Structural and biochemical characterisation of N_2-assimilation system and its regulation in thermophilic methanogenic archaea

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

Dr. rer. nat.

dem Fachbereich 2 Biologie/Chemie

der Universität Bremen

vorgelegt von

Nevena Maslać

Bremen, August 2023

Die vorliegende Doktorarbeit wurde in der Zeit von Juni 2019 bis August 2023 im Rahmen des Programms „International Max Planck Research School for Marine Microbiology: MarMic“, in den Abteilung Microbielle Metabolismen am Max Planck Institut für Marine Mikrobiologie in Bremen angefertigt.

This thesis was prepared in the framework of the „International Max Planck Research School of Marine Microbiology: MarMic“ in the Microbial Metabolism group at the Max Planck Institute for Marine Microbiology from June 2019 to August 2023.

Gutachter: Dr. Tristan Wagner

Gutachter: Prof. Dr. Ross D. Milton

Gutachter: Dr. Marco Bellinzoni

Prüfer: Prof. Dr. Jan-Hendrik Hehemann

Prüfer: Prof. Dr. Rudolf Amann

Datum des Promotionskolloquiums: 10. 08. 2023

"Insight must precede application"

-Max Planck

Nevena Maslać

ORCID: 0000-0002-8961-48

Contents

Summary

Nitrogen gas can only be biologically reduced to ammonia by diazotrophic microorganisms. Among these microbes, methanogens and anaerobic methanotrophic archaea represent the sole known archaeal diazotrophs, and they have been found to actively participate in nitrogen fixation within anaerobic environments such as marine and marsh sediments. The nitrogenase enzyme, catalysing the reduction, also serves as a model for developing effective and sustainable catalysts for industrial ammonia production in the Haber-Bosch process. However, numerous questions regarding the molecular mechanisms underlying this catalysis remain unanswered, and the structure of archaeal nitrogenases, proposed to be more ancient than bacterial, remains elusive.

In **Chapter 2**, we investigated the adaptation and cellular response of the thermophile *Methanothermococcus thermolithotrophicus* to N_2-fixation using differential transcriptomics. Our findings reveal that *M. thermolithotrophicus* would employ multiple pathways for nitrogen acquisition and a targeted response for energy saving through the downregulation of genes involved in transcription and translation. **Chapter 3** details the structures of nitrogenase reductase from three marine *Methanococcales.* Notably, the structures obtained from *M. thermolithotrophicus* and *Methanocaldococcus infernus* represent the first natively purified and solved nitrogenase reductases from methanogens. In **Chapter 4**, we present the first high-resolution structure of archaeal nitrogenase from *M. infernus* in both resting and mixed turnover states, supporting the existence of a common reduction mechanism shared by all nitrogenases. Additionally, we unveiled the molecular basis of the posttranslational regulation of archaeal N_2-fixation by characterising an inhibitory closed circular complex mediated by P_{II} family proteins.

To address two long-standing questions related to molybdenum nitrogenase, **Chapters 5 and 6** combine structural biology with the robust genetic system of *Azotobacter vinelandii*. By creating nitrogenase variants, we clarified the functional and structural role of an additional metal binding site at a dimeric protein interface and demonstrated negative cooperativity in the Mo-dependent nitrogenase.

In summary, this thesis delves into characterisation of diazotrophic microbes on multiple molecular levels, with an emphasis on nitrogen assimilation systems in (hyper)thermophilic methanogens. While shedding light on the intricate mechanisms underlying these processes, it offers valuable insights into nitrogen fixation in these ancient organisms.

Zusammenfassung

Stickstoffgas kann nur von diazotrophen Mikroorganismen biologisch zu Ammoniak reduziert werden. Unter diesen Mikroben sind Methanogene und anaerobe methanotrophe Archaeen die einzigen bekannten archaealen Diazotrophen, die nachweislich aktiv an der Stickstofffixierung in anaeroben Umgebungen wie Meeres- und Sumpfsedimenten beteiligt sind. Das Enzym Nitrogenase, das diese Reduktion katalysiert, dient auch als Modell für die Entwicklung effizienter und nachhaltiger Katalysatoren für die industrielle Ammoniakproduktion nach dem Haber-Bosch-Verfahren. Jedoch bleiben zahlreiche Fragen zu den molekularen Mechanismen, die dieser Katalyse zugrunde liegen, unbeantwortet, und die Struktur der archaealen Nitrogenasen, die vermutlich älter als die bakteriellen Nitrogenasen sind, ist immer noch unbekannt.

In **Kapitel 2** untersuchten wir die Anpassung und zelluläre Reaktion des thermophilen *Methanothermococcus thermolithotrophicus* an die N_2-Fixierung mittels differentieller Transkriptomik. Unsere Ergebnisse zeigen, dass *M. thermolithotrophicus* mehrere Wege zur Stickstoffakquisition und eine spezifische Reaktion zur Energieeinsparung durch Herunterregulierung von Transkriptions- und Translationsgenen nutzt. **Kapitel 3** beschreibt die Strukturen von Nitrogenase-Reduktasen aus drei marinen *Methanococcales*. Es ist hervorzuheben, dass die Strukturen von *M. thermolithotrophicus* und *Methanocaldococcus infernus* die ersten nativ gereinigten und gelösten Nitrogenase-Reduktasen aus Methanogenen sind. In **Kapitel 4** präsentieren wir die erste hochauflösende Struktur der archaealen Nitrogenase von *M. infernus* sowohl im Ruhe- als auch im gemischten Umsatzzustand, die die Existenz eines gemeinsamen Reduktionsmechanismus aller Nitrogenasen bestätigt. Darüber hinaus haben wir die molekulare Basis der posttranslationalen Regulation der archaealen N_2-Fixierung aufgedeckt, indem wir einen inhibierenden geschlossenen zirkulären Komplex charakterisiert haben, der durch Proteine der P_{II}-Familie vermittelt wird.

Um zwei seit langem bestehende Fragen zur Molybdän-Nitrogenase zu beantworten, werden in den **Kapiteln 5 und 6** strukturbiologische Erkenntnisse mit dem robusten genetischen System von *Azotobacter vinelandii* kombiniert. Durch die Erzeugung von Nitrogenase-Varianten konnten wir die funktionelle und strukturelle Rolle einer zusätzlichen Metallbindungsstelle an einer dimeren Proteinschnittstelle aufklären und negative Kooperativität in der Mo-abhängigen Nitrogenase nachweisen.

