functional are similar. Compared to TPSSh, the half-dissociate structures, as well as D26 (with two terminal hydrides), are stabilized relative to the other states. Consequently, the two half-dissociated H6S and H6M structures (which are essentially degenerate) are most stable, followed by B35, which is 16 kJ/mol less stable. The best structures with terminal hydrides (T23 and D26) are 43–44 kJ/mol less stable, and the best structure with two protonated sulfide ions (N2353) is 47 kJ/mol less stable than H6S.

As with TPSSh, we made a limited investigation of the most stable BS state with r^{2}SCAN (shown in Table S3). There are some variations compared to TPSS and TPSSh, but mainly within the BS7 and BS10–147 states. However, T23 and T25 are most stable in the BS6–157 state.

The Fe spin populations of r^{2}SCAN are very similar to those of TPSSh with an average difference of only 0.04 e. However, the spin population on Mo for the best BS state varies much more than for the other two methods. It is positive (0.4–0.9 e) for three structures, and it is −1.4 to −1.6 e for seven of the structures with two protonated sulfide ions.

Considering that at the TPSS level, relaxation of the surroundings had a quite large effect on the relative energies and destabilized the best half-dissociated structures by ~25 kJ/mol, we also ran QM/MM geometry optimizations with relaxed surroundings with the TPSSh and r^{2}SCAN functionals for the eight best structures. The results in Table S4 show that

in this case, relaxation of the surrounding has rather small effects, up to 10 kJ/mol, except that B35 is selectively destabilized by both methods (by 19−33 kJ/mol). Thus, the half-dissociated structures remain most favorable.

CONCLUSIONS

In this study, we have tried to explain the discrepancy between the results obtained by our group[27] and those obtained by T&B[38] regarding the relative energies of the E_2 state of nitrogenase and, in particular, the importance of half-dissociated states with the S2B ligand dissociated from either Fe2 or Fe6. Our results show that the BS state, the basis set, relativistic effects, the size of the QM model, and the relaxation of the surrounding have some influence on the relative stabilities (by up to 12, 22, 9, 20, and 37 kJ/mol, respectively). Our calculations also emphasize the importance of studying all conformations of the added protons, which may change the relative energies by up to 33 kJ/mol. However, neither of these variations changes the relative ordering of the different types of structures. Instead the difference between the two studies is caused by the use of different DFT methods:

- TPSS favors structures with both the hydride and S2B bridging Fe2 and Fe6 (B35, B33, and B53), which are 15−18 kJ/mol better than structures with a half-dissociated S2B (H6S), with two terminal hydride ions (D26) or with no hydride ions (N2353).
- B3LYP strongly favors the C2 structure with a doubly protonated carbide ion, which is 101 kJ/mol more stable than N2335. It strongly disfavors all structures with Fe-bound hydride ions.
- TPSSh also shows similar tendencies, but to a smaller extent (owing to the smaller amount of Hartree−Fock exchange). It also favors the half-dissociated structure so that H6S and H6M becomes 7 kJ/mol more stable than B35. This shows that the Hartree−Fock exchange weakens the Fe−S and Fe−H bonds.
- r^2SCAN selectively favors the half-dissociated structures and C2 (but the latter much less than the hybrid functionals). Therefore, H6S and H6M become 16 kJ/mol more stable than B35.

From Table 2, it can be seen that eight E_2 structures (B35, B33, B53, H6M, H6S, D26, N2353, and C2) are competitive in terms of energies (within 20 kJ/mol at least with some DFT method). However, two of them (C2 and N2353) do not contain iron-bound hydride ions and are therefore not compatible with the EPR data.[33,34] This calls in doubt the B3LYP calculations, which strongly prefer the C2 structure. The relative energies of the remaining five structures obtained with TPSS, TPSSh, and r^2SCAN are shown in Figure 3. It can be seen that the D26 structure is not preferred with any method and is relatively high in energy. The other structures all have a hydride ion bridging Fe2 and Fe6, showing that E_2 most likely contains such a bridging hydride ion. The low-energy structures are of two types: either with S2B also bridging Fe2 and Fe6 (B33, B35, and B53) or with S2B dissociated from Fe2 (H6M and H6S). Within these groups, the structures differ only in the direction of the proton on S2B and (for B-type structures) which side of S2B the hydride ion is. All three methods agree that there are at least one more low-energy structure of the same type within 1−5 kJ/mol, in agreement with the observation of two nearly degenerate structures in the EPR experiments.[33,34] Thus, the only
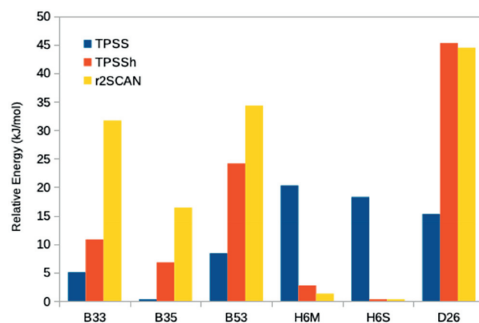
Figure 3. Relative stabilities of the best structures containing hydride ions (in agreement with experimental data[33,34]) for the TPSS, TPSSh, and r^2SCAN methods.

disagreement is that TPSS points out the nondissociated S2B as the best, whereas TPSSh and r^2SCAN prefer the half-dissociated structures, with energy difference of 7−18 kJ/mol toward the other structure.

A natural question is then which DFT functional to trust. Recent studies have indicated that r^2SCAN, TPSSh, and B3LYP* (with 15% Hartree−Fock exchange) give the best structures of Fe_2 and FeMo models.[70] On the other hand, another study indicated that pure GGA functionals, like PBE and PW91, gave the best results for structures and energies involving the binding of H_2 and N_2 to small transition-metal models with relation to nitrogenase.[71] Yet, further studies on the latter systems with dispersion-corrected DFT functionals indicated that pure GGA functionals give better structures, whereas hybrid functionals give more reliable energetic results.[72] Consequently, further investigation are required to decide which DFT functional gives the most reliable results for nitrogenase models. However, considering the small energy difference between the two types of structures with most methods, it is clear that both types need to be considered in investigations of the mechanism of nitrogenase.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.inorgchem.2c02488.

> Figure of the 144-atom T&B QM model; figure of all 26 structures; relative energies and Mulliken spin populations for the various BS states for the 26 structures, calculated with TPSS, TPSSh, and r^2SCAN (PDF)
> Coordinates of all studied structures (ZIP)

AUTHOR INFORMATION

Corresponding Author

   Ulf Ryde − *Department of Theoretical Chemistry, Lund University, Chemical Centre, SE-221 00 Lund, Sweden;* orcid.org/0000-0001-7653-8489; Phone: +46−46 2224502; Email: Ulf.Ryde@teokem.lu.se; Fax: +46−46 2228648

Authors

   Hao Jiang − *Department of Theoretical Chemistry, Lund University, Chemical Centre, SE-221 00 Lund, Sweden;* orcid.org/0000-0002-9641-1634

Oskar K. G. Svensson − *Department of Theoretical Chemistry, Lund University, Chemical Centre, SE-221 00 Lund, Sweden*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.inorgchem.2c02488

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Burgess, B. K.; Lowe, D. J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96*, 2983−3012.

(2) Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem. Rev.* **2014**, *114*, 4041−4062.

(3) Seefeldt, L. C.; Yang, Z.-Y.; Lukoyanov, D. A.; Harris, D. F.; Dean, D. R.; Raugei, S.; Hoffman, B. M. Reduction of Substrates by Nitrogenases. *Chem. Rev.* **2020**, *120*, 5082−5106.

(4) Kim, J.; Rees, D. C. Structural Models for the Metal Centers in the Nitrogenase Molybdenum-Iron Protein. *Science (80-.).* **1992**, *257* (5077), 1677−1682.

(5) Einsle, O.; Tezcan, F. A.; Andrade, S. L. A.; Schmid, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science (80-.).* **2002**, *297* (5587), 1696.

(6) Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S. L. A.; Schleicher, E.; Weber, S.; Rees, D. C.; Einsle, O. Evidence for Interstitial Carbon in Nitrogenase FeMo Cofactor. *Science (80-.).* **2011**, *334*, 940−940.

(7) Spatzal, T.; Perez, K. A.; Einsle, O.; Howard, J. B.; Rees, D. C. Ligand Binding to the FeMo-Cofactor: Structures of CO-Bound and Reactivated Nitrogenase. *Science (80-.).* **2014**, *345* (6204), 1620−1623.

(8) Einsle, O. Nitrogenase FeMo Cofactor: An Atomic Structure in Three Simple Steps. *J. Biol. Inorg. Chem.* **2014**, *19* (6), 737−745.

(9) Eady, R. R. Structure-Function Relationships of Alternative Nitrogenases. *Chem. Rev.* **1996**, *96*, 3013−3030.

(10) Thorneley, R. N. F.; Lowe, D. J.Kinetics and Mechanism of the Nitrogenase Enzyme System. In *Molybdenum Enzymes*, Spiro, T. G., Ed.; Wiley: New York, 1985; pp 221−284.

(11) Lancaster, K. M.; Roemelt, M.; Ettenhuber, P.; Hu, Y.; Ribbe, M. W.; Neese, F.; Bergmann, U.; DeBeer, S. X-Ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor. *Science (80-.).* **2011**, *334*, 974−977.

(12) Bjornsson, R.; Lima, F. A.; Spatzal, T.; Weyhermüller, T.; Glatzel, P.; Bill, E.; Einsle, O.; Neese, F.; DeBeer, S. Identification of a Spin-Coupled Mo(III) in the Nitrogenase Iron-Molybdenum Cofactor. *Chem. Sci.* **2014**, *5* (8), 3096−3103.

(13) Schmid, B.; Chiu, H.-J.; Ramakrishnan, V.; Howard, J. B.; Rees, D. C.Nitrogenase. In *Handbook of Metalloproteins*; John Wiley & Sons, Ltd, 2006; pp 1025−1036.

(14) Igarashi, R. Y.; Laryukhin, M.; Dos Santos, P. C.; Lee, H.-I.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Trapping H⁻ Bound to the Nitrogenase FeMo-Cofactor Active Site during H₂ Evolution: Characterization by ENDOR Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127*, 6231−6241.

(15) Hoeke, V.; Tociu, L.; Case, D. A.; Seefeldt, L. C.; Raugei, S.; Hoffman, B. M. High-Resolution ENDOR Spectroscopy Combined with Quantum Chemical Calculations Reveals the Structure of Nitrogenase Janus Intermediate E₄(4H). *J. Am. Chem. Soc.* **2019**, *141*, 11984−11996.

(16) Roth, L. E.; Tezcan, F. A. X-Ray Crystallography. *Methods Mol. Biol.* **2011**, *766*, 147−164.

(17) Van Stappen, C.; Decamps, L.; Cutsail, G. E.; Bjornsson, R.; Henthorn, J. T.; Birrell, J. A.; DeBeer, S. The Spectroscopy of Nitrogenases. *Chem. Rev.* **2020**, *120*, 5005−5081.

(18) Dance, I. Computational Investigations of the Chemical Mechanism of the Enzyme Nitrogenase. *ChemBioChem.* **2020**, *21*, 1671−1709.

(19) Rohde, M.; Sippel, D.; Trncik, C.; Andrade, S. L. A.; Einsle, O. The Critical E₄ State of Nitrogenase Catalysis. *Biochemistry* **2018**, *57*, 5497−5504.

(20) Siegbahn, P. E. M. Model Calculations Suggest That the Central Carbon in the FeMo-Cofactor of Nitrogenase Becomes Protonated in the Process of Nitrogen Fixation. *J. Am. Chem. Soc.* **2016**, *138* (33), 10485−10495.

(21) Siegbahn, P. E. M. Is There Computational Support for an Unprotonated Carbon in the E₄ State of Nitrogenase? *J. Comput. Chem.* **2018**, *39*, 743−747.

(22) Cao, L.; Caldararu, O.; Ryde, U. Protonation States of Homocitrate and Nearby Residues in Nitrogenase Studied by Computational Methods and Quantum Refinement. *J. Phys. Chem. B* **2017**, *121*, 8242−8262.

(23) Cao, L.; Ryde, U. Extremely Large Differences in DFT Energies for Nitrogenase Models. *Phys. Chem. Chem. Phys.* **2019**, *21*, 2480−2488.

(24) Cao, L.; Ryde, U. What Is the Structure of the E₄ Intermediate in Nitrogenase? *J. Chem. Theory Comput.* **2020**, *16* (3), 1936−1952.

(25) Lukoyanov, D.; Khadka, N.; Dean, D. R.; Raugei, S.; Seefeldt, L. C.; Hoffman, B. M. Photoinduced Reductive Elimination of H₂ from the Nitrogenase Dihydride (Janus) State Involves a FeMo-Cofactor-H₂ Intermediate. *Inorg. Chem.* **2017**, *56*, 2233−2240.

(26) Thorhallsson, A. T.; Benediktsson, B.; Bjornsson, R. A Model for Dinitrogen Binding in the E₄ State of Nitrogenase. *Chem. Sci.* **2019**, *10*, 11110−11124.

(27) Cao, L.; Caldararu, O.; Ryde, U. Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/ Molecular Mechanics (QM/MM) Calculations. *J. Chem. Theory Comput.* **2018**, *14*, 6653−6678.

(28) Van Stappen, C.; Thorhallsson, A. T.; Decamps, L.; Bjornsson, R.; DeBeer, S. Resolving the Structure of the E₁ State of Mo Nitrogenase through Mo and Fe K-Edge EXAFS and QM/MM Calculations. *Chem. Sci.* **2019**, *10* (42), 9807−9821.

(29) Van Stappen, C.; Davydov, R.; Yang, Z.-Y.; Fan, R.; Guo, Y.; Bill, E.; Seefeldt, L. C.; Hoffman, B. M.; DeBeer, S. Spectroscopic Description of the E₁ State of Mo Nitrogenase Based on Mo and Fe X-Ray Absorption and Mössbauer Studies. *Inorg. Chem.* **2019**, *58*, 12365−12376.

(30) Lukoyanov, D. A.; Harris, D. F.; Yang, Z.-Y.; Pérez-González, A.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. The One-Electron Reduced Active-Site FeFe-Cofactor of Fe-Nitrogenase Contains a Hydride Bound to a Formally Oxidized Metal-Ion Core. *Inorg. Chem.* **2022**, *61*, 5459−5464.

(31) Doan, P. E.; Telser, J.; Barney, B. M.; Igarashi, R. Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. ⁵⁷Fe ENDOR Spectroscopy and 'Electron Inventory' Analysis of the Nitrogenase E₄ Intermediate Suggest the Metal-Ion Core of FeMo-Cofactor Cycles Through Only One Redox Couple. *J. Am. Chem. Soc.* **2011**, *133*, 17329−17340.

(32) Lowe, D. J.; Eady, R. R.; Thorneley, R. N. F. Electron-Paramagnetic-Resonance Studies on Nitrogenase of *Klebsiella pneumoniae*. Evidence for Acetylene- and Ethylene-Nitrogenase Transient Complexes. *Biochem. J.* **1978**, *173*, 277−290.

(33) Lukoyanov, D.; Yang, Z.-Y.; Duval, S.; Danyal, K.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. A Confirmation of the Quench-Cryoannealing Relaxation Protocol for Identifying Reduction States of

Freeze-Trapped Nitrogenase Intermediates. *Inorg. Chem.* **2014**, *53*, 3688−3693.

(34) Lukoyanov, D. A.; Khadka, N.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Hydride Conformers of the Nitrogenase FeMo-Cofactor Two-Electron Reduced State $E_2(2H)$, Assigned Using Cryogenic Intra Electron Paramagnetic Resonance Cavity Photolysis. *Inorg. Chem.* **2018**, *57*, 6847−6852.

(35) Hallmen, P. P.; Kästner, J. N_2 Binding to the FeMo-Cofactor of Nitrogenase. *Zeitschrift fur Anorg. und Allg. Chemie* **2015**, *641* (1), 118−122.

(36) Dance, I. How Feasible Is the Reversible S-Dissociation Mechanism for the Activation of FeMo-Co, the Catalytic Site of Nitrogenase? *Dalt. Trans.* **2019**, *48*, 1251−1262.

(37) Lukoyanov, D. A.; Krzyaniak, M. D.; Dean, D. R.; Wasielewski, M. R.; Seefeldt, L. C.; Hoffman, B. M. Time-Resolved EPR Study of H_2 Reductive Elimination from the Photoexcited Nitrogenase Janus $E_4(4H)$ Intermediate. *J. Phys. Chem. B* **2019**, *123*, 8823−8828.

(38) Thorhallsson, A. T.; Bjornsson, R. The E_2 State of FeMoco: Hydride Formation versus Fe Reduction and a Mechanism for H_2 Evolution. *Chem. - A Eur. J.* **2021**, *27*, 16788−16800.

(39) Cao, L.; Ryde, U. Influence of the Protein and DFT Method on the Broken-Symmetry and Spin States in Nitrogenase. *Int. J. Quantum Chem.* **2018**, *118*, No. e25627.

(40) Bjornsson, R.; Neese, F.; DeBeer, S. Revisiting the Mössbauer Isomer Shifts of the FeMoco Cluster of Nitrogenase and the Cofactor Charge. *Inorg. Chem.* **2017**, *56* (3), 1470−1477.

(41) Barney, B. M.; McClead, J.; Lukoyanov, D.; Laryukhin, M.; Yang, T. C.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Diazene (HN = NH) Is a Substrate for Nitrogenase: Insights into the Pathway of N2 Reduction. *Biochemistry* **2007**, *46* (23), 6784−6794.

(42) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.

(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(44) Cao, L.; Ryde, U. N_2H_2 Binding to the Nitrogenase FeMo Cluster, Studied by QM/MM Methods. *J. Biol. Inorg. Chem.* **2020**, *25*, 521−540.

(45) Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7* (8), 2452−2463.

(46) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.

(47) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (2), 91−100.

(48) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(49) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119* (23), 12129−12137.

(50) Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic-Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785−789.

(52) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372−1377.

(53) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r²SCAN Meta-Generalized

Gradient Approximation. *J. Phys. Chem. Lett.* **2020**, *11* (19), 8208−8215.

(54) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97* (4), 2571−2577.

(55) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297−3305.

(56) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240* (4), 283−289.

(57) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97* (1−4), 119−124.

(58) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122.

(59) Spatzal, T.; Schlesier, J.; Burger, E.-M.; Sippel, D.; Zhang, L.; Andrade, S. L. A.; Rees, D. C.; Einsle, O. Nitrogenase FeMoco Investigated by Spatially Resolved Anomalous Dispersion Refinement. *Nat. Commun.* **2016**, *7*, 10902.

(60) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. FeMo Cofactor of Nitrogenase: A Density Functional Study of States M^N, M^{OX}, M^R, and M^I. *J. Am. Chem. Soc.* **2001**, *123* (49), 12392−12410.

(61) Greco, C.; Fantucci, P.; Ryde, U.; Gioia, L. de. Fast Generation of Broken-Symmetry States in a Large System Including Multiple Iron-Sulfur Assemblies: Investigation of QM/MM Energies, Clusters Charges, and Spin Populations. *Int. J. Quantum Chem.* **2010**, *111*, 3949−3960.

(62) Szilagyi, R. K.; Winslow, M. A. On the Accuracy of Density Functional Theory for Iron-Sulfur Clusters. *J. Comput. Chem.* **2006**, *27* (12), 1385−1397.

(63) Ryde, U. The Coordination of the Catalytic Zinc in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput. Aided. Mol. Des.* **1996**, *10*, 153−164.

(64) Ryde, U.; Olsson, M. H. M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.* **2001**, *81*, 335−347.

(65) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.

(66) Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761−777.

(67) Cao, L.; Ryde, U. On the Difference between Additive and Subtractive QM/MM Calculations. *Front. Chem.* **2018**, *6*, 89.

(68) Cao, L.; Ryde, U. Putative Reaction Mechanism of Nitrogenase after Dissociation of a Sulfide Ligand. *J. Catal.* **2020**, *391*, 247−259.

(69) Cao, L.; Ryde, U. Influence of the Protein and DFT Method on the Broken-Symmetry and Spin States in Nitrogenase. *Int. J. Quantum Chem.* **2018**, *118*, No. e25627.

(70) Benediktsson, B.; Bjornsson, R. Analysis of the Geometric and Electronic Structure of Spin-Coupled Iron-Sulfur Dimers with Broken-Symmetry DFT: Implications for FeMoco. *J. Chem. Theory Comput.* **2022**, *18* (3), 1437−1457.

(71) Dance, I. Evaluations of the Accuracies of DMol3 Density Functionals for Calculations of Experimental Binding Enthalpies of N_2, CO, H_2, C_2H_2 at Catalytic Metal Sites. *Mol. Simul.* **2018**, *44*, 568−581.

(72) Torbjörnsson, M.; Ryde, U. Comparison of the Accuracy of DFT Methods for Reactions with Relevance to Nitrogenase. *Electron. Struct.* **2021**, *3*, 34005.

Paper VI

Check for updates

Putative reaction mechanism of nitrogenase with a half-dissociated S2B ligand†

Hao Jiang [ID] and Ulf Ryde [ID] *

We have studied whether dissociation of the S2B sulfide ligand from one of its two coordinating Fe ions may affect the later parts of the reaction mechanism of nitrogenase. Such dissociation has been shown to be favourable for the E_2–E_4 states in the reaction mechanism, but previous studies have assumed that S2B either remains bridging or has fully dissociated from the active-site FeMo cluster. We employ combined quantum mechanical and molecular mechanical (QM/MM) calculations with two density-functional theory methods, r^2SCAN and TPSSh. To make dissociation of S2B possible, we have added a proton to this group throughout the reaction. We study the reaction starting from the E_4 state with N_2H_2 bound to the cluster. Our results indicate that half-dissociation of S2B is unfavourable in most steps of the reaction mechanism. We observe favourable half-dissociation of S2B only when NH or NH_2 is bound to the cluster, bridging Fe2 and Fe6. However, the former state is most likely not involved in the reaction mechanism and the latter state is only an intermittent intermediate of the E_7 state. Therefore, half-dissociation of S2B seems to play only a minor role in the later parts of the reaction mechanism of nitrogenase. Our suggested mechanism with a protonated S2B is alternating (the two N atoms of the substrate is protonated in an alternating manner) and the substrate prefers to bind to Fe2, in contrast to the preferred binding to Fe6 observed when S2B is unprotonated and bridging Fe2 and Fe6.

Introduction

Nitrogenase (EC 1.18/19.6.1) is the only enzyme that can cleave the triple bond in N_2, by converting it to ammonia.[1–4] Nitrogenase is produced by a few groups of bacteria but several higher plants live in symbiosis with such bacteria, *e.g.* legumes, rice and alder. Crystallographic studies have shown that the most common and active form of nitrogenase is a homodimer of heterodimers.[5–9] It contains two unusual iron–sulfur clusters, the P-cluster ($Fe_8S_7Cys_6$), which is used for electron transfer, and the FeMo cluster, which is the catalytic centre. The latter consists of a $MoFe_8S_9C$(homocitrate) cluster, which is connected to the protein by one His ligand to Mo and a Cys ligand to a Fe ion at the other end of the cluster (*cf.* Fig. 1a). There also exist alternative isoenzymes in which the Mo ion is replaced by V or Fe, but they have a lower activity.[10] During catalysis, electrons are delivered to nitrogenase by the Fe protein, which contains a Fe_4S_4 cluster.[3,4]

Department of Computational Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden. E-mail: Ulf.Ryde@compchem.lu.se; Tel: +46-46 2224502

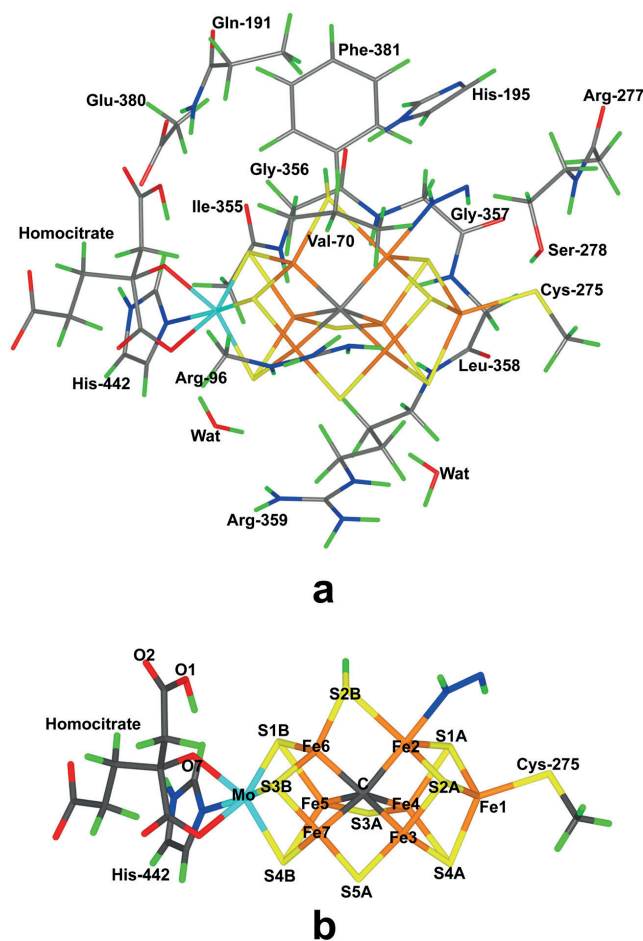†Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4dt00937a

Nitrogenase catalyses the reaction[3,4]

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \quad (1)$$

Thus, eight electrons and protons are needed to form two molecules of NH_3 from N_2. Consequently, the reaction mechanism is typically described by nine states, E_0–E_8, differing in the number of added electrons and protons.[11] Despite intensive biochemical, kinetic, spectroscopic and computational investigations,[1–9,12,13] many details of the nitrogenase reaction mechanism are still controversial, partly because of technical problems to isolate pure samples of the various reaction intermediates E_n. It is known that N_2 binds mainly to the E_4 state and that N_2 binding is coupled to the compulsory formation of H_2, in a reductive elimination reaction.[14–18] E_4 has been shown to contain two hydride ions bridging two pairs of Fe ions,[15,16,18] but the detailed structure of the E_4 state is highly controversial.[19–28]

The later part of the reaction mechanism, *i.e.* after N_2 binding has also been much discussed.[3] In particular, it has been debated whether nitrogenase follows a distal or an alternating mechanism, *i.e.* whether the protons are first added to one N atom, so that NH_3 dissociates already in the E_5 state, before the second N atom is protonated, or whether protons are added alternatively to the two N atoms, so that HNNH and H_2NNH_2 are intermediates and the first NH_3 molecule does

Fig. 1 The FeMo cluster in nitrogenase (*trans*(2;1A)–S2B(26;3) structure). (a) Illustrates the QM system used in all calculations, as well as the names of the nearby residues, (b) shows only the FeMo cluster with atom names indicated.

not dissociate until the E_7 state. A distal mechanism is supported by inorganic nitrogenase model complexes,[29–33] but also by some computational and crystallographic studies of nitrogenase.[25,34] An alternating mechanism is supported by the fact that H_2NNH_2 is a substrate of nitrogenase and can also be released from the enzyme by acid or base treatment during turnover.[1,3,35,36] It has also been shown that N_2, N_2H_2, CH_3N_2H and N_2H_4 all react *via* a common intermediate.[37] Several computational studies have suggested reaction mechanisms that are alternating, at least during the end of the reaction (*i.e.* involving H_2NNH_2 as an intermediate).[13,38,39]

Our group has presented several studies of the later part of the nitrogenase reaction mechanism. We have performed an exhaustive search of possible binding sites and binding modes of N_2H_2 to the FeMo cluster in the E_4 state.[40] The study showed that the most favourable binding sites were Fe2 and Fe6 (atom numbers are shown in Fig. 1b). *trans*-HNNH bound to Fe2 was most stable, but structures with the same molecule bound to Fe6, *cis*-HNNH bound to Fe2, or $HNNH_2$ bound to Fe6 (with the additional proton abstracted from homocitrate) were all competitive within 12 kJ mol^{-1}. With a somewhat larger quantum-mechanical (QM) model, including also Val-70 and Phe-381, which may restrict the binding of the substrate, the preferences change somewhat in favour of *trans*-HNNH or $HNNH_2$ binding to Fe6.[39] That study also considered the complete reaction mechanism and suggested an alternating

mechanism and that the homocitrate ligand may provide a proton buffer that may stabilise intermediates like H_2NNH_2 and NH_3 already at the E_5 and E_7 states, respectively. Based on suggestions from crystal structures of inhibited nitrogenase, we also investigated the corresponding reactions when S2B has dissociated from the FeMo cluster, providing a natural binding site for the substrate.[38] The results suggested that N_2H_2 binds as NNH_2 bridging Fe2 and Fe6 (i.e. an intermediate connected to a distal mechanism), but that the mechanism the becomes alternating, with H_2NNH_2 bound at the E_6 state and NH_3 formation in the E_7 state. Both mechanism seemed to be equally feasible, but a study of proton-transfer reactions in the cluster indicated that the proton transfer to the substrate is easier if S2B remains bound.[41]

Several recent studies have suggested that S2B may dissociate only from one of the two Fe ions, forming unhooked or half-dissociated structures.[21,42–44] In fact, such structures seem to be among the most likely candidates for the E_2–E_4 states of Mo-nitrogenase.[22,44,45] The question then naturally arises whether such structures are competitive also for the later part of the nitrogenase reaction, i.e. after binding of N_2. The aim of the present study is to investigate this possibility using combined QM and molecular mechanics (QM/MM) calculations.