Zusammenfassend befasst sich diese Arbeit mit der Charakterisierung diazotropher Mikroben auf mehreren molekularen Ebenen, wobei der Schwerpunkt auf Stickstoffassimilationssystemen in (hyper)thermophilen Methanogenen liegt. Sie beleuchtet die komplexen Mechanismen, die diesen Prozessen zugrunde liegen, und liefert wertvolle Einblicke in die Stickstofffixierung in diesen urzeitlichen Organismen.

Abbreviations

P_{450}NOR	NADH-dependent cytochrome P_{450} nitric oxide reductase
2-OG	2-oxoglutarate
ADP	Adenosine di-phosphate
AMO	Ammonia monooxygenase
Anammox	anaerobic ammonium oxidation
anf genes	alternative nitrogen fixation genes
ANME	anaerobic methanotrophic archaea
ATP	Adenosine tri-phosphate
Cas	CRISPR associated protein
ccNIR	dissimilatory periplasmic cytochrome c nitrite reductase
cd1-NIR	haem-containing nitrite reductase
c-di-GMP	cyclic dimeric guanosine monophosphate
CFCs	Chlorofluorocarbons
cNIR	assimilatory nitrite reductase
cNOR	cytochrome c-dependent nitric oxide reductase
CRISPR	clustered regularly interspaced short palindromic repeats
Cu_ANOR	copper-containing quinol-dependent nitric oxide reductase
Cu-NIR	copper-containing nitrite reductase
CYN	cyanase
DNAse	Deoxyribonuclease
DNRA	dissimilatory nitrate reduction
DRAG	dinitrogenase reductase-activating glycohydrolase
DRAT	dinitrogenase reductase ADP-ribosyltransferase
DT	dithionite
DTB	desthiobiotin
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance spectroscopy
FAD	Flavin adenine dinucleotide
FC	fold change
FeFe nitrogenase	iron only nitrogenase
FeMoco	FeMo cofactor
FeP	iron protein
fix genes	N_2 fixation genes
HAO	hydroxylamine oxidoreductase
HB	Haber-Bosch
HCP	hybrid cluster protein
HDH	hydrazine dehydrogenase
HIC	hydrophobic exchange
HOX	hydroxylamine oxidase
hrCN PAGE	high resolution clear native PAGE
HZS	hydrazine synthase

IEX	ionic exchange
IPTG	Isopropyl-β-d-thiogalactopyranoside
LUCA	Last Universal Common Ancstor
MMB	mononuclear metal binding site
mod genes	molybdenum transport genes
MoFe nitrogenase	molybdnum nitrogenase
mos genes	molybdenum storage genes
mRNA	messenger RNA
NAD	Nicotinamide adenine dinucleotide
naf genes	nitrogenase associated factors genes
NAP	periplasmic dissimilatory nitrate reductase
NAR	membrane-bound dissimilatory nitrate reductase
NAS	assimilatory nitrate reductase
nfa genes	nitrogen fixation associated genes
nif genes	nitrogen fixation genes
Ni-NTA	Ni-NTA Agarose
NOD	nitric oxide oxidase
NO-D	nitric oxide dismutase
NORvw	flavo-diiron nitric oxide reductase
NOS	nitrous oxide reductase
NXR	nitrite oxidoreductase
ONR	octahaem nitrite reductase
OTR	octahaem tetrathionate reductase
padj	adjusted p value
PBS	phosphate buffer slaine
PCA	principal component analysis
PHAs	Polyhydroxyalkanoates
pLDDT	per-residue confidence score
pMMO	particulate methane monooxygenase
qNOR	quinol-dependent nitric oxide reductase
re/*oa* equilibrium	reductive elimination/oxidative addition equilibrium
RNA	Ribonucleic acid
*rnf*genes	Rhodobacter nitrogen fixation genes
SDS-PAGE	sodium dodecyl sulfate poly(acrylamide) gel electrophoresis
sRNA	small RNA
TEW	polyoxotungstate
TL model	Thorneley-Lowe model
TPM	trancripts per million
tRNA	transport RNA
URE	urease
XFELs	X-Ray Free-Electron Lasers
VFe nitrogenases	vandium nitrogenase
vnf genes	vanadium nitrogen fixation genes
vod genes	vanadium transport genes
εHAO	ε-hydroxylamine oxidoreductase

CHAPTER 1: Introduction

1.1. The century of industrial nitrogen fixation: How are humans changing the nitrogen cycle on planet Earth? While the carbon cycle is the most often associated with climate change and environmental issues, the nitrogen cycle has been the most significantly altered by human activities (1-4). Nitrogen is an essential element for Life, as it composes biomolecules, including nucleic acids and proteins (4). Ironically, nitrogen on Earth is both overabundant and very limited: although the atmosphere itself is the largest nitrogen reservoir on the planet ($\sim 4 \times 10^9$ Tg N), being made of 78.1% dinitrogen gas (N_2) (5), nitrogen is the limiting nutrient in most of the terrestrial environments (6-8). This is because inert N_2 is not readily bioavailable and must be converted into more reactive forms (Nr) for use by living organisms (4, 5). These include, among others, most importantly, reduced nitrogen (NH_x), nitrogen oxides (NO_x), nitrous oxide (N_2O), and nitric acid (HNO_3) (4). The reduction of N_2 to ammonia (NH_3) is performed exclusively by *Bacteria* and *Archaea*, commonly referred to as diazotrophs. Microorganisms are the primary drivers of the other reactions in the nitrogen transformation network as well: they can transform nitrogen from the most reduced to the most oxidised states (-3 to +5) (Figure 1A) (5, 9). Once fixed, NH_3 can be sequentially oxidised to nitrite (NO_2^-) and nitrate (NO_3^-), which in return can be reduced back through dissimilatory nitrate reduction (DNRA) (Figure 1). Release of the fixed nitrogen can be done by denitrification (the reduction of NO_3^- to N_2) and anaerobic ammonium oxidation (the oxidation of NH_4^+ to N_2 using NO_2^-, anammox) (Figure 1B) (9). The simplified overview displayed in Figure 1 is enough to highlight the key differences between the nitrogen cycle and other major nutrient cycles. Notably, the nitrogen cycle stands apart from other cycles as its primary inputs and outputs are dominated by biological processes (5).

Until the 20th century, nitrogen fixation by microorganisms, along with a minor contribution from geochemical discharge through lightning, was the only way through which the nitrogen from the atmospheric pool could enter the bioavailable pool (10, 11). This changed with the invention of the Haber-Bosch (HB) process for industrial ammonia production (12-14). In just over a century since its invention, the HB process has dramatically transformed human population growth (through agricultural fertilisers) and the nitrogen cycle (3, 15). As accounted for, 40% of the world's population relies on nitrogen-based fertilisers for crop growth and survival (10).

Anthropogenic production of reactive nitrogen species through the HB process, when combined with other human activities like the expansion of legume cultivation, poor waste

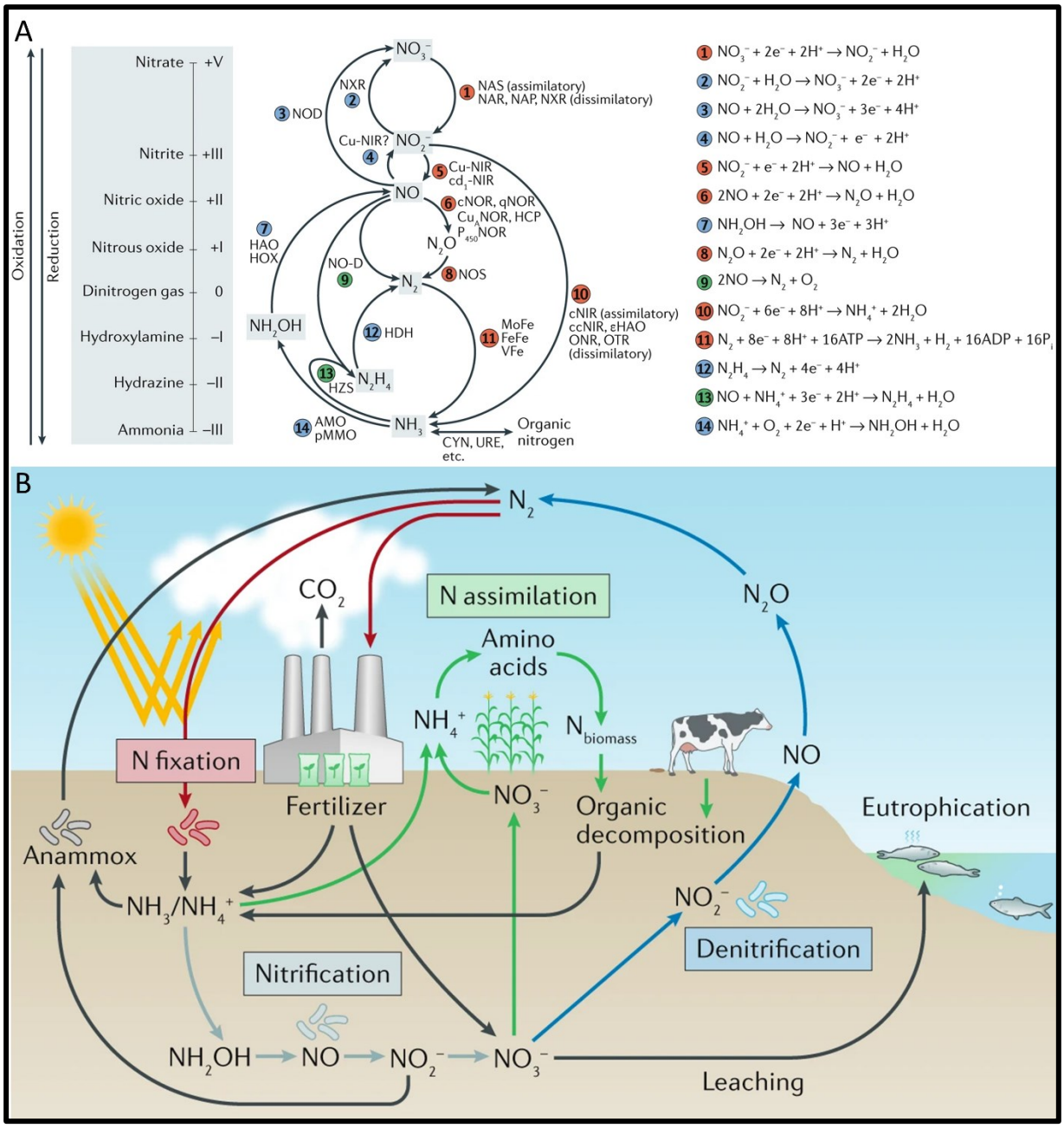


Figure 1. Microbial transformations and cycling of the main nitrogen species. A. Interconversion of key nitrogen species by microorganisms. Red, blue, and green circles correspond to reduction, oxidation, and disproportionation/comproportionation reactions, respectively. Abbreviated enzymes: assimilatory nitrate reductase (NAS, *nasA* and *nirA*); membrane-bound (NAR, *narGH*) and periplasmic (NAP, *napA*) dissimilatory nitrate reductases; nitrite oxidoreductase (NXR, *nxrAB*); nitric oxide oxidase (NOD, *hmp*); haem-containing (cd1-NIR, *nirS*) and copper-containing (Cu-NIR, *nirK*) nitrite reductases; cytochrome c-dependent (cNOR, *cnorB*), quinol-dependent (qNOR, *norZ*) and copper-containing quinol-dependent nitric oxide reductases (Cu_ANOR); NADH-dependent cytochrome P_{450} nitric oxide reductase (P_{450}NOR, *p450nor*); flavo-diiron nitric oxide reductase (NORvw, *norVW*); hybrid cluster protein (HCP, *hcp*); hydroxylamine oxidoreductase (HAO, *hao*); hydroxylamine oxidase (HOX; *hox*); nitrous oxide reductase (NOS, *nosZ*); nitric oxide dismutase (NO-D, *norZ*); assimilatory nitrite reductase (cNIR; *nasB* and *nirB*); dissimilatory periplasmic cytochrome c nitrite reductase (ccNIR, *nrfAH*); ε-hydroxylamine oxidoreductase (εHAO; *haoA*); octahaem nitrite reductase (ONR);

octahaem tetrathionate reductase (OTR); molybdenum-iron (MoFe, *nifHDK*), iron-iron (FeFe, *anfHGDK*) and vanadium-iron (VFe, *vnfHGDK*) nitrogenases; hydrazine dehydrogenase (HDH, *hdh*); hydrazine synthase (HZS, *hzsCBA*); ammonia monooxygenase (AMO, *amoCAB*); particulate methane monooxygenase (pMMO, *pmoCAB*); cyanase (CYN, *cynS*); and urease (URE, *ureABC*). **B.** NH_3 enters the nitrogen cycle by either biological or industrial N_2 reduction (red arrows), and it is supplied to the plants mostly in the form of fertiliser (green arrows). Microorganisms metabolise excess NH_3 in the soil which can be further nitrified (grey arrows). Finally, denitrification and leaching of NO_3^- (blue arrows) are the main source of the reactive nitrogen species, which are environmental pollutants. / Panel A of this figure is reprinted with permission granted under the licence ID 5543110417503 from Kuypers *et al.* (2018) (9). Panel B of this figure is reprinted with permission granted under the licence ID 5543030242070 from Lehnert *et al.* (2018) (16).