Methods

The protein

The calculations were based on the 1.0 Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q).[5] The setup of the protein is identical to that of our previous studies.[20,22,46–48] The entire heterotetramer was included in the calculations and the QM calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule rather close to the active site (~11 Å) in the A subunit. The two P clusters and the FeMo cluster in subunit A were modelled by molecular mechanics (MM) in the fully reduced and resting states, respectively, using a QM charge model.[47] The protonation states of all residues were the same as before,[47] and the homocitrate ligand was modelled in the singly protonated state with a proton shared between the hydroxyl group (O7 that coordinates to Mo) and the O1 carboxylate atom.[47,49] The protein was solvated in a sphere with a radius of 65 Å around the geometrical centre of the protein. Cl^- and Na^+ ions were added to yield an ionic strength of 0.2 M.[50] The final system contained 133 915 atoms. For the protein, we used the Amber ff14SB force field[51] and water molecules were described by the TIP3P model.[52] The metal sites were treated by a non-bonded model[53] and charges were obtained with the restrained electrostatic potential method.[54]

In the QM calculations, the FeMo cluster was modelled by $MoFe_7S_9C(homocitrate)(CH_3S)(imidazole)$, where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included, viz. Arg-96, Gln-191 and His-195 (sidechains), Ser-278 and Arg-359 (both backbone and sidechain, including

the CA and C and O atoms from Arg-277), Gly-356, Gly-357 and Leu-358 (backbone, including the CA and C and O atoms from Ile-355), as well as two water molecules. Finally, the sidechain of Glu-380 was included because it forms hydrogen bonds to Gln191 and His-442, as well as the sidechains of Val-70 and Phe-381 because they are close to S2B, Fe2 and Fe6, i.e. the expected reactive site. The QM system contained 192–197 atoms depending on the E_n state and the bound substrate. It is shown in Fig. 1a. The net charge of QM region was $-3e$.

QM calculations

All QM calculations were performed with the Turbomole software (versions 7.5 and 7.6).[55] All structures were studied with the r^2SCAN[56] density functional theory (DFT) method. To investigate the functional dependence, most structures were also studied with the TPSSh[57] functional. r^2SCAN is a *meta* generalised gradient approximation functional, whereas TPSSh is a hybrid functional with 10% Hartree–Fock exchange. Both functionals have been shown to give accurate structures of nitrogenase models.[58] All calculations involved the def2-SV(P) basis set.[59] The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation.[60,61] Empirical dispersion corrections were included with the DFT-D4 approach,[62] as implemented in Turbomole.

In this investigation we study the later part of the reaction mechanism of nitrogenase, starting after the binding and protonation of the substrate to N_2H_2 in the E_4 state. Like in our previous two studies,[38,39] we do this to avoid the problem that the nature of the E_4 state is highly controversial, with a very large number of possible locations of the added four H atoms (protons or hydride ions), and extremely large discrepancy in the prediction of various DFT methods.[13,20–23,46] The current consensus is that N_2 binds to the FeMo cluster together with reductive elimination of H_2.[4] The remaining two protons are then used to protonate N_2 to N_2H_2. As a consequence, the N_2H_2-bound E_4 state is in the same formal oxidation state as the resting E_0 state, viz. the $Mo^{III}Fe^{II}_5Fe^{III}_4$ oxidation state.[49,63,64]

The aim of the present study is to investigate whether dissociation of S2B from either Fe2 or Fe6 may facilitate the conversion of N_2H_2 to two NH_3 molecules. Previous studies have shown that such half-dissociation of S2B is energetically favourable only if S2B is protonated.[21,43,44] Therefore, we added an extra proton to S2B in all the current models. Consequently, the charge of the models is one step less negative than previously studied models.

The electronic structure in all QM calculations was obtained with the broken-symmetry (BS) approach:[65] Each of the seven Fe ions was modelled in the high-spin state, with either a surplus of α (four Fe ions) or β (three Fe ions) spin. Such a state can be selected in 35 different ways.[65,66] The various BS states were obtained either by swapping the coordinates of the Fe ions[67] or with the fragment approach by Szilagyi and Winslow.[68] The BS states are named by listing the numbers of the three Fe ions with minority spin, e.g. BS-235. At each E_n-level, we first optimised all possible structures with

one BS state (typically BS-235 or BS-147). For the most stable structure, a full investigation of all 35 BS was performed and if this was a different state, the best structures are reoptimised with that state. A similar procedure was performed with the spin state, which is also unknown for these intermediates.

QM/MM calculations

QM/MM calculations were performed with the ComQum software.[69,70] In this approach, the protein and solvent are split into two subsystems: system 1 (the QM region) was relaxed by QM methods. System 2 was kept fixed at the original coordinates (equilibrated crystal structure), to avoid the risk that different calculations end up in different local minima.

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup (electrostatic embedding). Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner. When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: the QM system was capped with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[69,71] All atoms were included in the point-charge model, except the CL atoms.[72] ComQum employs a subtractive scheme with van der Waals link-atom corrections.[73] No cut-off is used for the QM and QM–MM interactions. The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J mol^{-1} (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u.

In most structures where S2B binds to both Fe2 and Fe6, one Fe–S distance is ~2.3 Å and the other 0.1–0.3 Å longer. Thereby, S2B is in between the two Fe ions and both Fe–Fe–S2B angles are less than 90°. When S2B dissociates from one of the Fe ions, the substrate typically bridges Fe2 and Fe6 (with one or two N atoms) and one Fe–S2B distance increases to >3.4 Å and one Fe–Fe–S2B angle is larger than 90°. However, in some cases, the substrate binds only to one Fe ion and S2B to the other ion (often, but not always, receiving a hydrogen bond from the substrate). In those cases, it is less obvious whether S2B has dissociated or not. We used a cutoff of Fe–S2B > 2.9 Å to define a dissociated structure (no structure has a Fe–S2B bond length between 2.81 and 3.04 Å; likewise, there is a gap between 0.52 and 0.86 Å for the absolute difference between the two Fe–S2B distances). Mayer bond orders[74] (calculated with the Multiwfn package[75]) for the Fe–S interactions are correlated to the Fe–S distances (correlation coefficient −0.87) and show similar trends; a cutoff at 0.2 can be used to define bonds.

Results and discussion

In this investigation, we have studied possible paths for the second half of the mechanism of nitrogenase, allowing the S2B ligand to dissociate from one Fe ion. The results are com-

pared to our previous studies where S2B was either binding to both Fe2 and Fe6,[39] or was completely dissociated.[38] As in the previous studies, we start from a E_4 state, where H_2 has dissociated, N_2 is bound and is protonated to N_2H_2. This is done to avoid the severe problems that different DFT method give widely different predictions regarding the most stable protonation states of E_4 and the strength of the binding of N_2.[22,46]

For each new E_n state, we add one electron and a proton to the FeMo cluster. The electrons are provided by the Fe protein *via* the P-cluster.[2,3] Protons come ultimately from water solvent and two possible proton channels have been suggested, ending either at His-195 or at a water molecule close to S3B, S4B and S5A.[76–79] It has been suggested that His-195 can only provide a single proton, because rotation of the imidazole group is restricted in the protein.[76,80] Several groups have studied proton transfers within the FeMo cluster and have shown that the individual steps in general are facile,[76,78,81] although sometimes certain protonation states may act as thermodynamic sinks making the net barriers somewhat high.[41] However, Siegbahn has suggested that the barriers are strongly lowered by the employment of surrounding water molecules.[28,82] Therefore, we have not explicitly studied proton transfers within the FeMo cluster in this study. Instead we concentrate on determining the thermodynamically most stable structures at each E_n level and the cleavage of the N–N bond. We discuss the various E_n states in separate sections.

E_4 state

We started with the E_4 state with N_2H_2 bound. We tested four different isomers of N_2H_2, *viz.* NNH_2, *cis*-HNNH, *trans*-HNNH and $HNNH_2$ (in the latter case, the third proton was taken from homocitrate, which is nearby when the ligand binds to Fe6). We studied binding of N_2H_2 only to Fe2 and Fe6, because these two Fe ions have been pointed out by experimental studies[3,83,84] and it has also been shown to be the preferred binding sites by a systematic DFT search.[40] We studied both end-on and side-on binding, to one Fe ion as well as to both Fe ions. If relevant, we allowed the non-coordinating N atom or the protons to point in different directions. We use the following nomenclature: the four isomers are called NNH_2, *cis*, *trans* or $HNNH_2$, respectively, which is followed by the binding site in brackets and indicating only the number of the Fe atom. 2 and 6 means that it binds to only Fe2 or Fe6, whereas 26 means that it bridges Fe2 and Fe6. Two numbers with a comma between means that both N atoms bind to Fe, *e.g.* (26,2), indicating that one N atom bridges Fe2 and Fe6, whereas the other binds only to Fe2. A number or an atom after a semicolon indicates the direction of the non-coordinating atoms, where 3 and 5 indicate S3A and S5A, respectively.

In addition, we considered structures with the S2B ligand bridging Fe2 and Fe6 or binding only to one of the two Fe ions. In variance to the previous study,[39] we assumed that S2B is singly protonated, allowing it to dissociate from either Fe2 or Fe6. When it still bound to both ions, we tried to force it to dissociate from one of the Fe ions. Moreover, we also studied cases where the proton on S2B pointed in different directions,

typically towards S3A or S5A. The nomenclature is S2B with the coordinating Fe ions in brackets (2, 6, or 26) and with the direction indicated after a semicolon with numbers 3 or 5, or an atom.

We tested 61 different structures and managed to obtain 53 of them. They are listed in Table 1. It can be seen that the most stable structure is *trans*(2;1A)-S2B(26;3), *i.e.* with *trans*-HNNH binding to Fe2 with the non-bonding NH group point-

ing towards S1A (with a S–H hydrogen-bonding distance of 2.2 Å), and with S2B coordinating to both Fe2 and Fe6, with the proton pointing towards S3A. It is most stable in the quartet BS-235 state and it is shown in Fig. 2. Cleaving the Fe2–S2B bond is strongly unfavourable, so the best structure does not have any half-dissociated S2B group. Moving the proton on S2B to the opposite side, *trans*(2;1A)-S2B(26;5) (also in Fig. 2), changes the energy by only 3–4 kJ mol^{-1}. The corresponding structures with HNNH binding instead to Fe6 are only 6 kJ mol^{-1} less stable with the r^{2}SCAN functional, but 12–13 kJ mol^{-1} with TPSSh (*trans*(6;3B)-S2B(26;3) and *trans*(6;3B)-S2B(26;5) in Fig. 2). In this case, the non-coordinat-
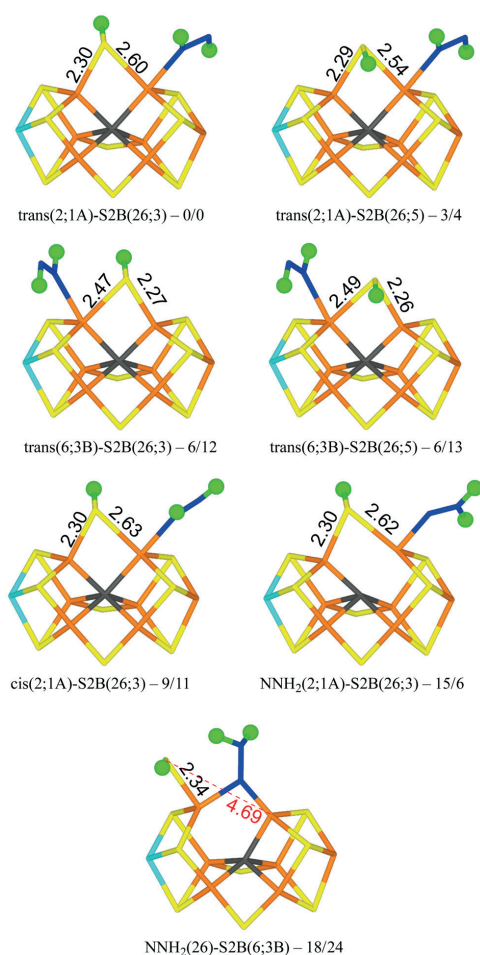
Table 1 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_4$ state. All structures are in the quartet BS-235 state unless otherwise noted

Structure	r^{2}SCAN	TPSSh
HNNH$_2$(6;HCA)-S2B(26;3)	45	39
HNNH$_2$(6;HCA)-S2B(26;5)	38	57
NNH$_2$(26)-S2B(2;Fe1)	21	31
NNH$_2$(26)-S2B(2;5)	34	39
NNH$_2$(26)-S2B(6;3B)	18	24
NNH$_2$(26)-S2B(6;1B)	18	40
NNH$_2$(2;1A)-S2B(26;3)	15	6
NNH$_2$(2;1A)-S2B(26;5)	25	33
NNH$_2$(2;2A)-S2B(6;3)	60	57
NNH$_2$(2;2A)-S2B(26;5)	76	71
NNH$_2$(6;1B)-S2B(2;Fe1)	80	96
NNH$_2$(6;2B)-S2B(2;4A)	91	
NNH$_2$(6;HCA)-S2B(26;3)	56	57
NNH$_2$(6;3)-S2B(26;5)	66	74
cis(26;3)-S2B(2;Fe1)	114	111
cis(26;3)-S2B(2;5)	129	139
cis(26,2;3)-S2B(6;N)	41	45
cis(26,2;5)-S2B(6;Mo)	34^a	52
cis(26,2;5)-S2B(6;3)	27^a	45
cis(26,6;3)-S2B(2;Fe1)	88	90
cis(26,6;3)-S2B(2;5)	102	104
cis(26,6;5)-S2B(2;Fe1)	71	73
cis(26,6;5)-S2B(2;3)	90	95
cis(26,6;5)-S2B(2;5)	78	81
cis(2,6)-S2B(2;Fe1)	103^b	108
cis(2,6)-S2B(2;5)	84^a	109
cis(2,6)-S2B(6;Mo)	92^a	112
cis(2,2)-S2B(6;3B)	81	99
cis(2,2)-S2B(6;3)	76	94
cis(2;1A)-S2B(26;3)	9	11
cis(2;1A)-S2B(26;5)	14	16
cis(2;2A)-S2B(26;3)	61	64
cis(2;2A)-S2B(26;5)	60	52
cis(6,6;5)-S2B(2;2A)	130	
cis(6,6)-S2B(2;Fe1)	130	
cis(6;1B)-S2B(26;3)	29	34
cis(6;1B)-S2B(26;5)	30	37
cis(6;3B)-S2B(26;3)	40	47
cis(6;3B)-S2B(26;5)	39	46
trans(26;5)-S2B(2;Cys)	83^b	86
trans(26;5)-S2B(2;N)	94	99
trans(26,2;5)-S2B(6;3B)	41	35
trans(26,2;5)-S2B(6;N)	33	28
trans(26;3)-S2B(2;Cys)	47	57
trans(2;1A)-S2B(26;3)	0	0
trans(2;1A)-S2B(26;5)	3	4
trans(2;2A)-S2B(6;3)	34	39
trans(2;2A)-S2B(26;5)	41	44
trans(6;1B)-S2B(26;3)	17	22
trans(6;1B)-S2B(26;5)	23	27
trans(6;3B)-S2B(26;3)	6	12
trans(6;3B)-S2B(26;5)	6	13

^{a} Quartet BS-147 state. ^{b} Doublet BS-147 state.

trans(2;1A)-S2B(26;3) – 0/0

trans(2;1A)-S2B(26;5) – 3/4

trans(6;3B)-S2B(26;3) – 6/12

trans(6;3B)-S2B(26;5) – 6/13

cis(2;1A)-S2B(26;3) – 9/11

NNH$_2$(2;1A)-S2B(26;3) – 15/6

NNH$_2$(26)-S2B(6;3B) – 18/24

Fig. 2 The best E$_4$ structures with relative energies in kJ mol^{-1} indicated (r^{2}SCAN/TPSSh).

ing N atom of HNNH forms a hydrogen bond to S3B (2.6 Å). On the other hand, structures with HNNH pointing in the opposite direction are appreciably less stable, by 34–44 kJ mol^{-1} when HNNH binds to Fe2 and 17–27 kJ mol^{-1} when it binds to Fe6. The non-coordinating N atom still forms hydrogen bonds to the cluster (2.4 Å to S2A or 2.5 Å to S1B), but Val-70 and Ser-278 are strongly restricting its movement.

Structures with *cis*-HNNH binding to Fe2 (*e.g. cis*(2;1A)-S2B (26;3) in Fig. 2) are also competitive, 9–16 kJ mol^{-1} less stable than the best one, and still with S2B binding to both Fe2 and Fe6. The non-bonding N atom is still pointing towards S1A, but no hydrogen bond can form for the *cis*-isomer of the ligand. The opposite orientation of HNNH is much less stable, owing to steric interactions with Val-70 and Ser-278. Likewise, *cis*-HNNH binding to Fe6 is quite unfavourable, 29–46 kJ mol^{-1} less stable than the best structure.

A structure with NNH$_2$ binding to Fe2 is also competitive, NNH$_2$(2;1A)-S2B(26,3) in Fig. 2, especially with TPSSh, being 6–15 kJ mol^{-1} less stable than the best structure. The non-bonding N atom forms a hydrogen bond to S1A (2.3 Å) and S2B binds to both Fe2 and Fe6. Structures with the substrate pointing instead towards S2A (2.3 Å hydrogen-bonding distance) is ~50 kJ mol^{-1} less stable.

The best structure with a half-dissociated S2B ligand is NNH$_2$(26)-S2B(6;3B) in Fig. 2 with r^{2}SCAN, *i.e.* with NNH$_2$ bridging Fe2 and Fe6, and with one of the H atoms forming a hydrogen bond to S2B (2.14 Å H–S distance). S2B is dissociated from Fe2, but binds to Fe6. Its proton is pointing towards S3B (2.9 Å distance). The structure with the proton on S2B pointing in the opposite direction (towards S3B) is essentially degenerate, 18 kJ mol^{-1} less stable than the best *trans*(2;1A)-S2B(26;3) structure. With TPSSh, the latter structure is most stable by 16 kJ mol^{-1}, 24 kJ mol^{-1} less stable than the best structure.

These results are quite similar to what has been observed in our previous studies: *trans*-HNNH binding to either Fe2 and Fe6 is nearly degenerate when S2B remains bound and bridging NNH$_2$ most stable without S2B.[38–40] The largest difference is that the HNNH$_2$ structures are not competitive, being 38–57 kJ mol^{-1} less stable than the best structures. The reason may be the extra proton on S2B, which makes the hydrogen bond between the substrate and S2B less favourable (2.42 Å, compared to 2.24 Å in our previous study[40]) and also orients the substrate so that the hydrogen bond to homocitrate is worse (1.97 Å, compared to 1.69 Å (ref. 40)). However, the most important conclusion is that there is no advantage of S2B dissociation in the N$_2$H$_2$-bond E$_4$ state.

E$_5$ state

Next, we added one proton and one electron to consider the E$_5$ state. We studied four variants of the substrate, HNNH$_2$, NNH$_3$, H$_2$NNH$_2$ (hydrazine) and HNNH$_3$. In the latter two cases, the substrate has abstracted a proton from homocitrate. As for E$_4$, we studied both end-on and side-on binding, to Fe2, Fe6 or both. Likewise, we considered structures with S2B bound to only Fe2 or Fe6, or to both. The structures are named as in the previous section and the results are listed in Table 2.

Table 2 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_5$ state. All structures are in the quintet BS-235 state unless otherwise noted

Structure	r^{2}SCAN	TPSSh
H$_2$NNH$_2$(6;1B)-S2B(26;3)	11	29
H$_2$NNH$_2$(6;1B)-S2B(26;5)	11	24
H$_2$NNH$_2$(6;2B)-S2B(26;3)	9	20
H$_2$NNH$_2$(6;2B)-S2B(26;5)	14	32
H$_2$NNH(26,6)-S2B(2;Cys)	90	82
H$_2$NNH(26,6)-S2B(2;2A)	96	91
H$_2$NNH(2,6)-S2B(6;3B)	113	120
H$_2$NNH(2,6)-S2B(6;1B)	117	
H$_2$NNH(2;2B)-S2B(6;3B)	79	93
H$_2$NNH(2,2)-S2B(6;1B)	40	51
HNNH$_2$(6,2)-S2B(2;Cys)	109	112
HNNH$_2$(6,2)-S2B(2;N)	121	114
HNNH$_2$(2,2)-S2B(6;3B)	47	42
HNNH$_2$(2,2)-S2B(6;1B)	37	49
HNNH$_2$(2;2A)-S2B(6;3B)	36	43
HNNH$_2$(2;2A)-S2B(6;1B)	28	35
HNNH$_2$(2;Cys)-S2B(6;1B)	9^a	15^b
HNNH$_2$(2;1A)-S2B(26;3)	0	0
HNNH$_2$(2;1A)-S2B(26;5)	7	8
HNNH$_2$(2;2A)-S2B(6;1B)	45	
HNNH$_2$(2;2A)-S2B(26;5)	67	68
HNNH$_2$(6,6)-S2B(2;Cys)	139	146
HNNH$_2$(6;2B)-S2B(26;3)	47	53
HNNH$_2$(6;2B)-S2B(26;5)	32	55
HNNH$_2$(6;1B)-S2B(26;3)	82	93
HNNH$_2$(6;1B)-S2B(26;5)	76	88
HNNH$_2$(6;3B)-S2B(26;3)	80	80
HNNH$_2$(6;3B)-S2B(26;5)	67	75
NNH$_3$(26)-S2B(2;1A)	140	139
NNH$_3$(26)-S2B(2;2A)	131	134
NNH$_3$(26)-S2B(6;3B)	59	79
NNH$_3$(26)-S2B(6;1B)	53	76
NNH$_3$(2;1A)-S2B(26;3)	137	131
NNH$_3$(2)-S2B(26;5)	143	137
NNH$_3$(6;2B)-S2B(2;2A)	218	243
NNH$_3$(6;1B)-S2B(26;3)	197	214
NNH$_3$(6;1B)-S2B(26;5)	186	197
NNH$_3$(6;3B)-S2B(26;3)	180	190
NNH$_3$(6;3B)-S2B(26;5)	185	191

^{a} BS-247 state. ^{b} BS-147 state.

We tried to optimise 45 structures and 39 of them were obtained. The best (HNNH$_2$(2;1A)-S2B(26;3)) has HNNH$_2$ bound to Fe2, with the non-coordinating NH$_2$ group pointing towards S1A, forming a hydrogen bond with a H–S1A distance of 2.22 Å. The structure is most stable in the quintet BS-235 state and is shown in Fig. 3. The other H atom of the NH$_2$ group points towards the backbone NH group of Ser-278 in a perpendicular manner (2.18 Å H–N distance). S2B binds both Fe2 and Fe6, with the proton on the S3A side. Dissociating S2B from Fe2 is strongly unfavourable. Moving the proton to the other side gives a structure that is 7–8 kJ mol^{-1} less stable (HNNH$_2$(2;1A)-S2B(26;5)) in Fig. 3. On the other hand, rotating the substrate around the Fe2–N bond so that the NH$_2$ group instead forms a hydrogen bond to S2A (2.37 Å), gives a structure that is 45 kJ mol^{-1} less stable and S2B dissociates from Fe2. Likewise, structures with HNNH$_2$ binding to Fe6 are quite unfavourable (by 32–55 kJ mol^{-1}).
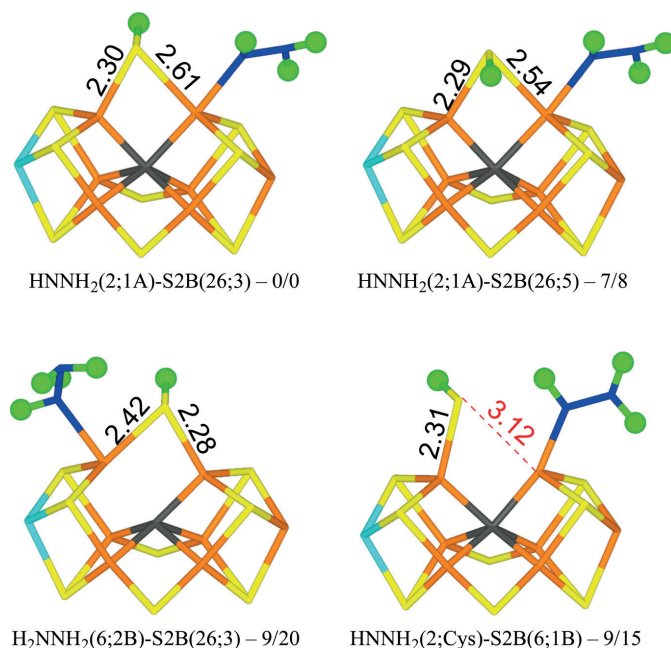
Fig. 3 The best E_5 structures with relative energies in kJ mol^{-1} indicated (r^{2}SCAN/TPSSh).

With r^{2}SCAN, structures with H$_2$NNH$_2$ bound to Fe6 are competitive, only 8–14 kJ mol^{-1} less stable than HNNH$_2$(2;1A)-S2B(26;3). The best is H$_2$NNH$_2$(6;2B)-S2B(26;3) in Fig. 3, in which one proton on the binding NH$_2$ group points to the alcohol oxygen of homocitrate (2.11 Å), whereas the two protons of the non-bonding NH$_2$ group point towards S2B (2.36 Å) and an acetate oxygen of homocitrate (3.13 Å) but with far from optimal hydrogen-bond geometries (explaining why the various conformations have similar energies). With TPSSh, these structures are 20–32 kJ mol^{-1} less stable than the best structure.

The best structure with S2B half-dissociated is HNNH$_2$(2; Cys)-S2B(6;1B) in Fig. 3, i.e. with HNNH$_2$ binding end-on to Fe2, the non-bonding NH$_2$ group forming a hydrogen bond to SG of Cys-275 (2.77 Å) and the proton of the bonding NH group pointing towards S2B (2.55 Å). S2B binds only to Fe6 with the proton pointing towards S1B. It is only 9–15 kJ mol^{-1} less stable than the best structure.

Structures with HNNH$_3$ binding to Fe6 are quite unfavourable, 67–93 kJ mol^{-1} less stable than the best structure. They all have S2B binding to both Fe2 and Fe6. Structures with NNH$_3$ are strongly unfavourable by 131–243 kJ mol^{-1}. The only exception are two structures with the unprotonated N atom bridging Fe2 and Fe6, S2B binding only to Fe6 and the proton on S2B pointing either two S1B or S3B. They are 53–59 (r^{2}SCAN) or 76–79 kJ mol^{-1} less stable than the best structure.

Table 3 Relative energies (kJ mol^{-1}) of the various structures optimised for the E_6 state with an intact N−N. bond. All structures are in the quartet BS-235 state unless otherwise noted

Structure	r^{2}SCAN	TPSSh
H$_2$NNH$_2$(2,6)-S2B(2;Cys)	133^a	131
H$_2$NNH$_2$(2,6)-S2B(6;3B)	161	148
H$_2$NNH$_2$(2,6)-S2B(6;1B)	173	160
H$_2$NNH$_2$(2;1A)-S2B(6;3B)	118	107
H$_2$NNH$_2$(2;1A)-S2B(26;3)	3	−1
H$_2$NNH$_2$(2;1A)-S2B(26;5)	17	9
H$_2$NNH$_2$(2;2A)-S2B(26;3)	33	27
H$_2$NNH$_2$(2;2A)-S2B(26;5)	25	20
H$_2$NNH$_2$(6;HCA)-S2B(2;2A)	47	52
H$_2$NNH$_2$(6;1B)-S2B(26;3)	13	13
H$_2$NNH$_2$(6;1B)-S2B(26;5)	11	11
H$_2$NNH$_2$(6;3B)-S2B(26;3)	0	0
H$_2$NNH$_2$(6;3B)-S2B(26;5)	1	2
H$_2$NNH$_2$(6;3B)-S2B(26;3)	24	21
H$_2$NNH$_2$(6;3B)-S2B(26;5)	25	23
HNNH$_3$(2;1A)-S2B(26;3)	63	58
HNNH$_3$(2;1A)-S2B(26;5)	71	67
HNNH$_3$(2;2A)-S2B(26;3)	113	103
HNNH$_3$(2;2A)-S2B(26;5)	118	108
HNNH$_3$(6;1B)-S2B(26;3)	95	97
HNNH$_3$(6;1B)-S2B(26;5)	97	100
HNNH$_3$(6;3B)-S2B(26;3)	102	107
HNNH$_3$(6;3B)-S2B(26;5)	103	101
NH$_3$ + NH$_2$(6)-S2B(26,3)	85	93
NH$_3$ + NH(26)-S2B(6,3)	49	−10

^{a} BS-147 state.

E_6 state

We next studied the E_6 state by adding another electron and proton to the E_5 state. We considered three variants of the substrate: H_2NNH_2, $HNNH_3$ or H_2NNH_3 (in the latter case with one proton abstracted from homocitrate). In total, 26 structures were tested and 24 of these were obtained. They are described in Table 3.

The best structure with r^2SCAN, $H_2NNH_2(6;3B)$-S2B(26;3) in Fig. 4, has H_2NNH_2 bound end-on to Fe2. The two H atoms of the non-coordinating NH_2 group point towards S2B and S3B with distances of 2.53 and 2.83 Å, respectively. One of the H atoms of the coordinating NH_2 group points towards the alcohol and acetate O atoms of homocitrate with distances of 2.28 and 2.27 Å, but the two O atoms are also involved in an internal hydrogen bond (1.42 Å) and a hydrogen bond to the sidechain NH_2 group of Gln-191. S2B bridges Fe2 and Fe6, and the proton points towards S3A. The structure is most stable in the quartet BS-235 state. The structure with the proton on S2B pointing in the opposite direction ($H_2NNH_2(6;3B)$-S2B(26;5) in Fig. 4) is only 1–2 kJ mol^{-1} less stable.