management during food production, and consumption and burning of fossil fuels, now surpasses the natural process of biological nitrogen fixation (4), resulting in a significant increase of bioavailable nitrogen input into terrestrial and coastal ecosystems. The amount of nitrogen cycling globally has more than doubled in the past century, and as a consequence of this massive perturbation, a variety of environmental problems emerged: from aquatic eutrophication and the development of the "dead zones" (17-19), soil acidification (20, 21), loss of biodiversity (22-24) and air pollution, to the greenhouse effect and stratospheric ozone layer damage from rising nitrous oxide (N_2O) levels (25). Current estimations are that the fields in industrialised agricultural systems are routinely over-fertilized: plants assimilate between 30 and 50% of all nitrogen fertiliser added to the soil (26), with the rest being oxidised by microbes to nitrite and nitrate. Not being retained in the soil, they are washed off to aquatic pools where they cause eutrophication (Figure 1B), such as, for example, the infamous dead zone in the Gulf of Mexico (27). Soil acidification results from increased rates of nitrogen cycling and higher NO_3^- leaching, which causes the loss of base cations, greater solubility of metals in soil solution, and decreased pH (28). The soil is also the primary source of N_2O release due to the excess Nr species reduction (25). N_2O is a very potent greenhouse gas with a long atmospheric lifetime, which allows it to accumulate in the stratosphere, where it can contribute to the destruction of the ozone layer. With chlorofluorocarbons (CFCs) being largely banned due to their role in ozone depletion, N_2O is now considered the main anthropogenic contributor to ozone depletion (25). Multiple studies have confirmed that ecosystem recovery, even if nitrogen input is reduced over a prolonged time, happens very slowly or not at all (28).

While these threats to biodiversity, health, and climate are the reason for concern, the global demand for food remains a pressing issue. To meet these demands, we are currently producing ~ 200 M tonnes of ammonia per year, with a predicted increase to ~ 270 M tonnes

by 2050 (29). Ensuring a sustainable food supply for the future while minimising environmental degradation and managing the impact of human activities on the global nitrogen cycle presents a significant challenge for sustainable development (30). Many environmental agencies and organisations have recognised this as the next major environmental issue, and this decade has been called the decade of sustainable nitrogen management (25).

In addition to implementing policies to reduce gas emissions from soils and water systems and developing better pollution and nitrogen waste management strategies, it is also crucial to advance fundamental research to tackle this issue. These efforts can be broadly grouped into three main directions of investigation: the development of a new generation of synthetic catalysts based on the chemical mechanism of biological nitrogen fixation (31), transferring the capacity for biological nitrogen fixation from microbes to plants (32, 33) and engineering the more efficient delivery of fixed nitrogen from microbes to plants (34, 35).

This thesis follows the first avenue of research, introducing (hyper)thermophilic methanogens as the new model organisms better suited for potential industrial applications. With the main focus of unveiling the particular properties of their nitrogen fixation machinery through structural analysis, we also aim to understand how the nitrogen fixation process intertwines with other metabolic pathways and the management of energy expenditure.

1.2. Nitrogenase: "The ultimate Swiss Army knife for reductive biocatalysis." Although the overall reaction of ammonia synthesis is thermodynamically favourable, extreme reaction conditions are necessary for the HB process because the uncatalysed reaction under ambient conditions is highly unfavourable (Figure 2) (36). It is why the energy requirements of the HB process are very high: it consumes up to 5% of the world's annual natural gas and about 2% of energy production to generate enough hydrogen and enough heat for the reaction (37). Since 95% of the hydrogen used for HB is derived from fossil fuels, 1.3% of global annual CO_2 emissions are attributed to it. Given the sensitivity of the iron or ruthenium-based catalysts used in the HB process, where hydrogen and nitrogen are reacted at high temperatures (400-500°C) and pressures (up to 200 bar) (37, 38) to gas impurities, sintering, and agglomeration, it is crucial to investigate how microbes can achieve efficient catalysis under ambient conditions.

In nature, diazotrophic microorganisms reduce N_2 to NH_3 via nitrogenase. The enzyme is able to activate one of the most thermodynamically stable chemical bonds: the triple bond in the inert molecule of N_2 (941 kJ.mol^{-1}) (39, 40). Nitrogenase is an impressive N_2-fixation catalyst

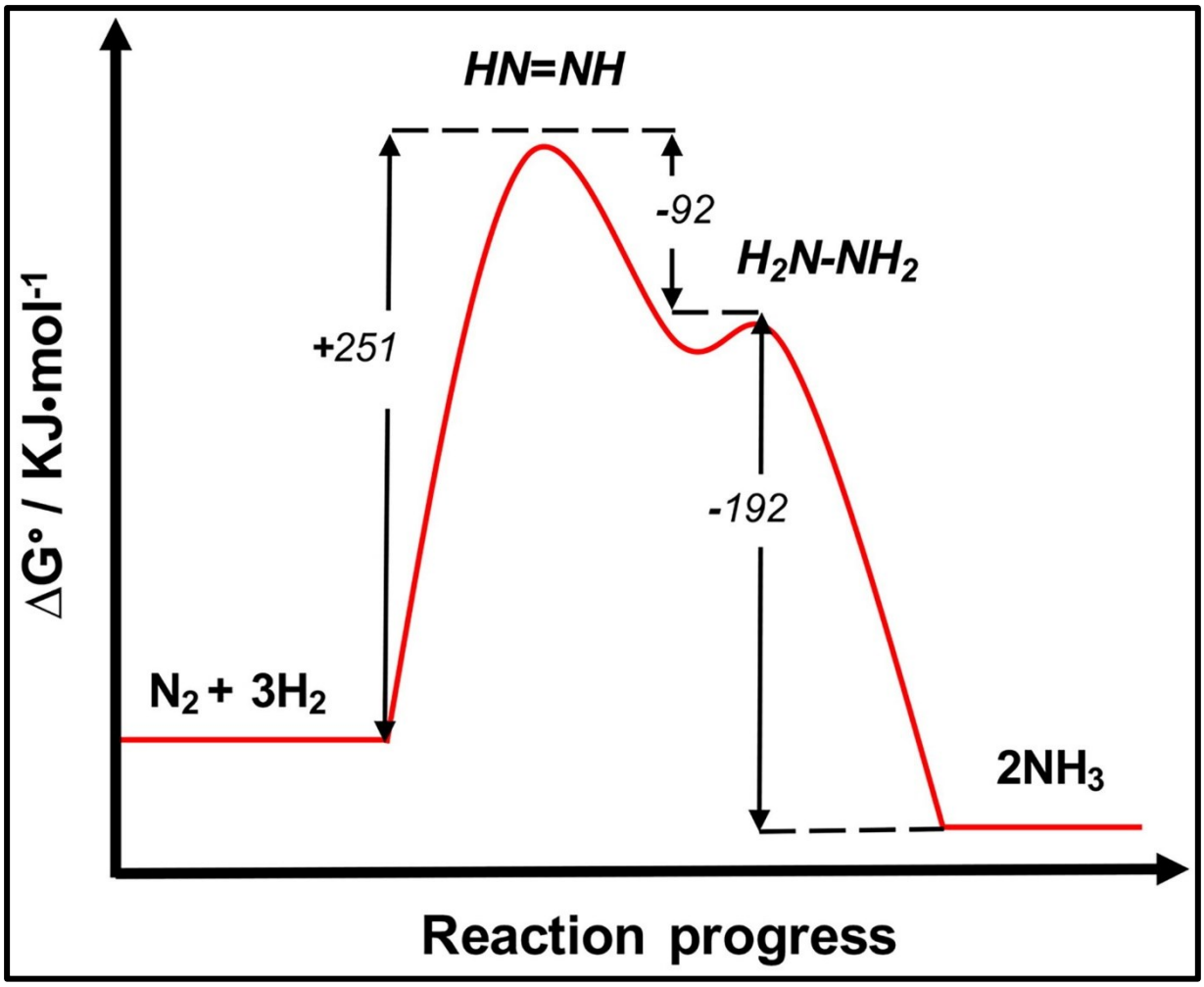


Figure 2. Standard Gibbs free energy changes during the reaction of N_2 reduction to NH_3. While the overall reaction of N_2 reduction to NH_3 is thermodynamically favourable, the formation of the first intermediate, diazene (HN=NH), from N_2 is endergonic and highly unfavourable. The main role of nitrogenase as a catalyst is the reduction of this activation energy and stabilisation of the transition state. / This figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*.

whose properties could be transposed for biomimicry (41-51).

Nitrogenase is a metalloprotein with three known isoforms, each consisting of two subunits:

- The dinitrogenase, also referred to as MoFe- (NifDK, molybdenum nitrogenase), VFe- (VnfDKG, vanadium nitrogenase), or FeFe- (AnfDKG, iron only nitrogenase) protein, depending on the isoform, or component I in early works.

- The dinitrogenase reductase, also called Fe-protein, because iron is the only metal in its clusters. The protein names are NifH, VnfH, and AnfH for the MoFe, VFe, or Fe-only isoforms, respectively. This subunit of the nitrogenase complex was also named component II in early works (Figure 3A, 3B) (40, 52).

The Molybdenum nitrogenase is the most common isoenzyme of nitrogenase, found in all