A structure with the non-bonding NH_2 group of the substrate pointing in the opposite direction ($H_2NNH_2(6;1B)$; with the H atoms pointing towards S2B and S1B with distances of
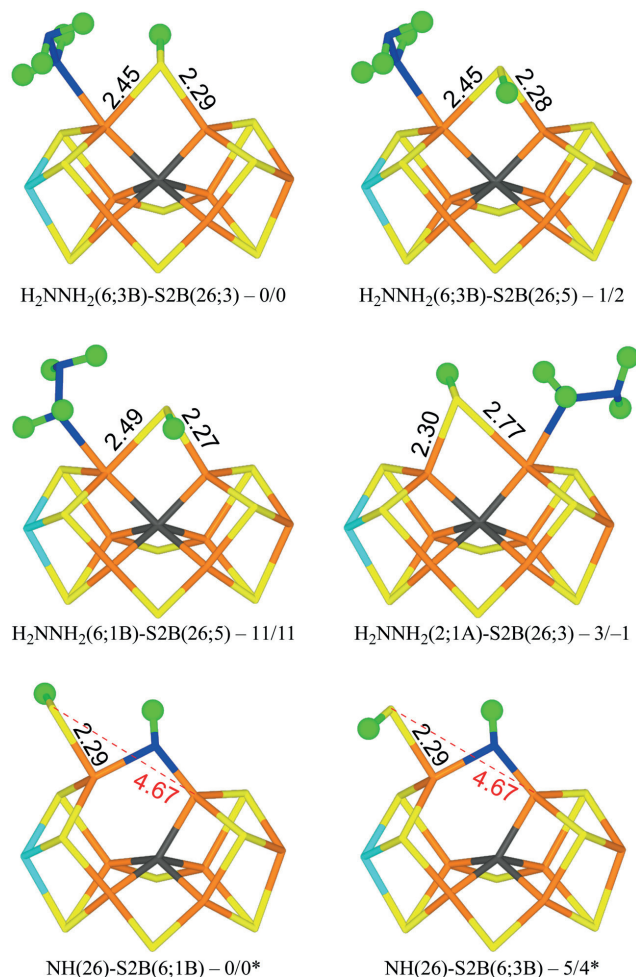
Fig. 4 The best E_6 structures with relative energies in kJ mol^{-1} indicated (r^2SCAN/TPSSh; * indicates a distinct reference state).

2.48 and 2.97 Å) is 11–13 kJ mol^{-1} less stable (*e.g.* H$_2$NNH$_2$(6;1B)-S2B(26;5) in Fig. 4). Still a third conformation, with the non-bonding NH$_2$ group pointing towards homocitrate, which leads to dissociation of S2B from Fe6 (H$_2$NNH$_2$(6; HCA)-S2B(2;2A)), is 47–52 kJ mol^{-1} less stable than the best structure and is the most stable structure with a half-dissociated S2B.

Structures with H$_2$NNH$_2$ binding to Fe2 are strongly competitive. The best is H$_2$NNH$_2$(2;1A)-S2B(26;3) in Fig. 4, *i.e.* with one H atom of the non-coordinating NH$_2$ group forming a hydrogen bond to S1A at a distance of 2.32 Å. Both the H atoms of the coordinating NH$_2$ group point towards S2B, but with poor hydrogen-bond geometries and distances of 2.61 and 3.00 Å. S2B binds to both Fe2 and Fe6, but with quite different distances of 2.77 and 2.30 Å. This structure is only 3 kJ mol^{-1} less stable than the best structure. With TPSSh, it is actually 1 kJ mol^{-1} better. In both cases, it is most stable in the quartet BS-235 state. This probably the most relevant isomer for the E$_6$ state, considering that the best structures of the E$_4$ and E$_5$ states had the substrate bound to Fe2. The corresponding S2B(26;3) structure is 10-14 kJ mol^{-1} less stable. Several other conformations of the substrate and the proton on S2B were also tested and they are 4–30 kJ mol^{-1} less stable. We also tested some structures with H$_2$NNH$_2$ bridging Fe2 and Fe6 and with S2B binding to only one of the Fe ions. However, they were all unfavourable by 131–173 kJ mol^{-1}.

Two structures with H$_2$NNH$_3$ bound to Fe6 were optimised. They were 21–25 kJ mol^{-1} less stable than the best one, indicating that proton transfer from homocitrate is slightly uphill.

Eight structures with HNNH$_3$ binding to either Fe2 or Fe6 were also tested, but they were all unfavourable by 58–118 kJ mol^{-1}. This shows that hydrazine is the most stable isomer of the substrate at the E$_6$ state. Again, structures with S2B half-dissociated seem to be of minor relevance.

We have tried to cleave the N–N bond in the two best structures (H$_2$NNH$_2$(2;1A)-S2B(26;3) and H$_2$NNH$_2$(2;1A)-S2B(26;3)). However, the reaction is quite prohibitive with barriers of 90–128 kJ mol^{-1} and the product is higher in energy than the starting structures.

For completeness, we also evaluated E$_6$ structures after dissociation of NH$_3$, *i.e.* structures with only NH bound to the FeMo cluster. Two structures with NH$_2$ bound to Fe6, after abstraction of the proton from homocitrate were also considered. The results are gathered in Table 4.

The most stable structure has NH bridging the Fe2 and Fe6 ions, and S2B binding only to Fe6 with the proton pointing towards S1B (NH(26)-S2B(6;1B) in Fig. 4). A structure with the proton pointing instead towards S3B is only 4–5 kJ mol^{-1} less stable (NH(26)-S2B(6;3B) in Fig. 4). Structures with S2B binding only to Fe2 are 37–44 kJ mol^{-1} less stable, whereas structures with NH binding only to Fe2 or Fe6 are 120–205 kJ mol^{-1} less stable. The two structures with NH$_2$ binding to Fe6 are 79–91 kJ mol^{-1} less stable. Thus, structures with S2B half-dissociated are important when NH binds to the cluster. However, compared to the structures with an intact N–N bond, the HN structures are 31–46 kJ mol^{-1} less stable (adding the

Table 4 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_6$ state with NH$_3$ dissociated. All structures are in the quartet BS-235 state unless otherwise noted

Structure	r^{2}SCAN	TPSSh
NH$_2$(6)-S2B(26;3)	86	79
NH$_2$(6)-S2B(26;5)	87^a	82
NH(26)-S2B(2;Cys)	40	37
NH(26)-S2B(2;5A)	44	41
NH(26)-S2B(6;3B)	5	4
NH(26)-S2B(6;1B)	0	0
NH(2;1A)-S2B(26;3)	120	99
NH(2;1A)-S2B(26;5)	127	108
NH(2)-S2B(6;HCA)	160	136
NH(6;1B)-S2B(2;Cys)	205	201
NH(6;3A)-S2B(2;Cys)	192	189
NH(6;3A)-S2B(26;3)	130^a	130
NH(6;3A)-S2B(26;5)	142^a	130

^{a} BS-147 state.

energy of an isolated NH$_3$ molecule in a water-like continuum solvent).

E$_7$ state

Next, we added an electron and a proton to reach state E$_7$. We studied first the H$_2$NNH$_3$ form of the substrate. The results are gathered in Table 5. The most stable structure has H$_2$NNH$_3$ bound end-on to Fe2 with the non-bonded NH$_3$ group forming a hydrogen bond to S1A (1.92 Å) and with another of the H atoms pointing transversely to the backbone N atom of Ser-278 (1.90 Å). S2B bridges Fe2 and Fe6 with the proton on the S3A side (H$_2$NNH$_3$(2;1A)-S2B(26;3) in Fig. 5). It is most stable in the triplet BS-147 state. A structure with the S2B proton on the other side is 10 kJ mol^{-1} less stable. Structures with the non-bonding NH$_3$ group pointing in the other direction, forming hydrogen bonds to S2A and S2B (2.13 and 2.16 Å) and with one of the NH$_2$ protons forming a hydrogen bond to SG of Cys-275 (2.68 Å) are 26–43 kJ mol^{-1} less stable.

Table 5 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_7$ state with an intact N–N. bond. All structures are in the triplet BS-147 state unless otherwise noted

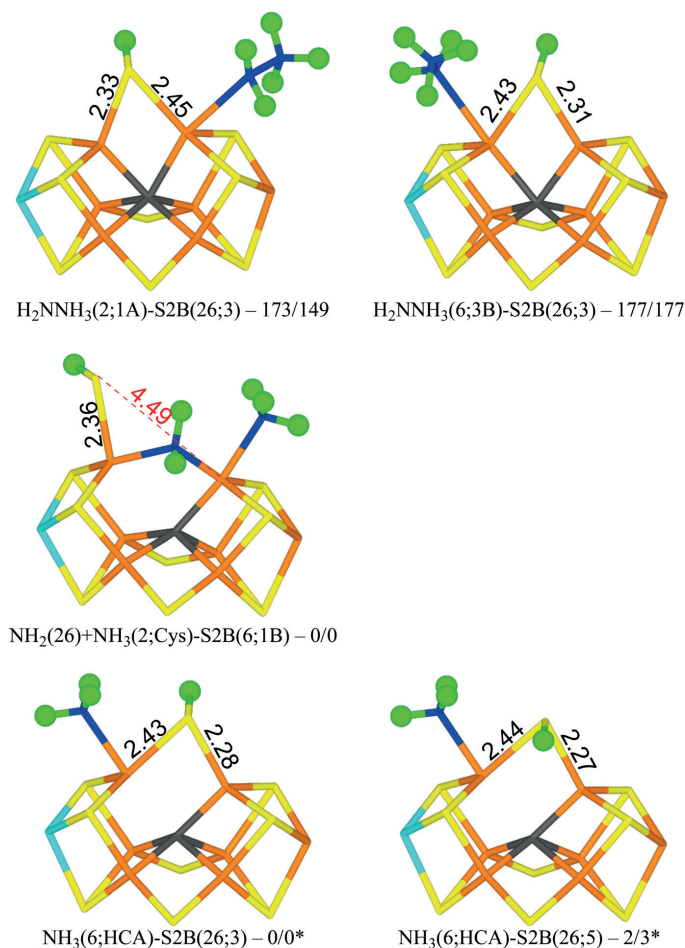| Structure	r^{2}SCAN	TPSSh
H$_2$NNH$_3$(2;1A)-S2B(26;3)	173	149
H$_2$NNH$_3$(2;1A)-S2B(26;5)	183	158
H$_2$NNH$_3$(2;2A)-S2B(26;3)	199	177
H$_2$NNH$_3$(2;2A)-S2B(26;5)	214	192
H$_2$NNH$_3$(6;1B-diss)-S2B(26;3)	168	154
H$_2$NNH$_3$(6;1B-diss)-S2B(26;5)	160^a	158
H$_2$NNH$_3$(6;3B)-S2B(26;3)	177^a	177
H$_2$NNH$_3$(6;3B)-S2B(26;5)	178	183
NH$_2$(2;1A) + NH$_3$(26)-S2B(6;HCA)	83^b	71
NH$_2$(26) + NH$_3$(6;HCA)-S2B(2;Cys)	23^b	14
NH$_2$(26) + NH$_3$(6;HCA)-S2B(2;N)	23^b	24
NH$_2$(26) + NH$_3$(2;Cys)-S2B(6;1B)	0^c	0^c
NH$_2$(26) + NH$_3$(diss)-S2B(6;1B)	29^b	
NH$_2$(2;NH$_3$) + NH$_3$(diss)-S2B(6;HCA)	58^c	44^b

^{a} BS-235 state. ^{b} BS-346 state. ^{c} BS-156 state.

Fig. 5 The E_7 structures with relative energies in kJ mol^{-1} indicated (r^{2}SCAN/TPSSh; * indicates another reference state).

Structures with H_2NNH_3 bound end-on to Fe6 are only 4–5 kJ mol^{-1} less stable with r^{2}SCAN, but 28–34 kJ mol^{-1} less stable with TPSSh. In the best structure ($H_2NNH_3(6;3B)$-S2B (26;3) in Fig. 5), the non-bonding NH_3 group forms hydrogen bonds to S3B and S2B (2.21 and 2.27 Å). S2B still bridges Fe2 and Fe6, with the proton on the S3A side. If the NH_3 group is moved to the other side, the ligand actually dissociates, giving a structure that is actually 9 kJ mol^{-1} more stable than the best state with r^{2}SCAN but 6 kJ mol^{-1} less stable with TPSSh.

From the $H_2NNH_3(2;1A)$-S2B(26;3) structure, the N–N bond can easily be cleaved with an activation energy of only 41–50 kJ mol^{-1}. Therefore, we studied several product (NH_2 + NH_3) structures. In the best structure ($NH_2(26) + NH_3(2;1A)$-S2B

(6;1B) in Fig. 5), NH_2 bridges Fe2 and Fe6, whereas NH_3 binds to Fe2, forming a hydrogen bond to S1A (2.67 Å). S2B binds only to Fe6, with the proton pointing towards S1B (3.29 Å). It is most stable in the triplet BS-156 state. This structure is 149–173 kJ mol^{-1} more stable than the best H_2NNH_3 structure. Structures with NH_3 instead binding to Fe6 is only 14–23 kJ mol^{-1} less stable.

NH_3 can dissociate from the FeMo cluster with an activation energy of only 32 kJ mol^{-1}. Therefore, we also studied structures without NH_3 (i.e. dissociated and excluded from the calculations). The results are shown in Table 6. The most stable structure has NH_3 bound to Fe6, i.e. with a proton abstracted from homocitrate ($NH_3(6;HCA)$-S2B(26;3) in Fig. 5).

Table 6 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_7$ state with NH$_3$ dissociated. All structures are in the quintet BS-235 state unless otherwise noted

Structure	r^{2}SCAN	TPSSh
NH$_2$(26)-S2B(2;Cys)	47	35
NH$_2$(26)-S2B(2;N)	54	42
NH$_2$(26)-S2B(6;3B)	24	26
NH$_2$(26)-S2B(6;1B)	26	19
NH$_2$(2)-S2B(26;3)	78^a	66
NH$_2$(2)-S2B(26;5)	75^a	66
NH$_2$(2;2A)-S2B(6;3B)	217	
NH$_2$(6;1B)-S2B(26;3)	77	79
NH$_2$(6;1B)-S2B(26;5)	76	80
NH$_3$(6;HCA)-S2B(26;3)	0	0
NH$_3$(6;HCA)-S2B(26;5)	2	3

^{a} BS-147 state.

Table 7 Relative energies (kJ mol^{-1}) of the various structures optimised for the E$_8$ state. All structures are in the quartet BS-235 state

Structure	r^{2}SCAN	TPSSh
NH$_3$(2)-S2B(26;3)	0.6	0
NH$_3$(2)-S2B(26;5)	3	3
NH$_3$(2)-S2B(6;3B)	26	^{a}
NH$_3$(2)-S2B(6;1B)	16	^{a}
NH$_3$(6)-S2B(26;3)	0.0	5
NH$_3$(6)-S2B(26;5)	0.1	5
NH$_4$(diss)-S2B(26;3)	13	22
NH$_4$(diss)-S2B(26;5)	11	21

^{a} Converged to another structure.

One of the protons forms a hydrogen bond to the alcohol O atom of homocitrate (2.23 Å), whereas another is directed towards S2B (2.81 Å). The latter atom bridges Fe2 and Fe6 with the proton on the S3A side. It is most stable in the quintet BS-235 state. A structure with the S2B proton on the other side is 2–3 kJ mol^{-1} less stable (NH$_3$(6;HCA)-S2B(26;5) in Fig. 5).

Structures with NH$_2$ bridging Fe2 and Fe6 are at least 19–24 kJ mol^{-1} less stable. They have S2B binding to only one Fe ion, preferably Fe6. Structures with NH$_2$ binding to either Fe2 or Fe6 (with S2B bridging Fe2 and Fe6) are at least 66–79 or 76–79 kJ mol^{-1} less stable than the NH$_3$(6;HCA)-S2B(26;3) structure.

E$_8$ state

Finally, we studied the E$_8$ state with NH$_3$ bound to the FeMo cluster. The results are collected in Table 7. Four structures are essentially degenerate (within 5 kJ mol^{-1}). Two have NH$_3$ bound to Fe6 with one of the protons forming a hydrogen bond to the acetate group of homocitrate (1.87–1.90 Å) and S2B bridging Fe2 and Fe6 with the proton either on the S3A or S5A side (NH$_3$(6)-S2B(26;3) and NH$_3$(6)-S2B(26;5) in Fig. 6). The other two have NH$_3$ bound to Fe2 with the three protons approximately in the directions of SG of Cys-275, S1A and S2A (all distances are 3.0–3.2 Å) and S2B bridging Fe2 and Fe6 with the proton either on the S3A or S5A side (NH$_3$(2)-S2B(26;3) and

NH$_3$(2)-S2B(26;5) in Fig. 6). Structures with S2B binding only to Fe6 and NH$_3$ to Fe2 are 16–26 kJ mol^{-1} less stable. Structures with NH$_4$, where the proton has been abstracted from homocitrate, are 11–13 (r^{2}SCAN) or 21–22 kJ mol^{-1} (TPSSh) less stable and in these NH$_4$ has dissociated from the cluster (and S2B bridges Fe2 and Fe6). All structures are most stable in the quartet BS-235 state.

Conclusions

We have investigated whether structures with S2B dissociated from either Fe2 or Fe6 may be involved in the second half of the reaction mechanism of nitrogenase. As mentioned in the Introduction, we have previously studied this part of the reaction mechanism assuming either that S2B binds both to Fe2 and Fe6 or that it is fully dissociated from the FeMo cluster.[38,39] However, recent studies have indicated that for the E$_2$–E$_4$ states, structures with S2B dissociated from either Fe2 or Fe6 are more stable than structures with a bridging S2B.[22,44,45] Therefore, it is of great interest to know how this finding affects the second half of the reaction mechanism. To make such a half-dissociation possible, we have added one extra proton on S2B, compared to the previous studies. Moreover, we employ two DFT functionals, r^{2}SCAN and TPSSh, which in previous studies have supported and favoured half-dissociation of S2B,[22,44,45] but also giving accurate results for the FeMo cluster of nitrogenase.[58]

Interestingly, we see little advantage of half-dissociation of S2B. For the E$_4$ and E$_5$ states, such structures are at least 16–24 and 9–15 kJ mol^{-1} less stable than the best structures with a bridging S2B, respectively. For the E$_6$ state, structures with a half-dissociated S2B are disfavoured by 47–52 kJ mol^{-1}. However, with NH$_3$ dissociated, the best E$_6$ structure has S2B bound only to Fe6, because the NH ligand takes the Fe2–Fe6 bridging position. On the other hand, our results indicate that such structures are not involved in the mechanism. The situation is similar for the E$_7$ state: With an intact N–N bond, only structures with a bridging S2B ligand is found, whereas after cleavage of N–N bond, NH$_2$ prefers to bridge between Fe2 and Fe6, forcing S2B to dissociate from Fe2 in the most stable state. However, the substrate may also extract a proton from homocitrate, giving NH$_3$, which prefers to bind to Fe6 and then S2B goes back to a bridging position. 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^{-1}.

Fig. 7 shows our suggested reaction mechanism for nitrogenase with an extra proton on S2B. It contains only a single half-dissociated structure, NH$_2$(26) + NH$_3$(2;1A)-S2B(6;1B), and only intermittently during the E$_7$ state. Therefore, we conclude that half-dissociation of S2B needs to be considered in the reaction mechanism of nitrogenase but seems to be of minor importance during the second half-reaction. The reason for this may be that there are no states with a bridging hydride ion in the second half-reaction.
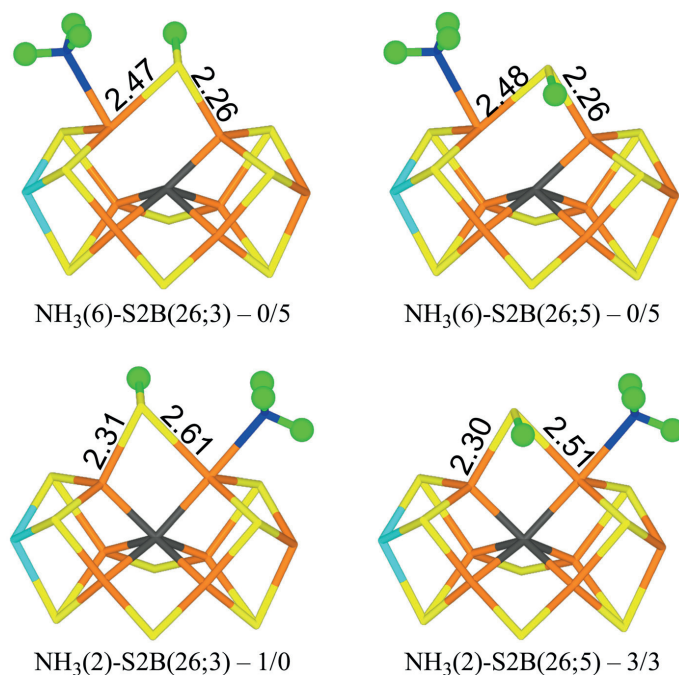
$NH_3(6)$-S2B(26;3) – 0/5

$NH_3(6)$-S2B(26;5) – 0/5

$NH_3(2)$-S2B(26;3) – 1/0

$NH_3(2)$-S2B(26;5) – 3/3

Fig. 6 The best E_8 structures with relative energies in kJ mol^{-1} indicated (r^{2}SCAN/TPSSh).
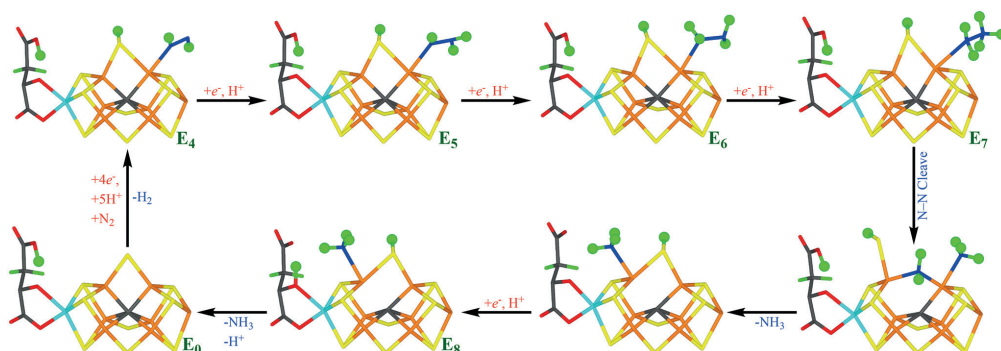

Fig. 7 Suggested reaction mechanism for nitrogenase, assuming that the S2B ligand is protonated.

In this study we have compared the results of two DFT functionals, r^{2}SCAN and TPSSh. The two functionals give similar results for both structures and relative energies. For example, mean absolute deviation of the relative energies calculated with the two methods in Tables 1–7 are 5–10 kJ mol^{-1} and for the most stable structures in Fig. 2–6, the maximum difference in the relative energies is 1–11 kJ mol^{-1}. The only exception is when different BS states are involved; then the difference can increase up to 25 kJ mol^{-1} (for example H_2NNH_3(2;1A)-S2B (26;3) in Fig. 5). Thus, the two functionals give quite similar results although one of them is a meta generalised gradient approximation functional, whereas TPSSh is a hybrid func-

tional with 10% Hartree–Fock exchange. The two functionals were selected because they give accurate structures of nitrogenase models[58] but they also give a stronger preference of half-dissociated structures, *e.g.* compared to pure generalised gradient approximation functionals like TPSS.[45]

Finally, we note that the suggested reaction mechanism is alternating, *i.e.* the protons are added alternating to the two N atoms of the substrate, so that HNNH and H_2NNH_2 are intermediates in the mechanism and that NH_3 does not dissociate until the E_7 state. This is in agreement with our previously suggested mechanism with S2B bridging or dissociated from the cluster.[38,39] On the other hand, the substrate binds preferably to Fe2, which is in contrast to the previously suggested mechanism with a bridging S2B,[39] where the substrate bound preferably to Fe6. The difference seems to come from the protonation of S2B. The protonation state of S2B remains to be settled, but the previous mechanism with unprotonated S2B has the advantage that it employs homocitrate as a proton buffer, which explains why this group is needed for the nitrogenase reaction.[4,85]

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 B. K. Burgess and D. J. Lowe, *Chem. Rev.*, 1996, **96**, 2983–3012.

2 B. Schmid, H.-J. Chiu, V. Ramakrishnan, J. B. Howard and D. C. Rees, in *Handbook of Metalloproteins*, John Wiley & Sons, Ltd, 2006, pp. 1025–1036, DOI: 10.1002/0470028637.met174.

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

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

5 T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees and O. Einsle, *Science*, 2011, **334**, 940–940.

6 J. Kim and D. C. Rees, *Science*, 1992, **257**, 1677–1682.

7 O. Einsle, F. A. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard and D. C. Rees, *Science*, 2002, **297**, 1696–1696.

8 T. Spatzal, K. A. Perez, O. Einsle, J. B. Howard and D. C. Rees, *Science*, 2014, **345**, 1620–1623.

9 O. Einsle, *J. Biol. Inorg. Chem.*, 2014, **19**, 737–745.

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

11 R. N. F. Thorneley and D. J. Lowe, in *Molybdenum Enzymes*, ed. T. G. Spiro, Wiley, New York, 1985, pp. 221–284.

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

13 I. Dance, *ChemBioChem*, 2020, **21**, 1671–1709.

14 Z. Y. Yang, N. Khadka, D. Lukoyanov, B. M. Hoffman, D. R. Dean and L. C. Seefeldt, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 16327–16332.

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

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

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

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

19 S. Raugei, L. C. Seefeldt and B. M. Hoffman, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 10521–10530.

20 L. Cao and U. Ryde, *J. Chem. Theory Comput.*, 2020, **16**, 1936–1952.

21 A. T. Thorhallsson, B. Benediktsson and R. Bjornsson, *Chem. Sci.*, 2019, **10**, 11110–11124.

22 H. Jiang and U. Ryde, *Dalton Trans.*, 2023, **52**, 9104–9120.

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

24 P. E. M. Siegbahn, *J. Comput. Chem.*, 2018, **39**, 743–747.

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

26 M. Rohde, D. Sippel, C. Trncik, S. L. A. Andrade and O. Einsle, *Biochemistry*, 2018, **57**, 5497–5504.

27 W.-L. Li, Y. Li, J. Li and T. Head-Gordon, *Chem Catal.*, 2023, **3**, 100662.

28 P. E. M. Siegbahn, *Phys. Chem. Chem. Phys.*, 2023, **25**, 23602–23613.

29 J. Chatt, J. R. Dilworth and R. L. Richards, *Chem. Rev.*, 1978, **78**, 589–625.

30 J. Chatt, *Annu. Proc. Phytochem. Soc. Eur.*, 1980, **18**, 1–18.

31 D. V. Yandulov and R. R. Schrock, *Science*, 2003, **301**, 76–78.

32 R. R. Schrock, *Acc. Chem. Res.*, 2005, **38**, 955–962.

33 R. R. Schrock, *Angew. Chem., Int. Ed.*, 2008, **47**, 5512–5522.

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

35 R. N. F. Thornely, R. R. Eady and D. J. Lowe, *Nature*, 1978, **272**, 557–558.

36 R. N. F. Thorneley and D. J. Lowe, *Biochem. J.*, 1984, **224**, 887–894.

37 D. Lukoyanov, S. A. Dikanov, Z.-Y. Yang, B. M. Barney, R. I. Samoilova, K. V. Narasimhulu, D. R. Dean, L. C. Seefeldt and B. M. Hoffman, *J. Am. Chem. Soc.*, 2011, **133**, 11655–11664.

38 L. Cao and U. Ryde, *J. Catal.*, 2020, **391**, 247–259.

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

40 L. Cao and U. Ryde, *J. Biol. Inorg. Chem.*, 2020, **25**, 521–540.

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

42 P. P. Hallmen and J. Kästner, *Z. Anorg. Allg. Chem.*, 2015, **641**, 118–122.

43 I. Dance, *Dalton Trans.*, 2019, **48**, 1251–1262.

44 A. T. Thorhallsson and R. Bjornsson, *Chem. – Eur. J.*, 2021, **27**, 16788–16800.

45 H. Jiang, O. K. G. Svensson and U. Ryde, *Inorg. Chem.*, 2022, **61**, 18067–18076.

46 L. Cao and U. Ryde, *Phys. Chem. Chem. Phys.*, 2019, **21**, 2480–2488.

47 L. Cao, O. Caldararu and U. Ryde, *J. Phys. Chem. B*, 2017, **121**, 8242–8262.

48 H. Jiang and U. Ryde, *Phys. Chem. Chem. Phys.*, 2024, **26**, 1364–1375.

49 R. Bjornsson, F. Neese and S. DeBeer, *Inorg. Chem.*, 2017, **56**, 1470–1477.

50 B. M. Barney, J. McClead, D. Lukoyanov, M. Laryukhin, T.-c. Yang, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, *Biochemistry*, 2007, **46**, 6784–6794.

51 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

52 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

53 L. Hu and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 2452–2463.

54 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

55 F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 91–100.

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

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

58 B. Benediktsson and R. Bjornsson, *J. Chem. Theory Comput.*, 2022, **18**, 1437–1457.

59 A. Schäfer, H. Horn and R. Ahlrichs, *J. Chem. Phys.*, 1992, **97**, 2571–2577.

60 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.*, 1995, **240**, 283–289.

61 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Acc.*, 1997, **97**, 119–124.

62 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *J. Chem. Phys.*, 2019, **150**, 154122–154122.