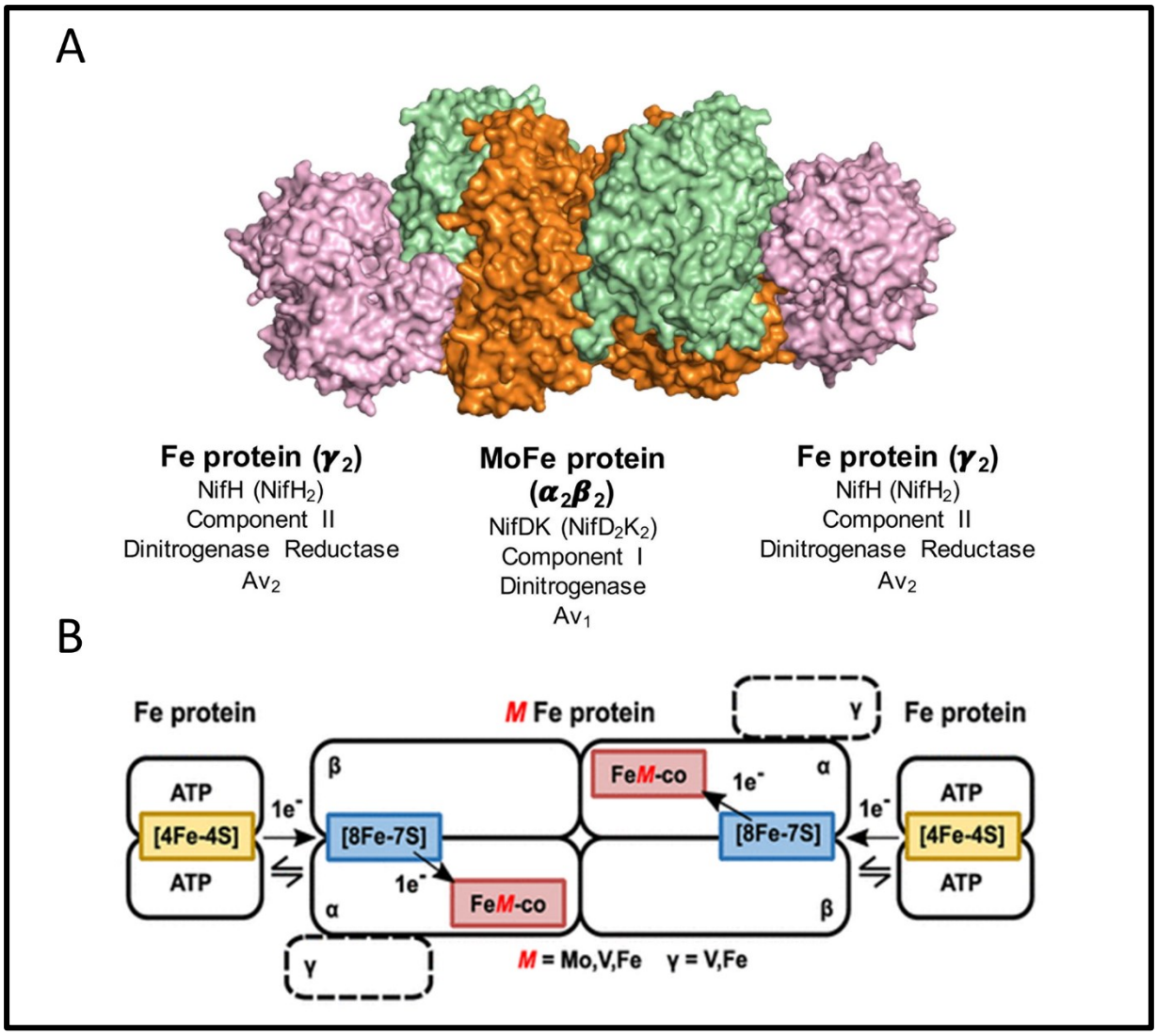


Figure 3. Nitrogenase enzyme organisation. A. Surface representation of the fully assembled nitrogenase complex and the overview of subunit nomenclature commonly used in literature. **B.** The scheme of the electron flow through nitrogenase enzyme. Homodimeric nitrogenase reductase (Fe protein) is an obligate electron donor to the heterotetrameric (2α2β) nitrogenase (M-Fe protein). Fe protein contains one [4Fe-4S] cluster. M-Fe protein contains one [8Fe-7S] cluster and one FeM cluster, where M=Mo, V, Fe. Nitrogenase isoforms containing FeV of FeFe cofactor are heterohexamers (2α2β2γ), containing an additional G subunit. / Panel A of this figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*. Panel B of this figure is reprinted with permission from Seefeldt LC, Yang ZY, Lukoyanov DA, Harris DF, Dean DR, Raugei S, Hoffman BM. 2020. Reduction of Substrates by Nitrogenases. Chemical Reviews 120:5082-5106 (52). Copyright 2020 *American Chemical Society*.

known diazotrophs. Some organisms additionally possess one or both "alternative" vanadium and iron-only isoforms. For example, all three nitrogenases are found in *Azotobacter vinelandii* (40), *Methanosarcina acetivorans* (53), and *Rhodopseudomonas palustris* (54). *Rhodobacter* species harbor the genes coding for Mo- and Fe-only nitrogenase in their genomes (55), while *Anabaena* species have Mo- and V-nitrogenases (56).

The architecture of Fe protein (FeP) in the three nitrogenase isoforms is highly conserved (Figure 4), with all the structural features described for NifH being present in VnfH and AnfH as well, including the structure and the position of the [4Fe-4S] cluster (57-59). FeP is a homodimer (2γ ~ 62 kDa) with a two-fold symmetry whose axis passes through one redox-active $[4Fe\text{-}4S]^{1+/2+}$ cluster at the interface of the two monomers. Each monomer contains an ATPMg binding site (39, 57, 60). It belongs to the P-loop NTPases, nucleotide-dependent switch proteins, and retains all three structural motifs characteristic for this protein family: P-loop (Walker A motif), switch I, and switch II (Figure 5) (61-64). The P-loop and switch I have a role in the coordination of the bound nucleotides and Mg^{2+}, while the switch II coordinates [4Fe-4S] cluster, and all three go through conformational changes upon ligand binding (Figure 5). The [4Fe-4S] cluster is coordinated by two cysteine residues (Cys97 and Cys132 in the case of *A. vinelandii* NifH) positioned on the surface of the protein (57, 65). The exposition of the cluster to the solvent is finely tuned by the conformational changes induced by nucleotide binding, which plays a critical role in electron transfer. In contrast to the canonical Cys-ligated [4Fe-4S] clusters of low-potential ferredoxins, which can access only one redox couple state under physiological conditions, the [4Fe-4S] cluster from NifH can access two different redox couple states: $[4Fe\text{-}4S]^{2+}/[4Fe\text{-}4S]^{1+}$ and $[4Fe\text{-}4S]^{1+}/[4Fe\text{-}4S]^{0}$ (60, 65). While $[4Fe\text{-}4S]^{2+}/[4Fe\text{-}4S]^{1+}$ couple is considered relevant for catalysis under physiological conditions, the possibility of reaching the $[4Fe\text{-}4S]^{1+}/[4Fe\text{-}4S]^{0}$ couple opens up the question of whether NifH can also act as a two-electron donor (discussed in more detail in section 1.4.) (60, 65). The redox potential of *A. vinelandii* NifH $[4Fe\text{-}4S]^{2+}/[4Fe\text{-}4S]^{1+}$ couple is ~ -300 mV in the apo state, and shifts to -430 mV upon MgATP binding (65). During catalysis, NifH sequentially transfers electrons to NifDK in an ATP-dependent manner, and it is the only known biological donor of electrons for dinitrogenase. In addition to supplying NifDK with electrons, NifH also has a role in the maturation of its metallocofactors (66). While the involvement of VnfH and AnfH in cofactor maturation remains empirically proven, it is assumed that they share analogous functions with NifH (67).

The overall structure of the dinitrogenase isoforms (*i.e.*, MoFe, VFe, and Fe-only) is also highly conserved (Figure 6) but nevertheless includes three notable differences. First, vanadium and Fe-only nitrogenases have an additional metallocofactor-free subunit named VnfG and AnfG, respectively. Both are bound to the subunit D (*i.e.*, VnfD/AnfD), making these dinitrogenases heterohexamer (2α2β2γ) (68, 69). The second significant structural difference is the presence of an extended N-terminus in NifK and an extended C-terminus in AnfD occupying a similar position on the protein surface. Both extensions are absent in VnfK