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

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

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

66 L. Cao and U. Ryde, *Int. J. Quantum Chem.*, 2018, **118**, e25627.

67 C. Greco, P. Fantucci, U. Ryde and L. Di Gioia, *Int. J. Quantum Chem.*, 2011, **111**, 3949–3960.

68 R. K. Szilagyi and M. A. Winslow, *J. Comput. Chem.*, 2006, **27**, 1385–1397.

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

70 U. Ryde and M. H. M. Olsson, *Int. J. Quantum Chem.*, 2001, **81**, 335–347.

71 N. Reuter, A. Dejaegere, B. Maigret and M. Karplus, *J. Phys. Chem. A*, 2000, **104**, 1720–1735.

72 L. Hu, P. Söderhjelm and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 761–777.

73 L. Cao and U. Ryde, *Front. Chem.*, 2018, **6**, 89–89.

74 I. Mayer, *Chem. Phys. Lett.*, 1983, **97**, 270–274.

75 T. Lu and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580–592.

76 I. Dance, *J. Am. Chem. Soc.*, 2005, **127**, 10925–10942.

77 I. Dance, *Dalton Trans.*, 2012, **41**, 7647–7659.

78 I. Dance, *Inorg. Chem.*, 2013, **52**, 13068–13077.

79 I. Dance, *Dalton Trans.*, 2015, **44**, 18167–18186.

80 I. Dance, *J. Inorg. Biochem.*, 2017, **169**, 32–43.

81 I. Dance, *Biochemistry*, 2006, **45**, 6328–6340.

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

83 L. C. Seefeldt, I. G. Dance and D. R. Dean, *Biochemistry*, 2004, **43**, 1401–1409.

84 S. J. George, B. M. Barney, D. Mitra, R. Y. Igarashi, Y. Guo, D. R. Dean, S. P. Cramer and L. C. Seefeldt, *J. Inorg. Biochem.*, 2012, **112**, 85–92.

85 J. Imperial, T. R. Hoover, M. S. Madden, P. W. Ludden and V. K. Shah, *Biochemistry*, 1989, **28**, 7796–7799.

Paper VII

Quantum Mechanical Calculations of Redox Potentials of the Metal Clusters in Nitrogenase

H. Jiang, O. K. G. Svensson and U. Ryde

Article

Quantum Mechanical Calculations of Redox Potentials of the Metal Clusters in Nitrogenase

Hao Jiang [ID], Oskar K. G. Svensson and Ulf Ryde *[ID]

Division of Theoretical Chemistry, Department of Chemistry, Lund University, P.O. Box 124, SE-221 00 Lund, Sweden

* Correspondence: ulf.ryde@teokem.lu.se; Tel.: +46-46-2224502; Fax: +46-46-2228648

Abstract: We have calculated redox potentials of the two metal clusters in Mo-nitrogenase with quantum mechanical (QM) calculations. We employ an approach calibrated for iron–sulfur clusters with 1–4 Fe ions, involving QM-cluster calculations in continuum solvent and large QM systems (400–500 atoms), based on structures from combined QM and molecular mechanics (QM/MM) geometry optimisations. Calculations on the P-cluster show that we can reproduce the experimental redox potentials within 0.33 V. This is similar to the accuracy obtained for the smaller clusters, although two of the redox reactions involve also proton transfer. The calculated P^{1+}/P^N redox potential is nearly the same independently of whether P^{1+} is protonated or deprotonated, explaining why redox titrations do not show any pH dependence. For the FeMo cluster, the calculations clearly show that the formal oxidation state of the cluster in the resting E_0 state is $Mo^{III}Fe_3^{II}Fe_4^{III}$, in agreement with previous experimental studies and QM calculations. Moreover, the redox potentials of the first five E_0–E_4 states are nearly constant, as is expected if the electrons are delivered by the same site (the P-cluster). However, the redox potentials are insensitive to the formal oxidation states of the Fe ion (i.e., whether the added protons bind to sulfide or Fe ions). Finally, we show that the later (E_4–E_8) states of the reaction mechanism have redox potential that are more positive (i.e., more exothermic) than that of the E_0/E_1 couple.

Keywords: nitrogenase; redox potential; formal oxidation states; FeMo cluster; P-cluster

1. Introduction

Gaseous nitrogen (N_2) is the main component of our atmosphere, but nitrogen is still a limiting element for plant life and a major ingredient in synthetic fertilisers. The reason for this is that the triple bond in N≡N is extremely strong and inert, making N_2 unavailable for most plants [1–3]. Industrially, N_2 is converted to NH_3 by the Born–Haber process, which was invited in the early 20th century and is considered as a major cause of the human population explosion [1]. It requires high temperature and pressure.

$$N_2 + 8\,e^- + 8\,H^+ + 16\,ATP \rightarrow 2\,NH_3 + H_2 + 16\,ADP + 16\,P_i \tag{1}$$
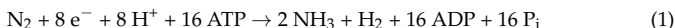
However, a few bacteria and archaea can perform the same conversion at ambient temperature and pressure. This is accomplished by the enzyme nitrogenase (EC 1.18/19.6.1) [2,4–10]. The reaction is still quite demanding, requiring 16 molecules of ATP, eight electrons and eight protons [2,4,5]: X-ray crystallographic studies have shown that nitrogenase contains two unusual iron–sulfur clusters [6–10]. The P-cluster, $Fe_8S_7Cys_6$, is essentially two merged [4Fe–4S] clusters with a central sulfide ligand coordinating to six iron ions (Figure 1a). It is employed for electron transfer from the Fe-protein, which donates the electrons, to the active site. The latter is the FeMo cluster, which is a complicated $MoFe_7S_9C$(homocitrate) cluster, with a central carbide ion (Figure 1b) and connected to the protein by one cysteine and one histidine residue at the opposite ends of the trigonal

prismatic cluster. There exist also alternative nitrogenases, in which the Mo ion is replaced by V or Fe [11].
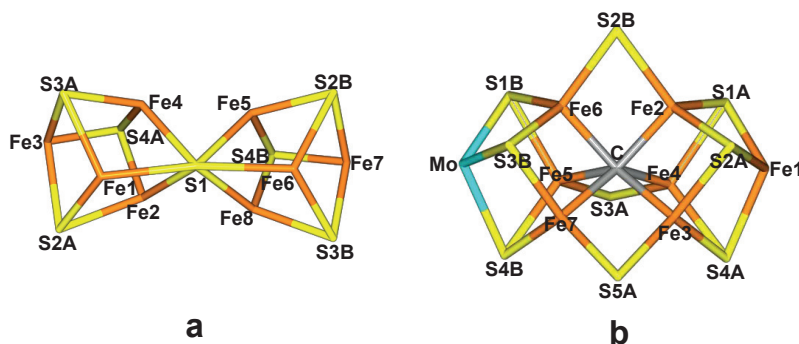

Figure 1. The cores of (**a**) the P-cluster (P^N state) and (**b**) the FeMo cluster (E_0 state) with atom names indicated.

Nitrogenase has been extensively studied by both experimental [2–5,12–17] and computational methods [18–22]. The reaction is normally described by the Lowe–Thorneley cycle [23], in which eight intermediates are recognised, E_0–E_8, based on the number of added electrons and protons. It has been shown that the resting E_0 state needs to be reduced by three or four electrons before the N_2 substrate can bind [2,3]. It is believed that the binding is facilitated by the reductive elimination of H_2 from the cluster, explaining why H_2 is a compulsory byproduct of the reaction. However, many details of the reaction are still unknown and many conflicting mechanisms have been suggested [3,18].

The Lowe–Thorneley cycle emphasizes the importance of electron and proton transfer in the reaction cycle of nitrogenase. The driving force of electron transfer is the redox potential. Unfortunately, it is hard to measure redox potentials of the FeMo cluster in nitrogenase because the reaction cannot be arrested at certain E_n states [24]. The only certain redox potential is between the resting E_0 state and a one-electron oxidised state, which is outside the Lowe–Thorneley cycle (we will denote it E_{-1}), -0.042 V [5,24–26]. For the reduction of the resting state, a redox potential of -0.45 to -0.49 V have been reported, but they may represent a mixture of reduced states [24,25,27–29].

Redox potentials can also be calculated by computational methods. However, the accuracy is limited. For redox sites of the same type in variants of the same protein (i.e., from different organisms or mutations), calculations based on the Poisson–Boltzmann equation or similar methods may give mean errors of 0.03–0.11 V for relative redox potentials [30–38]. For absolute potentials and sites of different types, quantum mechanical (QM) calculations are needed and the accuracy is appreciably worse. Typical errors are 0.2–0.6 V [30–34] and 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 [39]. Even all-atom QM molecular dynamics and free-energy calculations did not give an accuracy better than 0.26 V [40].

Recently, we performed a comparison and calibration of various combined QM and molecular mechanics (QM/MM) methods to estimate redox potentials of 13 iron–sulfur clusters with 1–4 Fe ions [41]. We showed that the best results were obtained by QM-cluster calculations in a continuum solvent with a high dielectric constant, using a large QM model (~300 atoms), based on QM/MM structures. With such an approach, we obtained a mean absolute error of 0.17 V, after removal of a systematic error of 0.62 V. The maximum error among the 13 studied potentials was 0.44 V. However, even if the accuracy is rather mediocre, it is enough to make useful predictions, e.g., what redox couple is employed by the [4Fe–4S] ferredoxins.

In this study, we employ this calibrated approach to study redox potentials in nitrogenase. We study four issues: First, we examine whether the computational method works also for the more complicated P- and FeMo clusters in nitrogenase, with redox potentials that also involve proton transfer. Second, we obtain an independent check of the redox state and charge state of the FeMo cluster. Third, we examine the recent suggestion that the E_0–E_4 states should operate at a nearly constant potential, employing only a single redox couple [42,43]. Fourth, we study the redox potentials of different reaction mechanisms after the binding of the substrate (i.e., for the E_4–E_8 states).

2. Result and Discussion

2.1. Redox Potentials of the P-Cluster

We started the investigation with testing our methodology [41] on the P-cluster to see if the results are reliable also for the large iron–sulfur clusters in nitrogenase and for redox reactions involving protonation of the clusters. The P-cluster (Figure 1a) contains eight Fe ions and the resting P_N state has been shown to be the fully reduced Fe_8^{II} state [44–46]. Three additional states have been experimentally observed, oxidised by 1–3 electrons [44–46]. They are denoted P^{1+}–P^{3+}. Only the first two states are believed to be involved in the catalytic mechanism, although there are some evidence that also the P^{2+} state may be used [46–49]. Crystallographic studies have shown that in the P^N state, the cluster is essentially two [4Fe–4S] clusters merged by the S1 sulfide ion that coordinates to six Fe ions (cf. Figure 1a). In the P^{2+} state, Ser-188D becomes deprotonated and coordinates to Fe6 [48,50,51]. Likewise, the backbone N atom of Cys-88 (the sidechain of which is one of the ligands to the P-cluster) also becomes deprotonated and coordinates to Fe5 [48,50,51]. This leads to cleavage of the Fe5–S1 and Fe6–S1 bonds. The same structure is expected for the P^{3+} state. The structure of the P^{1+} state is more uncertain, because redox titrations indicated that only the P^{2+}/P^{1+} redox-couple is pH-dependent, whereas no evidence was found for a coupled electron- and proton-transfer for the P^{1+}/P^N couple [52,53]. However, a recent crystal structure was interpreted to contain deprotonated Ser-188D and protonated Cys-88, although it is probably a mixture of the P^{1+} and the P^{2+} states [54–56].

Six states of the P-cluster were considered in this study as are described in Table 1. Each state was first QM/MM optimised with the small QM system and then subjected to a single-point energy calculation with the large QM system in a COSMO continuum solvent with a dielectric constant of 80. The QM/MM structures and the best BS states were taken from our previous publication [55]. The calculated redox potentials are listed in Table 2 and they are compared to experimentally measured redox potentials [5] in order to gauge the accuracy of the method when applied to the P-clusters of the nitrogenase. For the P^{1+} state, we tested three different protonation states.

Table 1. The various states studied for the P-cluster. The table lists an abbreviation (Abb.) for each state, the protonation status for Cys-88 and Ser-188D (P = protonated, D = deprotonated), the spin state (*S*), the BS state (specifying Fe ions with minority spin) and the net charge of the QM systems (ch; same for both for QM/MM and for the redox calculations).

State	Abb.	Cys-88	Ser-188	*S*	BS	ch
reduced	P^NH_2	P	P	0	1247	−4
one-electron oxidised	$P^{1+}H_2$	P	P	1/2	1247	−3
	$P^{1+}H$	P	D	1/2	1247	−4
	P^{1+}	D	D	1/2	1247	−5
two-electron oxidised	P^{2+}	D	D	4	358	−4
three-electron oxidised	P^{3+}	D	D	7/2	358	−3

Table 2. Redox potentials of five redox couples for the P-cluster in V. E^o_{calc} is the raw redox potentials from the COSMO calculations, E^o_{corr} are calculated redox potentials corrected by the mean signed error (MSE = -0.62 V) from our previous study [41] (i.e., $E^o_{corr} = E^o_{calc} + 0.62$ V), E^o_{exp} is the experimental redox potentials [5,24,25,45] and E^o_{err} is the error for the various calculations. Results with $E^o_{err} < 0.44$ V are marked in bold face.

Redox Couple	E^o_{calc}	E^o_{corr}	E^o_{exp}	E^o_{err}
$P^NH_2 \rightarrow P^{1+}H_2$	-0.62	0.00	-0.309	**0.31**
$P^NH_2 \rightarrow P^{1+}H$	-0.60	0.02	-0.309	**0.33**
$P^{1+}H \rightarrow P^{2+}$	-0.48	0.14	<-0.224 [a]	**<0.36**
$P^{1+} \rightarrow P^{2+}$	-1.69	-1.07	<-0.348 [a]	<0.72
$P^{2+} \rightarrow P^{3+}$	-0.58	0.04	0.09	-0.05

[a] The measured redox potential for the P^{2+}/P^{1+} couple is pH dependent, decreasing from -0.224 V at pH 6.0 to -0.348 V at pH 8.5 (cf. the text) [52].

It can be seen that the redox couples $P^{2+} \rightarrow P^{3+}$ and $P^NH_2 \rightarrow P^{1+}H$ give errors of 0.05 and 0.33 V, respectively, compared to experimental potentials. These are within the range of errors observed in our previous study (maximum error 0.44 V) [41]. This is quite satisfying, especially considering that one of the considered redox potentials for the P-cluster involves coupled redox and protonation reaction, whereas the calibration study involved only pure redox reactions. This gives us confidence to apply the method also to the FeMo cluster.

For the P^{1+}/P^N redox potential, the calculated result is essentially independent on whether we include the proton transfer or not in the calculations (the $P^NH_2 \rightarrow P^{1+}H_2$ and $P^NH_2 \rightarrow P^{1+}H$ give calculated redox potentials that differ by only 0.02 V). This is in agreement with the experimental observation that the P^{1+}/P^N redox potential is pH-independent [52,53] and may solve the enigma why redox titrations did not observe any pH-dependence although crystal structures indicate that a deprotonation should be involved in this redox reaction [54–56].

For the $P^{1+}H \rightarrow P^{2+}$ transition, the comparison with experimental results is somewhat harder, because the measured redox potential changes with pH, from -0.224 V at pH 6.0 to -0.348 V at pH 8.5 [52]. This has been interpreted to reflect the deprotonation of the backbone amide group of Cys-88: At very low pH, it is protonated in both the P^{1+} and P^{2+} states, whereas at high pH, it is deprotonated in both redox states. At intermediate pH (i.e., in the measured range), it is protonated in the P^{1+} state and deprotonated in the P^{2+} state [52]. Our calculations confirm these suggestions: We obtain a more positive potential for the $P^{2+}/P^{1+}H$ couple ($E^o_{corr} = 0.1$ V) than for the P^{2+}/P^{1+} couple ($E^o_{corr} = -1.1$ V). However, the restricted pH range of the measured potentials makes it hard to make a more detailed judgement of the calculated potentials. The experiments indicate that the redox potential of the pure $P^{1+}H \rightarrow P^{2+}$ transition is larger than -0.224 V, indicating an error of less than 0.36 V for our calculated potential. Likewise, the redox potential of the pure $P^{1+} \rightarrow P^{2+}$ transition is more negative than -0.348 V, indicating an error of less than 0.72 V. Thus, our calculations confirm the experimental observation that the $P^{1+} \rightarrow P^{2+}$ transition involves a proton transfer [52,53], but we cannot obtain any detailed estimate of the error for this redox couple.

2.2. Accuracy of the Redox Potential of the FeMo Cluster and Oxidation Level of the Cluster

As mentioned above, only two experimental redox potentials have been reported for the FeMo cluster [24,25]. A potential of -0.042 V has been measured between the resting state and a one-electron oxidised state [26,45] (not involved in the normal reaction mechanism). Redox potentials for more reduced states are harder to measure, because reduction starts substrate or proton reduction. No firm measurement is available, but the potential between the resting state and a steady-state reduced state (which my represent more than one reduced state) has been estimated in four studies between -0.45 and -0.49 V [24,25,27–29]. In two cases, another potential of -0.30 to -0.32 V was also reported, but it may come from the P-cluster [24].

For the FeMo cluster, we wanted to study two related questions, viz. whether we can reproduce these potentials with our QM calculations, within the accuracy of the method, and whether we can identify the correct redox couple of the FeMo cluster. Recent Mössbauer, anomalous dispersion and QM investigations have suggested that the resting E_0 state of the FeMo cluster is in the $Mo^{III}Fe_3^{II}Fe_4^{III}$ oxidation state [13,42,57,58]. However, this gives a large net negative charge of the cluster and its direct ligands ($-5\,e$), which is only partly neutralised by two nearby arginine residues. Moreover, protonation energies of various groups of the FeMo cluster are unexpectedly large [59,60]. Therefore, an independent confirmation of the oxidation-state assignment is desirable.

To this end, twelve states of the FeMo cluster were considered, described in Table 3. We studied the resting E_0 state, together with the one-electron oxidised and one-electron reduced states. For the latter, we considered both a structure with no protons added and a state with a proton added on the S2B μ_2 bridging sulfide ion, as has been suggested by several QM investigations [18,59] and is also supported by experimental studies [61,62]. For the standard oxidation-state assignment ($Mo^{III}Fe_3^{II}Fe_4^{III}$ for E_0), these states are denoted E_0, E_{-1}, E_1 and E_1H. In addition, we considered two alternative charge states, either with two more or two electrons less (experimentally the resting state is a quartet [2,3], i.e., with an odd number of electrons, so electrons need to be added or removed in pairs). These states are called A_0, A_{-1}, A_1 and A_1H when two electrons were added (i.e., giving $Mo^{III}Fe_5^{II}Fe_2^{III}$ for the A_0 state) and R_0, R_{-1}, R_1 and R_1H when two electrons were removed (i.e., giving $Mo^{III}Fe_1^{II}Fe_6^{III}$ for the R_0 state; note that $R_1 = E_{-1}$ and $A_{-1} = E_1$). The calculated redox potentials are listed in Table 4.

Table 3. States studied for the resting, one-electron reduced or one-electron oxidised states of the FeMo cluster. The table lists the protonation status of the S2B atom (P = protonated, D = deprotonated), the spin state (*S*), the BS state (specifying Fe ions with minority spin), and the net charge of the QM systems (ch; same for both for QM/MM and for the redox calculations).

State	S2B	*S*	BS	ch
R_{-1}	D	1	235	0
R_0	D	3/2	235	-1
R_1	D	2	235	-2
R_1H	P	2	235	-1
E_{-1}	D	1	235	-2
E_0	D	3/2	235	-3
E_1	D	2	235	-4
E_1H	P	2	235	-3
A_{-1}	D	2	235	-4
A_0	D	3/2	235	-5
A_1	D	2	235	-6
A_1H	P	2	235	-5

It can be seen that with the standard charge state ($Mo^{III}Fe_3^{II}Fe_4^{III}$ for E_0), our calculations reproduce the two experimental redox potentials with errors of 0.17 and 0.19 V, i.e., well within the error range observed in our previous study (maximum error 0.44 V) [41] and also for the P-cluster. However, the good results are observed only if it is assumed that the reduction of E_0 is accompanied by the uptake of by a proton (i.e., $E_0 \rightarrow E_1H$; for $E_0 \rightarrow E_1$, the error is 0.6 V), showing that the calculations confirm that a proton transfer is involved in the redox reaction.

If we instead consider a FeMo cluster with two electrons less (i.e., with a $Mo^{III}Fe_1^{II}Fe_6^{III}$ assignment for the resting state, here called R_0), the calculated potentials reproduce the experimental redox potentials worse: The $R_{-1} \rightarrow R_0$ transition gives and error of 2.0 V, much larger than the maximum error in our previous study [41], whereas $R_0 \rightarrow R_1H$ gives an error of 0.43 V.

Table 4. Redox potentials for various redox couples of the FeMo cluster in V. E^o_{calc} is the raw redox potentials, $E^o_{corr} = E^o_{calc} + 0.62$ V (the MSE in our previous study [41]), E^o_{exp} is the experimental redox potentials (for the E_0/E_1 couple, we used -0.47 V, which is in the middle of the range of reported values) [5,24,25,27–29] and E^o_{err} is the error for the various calculations. Results with $E^o_{err} < 0.44$ V are marked in bold face.

Transition	E^o_{calc}	E^o_{corr}	E^o_{exp}	E^o_{err}
$R_{-1} \to R_0$	1.31	1.93	-0.042	1.98
$R_0 \to R_1$	0.42	1.04	(-0.47)	1.51
$R_0 \to R_1H$	-0.66	-0.04	-0.47	**0.43**
$E_{-1} \to E_0$	-0.49	0.13	-0.042	**0.17**
$E_0 \to E_1$	-1.69	-1.07	(-0.47)	-0.60
$E_0 \to E_1H$	-1.28	-0.66	-0.47	**-0.19**
$A_{-1} \to A_0$	-6.88	-6.29	-0.042	-6.21
$A_0 \to A_1$	-3.46	-2.84	(-0.47)	-2.37
$A_0 \to A_1H$	-1.57	-0.95	-0.47	-0.48

Likewise, if we instead add two extra electrons ($Mo^{III}Fe_5^{II}Fe_2^{III}$ for the resting state), we get a very large error (over 6 V) for the $A_{-1} \to A_0$ potential. For the $A_0 \to A_1H$ transition, the error is smaller, -0.48 V, but it is still somewhat larger than the maximum error observed in our calibration study. Thus, our calculations confirm that $Mo^{III}Fe_3^{II}Fe_4^{III}$ is the proper redox assignment for E_0. Apparently, the E_{-1}/E_0 redox potential is more sensitive to the involved redox couple than the E_0/E_1H potential. The calculations also confirm that the E_1 state is protonated.

2.3. Redox Potentials of the E_0–E_4 States of the FeMo Cluster

Next, we studied also the E_0–E_4 states of the FeMo cluster from Lowe–Thorneley reaction cycle. The aim was to examine the suggestion that all the E_0–E_4 states should have similar redox potentials, because they use only two formal redox states, $Mo^{III}Fe_3^{II}Fe_4^{III}$ and $Mo^{III}Fe_4^{II}Fe_3^{III}$ [42,43]. This is suggested to be accomplished by the added protons, which bind to Fe ions in the E_2 and E_4 states, thereby formally becoming hydride ions and changing the oxidation state of the Fe ions by two. Thus, the five states E_0–E_4 would formally be $Fe_3^{II}Fe_4^{III}$, $Fe_4^{II}Fe_3^{III}H^+$, $Fe_3^{II}Fe_4^{III}H^+H^-$, $Fe_4^{II}Fe_3^{III}H_2^+H^-$ and $Fe_3^{II}Fe_4^{III}H_2^+H_2^-$ (leaving out Mo, which always is in the +III state).

We used mainly structures from previous studies [59,63–67] and included a few alternative structures for each state (except for E_0 and E_1H, for which there is a reasonable consensus), to see if we can discriminate between different possibilities using the redox potentials. The various structures are described in Table 5 (they are also shown in Figure S3) and the calculated redox potentials are listed in Table 6. We use the redox potential of the E_0/E_1H couple as a reference ($\Delta E^o_{calc} = E^o_{calc} - \Delta E^o_{calc}(E_0/E_1H)$) to judge if all transitions have similar redox potentials.

Table 5. Structures studied for the E_0–E_8 states of the FeMo cluster. The table lists the protonated atoms or the added substrate, the spin state (S), the BS state (specifying the Fe ions with minority spin), and the net charge of the QM systems (ch; same for both for QM/MM and for the redox calculations).

State	Protonated Atoms/Substrate	S	BS	ch
E_0	–	3/2	235	-3
E_1H	S2B(3)	2	235	-3
E_2H_2	S2B(3), Fe2/6(5)	3/2	247	-3
E_2H_2'	S2B (dissoc from Fe2), Fe2/6(5)	3/2	247	-3
E_2H_2''	S2B(3), S5A(3)	3/2	247	-3
E_3H_3	S2B(3), Fe2/6(5), Fe3/7(2)	1	14	-3
E_3H_3'	S2B(3), Fe2/6(5), Fe5	1	14	-3
E_4H_4	S2B(3), Fe2/6(3), Fe3/7(2), S5A(3)	1/2	14	-3

Table 5. *Cont.*

State	Protonated Atoms/Substrate	S	BS	ch
E_4H_4'	S2B(3), Fe2/6(5), Fe3/7(2), S5A(3)	1/2	14	−3
E_4H_4''	S2B(5), Fe2/6(5), Fe3/7(2), S5A(2)	1/2	14	−3
E_4H_4'''	S2B(3), Fe2/6(3), Fe6, Fe5	1/2	14	−3
With S2B still bound				
$E_4N_2H_2$	HNNH$_2$ (proton from HCA)	1/2	147	−3
$E_5N_2H_3$	H$_2$NNH$_2$ (proton from HCA)	1	147	−3
$E_6N_2H_4$	NH$_2$ + NH$_3$ (proton from HCA)	1/2	147	−3
E_6NH	NH$_2$ (proton from HCA)	1/2	147	−3
E_7NH_2	NH$_3$ (proton from HCA)	1	147	−3
E_8NH_3	NH$_3$	1/2	147	−3
With S2B dissociated				
$E_4N_2H_2'$	NNH$_2$	1/2	147	−1
$E_5N_2H_3'$	HNNH$_2$	1	147	−1
$E_6N_2H_4'$	H$_2$NNH$_2$	1/2	147	−1
$E_7N_2H_5'$	NH$_2$ + NH$_3$	0	147	−1
E_7NH_2'	NH$_2$	1	147	−1
E_8NH_3'	NH$_3$	1/2	147	−1

Table 6. Calculated redox potentials for the E_0–E_8 states of the FeMo cluster. The last column (ΔE^o_{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^o_{calc}	ΔE^o_{calc}
$E_0 \rightarrow E_1H$	**−1.28**	**0.00**
$E_1H \rightarrow E_2H_2$	**−1.20**	**0.08**
$E_1H \rightarrow E_2H_2'$	−1.45	−0.17
$E_1H \rightarrow E_2H_2''$	−1.29	0.00
$E_2H_2 \rightarrow E_3H_3$	**−1.47**	**−0.19**
$E_2H_2 \rightarrow E_3H_3'$	−1.81	−0.53
$E_3H_3 \rightarrow E_4H_4$	**−0.87**	**0.41**
$E_3H_3 \rightarrow E_4H_4'$	−1.10	0.18
$E_3H_3 \rightarrow E_4H_4''$	−1.34	−0.06
$E_3H_3 \rightarrow E_4H_4'''$	−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

In our previous studies, we found that with the TPSS functional, the best E_2H_2 structure has a proton on S2B and a hydride ion bridging Fe2 and Fe6 [59,64]. It can be seen that the estimated redox potential for the $E_1H \rightarrow E_2H_2$ transition is similar to that of the $E_0 \rightarrow E_1H$ transition, only 0.08 V less negative. This confirms the conjecture that the FeMo cluster in the early E_n states should have a constant redox potential.

If we instead use another structure for the E_2H_2 state with the extra two protons bound to S2B and as a hydride ion bridging Fe2 and Fe6, but with S2B dissociated from Fe2 (but not from Fe6; called E_2H_2'), the redox potential for the $E_1H \rightarrow E_2H_2'$ transition decreases by 0.25 V. This simply reflects that with TPSS-D4/def2-SV(P) and the 399-atom redox model, the E_2H_2' structure is 0.25 eV (24 kJ/mol) less stable than the E_2H_2 structure. For the 184-atom QM/MM model, the difference is 15 kJ/mol. In our previous study, we showed that the relative stability of the structures with the protonated S2B group bridging

Fe2 and Fe6 or dissociated from one of the two iron ions depends on what DFT functional is used [64]. For example, with the TPSSh functional, the E_2H_2' structure is instead 11 kJ/mol more stable. Consequently, the redox potentials will also depend on the QM method used, with differences of ~0.3 V.

If we instead use a structure for the E_2H_2 state with the two protons on S2B and S5A (E_2H_2''), the calculated redox potential is 0.08 V more negative than with the most stable E_2H_2 state. Again, this reflects that this protonation state is 9 kJ/mol less stable. However, it also shows that the formal oxidation states in the FeMo cluster have only a minor influence on the redox potential. The E_2H_2'' state involves two protons on sulfide groups and therefore represents a doubly reduced formal $Fe_5^{II}Fe_2^{III}H_2^+$ state in contrast to the $Fe_3^{II}Fe_4^{III}H^+H^-$ state for E_2H_2 and E_2H_2'. This shows that as long as the various protonation states have similar relative energies, they will also have similar redox potentials, showing that there is no major difference between different formal oxidation states of the Fe ions. This is in line with suggestions by Dance, who has pointed out that it is misleading to make a sharp contrast between protons on sulfides and hydride ions on Fe ions, because there is only a small difference in the charge populations on the H atom [18].