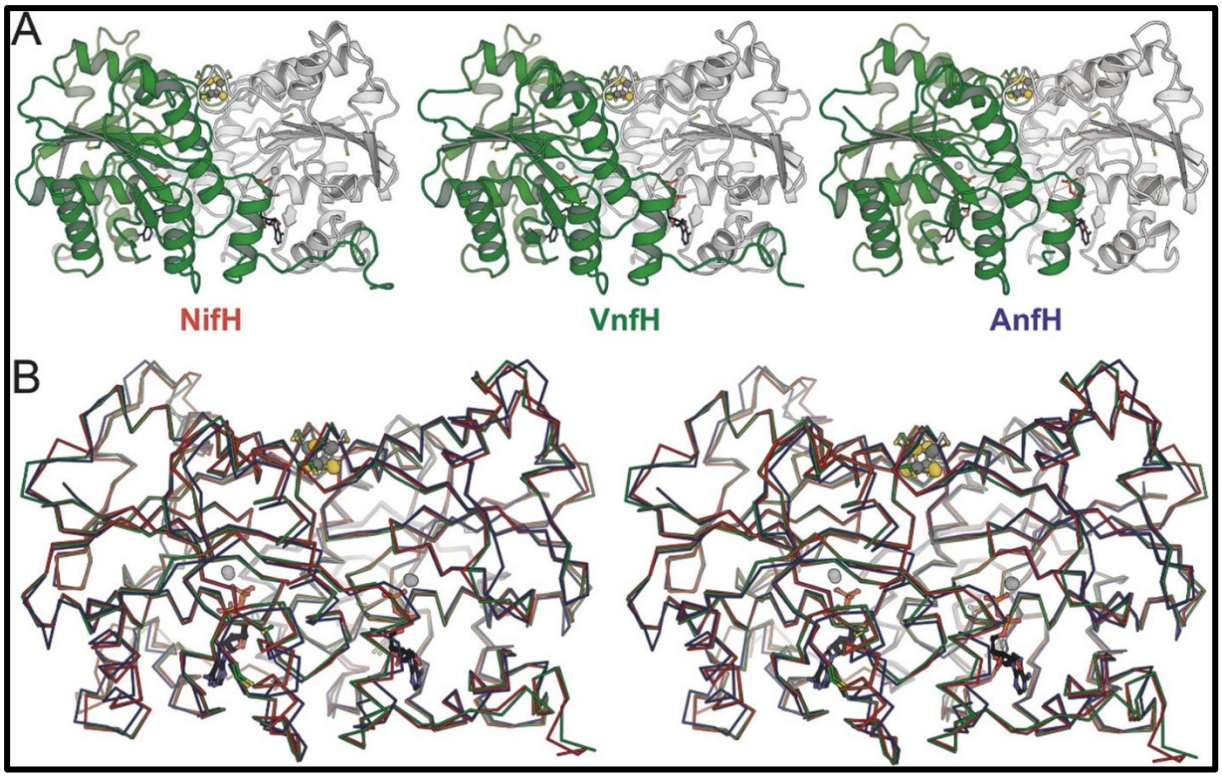


Figure 4. 3D structure and comparison of the three dinitrogenase reductase isoforms from *A. vinelandii*. **A.** NifH (Mo-nitrogenase, PDB 1NIP), VnfH (V-nitrogenase, PDB 6Q93) and AnfH (Fe-nitrogenase, PDB 7QQA) structures represented as cartoons. **B.** Superposition of the three nitrogenase reductase isoforms depicted in panel A represented as ribbons in stereo view. / This figure is reprinted with permission from Trncik *et al.* (2022) (59). Copyright 2022 *Elsevier.*

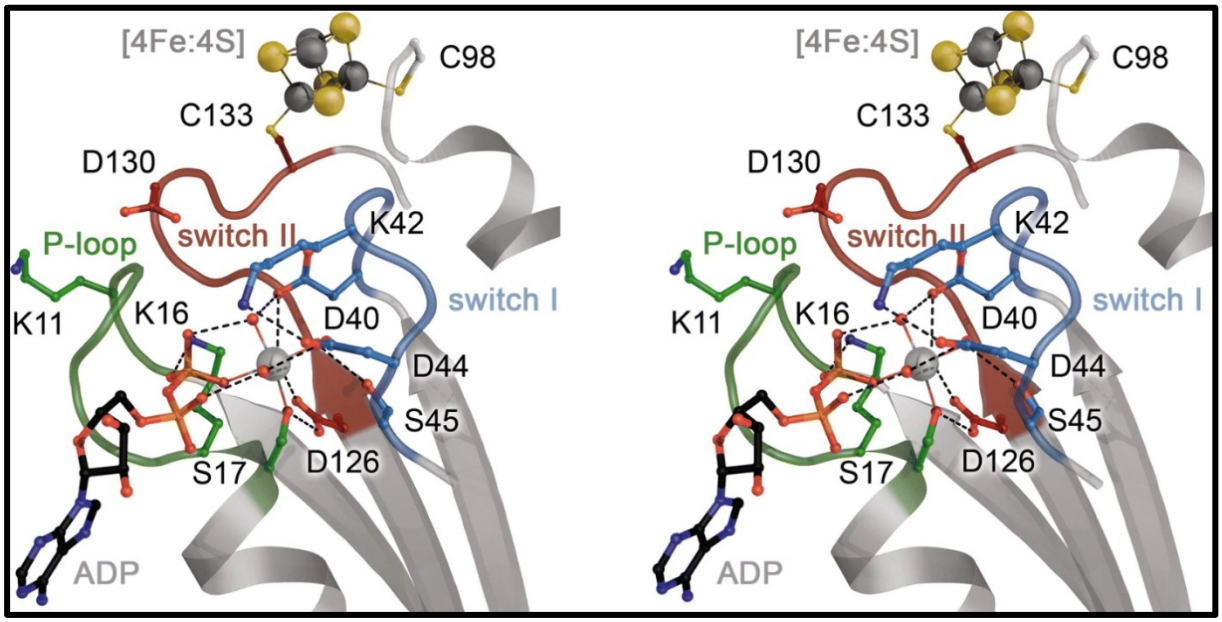


Figure 5. Nucleotide binding site and the flexible P-loop, switch I and switch II regions in the VnfH of *A. vinelandii*. The stereo figure displays the protein as cartoon (grey) and the [4Fe-4S] cluster and bound ligands as balls and sticks. The binding of ATP/ADP causes the conformational changes in the switch I (blue), switch II (red) and P-loop (green) regions, leading to the repositioning of the [4Fe-4S] cluster. / This figure is reprinted with permission granted under the licence ID 5565830949668 from Rohde *et al.* (2018) (58).

and replaced by a single β-sheet (68, 69). Finally, while the two NifK subunits bind a metal ion at their dimeric interface, traditionally modelled as Ca^{2+} (70), in all so far described structures of vanadium dinitrogenase, this site is occupied by Mg^{2+} (68, 71-73). However, it was also observed that this site, commonly referred to as the mononuclear metal binding site (MMB site, detailed in Chapter 5), could also be occupied by Fe, opening the debate about the structural, functional, and mechanistic involvement of this 16^{th} Fe atom in dinitrogenase (74).

NifDK protein is a heterotetramer (2α2β, ~230 kDa) containing two copies of both NifD (α-subunit) and NifK (β-subunit) (40, 52). NifD and NifK have three globular domains organised as a Rossmann fold (βαβ)(60). C terminal loops of these β sheets interact with the two harboured metalloclusters: [Mo:7Fe:9S:C]-homocitrate FeMoco (M cluster) and [8Fe-7S] P cluster (75, 76). To the best of our knowledge so far, both of these metalloclusters are completely unique to the nitrogenase enzyme (39, 60) and are thought to be among the largest and most complex enzymatic metallocofactors.