For the E_3H_3 state, we used a structure with the third proton bridging Fe3 and Fe7 (in addition a proton on S2B and a hydride ion bridging Fe2 and Fe6). The estimated redox potential for the $E_2H_2 \rightarrow E_3H_3$ transition is 0.19 V more negative than that of the $E_0 \rightarrow E_1H$ transition. This is well within the maximum error in our calibration study 0.44 V [41], and therefore still in agreement with the expectation that the electrons can be donated to the FeMo cluster at a constant redox potential. We tested also another structure, taken from our previous systematic study [59], in which the third proton bound to Fe5 instead (E_3H_3'). It gave a 0.35 V more negative redox potential (i.e., further away from that of the $E_0 \rightarrow E_1H$ transition), reflecting that this structure is less stable. Interestingly, we found in contrast to our previous study that the broken-symmetry BS-14 state was more favourable for both these structures with the large QM model used in the redox calculations (but only for the best structure with the smaller QM/MM-optimised model).

For the $E_3H_3 \rightarrow E_4H_4$ transition, our estimated redox potential is 0.41 V more positive than that of the $E_0 \rightarrow E_1H$ transition. This is within the maximum error in our calibration study [41], but considering that the redox potential of the E_2H_2/E_3H_3 couple was a bit to negative and this redox potential is a bit too positive, it might indicate that we have not yet found the best structure for the E_3 state. For E_4H_4, we employed the best structure in our previous investigation of this state [63], viz. a structure with two protons on S2B and S5A and two hydride ions bridging Fe2/6 and Fe3/7. There are several possible conformations of such a structure. The best one has all H atoms pointing towards S3A, except the Fe3/7 hydride, which points towards S2B (all structures are shown in Figure S3). If we instead use a structure with the Fe2/6 hydride on the other side of S2B, i.e., the same face as the Fe2/6 hydride, the redox potential becomes 0.23 V more negative, reflecting that such a structure is 22 kJ/mol less stable. If we instead use the structure suggested by Hoffman and coworkers [2,21], i.e., with all four H atoms on the same face of the cluster (i.e., the two protons on S2B and S5A pointing in the opposite direction compared to E_4H_4 and E_4H_4' structures), the redox potential becomes even more negative by 0.24 V. Likewise, if we use the best structure in our first investigation of the E_4H_4 state [59] (with H atoms on S2B, Fe2/6, Fe5 and Fe6), the redox potential becomes 0.15 V even more negative, reflecting that this structure is 59 kJ/mol less stable than the best state.

In conclusion, we find that for all four calculated redox potentials for the E_0–E_4 states are similar, within 0.41 V, i.e., within the accuracy of our method. This confirm the expectation that the redox potentials should be similar [42,43] so that they can accept electrons from the same source. However, the results are very sensitive to which structures are employed and which QM method and broken-symmetry state is used. On the other hand, we show that the formal oxidation states of the Fe ions (the number of protons on sulfide ions or hydride ions on Fe) is less important for the redox potentials, contrary to the suggestion that the redox potential depends on the formal oxidation state of the Fe ions [43,68].

2.4. Redox Potentials of the E_4–E_8 States of the FeMo Cluster

Finally, we studied also redox potentials of the FeMo cluster in the later part of the reaction, after binding of N_2 and its protonation to N_2H_2. In previous studies, we have suggested thermodynamically stable structures for the $E_4N_2H_2$ to E_8NH_3 states, either with S2B bound or dissociated from the cluster [65–67]. We use these structures also in this study. They are described in Table 5 and are shown in Figure S3. In Table 6, the calculated redox potentials for both scenarios are presented.

It can be seen that both with and without S2B, the calculated redox potentials are all less negative than that for the E_0/E_1H couple, by 0.2–2.65 V. This reflects that once N_2H_2 has been formed, the following reactions are quite facile. Electrons tend to move towards sites with a more positive redox potential. Therefore, redox potentials that are more positive than those of the E_0/E_1H couple indicate that the electron transfer is more exothermic than in the $E_0 \rightarrow E_1$ step. Thus, the electron transfers of the E_4–E_8 steps of the nitrogenase reaction are more downhill than those of the E_0–E_4 reactions. However, it might also indicate that the assumption that the bound N_2 directly is protonated may be incorrect. In fact, the first protonation of N_2 is the hardest step in the reduction of N_2 to ammonia [69], and it is possible that it requires further reduction of the FeMo cluster before it is feasible (this part of the reaction was not studied in our previous studies). Our results indicate that this should be further studied.

The relative sizes of the four redox potentials for the E_4–E_8 states are also rather independent on whether S2B remains bound or is dissociated: The potentials of the first and third steps $E_4 \rightarrow E_5$ and $E_6 \rightarrow E_7$ are most positive, especially the latter, whereas the redox potentials of the other two steps are closer to that of the E_0/E_1H couple. For the four reductions with S2B still bound, our calculated redox potentials are 1.13, 0.41, 2.02 and 0.57 V, whereas with S2B dissociated, the four calculated redox potentials are 0.97, 0.20, 2.65 and 0.19 V (i.e., with a somewhat larger variation). The similarity of the trends for the two sets of potentials is conspicuous considering that the N–N bond is cleaved in the $E_5N_2H_3 \rightarrow E_6N_2H_4$ transition with S2B bound, but in the $E_6N_2H_4 \rightarrow E_7N_2H_5$ transition when S2B has dissociated.

In conclusion, our results show that later part of the reaction mechanism of Mo-nitrogenase give redox reactions that are more exothermic than that of the E_0/E_1H redox couple. Thus, we see no evidence that a stronger driving force is needed for the reaction, as has been suggested by Siegbahn [22]. Moreover, there is no large difference between the mechanisms with S2B bound or dissociated.

3. Methods

3.1. The Protein

The calculations were based on the 1.0-Å crystal structure of Mo nitrogenase from *Azotobacter vinelandii* (PDB code 3U7Q) [8]. The setup of the protein is identical to that of our previous studies [60,63,69,70]. The entire heterotetramer was considered in the calculations and the quantum mechanical (QM) calculations were concentrated on the FeMo clusters in the C subunit because there is a buried imidazole molecule from the solvent rather close to the active site (~11 Å) in the A subunit. The metal clusters not involved in the QM calculations were modelled by MM in the fully reduced and E_0 resting states, respectively, using a QM charge model [70]. The protonation states of all residues were the same as before [70], and the homocitrate ligand was modelled in the singly protonated state with a proton shared between the hydroxyl group (O7 that coordinates to Mo) and the O1 carboxylate atom [57,70]. The protein was solvated in a sphere with a radius of 65 Å around the geometrical centre of the protein. Cl^- and Na^+ ions were added to an ionic strength of 0.2 M [71]. The final system contained 133 915 atoms. For the protein, we used the Amber ff14SB force field [72] and water molecules were described by the TIP3P model [73]. The metal sites [70,74] were treated by a non-bonded model [75] and charges were obtained with the restrained electrostatic potential method [76].

3.2. QM Calculations

All QM calculations were performed with the Turbomole software (version 7.5) [77]. All calculations were performed with the TPSS [78] functional with the def2-SV(P) basis set [79], a combination that gave the best relative redox potentials in our previous study (lowest mean absolute deviation and maximum error, after removal of the mean signed error; B3LYP or larger basis sets gave more than twice as large mean absolute deviations and maximum errors, but a slightly smaller mean signed error) [41]. The calculations were sped up by expanding the Coulomb interactions in an auxiliary basis set, the resolution-of-identity (RI) approximation [80,81]. Empirical dispersion corrections were included with the DFT-D4 approach [82], as implemented in Turbomole. QM calculations were performed on both the FeMo cluster and the P-cluster, and two different sizes of the QM systems were employed, one smaller for QM/MM geometry optimisations and one larger for the redox-potential calculations.

In the QM/MM geometry optimisations, the P-cluster was modelled as $Fe_8S_7Cys_6$, with the five of the Cys ligands modelled by CH_3S^-, whereas Cys–88 was modelled by $CH_3CONHCH_2CH_2S^-$, because the backbone amide group is deprotonated and coordinates to Fe5 in some the more oxidised states. Likewise, Ser-188D (i.e., belonging to subunit D, rather than C for all the other numbered residues) was included in the model as CH_3OH because it is deprototated and coordinates to Fe6 in the oxidised states. The model contained 64 atoms for the fully protonated state and it is shown in Figure S1a in the Supporting Information. All QM/MM structures of the P-cluster were taken from our previous study [55].

The FeMo cluster was modelled by $MoFe_7S_9C$(homocitrate)(CH_3S) (imidazole) in the QM/MM calculations, where the two last groups are models of Cys-275 and His-442. In addition, all groups that form hydrogen bonds to the FeMo cluster were also included in the QM model, viz. Arg-96, Gln-191 and His-195 (sidechains), Ser-278 and Arg-359 (both backbone and sidechain, including the CA and C and O atoms from Arg-277), Gly-356, Gly-357 and Leu-358 (backbone, including the CA and C and O atoms from Ile-355), as well as two water molecules. Finally, the sidechains of Glu-380, Val-70 and Phe-381 were also included. This QM system involved 191 atoms and is shown in Figure S1b.

For the redox-potential calculations, we used the largest QM system suggested in our previous investigation [41]. It included all functional groups in the proteins with any atom within 3.5 Å of a minimal QM system, consisting of all metal and sulfide ions together with all direct ligands (Cys, His and homocitrate). These QM systems were set up using our local program for BigQM calculations (changepdb) [83]. They contained ~400 atoms for the FeMo cluster and ~500 atoms for the P-cluster, and are shown in Figure S2. The calculations were based on the QM/MM-optimised structures.

In the redox calculations, the QM system was immersed into a continuum solvent, employing the conductor-like screening model (COSMO) [84,85] implemented in Turbomole. The default optimised COSMO atomic radii and a water solvent radius of 1.3 Å were employed to construct the solvent-accessible surface cavity [86], whereas a radius of 2.0 Å was used for Fe and Mo [87]. Structures for the QM + COSMO calculations were taken directly from the QM/MM calculations without further optimisation. The dielectric constant was 80, which gave the best redox potentials in our previous study [41].

Redox potentials (E^o) were calculated according to

$$E^o = E(\text{ox}) - E(\text{red}) - c \qquad (2)$$

where $E(\text{ox})$ and $E(\text{red})$ are the energies of the oxidised and reduced states, and c is a correction factor (4.28 eV) to place the potentials on the scale of the standard hydrogen electrode [88]. The actual value of this factor has been much discussed and values between 4.05–4.44 eV have been suggested [88,89]. To avoid this problem, we use the method calibrated in our previous study to 13 different iron–sulfur clusters [41]. Therefore, we subtract from the c constant the mean signed error (MSE) obtained in this calibration study,

MSE = −0.62 V or we subtract the calculated potential of the $E_0 \rightarrow E_1$ transition from the redox potentials of the other states.

In some cases, it is known or assumed that the cluster takes up a proton during or after the electron transfer. In those cases, it was assumed that the proton comes from an imidazole molecule, studied with the same QM method in a COSMO continuum solvation solvent with the dielectric constant of water (80). Thus, we assume that the proton comes from a group with a pK_a close to 7 (6.95). This changes the calculated redox potentials by −12.47 V.

The electronic structure of all QM systems was obtained with the broken-symmetry (BS) approach [90]: Each of the seven or eight Fe ions was modelled in the high-spin state, with either a surplus of α or β spin. Such a state can be selected in many different ways, giving rise to different BS states, which are specified by giving the number of the Fe ions with minority spin (the numbering of the Fe ions is shown in Figure 1) [55,60]. The various BS states were obtained either by swapping the coordinates of the Fe ions [91] or with the fragment approach by Szilagyi and Winslow [92].

3.3. QM/MM Calculations

QM/MM calculations were performed with the COMQUM software [93,94]. In this approach, the protein and solvent are split into two subsystems: System 1 (the QM region) was relaxed by QM methods. System 2 contained the remaining part of the protein and the solvent, and it was kept fixed at the original coordinates (equilibrated crystal structure [70], to avoid the risk that different calculations end up in different local minima).

In the QM calculations, system 1 was represented by a wavefunction, whereas all the other atoms were represented by an array of partial point charges, one for each atom, taken from the MM setup. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: The QM system was capped with hydrogen atoms, the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system [93,95]. All atoms were included in the point-charge model, except the CL atoms [96]. ComQum employs a subtractive scheme with van der Waals link-atom corrections [97]. No cut-off is used for the QM and QM–MM interactions. The geometry optimisations were continued until the energy change between two iterations was less than 2.6 J/mol (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u.

4. Conclusions

In this study, we have investigated what information we can get from calculated redox potentials of the two metal clusters in Mo-nitrogenase. We employ our calibrated approach to calculate redox potentials for iron–sulfur clusters involving QM-cluster calculations in a continuum solvent with large QM models (400–500 atoms), based on structures from QM/MM optimisations [41]. We obtain several interesting results:

The calculations on the P-cluster show that our method gives approximately the same accuracy as for the simple iron–sulfur clusters with 1–4 Fe ions, with a maximum error of 0.33 V (0.44 V in our previous study [41]). This shows that the calculations are accurate also for the larger P-cluster and for redox reactions that include proton transfers.

The calculations confirm that the $P^{1+} \rightarrow P^{2+}$ transition involves a proton transfer (i.e., $P^{1+}H \rightarrow P^{2+}$), as is also suggested by electrochemical and crystallographic studies [8,48,50,51,54–56].

The calculations show that the $P^{1+}H_2/P^N H_2$ and $P^{1+}H/P^N H_2$ redox couples give very similar redox potentials, which may explain the experimental enigma that redox titrations do not show any pH dependence of the P^{1+}/P^N redox potential [52,53], although crystal structures indicate that also the $P^N \rightarrow P^{1+}$ transition should involve a proton transfer [54–56].

For the FeMo-cluster, the calculations unambiguously identify $Mo^{III}Fe_3^{II}Fe_4^{III}$ as the proper formal oxidation state for the resting E_0 state of the protein. This provides an independent confirmation of this oxidation state, also suggested by previous experimental and QM studies [13,42,57,58].

The calculations agree with experiments only if it is assumed that a proton is taken up together with the electron in the $E_0 \rightarrow E_1$ reaction.

The calculations confirm that the E_0 to E_4H_4 states all have similar redox potentials (within 0.41 V, i.e., lower than the estimated maximum error of the method), as expected for a site that should receive electrons from the same donor.

However, there is no major difference in the redox potentials between structures with protons on the μ_2-bridging sulfide ions or hydride ions on the Fe ions (for E_2H_2 the difference is only 0.08 V). This shows that there is only minor differences between hydrogen atoms bound to S or Fe ions, as previously has been advocated by Dance [18] and that the formal oxidation states of the Fe ions are no good indicators of the redox potentials.

The redox potentials of the later steps of the reaction mechanism ($E_4N_2H_2$ to E_8) are more positive than that of the resting state (E_0/E_1H), showing that the reactions are more exothermic. The trends in the potentials do not change if S2B remains bound to the cluster or if it dissociates. This shows that there is no need of a significantly more negative potential for the Fe protein than measured.

The calculated redox potentials strongly depend on the structures used (the positions of the added protons), the broken-symmetry states and the QM method employed.

Thus, our calculations show that quite strong predictions can be provided by redox-potential calculations, even if the accuracy is rather poor compared to experimental measurements (a maximum error of 0.44 V), provided that the calculations are calibrated so that the expected errors are known. Such calculations can also identify possible problems in suggested reaction mechanisms. Our calculations also allow us to identify what redox reactions involve coupled electron and proton transfer, which is crucial to identify the detailed reaction mechanism of nitrogenase. Moreover, we have been able to explain the enigma why the P^{1+}/P^N redox potential is not pH dependent [52,53], although the crystal structures show that the a deprotonation reaction is involved [54–56].

References

1. Smith, B.E. Nitrogenase Reveals Its Inner Secrets. *Science* **2002**, *297*, 1654–1655. [CrossRef] [PubMed]
2. Hoffman, B.M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D.R.; Seefeldt, L.C. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem. Rev.* **2014**, *114*, 4041–4062. [CrossRef]
3. Seefeldt, L.C.; Yang, Z.-Y.; Lukoyanov, D.A.; Harris, D.F.; Dean, D.R.; Raugei, S.; Hoffman, B.M. Reduction of Substrates by Nitrogenases. *Chem. Rev.* **2020**, *120*, 5082–5106. [CrossRef] [PubMed]
4. Burgess, B.K.; Lowe, D.J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96*, 2983–3012. [CrossRef] [PubMed]
5. Schmid, B.; Chiu, H.-J.; Ramakrishnan, V.; Howard, J.B.; Rees, D.C. Nitrogenase. In *Handbook of Metalloproteins*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2006; pp. 1025–1036. ISBN 9780470028636.
6. Kim, J.; Rees, D.C. Structural Models for the Metal Centers in the Nitrogenase Molybdenum-Iron Protein. *Science* **1992**, *257*, 1677–1682. [CrossRef] [PubMed]
7. Einsle, O.; Tezcan, F.A.; Andrade, S.L.A.; Schmid, B.; Yoshida, M.; Howard, J.B.; Rees, D.C. Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science* **2002**, *297*, 1696. [CrossRef]
8. Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S.L.A.; Schleicher, E.; Weber, S.; Rees, D.C.; Einsle, O. Evidence for Interstitial Carbon in Nitrogenase FeMo Cofactor. *Science* **2011**, *334*, 940. [CrossRef]
9. Spatzal, T.; Perez, K.A.; Einsle, O.; Howard, J.B.; Rees, D.C. Ligand Binding to the FeMo-Cofactor: Structures of CO-Bound and Reactivated Nitrogenase. *Science* **2014**, *345*, 1620–1623. [CrossRef]
10. Einsle, O. Nitrogenase FeMo Cofactor: An Atomic Structure in Three Simple Steps. *J. Biol. Inorg. Chem.* **2014**, *19*, 737–745. [CrossRef]
11. Eady, R.R. Structure−Function Relationships of Alternative Nitrogenases. *Chem. Rev.* **1996**, *96*, 3013–3030. [CrossRef]
12. Lancaster, K.M.; Roemelt, M.; Ettenhuber, P.; Hu, Y.; Ribbe, M.W.; Neese, F.; Bergmann, U.; DeBeer, S. X-Ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor. *Science* **2011**, *334*, 974–977. [CrossRef] [PubMed]
13. Bjornsson, R.; Lima, F.A.; Spatzal, T.; Weyhermüller, T.; Glatzel, P.; Bill, E.; Einsle, O.; Neese, F.; DeBeer, S. Identification of a Spin-Coupled Mo(III) in the Nitrogenase Iron–Molybdenum Cofactor. *Chem. Sci.* **2014**, *5*, 3096–3103. [CrossRef]
14. Igarashi, R.Y.; Laryukhin, M.; Dos Santos, P.C.; Lee, H.-I.; Dean, D.R.; Seefeldt, L.C.; Hoffman, B.M. Trapping H⁻ Bound to the Nitrogenase FeMo-Cofactor Active Site during H_2 Evolution: Characterization by ENDOR Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127*, 6231–6241. [CrossRef] [PubMed]
15. Hoeke, V.; Tociu, L.; Case, D.A.; Seefeldt, L.C.; Raugei, S.; Hoffman, B.M. High-Resolution ENDOR Spectroscopy Combined with Quantum Chemical Calculations Reveals the Structure of Nitrogenase Janus Intermediate E_4(4H). *J. Am. Chem. Soc.* **2019**, *141*, 11984–11996. [CrossRef]
16. Roth, L.E.; Tezcan, F.A. X-ray Crystallography. *Methods Mol. Biol.* **2011**, *766*, 147–164. [CrossRef]
17. Van Stappen, C.; Decamps, L.; Cutsail, G.E.; Bjornsson, R.; Henthorn, J.T.; Birrell, J.A.; DeBeer, S. The Spectroscopy of Nitrogenases. *Chem. Rev.* **2020**, *120*, 5005–5081. [CrossRef] [PubMed]
18. Dance, I. Computational Investigations of the Chemical Mechanism of the Enzyme Nitrogenase. *ChemBioChem* **2020**, *21*, 1671–1709. [CrossRef]
19. Kästner, J.; Blöchl, P.E. Ammonia Production at the FeMo Cofactor of Nitrogenase: Results from Density Functional Theory. *J. Am. Chem. Soc.* **2007**, *129*, 2998–3006. [CrossRef]
20. Varley, J.B.; Wang, Y.; Chan, K.; Studt, F.; Nørskov, J.K. Mechanistic Insights into Nitrogen Fixation by Nitrogenase Enzymes. *Phys. Chem. Chem. Phys.* **2015**, *17*, 29541–29547. [CrossRef]
21. Raugei, S.; Seefeldt, L.C.; Hoffman, B.M. Critical Computational Analysis Illuminates the Reductive-Elimination Mechanism That Activates Nitrogenase for N_2 Reduction. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 10521–10530. [CrossRef]
22. Siegbahn, P.E.M. The Mechanism for Nitrogenase Including All Steps. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15747–15759. [CrossRef] [PubMed]
23. Thorneley, R.N.F.; Lowe, D.J. Kinetics and Mechanism of the Nitrogenase Enzyme System. In *Molybdenum Enzymes*; Spiro, T.G., Ed.; Wiley: New York, NY, USA, 1985; pp. 221–284.
24. Rutledge, H.L.; Tezcan, F.A. Electron Transfer in Nitrogenase. *Chem. Rev.* **2020**, *120*, 5158–5193. [CrossRef] [PubMed]
25. Igarashi, R.Y.; Seefeldt, L.C. Nitrogen Fixation: The Mechanism of the Mo-Dependent Nitrogenase. *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 351–384. [CrossRef] [PubMed]
26. O'Donnell, M.J.; Smith, B.E. Electron-Paramagnetic-Resonance Studies on the Redox Properties of the Molybdenum-Iron Protein of Nitrogenase between +50 and −450 MV. *Biochem. J.* **1978**, *173*, 831–838. [CrossRef] [PubMed]
27. Watt, G.D.; Burns, A.; Lough, S.; Tennent, D.L. Redox and Spectroscopic Properties of Oxidized MoFe Protein from Azotobacter Vinelandii. *Biochemistry* **1980**, *19*, 4926–4932. [CrossRef]
28. Watt, G.D.; Bulen, W.A. Redox Potentials of FeMo Cofactor in Nitrogenase. In Proceedings of the First International Symposium on Nitrogen Fixation; Newton, W.E., Nymans, C.J., Eds.; Washington State University Press: Washington, DC, USA, 1976; Volume 1, p. 123.
29. Lough, S.; Burns, A.; Watt, G.D. Redox Reactions of the Nitrogenase Complex from Azotobacter Vinelandii. *Biochemistry* **1983**, *22*, 4062–4066. [CrossRef]
30. Torres, R.A.; Lovell, T.; Noodleman, L.; Case, D.A. Density Functional and Reduction Potential Calculations of Fe4S4 Clusters. *J. Am. Chem. Soc.* **2003**, *125*, 1923–1936. [CrossRef]

31. Ullmann, M.G.; Noodleman, L.; Case, D.A. Density Functional Calculation of PKa Values and Redox Potentials in the Bovine Rieske Iron-Sulfur Protein. *JBIC J. Biol. Inorg. Chem.* **2002**, *7*, 632–639. [CrossRef]

32. Mouesca, J.-M.; Chen, J.L.; Noodleman, L.; Bashford, D.; Case, D.A. Density Functional/Poisson-Boltzmann Calculations of Redox Potentials for Iron-Sulfur Clusters. *J. Am. Chem. Soc.* **1994**, *116*, 11898–11914. [CrossRef]

33. Noodleman, L.; Lovell, T.; Liu, T.; Himo, F.; Torres, R.A. Insights into Properties and Energetics of Iron–Sulfur Proteins from Simple Clusters to Nitrogenase. *Curr. Opin. Chem. Biol.* **2002**, *6*, 259–273. [CrossRef]

34. Noodleman, L.; Han, W.-G. Structure, Redox, PKa, Spin. A Golden Tetrad for Understanding Metalloenzyme Energetics and Reaction Pathways. *J. Biol. Inorg. Chem.* **2006**, *11*, 674–694. [CrossRef]

35. Kim, C.-H.; Newton, W.E.; Dean, D.R. Role of the MoFe Protein.Alpha.-Subunit Histidine-195 Residue in FeMo-Cofactor Binding and Nitrogenase Catalysis. *Biochemistry* **1995**, *34*, 2798–2808. [CrossRef] [PubMed]

36. Perrin, B.S., Jr.; Niu, S.; Ichiye, T. Calculating Standard Reduction Potentials of [4Fe–4S] Proteins. *J. Comput. Chem.* **2013**, *34*, 576–582. [CrossRef] [PubMed]

37. Perrin, B.S., Jr.; Miller, B.T.; Schalk, V.; Woodcock, H.L.; Brooks, B.R.; Ichiye, T. Web-Based Computational Chemistry Education with CHARMMing III: Reduction Potentials of Electron Transfer Proteins. *PLOS Comput. Biol.* **2014**, *10*, e1003739. [CrossRef]

38. Stephens, P.J.; Jollie, D.R.; Warshel, A. Protein Control of Redox Potentials of Iron−Sulfur Proteins. *Chem. Rev.* **1996**, *96*, 2491–2514. [CrossRef]

39. Lovell, T.; Liu, T.; Case, D.A.; Noodleman, L. Structural, Spectroscopic, and Redox Consequences of a Central Ligand in the FeMoco of Nitrogenase: A Density Functional Theoretical Study. *J. Am. Chem. Soc.* **2003**, *125*, 8377–8383. [CrossRef]

40. Cheng, J.; Liu, X.; VandeVondele, J.; Sulpizi, M.; Sprik, M. Redox Potentials and Acidity Constants from Density Functional Theory Based Molecular Dynamics. *Acc. Chem. Res.* **2014**, *47*, 3522–3529. [CrossRef]

41. Jafari, S.; Tavares Santos, Y.A.; Bergmann, J.; Irani, M.; Ryde, U. Benchmark Study of Redox Potentials Calculation for Iron–Sulfur Clusters in Proteins. *Inorg. Chem.* **2022**, *61*, 5991–6007. [CrossRef]

42. Spatzal, T.; Schlesier, J.; Burger, E.-M.; Sippel, D.; Zhang, L.; Andrade, S.L.A.; Rees, D.C.; Einsle, O. Nitrogenase FeMoco Investigated by Spatially Resolved Anomalous Dispersion Refinement. *Nat. Commun.* **2016**, *7*, 10902. [CrossRef]

43. Doan, P.E.; Telser, J.; Barney, B.M.; Igarashi, R.Y.; Dean, D.R.; Seefeldt, L.C.; Hoffman, B.M. ^{57}Fe ENDOR Spectroscopy and 'Electron Inventory' Analysis of the Nitrogenase E$_4$ Intermediate Suggest the Metal-Ion Core of FeMo-Cofactor Cycles Through Only One Redox Couple. *J. Am. Chem. Soc.* **2011**, *133*, 17329–17340. [CrossRef]

44. Yang, Z.-Y.; Danyal, K.; Seefeldt, L.C. Nitrogen Fixation. *Methods Mol. Biol.* **2011**, *766*, 267–292.

45. Pierik, A.J.; Wassink, H.; Haaker, H.; Hagen, W.R. Redox Properties and EPR Spectroscopy of the P Clusters of Azotobacter Vinelandii MoFe Protein. *Eur. J. Biochem.* **1993**, *212*, 51–61. [CrossRef] [PubMed]

46. Rupnik, K.; Hu, Y.; Lee, C.C.; Wiig, J.A.; Ribbe, M.W.; Hales, B.J. P$^+$ State of Nitrogenase P-Cluster Exhibits Electronic Structure of a [Fe$_4$S$_4$]+ Cluster. *J. Am. Chem. Soc.* **2012**, *134*, 13749–13754. [CrossRef]

47. Katz, F.E.H.; Owens, C.P.; Tezcan, F.A. Electron Transfer Reactions in Biological Nitrogen Fixation. *Isr. J. Chem.* **2016**, *56*, 682–692. [CrossRef]

48. Owens, C.P.; Katz, F.E.H.; Carter, C.H.; Oswald, V.F.; Tezcan, F.A. Tyrosine-Coordinated P-Cluster in G. Diazotrophicus Nitrogenase: Evidence for the Importance of O-Based Ligands in Conformationally Gated Electron Transfer. *J. Am. Chem. Soc.* **2016**, *138*, 10124–10127. [CrossRef] [PubMed]

49. Seefeldt, L.C.; Hoffman, B.M.; Dean, D.R. Electron Transfer in Nitrogenase Catalysis. *Curr. Opin. Chem. Biol.* **2012**, *16*, 19–25. [CrossRef]

50. Peters, J.W.; Stowell, M.H.B.; Soltis, S.M.; Finnegan, M.G.; Johnson, M.K.; Rees, D.C. Redox-Dependent Structural Changes in the Nitrogenase P-Cluster. *Biochemistry* **1997**, *36*, 1181–1187. [CrossRef]

51. Rupnik, K.; Lee, C.C.; Wiig, J.A.; Hu, Y.; Ribbe, M.W.; Hales, B.J. Nonenzymatic Synthesis of the P-Cluster in the Nitrogenase MoFe Protein: Evidence of the Involvement of All-Ferrous [Fe4S4]0 Intermediates. *Biochemistry* **2014**, *53*, 1108–1116. [CrossRef]

52. Lanzilotta, W.N.; Christiansen, J.; Dean, D.R.; Seefeldt, L.C. Evidence for Coupled Electron and Proton Transfer in the [8Fe-7S] Cluster of Nitrogenase. *Biochemistry* **1998**, *37*, 11376–11384. [CrossRef]

53. Chan, J.M.; Christiansen, J.; Dean, D.R.; Seefeldt, L.C. Spectroscopic Evidence for Changes in the Redox State of the Nitrogenase P-Cluster during Turnover. *Biochemistry* **1999**, *38*, 5779–5785. [CrossRef]

54. Keable, S.M.; Zadvornyy, O.A.; Johnson, L.E.; Ginovska, B.; Rasmussen, A.J.; Danyal, K.; Eilers, B.J.; Prussia, G.A.; LeVan, A.X.; Raugei, S.; et al. Structural Characterization of the P1+ Intermediate State of the P-Cluster of Nitrogenase. *J. Biol. Chem.* **2018**, *293*, 9629–9635. [CrossRef] [PubMed]

55. Cao, L.; Börner, M.C.; Bergmann, J.; Caldararu, O.; Ryde, U. Geometry and Electronic Structure of the P-Cluster in Nitrogenase Studied by Combined Quantum Mechanical and Molecular Mechanical Calculations and Quantum Refinement. *Inorg. Chem.* **2019**, *58*, 9672–9690. [CrossRef] [PubMed]

56. Cao, L.; Ryde, U. Quantum Refinement with Multiple Conformations: Application to the P-Cluster in Nitrogenase. *Acta Crystallogr. Sect. D* **2020**, *76*, 1145–1156. [CrossRef] [PubMed]

57. Bjornsson, R.; Neese, F.; DeBeer, S. Revisiting the Mössbauer Isomer Shifts of the FeMoco Cluster of Nitrogenase and the Cofactor Charge. *Inorg. Chem.* **2017**, *56*, 1470–1477. [CrossRef] [PubMed]

58. Siegbahn, P.E.M. Model Calculations Suggest That the Central Carbon in the FeMo-Cofactor of Nitrogenase Becomes Protonated in the Process of Nitrogen Fixation. *J. Am. Chem. Soc.* **2016**, *138*, 10485–10495. [CrossRef]

59. Cao, L.; Caldararu, O.; Ryde, U. Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations. *J. Chem. Theory Comput.* **2018**, *14*, 6653–6678. [CrossRef]
60. Cao, L.; Ryde, U. Influence of the Protein and DFT Method on the Broken-Symmetry and Spin States in Nitrogenase. *Int. J. Quantum Chem.* **2018**, *118*, e25627. [CrossRef]
61. Van Stappen, C.; Thorhallsson, A.T.; Decamps, L.; Bjornsson, R.; DeBeer, S. Resolving the Structure of the E_1 State of Mo Nitrogenase through Mo and Fe K-Edge EXAFS and QM/MM Calculations. *Chem. Sci.* **2019**, *10*, 9807–9821. [CrossRef]
62. Van Stappen, C.; Davydov, R.; Yang, Z.-Y.; Fan, R.; Guo, Y.; Bill, E.; Seefeldt, L.C.; Hoffman, B.M.; DeBeer, S. Spectroscopic Description of the E_1 State of Mo Nitrogenase Based on Mo and Fe X-Ray Absorption and Mössbauer Studies. *Inorg. Chem.* **2019**, *58*, 12365–12376. [CrossRef]
63. Cao, L.; Ryde, U. What Is the Structure of the E_4 Intermediate in Nitrogenase? *J. Chem. Theory Comput.* **2020**, *16*, 1936–1952. [CrossRef]
64. Jiang, H.; Svensson, O.K.G.; Ryde, U. QM/MM Study of Partial Dissociation of S2B for the E_2 Intermediate of Nitrogenase. *Inorg. Chem.* **2022**, *61*, 18067–18076. [CrossRef] [PubMed]
65. Cao, L.; Ryde, U. Putative Reaction Mechanism of Nitrogenase after Dissociation of a Sulfide Ligand. *J. Catal.* **2020**, *391*, 247–259. [CrossRef]
66. Jiang, H.; Ryde, U. Thermodynamically Favourable States in the Reaction of Nitrogenase without Dissociation of Any Sulfide Ligand. *Chem.-Eur. J.* **2022**, *28*, e202103933. [CrossRef] [PubMed]
67. Jiang, H.; Svensson, O.K.G.; Cao, L.; Ryde, U. Proton Transfer Pathways in Nitrogenase with and without Dissociated S2B. *Angew. Chemie Int. Ed.* **2022**, *61*, e202208544. [CrossRef] [PubMed]
68. Rohde, M.; Sippel, D.; Trncik, C.; Andrade, S.L.A.; Einsle, O. The Critical E_4 State of Nitrogenase Catalysis. *Biochemistry* **2018**, *57*, 5497–5504. [CrossRef]
69. Cao, L.; Ryde, U. Extremely Large Differences in DFT Energies for Nitrogenase Models. *Phys. Chem. Chem. Phys.* **2019**, *21*, 2480–2488. [CrossRef]
70. Cao, L.; Caldararu, O.; Ryde, U. Protonation States of Homocitrate and Nearby Residues in Nitrogenase Studied by Computational Methods and Quantum Refinement. *J. Phys. Chem. B* **2017**, *121*, 8242–8262. [CrossRef]
71. Barney, B.M.; McClead, J.; Lukoyanov, D.; Laryukhin, M.; Yang, T.C.; Dean, D.R.; Hoffman, B.M.; Seefeldt, L.C. Diazene (HN=NH) Is a Substrate for Nitrogenase: Insights into the Pathway of N2 Reduction. *Biochemistry* **2007**, *46*, 6784–6794. [CrossRef]
72. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef]
73. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]
74. Cao, L.; Ryde, U. N_2H_2 Binding to the Nitrogenase FeMo Cluster, Studied by QM/MM Methods. *J. Biol. Inorg. Chem.* **2020**, *25*, 521–540. [CrossRef] [PubMed]
75. Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7*, 2452–2463. [CrossRef] [PubMed]
76. Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280. [CrossRef]
77. Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdisc. Rev. Comput. Mol. Sci.* **2014**, *4*, 91–100. [CrossRef]
78. Tao, J.; Perdew, J.P.; Staroverov, V.N.; Scuseria, G.E. Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401. [CrossRef]
79. Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577. [CrossRef]
80. Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240*, 283–289. [CrossRef]
81. Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124. [CrossRef]
82. Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150*, 154122. [CrossRef]
83. Hu, L.; Söderhjelm, P.; Ryde, U. Accurate Reaction Energies in Proteins Obtained by Combining QM/MM and Large QM Calculations. *J. Chem. Theory Comput.* **2013**, *9*, 640–649. [CrossRef]
84. Klamt, A.; Schüürmann, G. Cosmo—A New Approach To Dielectric Screening in Solvents With Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc.-Perkin Trans. 2* **1993**, *1993*, 799–805. [CrossRef]
85. Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J.C.W.; Eckert, F. COSMO Implementation in TURBOMOLE: Extension of an Efficient Quantum Chemical Code towards Liquid Systems. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187–2193. [CrossRef]
86. Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J.C.W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074–5085. [CrossRef]
87. Sigfridsson, E.; Ryde, U. Comparison of Methods for Deriving Atomic Charges from the Electrostatic Potential and Moments. *J. Comput. Chem.* **1998**, *19*, 377–395. [CrossRef]

88.   Kelly, C.P.; Cramer, C.J.; Truhlar, D.G. Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006**, *110*, 16066–16081. [CrossRef] [PubMed]

89.   Cheng, C.; Hayashi, S. Ab Initio Evaluation of the Redox Potential of Cytochrome C. *J. Chem. Theory Comput.* **2021**, *17*, 1194–1207. [CrossRef]

90.   Lovell, T.; Li, J.; Liu, T.; Case, D.A.; Noodleman, L. FeMo Cofactor of Nitrogenase: A Density Functional Study of States M^{N}, M^{OX}, M^{R}, and M^{I}. *J. Am. Chem. Soc.* **2001**, *123*, 12392–12410. [CrossRef]

91.   Greco, C.; Fantucci, P.; Ryde, U.; de Gioia, L. Fast Generation of Broken-Symmetry States in a Large System Including Multiple Iron–Sulfur Assemblies: Investigation of QM/MM Energies, Clusters Charges, and Spin Populations. *Int. J. Quantum Chem.* **2011**, *111*, 3949–3960. [CrossRef]

92.   Szilagyi, R.K.; Winslow, M.A. On the Accuracy of Density Functional Theory for Iron-Sulfur Clusters. *J. Comput. Chem.* **2006**, *27*, 1385–1397. [CrossRef]

93.   Ryde, U. The Coordination of the Catalytic Zinc in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput. Aided. Mol. Des.* **1996**, *10*, 153–164. [CrossRef]

94.   Ryde, U.; Olsson, M.H.M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.* **2001**, *81*, 335–347. [CrossRef]

95.   Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104*, 1720–1735. [CrossRef]

96.   Hu, L.; Söderhjelm, P.; Ryde, U. On the Convergence of QM/MM Energies. *J. Chem. Theory Comput.* **2011**, *7*, 761–777. [CrossRef] [PubMed]

97.   Cao, L.; Ryde, U. On the Difference between Additive and Subtractive QM/MM Calculations. *Front. Chem.* **2018**, *6*, 89. [CrossRef] [PubMed]

Paper VIII

Article

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

Hao Jiang, Kristoffer J. M. Lundgren, and Ulf Ryde*

Cite This: Inorg. Chem. 2023, 62, 19433−19445

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Nitrogenase is the only enzyme that can cleave the strong triple bond in N_2, making nitrogen available for biological life. There are three isozymes of nitrogenase, differing in the composition of the active site, viz., Mo, V, and Fe-nitrogenase. Recently, the first crystal structure of Fe-nitrogenase was presented. We have performed the first combined quantum mechanical and molecular mechanical (QM/MM) study of Fe-nitrogenase. We show with QM/MM and quantum-refinement calculations that the homocitrate ligand is most likely protonated on the alcohol oxygen in the resting E_0 state. The most stable broken-symmetry (BS) states are the same as for Mo-nitrogenase, i.e., the three Noodleman BS7-type states (with a surplus of β spin on the eighth Fe ion), which maximize the number of nearby antiferromagnetically coupled Fe−Fe pairs. For the E_1 state, we find that protonation of the S2B μ_2 belt sulfide ion is most favorable, 14−117 kJ/mol more stable than structures with a Fe-bound hydride ion (the best has a hydride ion on the Fe2 ion) calculated with four different density-functional theory methods. This is similar to what was found for Mo-nitrogenase, but it does not explain the recent EPR observation that the E_1 state of Fe-nitrogenase should contain a photolyzable hydride ion. For the E_1 state, many BS states are close in energy, and the preferred BS state differs depending on the position of the extra proton and which density functional is used.
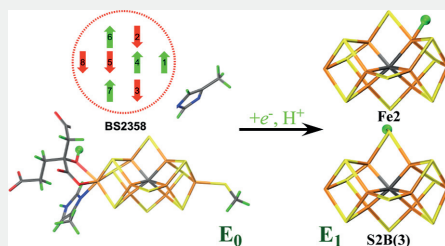
■ INTRODUCTION

Nitrogen is crucial in 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] The reason is that plants cannot metabolize N_2 because it involves a strong and inert triple bond. The industrial conversion of nitrogen to ammonia occurs through the energy-intensive Haber−Bosch process, which involves high temperatures and pressures and accounts for almost 2% of the world's total energy consumption.[2]

Nitrogenase (EC 1.18/19.6.1) is the only enzyme that can cleave the N−N bond in N_2 and convert it to ammonia. It functions under ambient temperature and pressure. Nitrogenase exists in three forms: Mo-nitrogenase, V-nitrogenase, and Fe-only nitrogenase. Mo-nitrogenase is the most prevalent form, with the highest N_2-reducing activity.[1,3−9]

Crystal structures of Mo-nitrogenase have been known since 1992[10,11] and of V-nitrogenase since 2017.[12] However, the first crystal structure of Fe-only nitrogenase was published this year,[13] and a cryogenic electron microscopy structure has also been presented.[14] The studies have shown that all nitrogenases involve two proteins: the Fe protein and the Mo/V/FeFe protein. Electrons are supplied by the Fe protein, which also binds two ATP molecules. This binding triggers docking to the other protein and facilitates electron transfer. Hydrolysis of the

ATP molecules induces the dissociation of Fe protein, thereby enabling further electron transfers. The MoFe protein is a $\alpha_2\beta_2$ heterotetramer, whereas the VFe and FeFe proteins are $\alpha_2\beta_2\gamma_2$ heterohexamers, with one extra subunit. These proteins contain an $Fe_8S_7Cys_6$ cluster called the P-cluster, which is used for electron transfer. In addition, they contain the active site, which is slightly different for the three types of nitrogenases. 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), whereas Fe-only nitrogenase contains a Fe_8S_9C(homocitrate) cluster (the FeFe cluster, as shown in Figure 1a). In all three cases, the active-site cluster is coordinated to the protein via a cysteine and a histidine residue.[10,11,15−17]

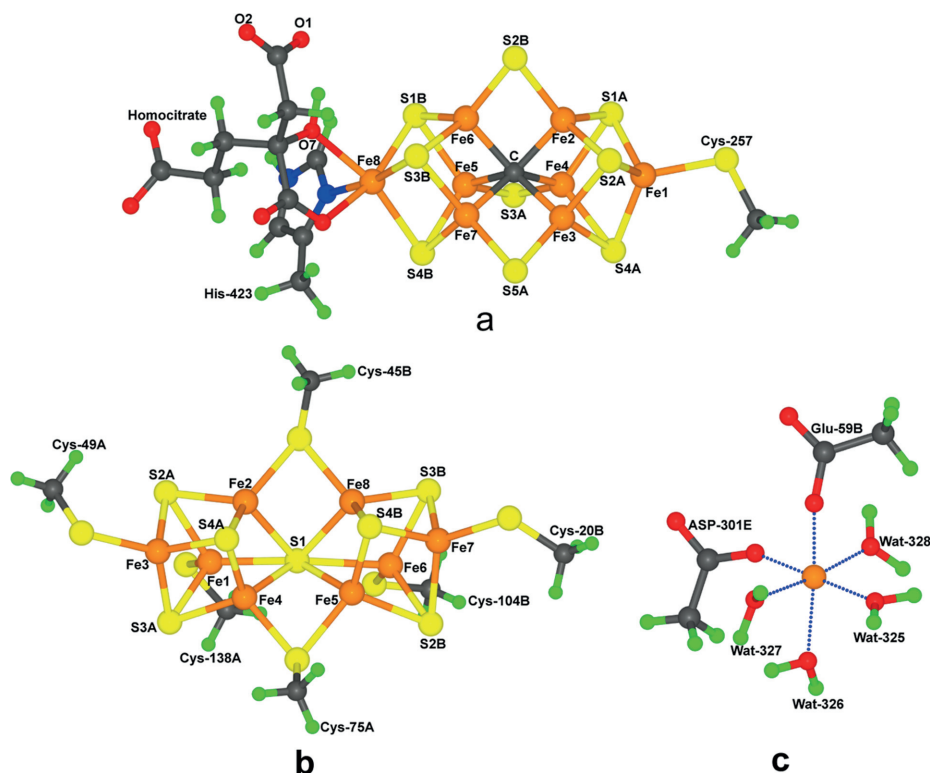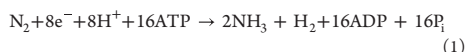The nitrogenases catalyze the reaction

Figure 1. (a) FeFe cluster, (b) P cluster, and (c) Mg site in nitrogenase, also showing the atom names and the models used to calculate charges for the MM force field.

$$N_2 + 8e^- + 8H^+ + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \tag{1}$$

although all three enzymes under normal N_2 pressure produce more H_2 and therefore consume more electrons, protons, and ATP molecules. The mechanism is commonly described by the Lowe–Thorneley scheme, which involves nine intermediates denoted as E_0 to E_8. These intermediates differ in the number of electrons and protons accumulated. Thorough biochemical, kinetic, and spectroscopic studies have demonstrated the necessity of reducing the E_0 state to the E_4 state before nitrogen binding can occur.[1,4−7,9,18]

It is generally believed that the three types of nitrogenases follow similar reaction mechanisms.[19,20] However, recently, an EPR study of the one-electron reduced E_1 state in Fe-nitrogenase (the E_1 state is EPR active in this enzyme, in contrast to Mo-nitrogenase) suggested that it contains a Fe-bound hydride ion (based on the fact that the ligand is photolyzable) rather than a sulfur-bound proton.[21] This is in contrast to Mo-nitrogenase, for which EXAFS measurements and combined quantum mechanical (QM) and molecular mechanical (MM) calculations have indicated that the E_1 intermediate involves a protonated μ_2 belt sulfide, probably S2B (atom names are shown in Figure 1a).[22] This also agrees with previous QM and QM/MM studies, pointing out S2B as

the energetically most favorable protonation site in the E_1 state.[23,24]

Therefore, it is of great interest to examine whether there is an intrinsic difference in the protonation preferences of Mo- and Fe-nitrogenase in the E_1 state. The recent crystal structure of Fe-nitrogenase makes such an investigation possible. In this article, we set up the first QM/MM calculations of Fe-nitrogenase, determining the proper protonation states of homocitrate and His-180, as well as the broken-symmetry (BS) state of the resting E_0 state. Then, we evaluated the protonation preferences of the E_1 state and discussed the implications of the findings.

■ METHODS

Protein. All calculations were based on the recent crystal structure of Fe-only nitrogenase from *Azotobacter vinelandii* (PDB code 8BOQ), with a resolution of 1.55 Å.[13] The calculations encompassed the entire $\alpha_2\beta_2\gamma_2$ heterohexamer as the subunits are intertwined. Likewise, two Mg^{2+} ions were retained because they are deeply buried in the protein, stabilizing the subunit interface. All crystal-water molecules were also kept, except eight that overlapped with each other or with protein atoms: HOH-223, 878, 976, 1086, 1122, 1173, 1255, and 1268.

The crystal structure is a mixture of the resting state (E_0) and a turnover state in which the S2B ion is replaced by a light atom, modeled by O in the PDB file.[13] The two states have almost equal

occupancy (0.5/0.5 in subunit A and 0.4/0.6 in subunit D). Gln-176 also shows two conformations. In one (connected to the resting state), it points away from the FeFe cluster. In the other, it forms a hydrogen bond to His-180 (2.8 Å) and to the light atom, replacing S2B (2.5 Å). The OE1 atom in the resting-state conformation of Gln-176 is replaced by a stronger density, which is interpreted as the storage position of the replaced S2B. In our QM/MM calculations, we studied only the resting-state conformation so the extra O atom was deleted as well as the corresponding conformation of Gln-176 and the storage conformation of S2B.

The protonation states of all of the residues were determined through a thorough analysis of the hydrogen-bond pattern and the solvent accessibility. It was checked by the PROPKA[25] and Maestro[26] software. All Arg, Lys, Asp, and Glu residues were assumed to be charged, with the exceptions of Asp-78, Glu-62B, 245B, Lys-22, 55, 83, 339, 361, and 398B (residues without any letter after the residue number belong to subunit A, whereas those belonging to subunits B or C end with that letter; subunits D, E, and F were treated identically to subunits A, B, and C, respectively, and are not explicitly mentioned). Cys residues coordinating to Fe ions were considered deprotonated. A thorough manual investigation of all of the His residues gave the following protonation assignment: His-3, 4, 18, 248, 345, 364, 426, 452, 69B, 221B, and 41C were assumed to be protonated on the ND1 atom, His-181, 342, 140B, and 190B were presumed to be protonated on both the ND1 and NE2 atoms (and therefore positively charged), whereas the remaining 20 His residues were modeled with a proton on the NE2 atom. Furthermore, residue His-452D was flipped (i.e., the C and N atoms in the imidazole ring were exchanged). Protons were added by Maestro software, optimizing the hydrogen-bond network.

Molecular Dynamics Simulations. All molecular dynamics (MD) simulations were performed with the Amber22 software.[27] For the protein, we used the Amber ff14SB force field,[28] and water molecules were described by the TIP3P model.[29] For the metal sites, restrained electrostatic potential charges were employed, using electrostatic potentials calculated at the TPSS/def2-SV(P) level of theory[30,31] and sampled with the Merz−Kollman scheme,[32] although at a higher-than-default point density (∼2000/atom).[33] The charge calculations were performed on a minimal model of the E_0 resting state for the FeFe cluster, $[Fe_8S_7(CH_3S)_6]^{4-}$ for the P-cluster (in the fully reduced state) and $[Mg(CH_3COO)_2(H_2O)_4]$ for the Mg site, all shown in Figure 1. The charges are listed in Tables S1−S8. The metal sites were treated by a nonbonded model,[34] and the positions of all heavy atoms were strongly restrained toward the crystal structure in the MD simulations (like all other heavy atoms in the protein; see below).

For the MD simulations, the protein was solvated in a periodic rectangular box of explicit water molecules, extending at least 10 Å from the solute using the leap program in the Amber suite. 234 Cl⁻ ions and 310 Na⁺ ions were added to neutralize the protein and obtain an ionic strength of 0.2 M.[35] The ions were added by replacing random water molecules with the leap program in Amber software. However, some of them then end up inside the protein. This was avoided by using local software, ensuring that all of the counterions were in the solvent. The final system contained 251 429 atoms.

After the solvation, we performed 1000 cycles of minimization. This was followed by 1 ns constant-volume equilibration. Finally, the system was subjected to a 10 ns simulated annealing with a temperature of up to 373 K at constant pressure, followed by 1000 cycles of minimization. In all these calculations, the heavy atoms of the protein and the oxygen atoms of crystal-water molecules were restrained toward the crystal structure with a force constant of 10,000 kcal/mol/Å².

The temperature was kept constant at 300 K using Langevin dynamics, with a collision frequency of 2 ps⁻¹.[36] The pressure was kept constant at 1 atm using Berendsen's weak-coupling isotropic algorithm with a relaxation time of 1 ps.[37] Long-range electrostatics were handled by particle-mesh Ewald summation[38] with a fourth-order B spline interpolation and a tolerance of 10⁻⁵. The cutoff radius for Lennard-Jones interactions was set to 8 Å. All bonds involving

hydrogen atoms were constrained to their equilibrium values using the SHAKE algorithm (except in the constant-volume simulations),[39] allowing for a time step of 2 fs during the simulations. The final structure was used for QM/MM calculations.

In addition, we set up eight MD simulations of the protein in different protonation states of homocitrate, His-180, and the FeFe cluster. In these, no restraints toward the crystal structure were used in the first steps. For the metal sites, we used restraint for all metal−ligand bonds with the average distance in the two subunits of the crystal structure as the target (but for the FeFe cluster, distances from the QM/MM calculations with the various protonation states were used) and a force constant of 50 kcal/mol/Å². This ensures that the metal sites are kept intact, with a structure close to the crystal structure, but it also allows for some dynamics and avoids problems with water molecules that are often encountered with a bonded potential.[34] The Fe, S, and carbide ions of the FeFe and P-clusters were restrained to the crystal structure with a force constant of 1000 kcal/mol/Å².

The same simulations were performed as for the equilibration of the crystal structure (but with a force constant of 1000 kcal/mol/Å²), except for the final two steps, which were replaced by a 1 ns equilibration, and a 100 ns production simulation in which no restraints were applied (except for the Fe, S and carbide ions). 1000 snapshots were collected during the production simulation.

QM/MM Calculations. QM/MM calculations were performed with the ComQum software.[40,41] In this approach, the protein and solvent were split into three subsystems: system 1 (the QM region) was relaxed by the QM methods. System 2 contained all residues or water molecules with any atom within 6 Å of any atom in system 1. It was optionally relaxed by a MM minimization in each cycle of the QM/MM optimization, using updated charges for the QM system. System 3 contained the remaining part of the protein and the solvent, and it was kept fixed at the original coordinates (equilibrated crystal structure to reduce the risk that different calculations end up at different local minima).

In the QM calculations, system 1 was represented by a wave function, whereas all the other atoms were represented by an array of partial charges, one for each atom, taken from the MM setup. Thereby, the polarization of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding). When there is a bond between systems 1 and 2 (a junction), the hydrogen link-atom approach was employed: the QM system was capped with hydrogen atoms (hydrogen link atoms, HL), the positions of which are linearly related to the corresponding carbon atoms (carbon link atoms, CL) in the full system.[41,42] All atoms were included in the point-charge model, except the CL atoms.[34]

The total QM/MM energy in ComQum was calculated as[40,41]

$$E_{QM/MM} = E_{QM1+ptch23}^{HL} + E_{MM123,q_1=0}^{CL} - E_{MM1,q_1=0}^{HL} \qquad (2)$$

where $E_{QM1+ptch23}^{HL}$ is the QM energy of the QM system truncated by HL atoms and embedded in the set of point charges modeling systems 2 and 3 (but excluding the self-energy of the point charges). $E_{MM1,q_1=0}^{HL}$ is the MM energy of the QM system, still truncated by HL atoms but without any electrostatic interactions. Finally, $E_{MM123,q_1=0}^{CL}$ is the classical energy of all atoms with CL atoms and with the charges of the QM region set to zero (to avoid double-counting of the electrostatic interactions). Thus, ComQum employs a subtractive scheme with electrostatic embedding and van der Waals link-atom corrections.[43] No cutoff is used for any of the interactions in the three energy terms in eq 2.

The QM calculations for QM/MM were performed using the Turbomole software (version 7.7).[44] We employed four density functional theory (DFT) methods, TPSSh,[45] r²SCAN,[46] B3LYP,[47−49] and TPSS.[31] The former two were selected because they have been shown to give the best structures for nitrogenase models.[50] B3LYP gave the best results in two recent calibration studies on simple nitrogenase model systems with one or two Fe ions,[51,52] whereas TPSS has been used in most of our previous studies.[23,53,54] The

calculations involved either the def2-SV(P)[30] basis set for all atoms or the def2-TZVP[30] basis set for the FeFe cluster (including the added proton), homocitrate, Cys-257 and His-423, and the def2-SV(P) basis set for other groups. To enhance computational efficiency, Coulomb interactions were expanded in an auxiliary basis set by using the resolution-of-identity (RI) approximation.[55,56] Empirical dispersion corrections were applied using DFT-D4,[57] as implemented in Turbomole.

Two sizes of the QM system were used. In the minimal model, the FeFe cluster was represented by Fe$_8$S$_9$C(homocitrate)(CH$_3$S)-(methylimidazole) (56 atoms, see Figure 1a), where the last two groups model Cys-257 and His-423, taken from the A subunit of the protein (the two P clusters and the FeFe cluster in subunit D were modeled by MM in the fully reduced and the E$_0$ resting states). In the large model, we added all groups that form steric or hydrogen-bond interactions with the FeFe cluster: Val-57, Lys-83, Gln-176, His-180, and Phe-362 (side chains), four water molecules, the whole Ser-260 (except the O atom, but including CH$_3$CO− from the previous residue), as well as the backbone from Pro-335 to Lys-339, including the full side chain of the latter residue (187 atoms in total, as shown in Figure S1). Following experiment data,[58] we used the oxidation state-assignment Fe$_4^{III}$Fe$_4^{II}$ and a singlet spin state, $S = 0$[59] for the resting E$_0$ state and a doublet state for the E$_1$ state.[21]

Four different protonation states of homocitrate were tested: fully deprotonated (0H; net charge −4), with one proton either on the alcohol (1Ha) or on the O2 carboxylate atom (1Hc; both with a −3 net charge), or with protons on both of these groups (2H; net charge −2). These protonation states are shown in Figure 2. They give net charges for the large QM region of −5 to −3.
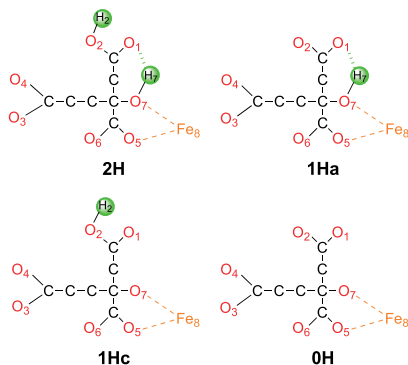

Figure 2. Four considered protonation states of homocitrate, 2H, 1Ha, 1Hc, and 0H. Atom numbers are also shown. Nonpolar H atoms are omitted. Charge of homocitrate is −2, −3, −3, and −4, respectively, in these four protonation states.

The electronic structure in all QM calculations was described using the BS approach.[60] Each of the eight Fe ions was modeled in the high-spin state, and these spins were then coupled antiferromagnetically to a singlet (E$_0$) or doublet (E$_1$) state. This means that the eight Fe ions should have either a surplus of α (four Fe ions) or β (four Fe ions) spin. Such a state can be selected in 70 different ways. The various BS states were obtained either by swapping the coordinates of the Fe ions[61] or with the fragment approach by Szilagyi and Winslow.[62] The various BS states are denoted simply by giving the number of the four Fe ions with β spin (Fe ion numbers are shown in Figure 1a), e.g., BS-2358, as shown in Figure 3.

Quantum Refinement. In crystallographic refinement, the goal is to find the model (coordinates, B-factors, and occupancies) that best explains the observed structure factors. This is done by minimizing the difference between the experimentally obtained structure factors
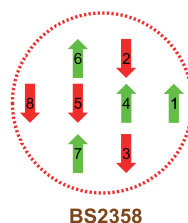

Figure 3. Energetically lowest BS state for the resting E$_0$ state, showing the local spin surplus on each Fe ion.

and those calculated from the current model. Due to the limited resolution obtained in protein crystallography, it is often necessary to supplement the refinement target function with empirical restraints that encode chemical knowledge. These restraints are obtained from either high-resolution crystallography of small molecules or quantum chemical calculations. In terms of computational chemistry, this is a MM force field. The refinement target then becomes a pseudoenergy function of the form

$$E_{tot} = w_A E_{Xray} + E_{MM} \qquad (3)$$

where E_{Xray} is a crystallographic goodness-of-fit criterion (typically a least-squares or a likelihood function), E_{MM} is the empirical restraints, and w_A[63] is a weight factor determining the relative importance of the two terms.

This approach works well when restraints of high accuracy are available, which is the case for amino acid residues and nucleic acids. However, for cofactors, substrates, inhibitors, and metal sites[34] (which often are found in the most interesting part of the structure), less experimental information is available and the restraints are therefore less accurate. One solution to this problem is to use more accurate QM calculations for a small but interesting part of the structure (system 1), i.e., an approach similar to QM/MM calculations.[40,41] This leads to a refinement target of the form

$$E_{tot} = w_A E_{Xray} + E_{MM} - E_{MM1} + w_{QM} E_{QM1} \qquad (4)$$

where index 1 indicates calculations of system 1. This represents the energy function of the quantum refinement. As crystallographic force fields are of a statistical nature, whereas the QM calculations are in energy units, another scaling factor, w_{QM}, needs to be introduced to put the restraints on a similar level. We recently implemented[64] eq 4 in the *cctbx*[65] layer of the crystallographic refinement software *phenix.refine*,[66] utilizing the free QM software ORCA 5.0.4[67,68] to calculate E_{QM1}.

In this study, we have used this new version of quantum refinement to study the protonation state of the homocitrate ligand in the crystal structure of Fe-nitrogenase (8BOQ).[13] Coordinates, occupancies, B-factors, and structure factors were obtained from this structure. We used the minimal QM model in Figure 1a for the QM calculations with the TPSS-D4/def2-SV(P) method[30,31,57] in the singlet state and BS state 2358 (obtained with the Flipspin approach in ORCA). Depending on the protonation state of the homocitrate ligand (see Figure 2), the net charge of the QM system was either −7 (0H), −6 (1Ha and 1Hc), or −5 (2H). The occupancy of S2B was set to 1.00, while the oxygen atom replacing S2B in the other conformation was discarded. Gln-176 was modeled with dual conformations as in the PDB file. Protonation of system 1 was done with *phenix.ready_set*. Restraint files for the nonstandard ligands (homocitrate, the P-cluster, and the FeFe cluster) were generated using *phenix.elbow*.[69] Three macrocycles of combined coordinate and standard individual B-factor refinement in *phenix.refine* were performed, in which only Cys-257, His-423, homocitrate, and the FeFe cluster were allowed to move, after which the real-space Z-scores based on the difference maps (RSZD), the real-space R factors (RSR), and the real-space correlation coefficients (RSCCs) for homocitrate were calculated by

Table 1. Relative Energy (ΔE in kJ/mol), Average (av), and Maximum (max) Metal−Metal and Meta−Ligand Distance Deviation from the Crystal Structure for the Various BS States of the Resting State[a]

BS	N	r²SCAN							TPSSh						
		ΔE	metal−metal			metal−ligand			ΔE	metal−metal			metal−ligand		
			av	max	bond	av	max	bond		av	max	bond	av	max	bond
1238	3	91.9	0.096	0.220	Fe5−Fe8	0.065	0.202	Fe8−S4B	96.7	0.081	0.177	Fe5−Fe8	0.061	0.131	Fe5−S4B
1248	3	81.1	0.103	0.179	Fe1−Fe4	0.067	0.199	Fe8−S4B	84.4	0.102	0.250	Fe5−Fe8	0.061	0.154	Fe8−S4B
1258	10	19.3	0.059	0.130	Fe6−Fe8	0.061	0.168	Fe8−S4B	20.4	0.065	0.136	Fe6−Fe8	0.059	0.148	Fe8−S4B
1268	9	49.7	0.104	0.216	Fe5−Fe7	0.065	0.149	Fe8−S1B	50.3	0.109	0.221	Fe5−Fe7	0.062	0.142	Fe5−S4B
1278	10	21.5	0.053	0.134	Fe5−Fe8	0.060	0.156	Fe8−S4B	22.7	0.060	0.161	Fe5−Fe8	0.057	0.139	Fe8−S4B
1348	3	87.4	0.097	0.206	Fe6−Fe8	0.061	0.176	Fe8−S1B	74.5	0.083	0.237	Fe5−Fe8	0.060	0.153	Fe5−S4B
1358	10	27.9	0.052	0.129	Fe6−Fe8	0.064	0.174	Fe8−S4B	27.0	0.057	0.131	Fe6−Fe8	0.061	0.155	Fe8−S4B
1368	10	26.5	0.057	0.146	Fe6−Fe8	0.064	0.145	Fe8−S1B	26.8	0.064	0.161	Fe5−Fe8	0.062	0.132	Fe8−S4B
1378	2	52.6	0.096	0.236	Fe2−Fe4	0.065	0.167	Fe8−S4B	49.2	0.098	0.237	Fe2−Fe4	0.063	0.150	Fe8−S4B
1458	3	43.0	0.110	0.222	Fe2−Fe3	0.063	0.179	Fe8−S4B	41.1	0.114	0.222	Fe2−Fe3	0.062	0.164	Fe8−S4B
1468	10	23.4	0.068	0.151	Fe5−Fe8	0.062	0.146	Fe8−S1B	22.8	0.079	0.175	Fe5−Fe8	0.060	0.128	Fe8−S4B
1478	10	28.6	0.054	0.138	Fe5−Fe8	0.059	0.158	Fe8−S4B	26.6	0.064	0.158	Fe5−Fe8	0.057	0.140	Fe8−S4B
1568	6	47.8	0.100	0.244	Fe6−Fe8	0.068	0.165	Fe8−S4B	57.3	0.102	0.241	Fe6−Fe8	0.066	0.148	Fe8−S4B
1578	6	53.9	0.102	0.200	Fe6−Fe8	0.068	0.197	Fe8−S4B	63.8	0.100	0.190	Fe6−Fe8	0.066	0.178	Fe8−S4B
1678	6	58,0	0.097	0.235	Fe6−Fe8	0.065	0.173	Fe4−C	66.7	0.098	0.235	Fe6−Fe8	0.062	0.156	Fe4−C
2348	2	59.4	0.099	0.201	Fe5−Fe8	0.059	0.184	Fe8−S4B	81.9	0.090	0.240	Fe5−Fe8	0.056	0.170	Fe8−S4B
2358	7	0.8	0.040	0.119	Fe5−Fe8	0.049	0.169	Fe8−S4B	0.0	0.042	0.122	Fe5−Fe8	0.047	0.152	Fe8−S4B
2368	8	29.3	0.072	0.167	Fe5−Fe8	0.062	0.155	Fe8−S1B	28.7	0.074	0.175	Fe5−Fe8	0.058	0.137	Fe8−S4B
2378	8	15.6	0.088	0.162	Fe2−Fe3	0.061	0.168	Fe8−S4B	20.8	0.089	0.165	Fe2−Fe3	0.057	0.152	Fe8−S4B
2458	8	25.8	0.070	0.142	Fe2−Fe4	0.063	0.182	Fe8−S4B	23.2	0.069	0.139	Fe2−Fe4	0.059	0.163	Fe8−S4B
2468	8	33.6	0.070	0.142	Fe6−Fe8	0.062	0.140	Fe8−S1B	32.0	0.070	0.147	Fe7−Fe8	0.058	0.134	Fe8−S4B
2478	7	3.9	0.033	0.088	Fe6−Fe8	0.051	0.152	Fe8−S4B	2.2	0.034	0.101	Fe5−Fe8	0.048	0.135	Fe8−S4B
2568	5	55.5	0.099	0.235	Fe1−Fe3	0.068	0.158	Fe8−S4B	61.2	0.096	0.211	Fe6−Fe8	0.065	0.144	Fe3−S4A
2578	4	45.0	0.079	0.197	Fe6−Fe8	0.058	0.173	Fe8−S4B	37.9	0.081	0.194	Fe5−Fe8	0.056	0.158	Fe8−S4B
2678	5	61.7	0.085	0.248	Fe5−Fe8	0.066	0.172	Fe8−S1B	57.2	0.090	0.248	Fe5−Fe8	0.063	0.163	Fe8−S1B
3458	8	28.9	0.055	0.138	Fe6−Fe8	0.061	0.168	Fe8−S4B	27.8	0.056	0.139	Fe6−Fe8	0.057	0.150	Fe8−S4B
3468	7	0.0	0.042	0.158	Fe6−Fe8	0.049	0.144	Fe8−S1B	2.6	0.042	0.139	Fe6−Fe8	0.045	0.114	Fe8−S4B
3478	8	28.9	0.062	0.129	Fe6−Fe8	0.058	0.151	Fe8−S4B	27.6	0.062	0.133	Fe6−Fe8	0.054	0.135	Fe8−S4B
3568	4	40.6	0.077	0.179	Fe6−Fe8	0.062	0.173	Fe8−S4B	50.5	0.080	0.171	Fe6−Fe8	0.060	0.163	Fe4−S4A
3578	5	57.1	0.099	0.227	Fe7−Fe8	0.071	0.194	Fe8−S4B	63.2	0.100	0.225	Fe7−Fe8	0.067	0.177	Fe8−S4B
3678	5	52.6	0.085	0.220	Fe5−Fe8	0.067	0.160	Fe8−S4B	58.4	0.082	0.237	Fe5−Fe8	0.063	0.144	Fe8−S1B
4568	5	62.4	0.103	0.259	Fe1−Fe3	0.068	0.177	Fe8−S4B	70.5	0.104	0.266	Fe1−Fe3	0.064	0.160	Fe8−S4B
4578	5	73.7	0.093	0.218	Fe6−Fe8	0.070	0.192	Fe8−S4B	79.8	0.089	0.225	Fe6−Fe8	0.065	0.173	Fe8−S4B
4678	4	61.6	0.077	0.198	Fe6−Fe8	0.057	0.159	Fe8−S1B	59.5	0.066	0.195	Fe6−Fe8	0.051	0.144	Fe8−S1B
5678	1	121.4	0.097	0.215	Fe1−Fe2	0.074	0.186	Fe8−S4B	122.6	0.091	0.206	Fe1−Fe2	0.072	0.170	Fe8−S4B

[a]Bond lists the bond that gives the maximum deviation. N is the type of the BS state in Noodleman's nomenclature.

the use of *EDSTATS*.[70] We tested different values of the w_A weight factor in eq 4 and selected a value for which the structures were significantly affected by both the crystallographic and QM data, $w_A = 1.0$, cf. Table S9.

RESULTS AND DISCUSSION

We here present the first QM/MM study of Fe-nitrogenase, based on the recent crystal structure.[13] We determine the most stable BS states for the resting state and the proper protonation states for the homocitrate ligand and for the catalytic His-180 residue. In addition, we have studied the protonation of the one-electron reduced E_1 state.

BS States of the E_0 Resting State. We started by investigating which is the most stable BS state of the FeFe cluster in the resting E_0 state. The cluster contains eight Fe ions, all of which have a high-spin configuration. However, the spins couple antiferromagnetically to a singlet state.[71] Thus, four of the Fe ions have a surplus of α spin, and the other four have a surplus of β spin. Four Fe ions can be selected out of

eight in 70 different ways $\left(\frac{8!}{4!\cdot4!}\right)$, twice as many as for Mo- and V-nitrogenase. However, for the resting state, with an equal number of Fe(II) and Fe(III) ions, there is no distinction between the α and β electrons, and therefore, only 35 states are distinct (i.e., state 1234 is equivalent to the 5678 state, and so on; this was confirmed by explicit calculations, cf. Table S10). In the following, we present the results only for states with a surplus of β spin on Fe8. The relative energies of all these states were studied using the minimal 56-atom QM model (Figure 1a) with two DFT functionals: TPSSh and r²SCAN.

The relative energies are presented in Table 1. It can be seen that the results with the two functionals are quite consistent, with mean signed and mean absolute differences (MAD) for the relative energies of only 2 and 4 kJ/mol. Three states are lowest in energy and almost degenerate (within 3−4 kJ/mol): BS-2358, 3468, and 2478 (the numbers denote the Fe ions with β spin; atom numbering shown in Figure 1a). These correspond to the three BS7 in the Noodleman nomenclature for the MoFe cluster,[13] with the extra Fe8 ion having a surplus
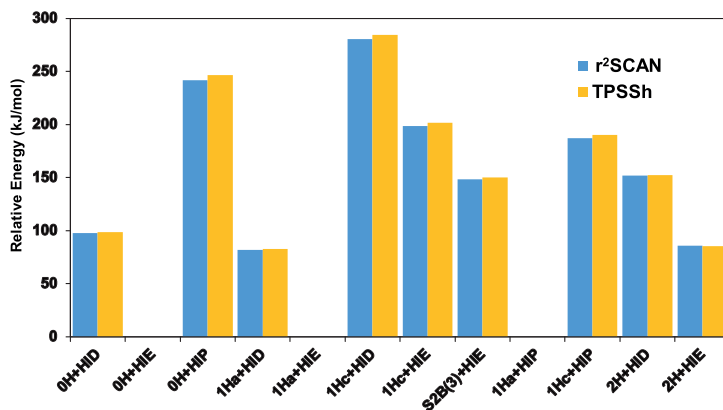
Figure 4. Relative energies (kJ/mol) of the various protonation states tested for Fe-nitrogenase in the E_0 state.

of β spin. As can be seen from the schematic picture in Figure 3, these are the spin configurations that give the largest number of antiferromagnetically coupled pairs of nearby Fe ions (always for the Fe2—Fe6, Fe3—Fe7, and Fe4—Fe5 pairs, connecting the two Fe_4S_4 subclusters and two of the three pairs within each subcluster, Fe1—Fe2/3/4 and Fe8—Fe5/6/7). The fourth-best state is 16—20 kJ/mol less stable than the best state, BS-2378 with r^2SCAN or BS-1258 with TPSSh. The ordering of the states follows approximately the Noodleman nomenclature[13] with the order BS7 < BS10 ≈ BS8 < BS9 ≈ BS4 < BS6 ≈ BS5 < BS3 < BS1, with an ambiguity only for the BS2 state, which is similar to BS3 for TPSSh but similar to BS5 with r^2SCAN. This order is rather different from what was observed for the E_0 state of Mo-nitrogenase: BS7 < BS6 < BS2 < BS8 < BS4 ≈ BS10 < BS9 < BS5 < BS3 < BS1.[53]

Fe spin populations are shown in Table S10. They are (in absolute terms) 3.4—3.8 e with TPSSh and 3.3—3.9 e for r^2SCAN, with little variation among the eight Fe ions, besides that, Fe8 (which has six rather than four ligands) always has the highest spin population, by 0.2—0.3 e.

The various BS states give slightly different geometries of the FeFe cluster. Therefore, we can compare which of the BS states reproduces the crystal structure best in terms of the short Fe—Fe and Fe—ligand distances. The results in Table 1 show that the three BS7 states also reproduce the structure best, with mean absolute deviations of 0.03—0.04 and 0.05 Å for the Fe—Fe and Fe—ligand distances.

Based on these findings, we rather arbitrarily decided to make further studies of the E_0 state with the BS-2358 state (BS-235), which has been argued to be the proper BS state for the E_0 of Mo-nitrogenase.[72]

Protonation State of Homocitrate and His-180. Next, we studied the protonation of the resting E_0 state using QM/MM calculations. This was done with the larger 186-atom QM region, as shown in Figure S1. We investigated four different protonation states for HCA (shown in Figure 2) and three different protonation states for His-180. We also tested to protonate S2B.

In the crystal structure, all oxygen atoms of homocitrate are involved in hydrogen bonds with water molecules or with the surrounding residues. O1 receives hydrogen bonds from the NZ atom Lys-361 and NE2 of Gln-176 (in both of its

conformations), whereas O2 forms two hydrogen bonds to water molecules (which form hydrogen bonds to other water molecules), and it is 2.99—3.17 Å from SG of Cys-52. O3 receives a hydrogen bond from the backbone N atom of Lys-406 and forms hydrogen bonds to two water molecules. O4 forms hydrogen bonds to three water molecules. O5 is an Fe ligand (the Fe—O distance is 2.08 or 2.36 Å in the two subunits) and forms a hydrogen bond to a water molecule, whereas O6 forms hydrogen bonds to two water molecules. Finally, the alcohol atom O7 is also a Fe ligand (2.09 or 2.17 Å distance), and it is 2.69—2.70 Å from the O1 atom. The NE2 atom of His-180 forms a hydrogen bond with the OE1 atom of Gln-176 in one of its two conformations (both have occupancies of 0.5). It is 3.33—3.45 Å from S2B which is also half-occupied and 3.54—3.62 Å from the alternative O atom. The ND1 atom forms a hydrogen bond with a water molecule, which forms hydrogen bonds to the OG of Ser-176 and the OH of Tyr-262 (which both can be either donors or acceptors).

In the MD equilibration of the protein for the QM/MM calculations, the O atom replacing S2B was deleted, together with the corresponding alternative conformation of Gln-176 and the storage position of S2B. After the equilibration, all water molecules donate hydrogen bonds to the carboxylate oxygen atoms of homocitrate (which were modeled unprotonated), giving H···O distances of 1.70—1.91 Å. The other hydrogen bonds also fall in the same range, except that between O1 and HE2 of Gln-175, which is 2.25 (all in the A subunit, used for the QM/MM calculations; 1.67—1.97 and 2.41 Å; and 2.06 Å to Lys in the D subunit). Adding a proton on the alcoholic O7 has no impact on the surroundings as it points toward O1 and does not interact with anything else in the surroundings. Likewise, a proton added to O2 does not interfere with the two water molecules forming hydrogen bonds to it. However, protonation of the ND1 of His-180 would require the imidazole group to rotate or a change in the hydrogen-bond network involving a water molecule, Ser-176 and Tyr-262.

The relative QM/MM energies of the various protonation states are presented in Figure 4. It can be seen that the two DFT methods give very similar results with a MAD of only 2 kJ/mol. Of course, energies are comparable only for structures

Table 2. Relative Energy (ΔE in kJ/mol), Average (av), and Maximum (max) Metal−Metal and Meta−Ligand Distance Deviations from the Crystal Structure for the QM/MM Structures of the Various Protonation States of the E_0 Resting State[a]

	r²SCAN							TPSSh						
	ΔE	metal−metal			metal−ligand			ΔE	metal−metal			metal−ligand		
		av	max	bond	av	max	bond		av	max	bond	av	max	bond
0H + HID	98.0	0.065	0.25	Fe5−Fe8	0.061	0.30	Fe8−O1	98.7	0.058	0.20	Fe5−Fe8	0.060	0.29	Fe8−O1
0H + HIE	0.0	0.062	0.22	Fe5−Fe8	0.060	0.30	Fe8−O1	0.0	0.058	0.19	Fe7−Fe8	0.058	0.29	Fe8−O1
0H + HIP	241.6	0.069	0.26	Fe5−Fe8	0.059	0.31	Fe8−O1	246.6	0.062	0.22	Fe5−Fe8	0.061	0.31	Fe8−O1
1Ha + HID	81.7	0.039	0.08	Fe6−Fe8	0.045	0.12	Fe8−S4B	82.8	0.038	0.07	Fe5−Fe8	0.048	0.14	Fe8−S4B
1Ha + HIE	0.0	0.040	0.09	Fe6−Fe8	0.045	0.12	Fe8−S4B	0.0	0.039	0.09	Fe6−Fe8	0.047	0.13	Fe8−S4B
1Hc + HID	280.7	0.054	0.18	Fe7−Fe8	0.056	0.23	Fe8−O1	284.4	0.050	0.15	Fe7−Fe8	0.056	0.22	Fe8−O1
1Hc + HIE	198.8	0.056	0.18	Fe7−Fe8	0.055	0.23	Fe8−O1	201.8	0.051	0.15	Fe7−Fe8	0.056	0.22	Fe8−O1
S2B + HIE	148.2	0.068	0.27	Fe5−Fe8	0.061	0.32	Fe8−O1	150.1	0.064	0.23	Fe5−Fe8	0.062	0.32	Fe8−O1
1Ha + HIP	0.0	0.044	0.13	Fe6−Fe8	0.046	0.12	Fe8−S4B	0.0	0.043	0.12	Fe6−Fe8	0.048	0.14	Fe8−S1B
1Hc + HIP	187.3	0.061	0.18	Fe7−Fe8	0.056	0.24	Fe8−O1	190.0	0.055	0.15	Fe7−Fe8	0.057	0.24	Fe8−O1
2H + HID	151.8	0.040	0.12	Fe6−Fe8	0.049	0.18	Fe8−S3B	152.5	0.039	0.123	Fe6−Fe8	0.052	0.15	Fe7−S3B
2H + HIE	86.1	0.041	0.13	Fe6−Fe8	0.051	0.14	Fe8−S3B	85.6	0.040	0.14	Fe6−Fe8	0.053	0.15	Fe7−S3B

[a]All calculations were performed in the 2358 BS state. Bond shows the bond that gives the maximum deviation.

Table 3. Quality Measures and Fe−O Distances (d in Å) for the Four Quantum-Refined Structures with Varying Protonation States of Homocitrate (cf. Figure 2)[a]

	RSZD	RSR	RSCC	ΔE_{str} (kJ/mol)	$\Sigma\Delta d$ (Å)	d(Fe−O)$_{QR}$		d(Fe−O)$_{QM}$	
						O5	O7	O5	O7
2H	1.1	0.046	0.962	35	0.08	2.22	2.21	2.23	2.28
1Ha	**0.5**	**0.044**	**0.967**	41	**0.07**	2.17	2.20	2.13	2.23
1Hc	1.7	0.050	0.955	54	0.08	2.23	2.13	2.22	2.06
0H	2.8	0.055	0.944	38	0.10	2.22	2.06	2.19	1.99

[a]The quality measures are the real-space Z-scores based on the difference maps (RSZD), the real-space R factors (RSR), the real-space correlation coefficients (RSCC) for homocitrate, the strain energy of the QM region (ΔE_{str}), and the sum of the difference in the Fe−O bond lengths ($\Sigma\Delta d$) to the homocitrate O5 (carboxylate) and O7 (alcohol) atoms in the quantum-refined structure [d(Fe−O)$_{QR}$] and in a structure optimized by QM/MM without any crystallographic information [d(Fe−O)$_{QM}$]. The best results of each quality measure are marked in bold face.

with the same number of protons. Therefore, the data must be divided into three groups with one, two, or three protons that are moved around.

The first group involves only two structures, with no extra proton on homocitrate (0H) and with the proton on His-180 either on ND1 (called HID) or on NE2 (HIE). It can be seen that the latter state is most stable, ~99 kJ/mol.

In the second group, involving six structures, an extra proton is put either on His-180 (called HIP) on S2B or on homocitrate, on either the O2 carboxylate atom (called 1Hc), or on the O7 alcohol atom (called 1Ha). The results show that the HIE state is always ~82 kJ/mol more stable than the HID state. Moreover, protonation of the homocitrate alcohol atom is 202, 247, and 150 kJ/mol more favorable than protonating the carboxylate atom, His-180 or S2B. Thus, the 1Ha + HIE structure is by far the most stable structure in this group.

Finally, we also tested adding a third proton to the structures, which could either go to His-180 or to homocitrate (2H), giving four tested structures. The results show that it is more favorable to protonate His-180 than homocitrate. Moreover, HIE is still ~67 kJ/mol more stable than HID and 1Ha is ~190 kJ/mol more stable than 1Hc. Consequently, the 1Ha + HIP structure is the most stable for this protonation level.

The relative energies give no indication of which of the three groups is most favorable, i.e., how many protons the structure should contain. However, we can again compare the optimized structures with the crystal structure. This is done in Table 2,

from which it can be seen that the 1Ha + HID and 1Ha + HIE structures reproduce the crystal structure best, with MADs of 0.04 and 0.05 Å for the Fe−Fe and Fe−ligand distances. The third-best structure is 1Ha + HIP, showing that the structure of the FeFe cluster is insensitive to the protonation of His-180, but that homocitrate is most likely singly protonated on the O7 alcohol atom (1Ha).

The Fe−O7 distance is 2.13−2.15 Å when it is protonated in the QM/MM structures (2.32−2.35 Å when O2 is also protonated), but 1.90−1.93 (0H) or 1.98−1.99 Å (H1c) when it is deprotonated. The Fe−O5 distance is 2.14−2.22 Å in the various structures. As mentioned above, there is a large variation in the distances in the two subunits of the crystal structure (2.08−2.36 Å), but it is never close to the distance of the deprotonated alcoholate ligand.

To further strengthen this important conclusion, we also performed quantum refinement of the E_0 state of Fe-nitrogenase with the four protonation states of homocitrate. The results of these refinements are shown in Table 3. It can be seen that all three crystallographic quality measures (RSZD, RSR, and RSCC) for homocitrate are best for the 1Ha protonation state, in agreement with the QM/MM data. This is also confirmed by the electron-density difference maps around the homocitrate ligand for the four quantum refinements, as shown in Figure S2. The difference in the Fe−O bond lengths to the two O atoms of homocitrate between the quantum-refined structure and a structure optimized without any restraints to the crystal structures (i.e., a QM/MM structure with the MM force field used by
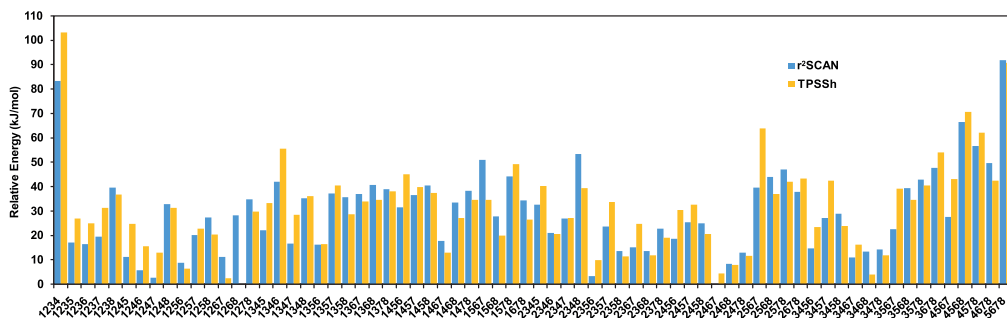
Figure 5. Relative energies (kJ/mol) of the various BS states of Fe-nitrogenase in the E_1 state, protonated on S2B(3), using the minimal QM region.
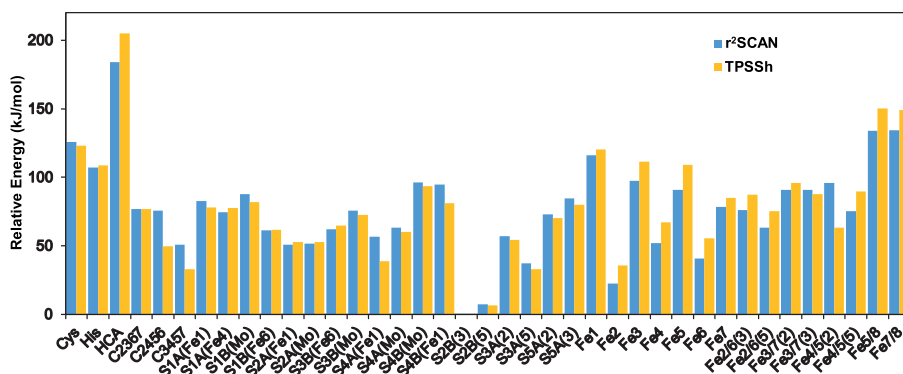

Figure 6. Relative energies (kJ/mol) of the various protonation states of Fe-nitrogenase in the E_1 state using the large QM region and the BS-2468 state.

Phenix) is also the smallest for the 1Ha protonation state. In particular, it can be seen that the Fe–O7 distance to the alcohol atom in the quantum-refined structures is always longer than the expected bond length if it is deprotonated, 1.99–2.06 Å. However, the strain energy (ΔE_{str}), i.e., the difference in the QM energy of the QM region between the quantum-refined structure and the structure optimized without any crystallographic restraints, is lowest for the 2H protonation state, but ΔE_{str} is comparable only for structures with the same net charge, i.e., in this case only for 1Ha and 1Hc, for which 1Ha gives the better results. In conclusion, the quantum-refinement calculations also quite conclusively point out 1Ha as the protonation state observed in the crystal structure of Fe-nitrogenase in the E_0 state.

Finally, we also performed MD simulations of Fe-nitrogenase with different protonation states of homocitrate and His-180. The hydrogen-bond pattern in these simulations is shown in Tables S11–S16. In the preferred protonation state (1Ha for homocitrate and His-180 protonated on NE2), the proton on O7 of homocitrate forms an internal hydrogen bond to O1 (like in the QM/MM and quantum-refined structures). O1 also receives hydrogen bonds from the HZ atoms Lys-361, from one of the HE2 atoms of Gln-176 and from water molecules, with rather large variation between the two subunits of the protein. O2 forms hydrogen bonds to 2–3 water

molecules. O3 and O4 receive a hydrogen bond from the backbone N atom of Lys-406 and form hydrogen bonds to 2–3 water molecules (they show similar patterns and thus rotate during the MD simulation). O5 forms a hydrogen bond to a water molecule, whereas O6 forms hydrogen bonds to two water molecules. The HE2 atom of His-180 forms a hydrogen bond to S2B in most snapshots (~2.4 Å average distance). The ND1 atom forms a hydrogen bond with a water molecule in 54–71 of the snapshots and occasionally with the backbone H atom of the same residue.

When instead ND1 of His-180 is protonated, the HD1 proton donates hydrogen bonds to either the backbone O atom or to a water molecule (observed in 46–51 and 30–40% of the MD snapshots, respectively). The NE2 atom sometimes receives a hydrogen bond from a water molecule. If both the ND1 and NE2 atoms are protonated, the same hydrogen bonds are observed, but with higher occurrences (72–87 and 34–64%).

If instead, O2 is protonated in homocitrate, the proton forms occasional hydrogen bonds to a water molecule (23–35% occurrences). The unprotonated O7 receives hydrogen bonds from the HZ atoms of Lys-361 and occasionally from water.

The assignment of the 1Ha protonation state for homocitrate (i.e., singly protonated on the alcohol oxygen)
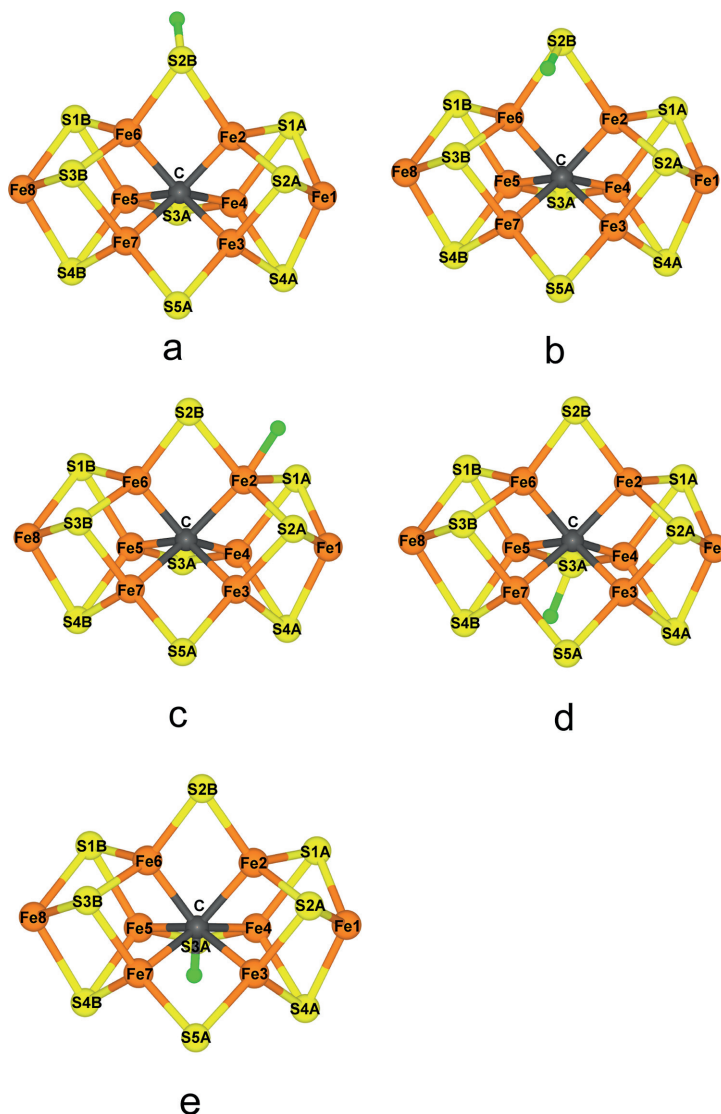
Figure 7. Best QM/MM structures of the E_1 state of Fe-nitrogenase, protonated on (a) S2B(3), (b) S2B(5), (c) Fe2, (d) S3A(5), and (e) C3457, all optimized with the r^2SCAN functional (with the 1Ha state of homocitrate and the HIE state of His-180).

agrees with a previous quantum-refinement study of Mo-nitrogenase,[54] as well as a comparison of the QM/MM and crystal structures of this enzyme.[73] Comparisons of vibrational and CD spectra, as well as crystallographic structures of model compounds and the extracted cofactor of Mo- and V-nitrogenase, have given the same results.[74−76]

Protonation of the E_1 State. Once we have settled the proper protonation states of homocitrate and His-180 and the

BS state for E_0, we can turn to the main subject of the present investigation, viz., the protonation of the E_1 state.

First, we tested again what BS state is most favorable (with the extra proton on S2B). The results are shown in Figure 5 and Table S17. For E_1, the 70 BS states are distinct, with mean absolute deviations of 16−17 kJ/mol between states for which the Fe ions with majority α and β spin have been swapped (e.g., BS-1234 and BS-5678; for E_0, the difference was in general less than 1 kJ/mol). It can be seen that there is a larger

difference between the r^{2}SCAN and TPSSh functionals than that for E$_0$, with a MAD of 7 kJ/mol (4 kJ/mol for E$_0$). With r^{2}SCAN, seven BS states give energies that are degenerate within 9 kJ/mol, viz., BS-2467, 2356, 1247, 1246, 2468, 1256, and 1268 (in this order). They belong to Noodleman's BS4, 5, 8, 9, and 10 types but do not include the BS7 states that were lowest for E$_0$ (they are 13−20 kJ/mol less stable than the best BS-2467 state). With TPSSh, the most stable BS state is 1268 (which is 4 kJ/mol more stable than the 2467 state), and the same states are among the most stable ones (although the ordering is different), including also BS-1267 and 3468 (the latter one of the BS7-type states).

Next, we calculated the relative energies of 50 different protonation states for the E$_1$ state of the FeFe cluster with both functionals (using BS-2468). This involved protonation of Cys-257, His-180, homocitrate, the central carbide (with the proton on the three different faces of the cluster), sulfide ions (at least two different conformations of each), Fe ions, or bridging two Fe ions (typically two different conformations).[23]

The results are collected in Figure 6 and Table S18. With both functionals, protonation of S2B is most favorable. We tested two directions of this proton, and it turns out that it is 6−7 kJ/mol more favorable if the proton points toward S3A (S2B(3)), rather than toward S5A (S2B(5)), cf. Figure 7a,b. In this conformation, the added proton is rather close to Phe-361 (2.21 Å between the proton and CZ; however, the SB2−CZ distance has only increased by 0.1 Å compared to the other structures). Besides this, there is some discrepancy between the two functionals. In general, TPSSh shows a tendency to favor protonation of the central carbide and disfavor protonation of one or two Fe ions. Consequently, with r^{2}SCAN, protonation of Fe2 (Figure 7c) is third best, 22 kJ/mol less stable than S2B(3), and protonation of Fe6 and Fe4 is among the nine best structures, whereas these structures rank 5, 11, and 15 with TPSSh. Instead, protonation of S3A(5) (also a μ_2 belt sulfide; Figure 7d) or the central hydride on the Fe3/4/5/7 face (called C3457; Figure 7e) is the third and fourth best structure, both 33 kJ/mol less stable S2B(3). These two structures rank four and seven with r^{2}SCAN. Next come several structures with μ_3 cubane sulfide ions protonated, especially S2A and S4A (both in the Fe1−4 subcluster). The best structure with a bridging hydride ion is Fe2/6(5) with both functionals, 63 or 75 kJ/mol less stable than the best structure. The least favorable structures have the proton on His-180, Cys-257, homocitrate, or bridging two Fe ions within the same subcluster (most structures of the latter type were not found).

We also repeated the complete BS investigation using the large 188-atom model for both the S2B(3) and Fe2-protonated structures and the r^{2}SCAN, TPSSh, B3LYP, and TPSS functionals. The results are shown in Tables S19−S22. The MAD in the relative BS energies obtained with the large and minimal models is 5−7 kJ/mol for the r^{2}SCAN and TPSSh functionals. With all four functionals, S2B(3) is more stable than the Fe2 structure by 26, 32, 117, and 14 kJ/mol, respectively, reflecting that iron-bound hydride ions are increasingly disfavored as the amount of Hartree−Fock exchange increases in the hybrid functionals. The most stable BS state also varies among the four functionals, BS-2467, 1256, 2368, and 3468 for the S2B(3) structure, and BS-2458 (B3LYP) or 2358 (the other functionals) for the structure with a hydride ion on Fe2. Many of the Fe2 structures reorganized into other structures during the geometry optimization,

especially with B3LYP. There are many BS states with similar energies, especially with the pure functionals and the S2B(3) state (e.g., 9, 9, 2, and 5 states within 10 kJ/mol of the best one for S2B(3) and 6, 2, 2, and 7 for Fe2 with the four functionals, respectively).

Finally, we calculated the relative energies of five of the best protonation states also with the larger def2-TZVP basis set (full geometry optimizations) and with the surrounding relaxed, with both the r^{2}SCAN and TPSSh functionals. The results are collected in Table 4. It can be seen that increasing

Table 4. Relative Energies (kJ/mol) of the Five Structures in Figure 7 Calculated with Either r^{2}SCAN or TPSSh and the Larger def2-TZVP Basis Set (TZ) or with the Original def2-SV(P) Basis Set (SV) and with the Surrounding Protein and Water Molecules Relaxed (Relax)

DFT	protonation	BS	SV	TZ	SV-relax
r^{2}SCAN	S2B(3)	2467	0	0	0
	S2B(5)	2467	9	7	20
	Fe2	2358	26	26	29
	S3A(5)	2467	43	41	54
	C3457	2468	61	64	62
TPSSh	S2B(3)	1256	0	0	0
	S2B(5)	1256	6	−1	12
	Fe2	2358	32	30	37
	S3A(5)	2468	42	31	46
	C3457	1256	35	32	36

the basis set has a minor influence on the relative energies, by up to 3 kJ/mol for r^{2}SCAN and up to 11 kJ/mol with TPSSh. However, this leads to slight changes in the ranking of the five states with TPSSh (S2B(5) is now 1 kJ/mol more stable than the S3B(3) structure and the Fe2, S3A(5), and C3457 structures become essentially degenerate). If the surrounding protein and solvent are allowed to relax, similar restricted changes in the relative energies are observed (but this time larger for r^{2}SCAN than for TPSSh, up to 12 and 5 kJ/mol, respectively). With both functionals, the S2B(3) state is stabilized compared to the other states, and with r^{2}SCAN, also the Fe2 and C3457 states. However, neither the basis set nor the relaxation of the surroundings change the conclusion that protonation of S2B is 26−37 kJ/mol more favorable than a hydride ion on Fe2.

With Mo-nitrogenase,[23] protonation of S2B(3) was found to be most stable, 7 (TPSS) or 35 (B3LYP) kJ/mol more stable than S2B(5). However, the most stable structure with a hydride ion was on Fe4, which was 27 or 96 kJ/mol (TPSS and B3LYP), less stable than that protonated on S2B(3). The Fe2 structure was 51 or 131 kJ/mol less stable than S2B(3). Moreover, the best BS state for S2B(3) was BS7-346, followed by the other two BS7 states, 24−33 kJ/mol higher in energy. Thus, there are significant differences between Fe and Mo-nitrogenases, especially regarding the preferred BS states.

■ CONCLUSIONS

In this study, we have set up the first QM/MM calculations of Fe-nitrogenase. This involves determining the proper protonation states of all residues and MD relaxation of the added protons and the surrounding water molecules.

Then, we have examined all 70 BS states of the cluster in the resting E$_0$ state, showing that the relative stabilities of the states are rather similar to those obtained for the FeMo cluster,

although the order of the various BS states has some conspicuous differences. However, the most stable BS states are still of Noodleman's BS7 type, together with a surplus of β spin on the eighth Fe ion, viz., BS-2358, 2478, and 3468, which are degenerate within 4 kJ/mol. These are the states that give the largest number of antiferromagnetically coupled close Fe–Fe pairs (Figure 3), and they also reproduce Fe–Fe and Fe–ligand distances best compared to the crystal structure. For the E_0 state, there is no difference between α and β spins so there are only 35 distinct BS states.

Next, we investigated the protonation states of homocitrate and His-180. The relative QM/MM energies show that His-180 strongly prefers protonation on NE2 rather than on ND1. Moreover, homocitrate seems to be most stable when it is singly protonated on the alcohol atom, O7 atom. This finding is also supported by the quantum refinement of the crystal structure. In both cases, the preferred protonation states are the same as those found for Mo-nitrogenase.

Finally, we turned to the E_1 state. For this state, the 70 BS states are distinct, and many BS states are close in energy. There are significant differences in the preferred BS states in the E_1 state compared to those observed for Mo-nitrogenase. We optimized the structures of 50 different protonation states. The relative energies depend somewhat on what DFT method is used, but with all four functionals tested, protonation of the μ_2 belt sulfide ion S2B is more favorable than the formation of a Fe-bound hydride ion. A hydride bound terminally to Fe2 in the exo position (trans to the carbide ion) is the best hydride-bound structure, but it is 14, 26, 32, and 117 kJ/mol less stable than the structure protonated on S2B with the TPSS, r²SCAN, TPSSh, and B3LYP functionals, respectively. This does not change if a larger basis set is used or if the surroundings are relaxed during the geometry optimization. Thus, our results indicate that the E_1 state does not contain any Fe-bound hydride ion, in agreement with what has been found for Mo-nitrogenase,[22–24] but contrary to a recent EPR results suggesting that the E_1 state of Fe-nitrogenase involves a hydride ion.[21] However, it should be noted that the experimental evidence is only indirect: it is observed that the EPR signal of the E_1 state at 12 K is partly converted to another signal if illuminated with 450 nm light, leading to a ~70/30% equilibrium at long times. The conversion has a kinetic isotope effect of 2.0–2.8. If the temperature is increased to 145 K, then the structure relaxes back to the original state. This is interpreted as a conversion between two hydride-bound states, which the authors suggest to be Fe2/6 and Fe3/7 structures. Clearly, these two structures are high in energy in our calculations, 63–96 kJ/mol less stable than the S2B(3) structure. Undoubtedly, more investigations are needed to settle the nature of the E_1 state in Fe-nitrogenase.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.inorgchem.3c02329.

Large QM system with residue labels, difference maps for the four quantum-refined structures of homocitrate, RESP charges used for the FeFe cluster, P-cluster and the Mg site, relative energies and spin densities for the various BS states for the E_0 and E_1 state in the S2B(3) and Fe2 protonation states, and various protonation states of E_1 in the BS-2468 state (PDF)

Coordinates of the most stable states (ZIP)

AUTHOR INFORMATION

Corresponding Author

Ulf Ryde − *Department of Computational Chemistry, Lund University, Lund SE-221 00, Sweden;* ⓘ orcid.org/0000-0001-7653-8489; Phone: +46-46 2224502; Email: Ulf.Ryde@compchem.lu.se

Authors

Hao Jiang − *Department of Computational Chemistry, Lund University, Lund SE-221 00, Sweden;* ⓘ orcid.org/0000-0002-9641-1634

Kristoffer J. M. Lundgren − *Department of Computational Chemistry, Lund University, Lund SE-221 00, Sweden;* ⓘ orcid.org/0009-0002-2611-7883

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.inorgchem.3c02329

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Hoffman, B. M.; Lukoyanov, D.; Yang, Z.-Y.; Dean, D. R.; Seefeldt, L. C. Mechanism of Nitrogen Fixation by Nitrogenase: The Next Stage. *Chem. Rev.* **2014**, *114* (8), 4041–4062.

(2) Zhang, X.; Ward, B. B.; Sigman, D. M. Global Nitrogen Cycle: Critical Enzymes, Organisms, and Processes for Nitrogen Budgets and Dynamics. *Chem. Rev.* **2020**, *120* (12), 5308–5351.

(3) Seefeldt, L. C.; Yang, Z.-Y.; Lukoyanov, D. A.; Harris, D. F.; Dean, D. R.; Raugei, S.; Hoffman, B. M. Reduction of Substrates by Nitrogenases. *Chem. Rev.* **2020**, *120* (12), 5082–5106.

(4) Burgess, B. K.; Lowe, D. J. Mechanism of Molybdenum Nitrogenase. *Chem. Rev.* **1996**, *96* (7), 2983–3012.

(5) Schmid, B.; Chiu, H.-J.; Ramakrishnan, V.; Howard, J. B.; Rees, D. C. Nitrogenase. *Handbook of Metalloproteins*; John Wiley & Sons, Ltd, 2006.

(6) Lancaster, K. M.; Roemelt, M.; Ettenhuber, P.; Hu, Y.; Ribbe, M. W.; Neese, F.; Bergmann, U.; DeBeer, S. X-Ray Emission Spectroscopy Evidences a Central Carbon in the Nitrogenase Iron-Molybdenum Cofactor. *Science* **2011**, *334* (6058), 974–977.

(7) Bjornsson, R.; Lima, F. A.; Spatzal, T.; Weyhermüller, T.; Glatzel, P.; Bill, E.; Einsle, O.; Neese, F.; DeBeer, S. Identification of a Spin-Coupled Mo(III) in the Nitrogenase Iron-Molybdenum Cofactor. *Chem. Sci.* **2014**, *5* (8), 3096–3103.

(8) Igarashi, R. Y.; Laryukhin, M.; Dos Santos, P. C.; Lee, H.-I.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Trapping H⁻ Bound to the Nitrogenase FeMo-Cofactor Active Site during H₂ Evolution: Characterization by ENDOR Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127* (17), 6231–6241.

(9) Hoeke, V.; Tociu, L.; Case, D. A.; Seefeldt, L. C.; Raugei, S.; Hoffman, B. M. High-Resolution ENDOR Spectroscopy Combined with Quantum Chemical Calculations Reveals the Structure of Nitrogenase Janus Intermediate E$_4$ (4H). *J. Am. Chem. Soc.* **2019**, *141* (30), 11984−11996.

(10) Kim, J.; Rees, D. C. Structural Models for the Metal Centers in the Nitrogenase Molybdenum-Iron Protein. *Science* **1992**, *257* (5077), 1677−1682.

(11) Einsle, O. Nitrogenase FeMo Cofactor: An Atomic Structure in Three Simple Steps. *J. Biol. Inorg. Chem.* **2014**, *19* (6), 737−745.

(12) Sippel, D.; Einsle, O. The Structure of Vanadium Nitrogenase Reveals an Unusual Bridging Ligand. *Nat. Chem. Biol.* **2017**, *13* (9), 956−960.

(13) Trncik, C.; Detemple, F.; Einsle, O. Iron-Only Fe-Nitrogenase Underscores Common Catalytic Principles in Biological Nitrogen Fixation. *Nat. Catal.* **2023**, *6* (5), 415−424.

(14) Schmidt, F. V.; Schulz, L.; Zarzycki, J.; Oehlmann, N. N.; Prinz, S.; Erb, T. J.; Rebelein, J. G. Structural Insights into the Iron Nitrogenase Complex. *bioRxiv* **2023**.

(15) Einsle, O.; Tezcan, F. A.; Andrade, S. L. A.; Schmid, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor. *Science* **2002**, *297* (5587), 1696−1700.

(16) Spatzal, T.; Aksoyoglu, M.; Zhang, L.; Andrade, S. L. A.; Schleicher, E.; Weber, S.; Rees, D. C.; Einsle, O. Evidence for Interstitial Carbon in Nitrogenase FeMo Cofactor. *Science* **2011**, *334* (6058), 940.

(17) Spatzal, T.; Perez, K. A.; Einsle, O.; Howard, J. B.; Rees, D. C. Ligand Binding to the FeMo-Cofactor: Structures of CO-Bound and Reactivated Nitrogenase. *Science* **2014**, *345* (6204), 1620−1623.

(18) Van Stappen, C.; Decamps, L.; Cutsail, G. E.; Bjornsson, R.; Henthorn, J. T.; Birrell, J. A.; DeBeer, S. The Spectroscopy of Nitrogenases. *Chem. Rev.* **2020**, *120* (12), 5005−5081.

(19) Jasniewski, A. J.; Lee, C. C.; Ribbe, M. W.; Hu, Y. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. *Chem. Rev.* **2020**, *120* (12), 5107−5157.

(20) Eady, R. R. Structure-Function Relationships of Alternative Nitrogenases. *Chem. Rev.* **1996**, *96* (7), 3013−3030.

(21) Lukoyanov, D. A.; Harris, D. F.; Yang, Z.-Y.; Pérez-González, A.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. The One-Electron Reduced Active-Site FeFe-Cofactor of Fe-Nitrogenase Contains a Hydride Bound to a Formally Oxidized Metal-Ion Core. *Inorg. Chem.* **2022**, *61* (14), 5459−5464.

(22) Van Stappen, C.; Thorhallsson, A. T.; Decamps, L.; Bjornsson, R.; DeBeer, S. Resolving the Structure of the E1 State of Mo Nitrogenase through Mo and Fe K-Edge EXAFS and QM/MM Calculations. *Chem. Sci.* **2019**, *10* (42), 9807−9821.

(23) Cao, L.; Caldararu, O.; Ryde, U. Protonation and Reduction of the FeMo Cluster in Nitrogenase Studied by Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations. *J. Chem. Theory Comput.* **2018**, *14* (12), 6653−6678.

(24) Dance, I. Survey of the Geometric and Electronic Structures of the Key Hydrogenated Forms of FeMo-Co, the Active Site of the Enzyme Nitrogenase: Principles of the Mechanistically Significant Coordination Chemistry. *Inorganics* **2019**, *7* (1), 8.

(25) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525−537.

(26) *Schrödinger Release 2023-2: MacroModel*; Schrödinger, LLC: New York, NY, 2023.

(27) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Forouzesh, N.; Giambaşu, G.; Giese, T.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, J.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shajan, A.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wang, J.; Wei, H.; Wu, X.; Wu, Y.; Xiong, Y.; Xue, Y.; York, D. M.; Zhao, S.; Zhu, Q.; Kollman, P. A. *Amber 2023*; University of California: San Francisco, 2023.

(28) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.

(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(30) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97* (4), 2571−2577.

(31) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401.

(32) Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11* (4), 431−439.

(33) Sigfridsson, E.; Ryde, U. Comparison of methods for deriving atomic charges from the electrostatic potential and moments. *J. Comput. Chem.* **1998**, *19* (4), 377−395.

(34) Hu, L.; Ryde, U. Comparison of Methods to Obtain Force-Field Parameters for Metal Sites. *J. Chem. Theory Comput.* **2011**, *7* (8), 2452−2463.

(35) Barney, B. M.; McClead, J.; Lukoyanov, D.; Laryukhin, M.; Yang, T.-C.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Diazene (HNNH) Is a Substrate for Nitrogenase: Insights into the Pathway of N2 Reduction. *Biochemistry* **2007**, *46* (21), 6784−6794.

(36) Wu, X.; Brooks, B. R. Self-Guided Langevin Dynamics Simulation Method. *Chem. Phys. Lett.* **2003**, *381* (3−4), 512−518.

(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81* (8), 3684−3690.

(38) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.

(39) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.

(40) Ryde, U.; Olsson, M. H. M. Structure, Strain, and Reorganization Energy of Blue Copper Models in the Protein. *Int. J. Quantum Chem.* **2001**, *81* (5), 335−347.

(41) Ryde, U. The Coordination of the Catalytic Zinc Ion in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput.-Aided Mol. Des.* **1996**, *10* (1), 153−164.

(42) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104* (8), 1720−1735.

(43) Dance, I. The Stereochemistry and Dynamics of the Introduction of Hydrogen Atoms onto FeMo-Co, the Active Site of Nitrogenase. *Inorg. Chem.* **2013**, *52* (22), 13068−13077.

(44) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (2), 91−100.

(45) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119* (23), 12129−12137.

(46) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J. Accurate and Numerically Efficient r2SCAN Meta-Generalized

Gradient Approximation. *J. Phys. Chem. Lett.* **2020**, *11* (19), 8208−8215.

(47) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098−3100.

(48) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785−789.

(49) Becke, A. D. A New Mixing of Hartree-Fock and Local Density-functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372−1377.

(50) Benediktsson, B.; Bjornsson, R. Analysis of the Geometric and Electronic Structure of Spin-Coupled Iron-Sulfur Dimers with Broken-Symmetry DFT: Implications for FeMoco. *J. Chem. Theory Comput.* **2022**, *18* (3), 1437−1457.

(51) Vysotskiy, V. P.; Torbjörnsson, M.; Jiang, H.; Larsson, E. D.; Cao, L.; Ryde, U.; Zhai, H.; Lee, S.; Chan, G. K.-L. Assessment of DFT Functionals for a Minimal Nitrogenase [Fe(SH)4H]- Model Employing State-of-the-Art Ab Initio Methods. *J. Chem. Phys.* **2023**, *159* (4), 044106.

(52) Zhai, H.; Lee, S.; Cui, Z.-H.; Cao, L.; Ryde, U.; Chan, G. K.-L. Multireference Protonation Energetics of a Dimeric Model of Nitrogenase Iron-Sulfur Clusters. *J. Phys. Chem. A* **2023**, in press, DOI: 10.1021/acs.jpca.3c06142.

(53) Cao, L.; Ryde, U. Influence of the Protein and DFT Method on the Broken-Symmetry and Spin States in Nitrogenase. *Int. J. Quantum Chem.* **2018**, *118* (15), No. e25627.

(54) Cao, L.; Caldararu, O.; Ryde, U. Protonation States of Homocitrate and Nearby Residues in Nitrogenase Studied by Computational Methods and Quantum Refinement. *J. Phys. Chem. B* **2017**, *121* (35), 8242−8262.

(55) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials. *Theor. Chim. Acta* **1997**, *97* (1−4), 119−124.

(56) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary Basis Sets to Approximate Coulomb Potentials. *Chem. Phys. Lett.* **1995**, *240* (4), 283−290.

(57) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A Generally Applicable Atomic-Charge Dependent London Dispersion Correction. *J. Chem. Phys.* **2019**, *150* (15), 154122.

(58) Schneider, K.; Müller, A. Iron-Only Nitrogenase: Exceptional Catalytic, Structural and Spectroscopic Features. In *Catalysts for Nitrogen Fixation: Nitrogenases, Relevant Chemical Models and Commercial Processes*; Smith, B. E., Richards, R. L., Newton, W. E., Eds.; *Nitrogen Fixation: Origins, Applications, and Research Progress*; Springer Netherlands: Dordrecht, 2004; pp 281−307..

(59) Krahn, E.; Weiss, B.; Kröckel, M.; Groppe, J.; Henkel, G.; Cramer, S.; Trautwein, A.; Schneider, K.; Müller, A. The Fe-Only Nitrogenase from Rhodobactercapsulatus: Identification of the Cofactor, an Unusual, High-Nuclearity Iron-Sulfur Cluster, by Fe K-Edge EXAFS and 57Fe Mössbauer Spectroscopy. *J. Biol. Inorg. Chem.* **2002**, *7* (1−2), 37−45.

(60) Lovell, T.; Li, J.; Liu, T.; Case, D. A.; Noodleman, L. FeMo Cofactor of Nitrogenase: A Density Functional Study of States M^N, M^{OX}, M^R, and M^I. *J. Am. Chem. Soc.* **2001**, *123* (49), 12392−12410.

(61) Greco, C.; Fantucci, P.; Ryde, U.; Gioia, L. D. Fast Generation of Broken-Symmetry States in a Large System Including Multiple Iron-Sulfur Assemblies: Investigation of QM/MM Energies, Clusters Charges, and Spin Populations. *Int. J. Quantum Chem.* **2011**, *111* (14), 3949−3960.

(62) Szilagyi, R. K.; Winslow, M. A. On the Accuracy of Density Functional Theory for Iron-Sulfur Clusters. *J. Comput. Chem.* **2006**, *27* (12), 1385−1397.

(63) Adams, P. D.; Pannu, N. S.; Read, R. J.; Brünger, A. T. Cross-Validated Maximum Likelihood Enhances Crystallographic Simulated Annealing Refinement. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94* (10), 5018−5023.

(64) Lundgren, K. J. M.; Veenman, E. A.; Caldararu, O.; Oksanen, E.; Ryde, U. Quantum Refinement in Real and Reciprocal Space. Manuscript in Preparation.

(65) Grosse-Kunstleve, R. W.; Sauter, N. K.; Moriarty, N. W.; Adams, P. D. The Computational Crystallography Toolbox: Crystallographic Algorithms in a Reusable Software Framework. *J. Appl. Crystallogr.* **2002**, *35* (1), 126−136.

(66) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix. *Acta Crystallogr., Sect. D: Struct. Biol.* **2019**, *75* (10), 861−877.

(67) Neese, F. Software Update: The ORCA Program System-Version 5.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2022**, *12* (5), No. e1606.

(68) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152* (22), 224108.

(69) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. Electronic Ligand Builder and Optimization Workbench (eLBOW): A Tool for Ligand Coordinate and Restraint Generation. *Acta Crystallogr., Sect. D: Struct. Biol.* **2009**, *65* (10), 1074−1080.

(70) Tickle, I. J. Statistical Quality Indicators for Electron-Density Maps. *Acta Crystallogr., Sect. D: Struct. Biol.* **2012**, *68* (4), 454−467.

(71) Yang, Z.-Y.; Jimenez-Vicente, E.; Kallas, H.; Lukoyanov, D. A.; Yang, H.; Martin del Campo, J. S.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. The Electronic Structure of FeV-Cofactor in Vanadium-Dependent Nitrogenase. *Chem. Sci.* **2021**, *12* (20), 6913−6922.

(72) Benediktsson, B.; Bjornsson, R. QM/MM Study of the Nitrogenase MoFe Protein Resting State: Broken-Symmetry States, Protonation States, and QM Region Convergence in the FeMoco Active Site. *Inorg. Chem.* **2017**, *56* (21), 13417−13429.

(73) Bjornsson, R.; Neese, F.; DeBeer, S. Revisiting the Mössbauer Isomer Shifts of the FeMoco Cluster of Nitrogenase and the Cofactor Charge. *Inorg. Chem.* **2017**, *56* (3), 1470−1477.

(74) Deng, L.; Wang, H.; Dapper, C. H.; Newton, W. E.; Shilov, S.; Wang, S.; Cramer, S. P.; Zhou, Z.-H. Assignment of Protonated R-Homocitrate in Extracted FeMo-Cofactor of Nitrogenase via Vibrational Circular Dichroism Spectroscopy. *Commun. Chem.* **2020**, *3* (1), 145−147.

(75) Jin, W.-T.; Wang, H.; Wang, S.-Y.; Dapper, C. H.; Li, X.; Newton, W. E.; Zhou, Z.-H.; Cramer, S. P. Preliminary Assignment of Protonated and Deprotonated Homocitrates in Extracted FeMo-Cofactors by Comparisons with Molybdenum(IV) Lactates and Oxidovanadium Glycolates. *Inorg. Chem.* **2019**, *58* (4), 2523−2532.

(76) Wang, S.-Y.; Jin, W.-T.; Chen, H.-B.; Zhou, Z.-H. Comparison of Hydroxycarboxylato Imidazole Molybdenum(IV) Complexes and Nitrogenase Protein Structures: Indirect Evidence for the Protonation of Homocitrato FeMo-Cofactors. *Dalton Trans.* **2018**, *47* (22), 7412−7421.

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.

Computational Chemistry
Department of Chemistry
Faculty of Science

LUND UNIVERSITY