The [8Fe-7S] P cluster consists of two symmetric [4Fe-4S] cubanes that share a central μ6-sulfur (S) atom in its reduced state, a structural feature unique to the P cluster (Figures 7 and 8) (40, 60, 65). The coordination of the P cluster is unique compared to the canonical [4Fe-4S] clusters. Typically, [4Fe-4S] clusters are coordinated only by terminal cysteine residues, but the P cluster is coordinated by two bridging cysteines in addition to the four terminal ones, with each NifDK subunit contributing two terminal cysteines and one bridging cysteine (Figures 7 and 8) (77). Located at the NifD-NifK interface, the P cluster bridges the interprotein electron transfer from the [4Fe-4S] cluster of NifH to the FeMo cofactor of NifDK, where nitrogen reduction occurs.

In the reduced, resting state (P^N), all irons in the P cluster are in the 2^+ state (78). Upon oxidation, the P cluster switches consecutively to P^{1+} and P^{2+} states (79-81), and undergoes significant structural changes, breaking its symmetrical geometry. Under physiological conditions, the P^N/P^{1+} transition causes the displacement of the Fe6 out of the bonding distance from the central sulfide, and Fe6 is coordinated by the nearby serine residue in *Azotobacter* ($Ser188^K$ in the case of *A. vinelandii*) instead (Figure 8) (79). The non-physiological P^{1+}/P^{2+} transition provokes a break of the Fe5 coordination to the central sulfide and engages its interaction with the main chain of the bridging cysteine residue ($Cys88^D$ in the case of *A. vinelandii*) (Figure 8) (77). Such changes in the coordination have been shown to protect the metals during an oxidation event (82). The midpoint potentials of P^N/P^{1+} and P^{1+}/P^{2+} couples are the same (-310 mV) at pH 7.5-8.0 (65). While the P cluster can be further

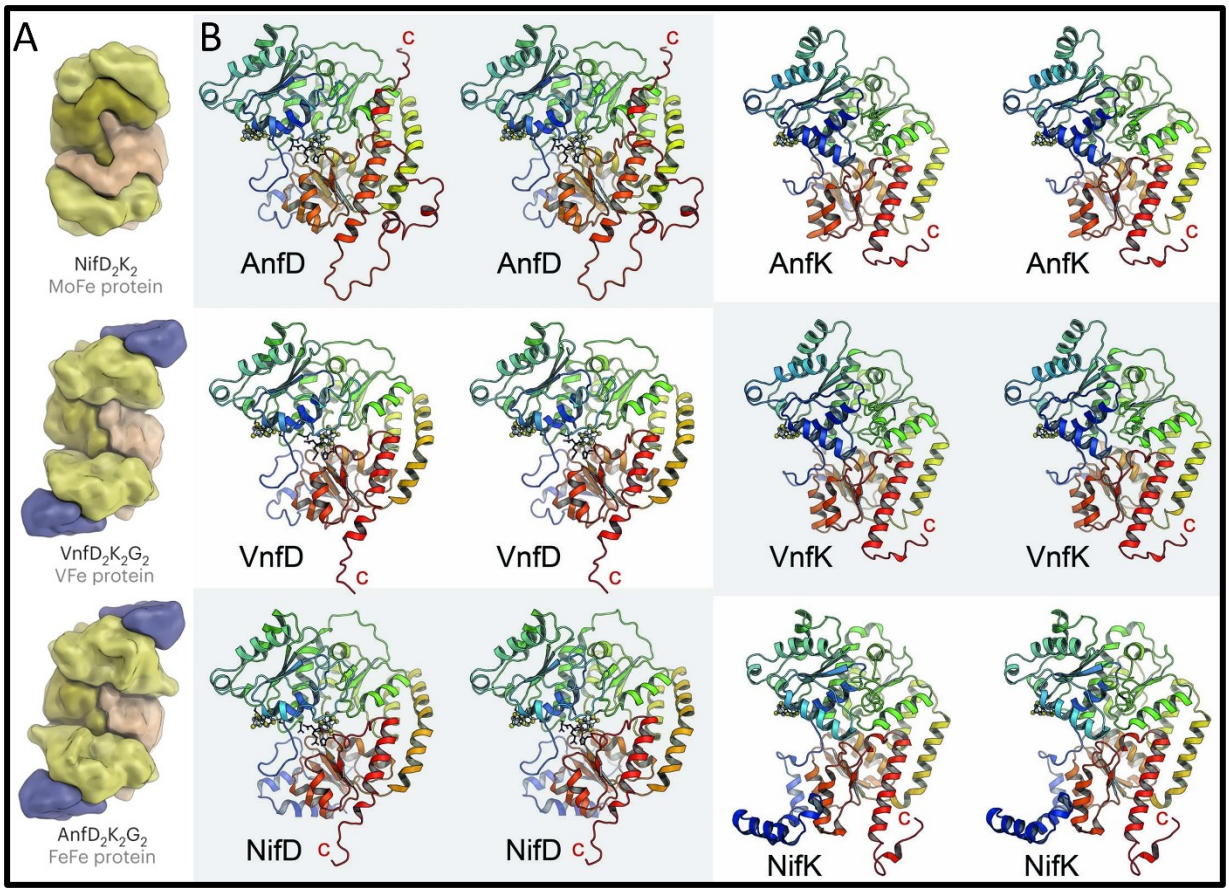


Figure 6. 3D structure and comparison of the core subunits of three nitrogenase isoforms from *Azotobacter vinelandii*. A. Subunit organisation of molybdenum (PDB 3U7Q), vanadium (PDB 5N6Y), and iron only (PDB 8BOQ) nitrogenases from *A. vinelandii*. Colour code: yellow (D subunit), olive (K subunit), blue (G subunit). **B.** The structures displayed in the cartoon are shown as stereo images and coloured in a rainbow from blue (N-terminus) to red (C-terminus). / Both panels of this figure are reprinted with permission granted under the licence ID 5553591206909 from Trncik *et al.* (2023) (69).

oxidised to the P^{3+} state, there is no evidence that this happens during the catalytic cycle (79, 83).

The FeMoco is located in the NifD core, where the substrate reduction occurs (70, 75). FeMoco is asymmetric, with a characteristic trigonal prismatic geometry (*D*3h) with an apical molybdenum atom breaking the point group symmetry (Figures 7 and 8) (60, 70, 75). The centre of the prism harbours a μ6-carbide (Figure 8) (84, 85), another unprecedented feature in biology. Two halves of the prism are finally joined through three bridging sulfides (S2B, S3A, S5A) (Figure 8). The seven Fe atoms are all tetrahedrally four-coordinated. In contrast to the P cluster, FeMoco is coordinated to the protein only by two residues: a conserved cysteine residue ($Cys275^{D}$ in case of *A. vinelandii*) coordinating Fe1 and a histidine ($His442^{D}$ in case of *A. vinelandii*), coordinating Mo atom along with the homocitrate (Figure 8) (40). The homocitrate is produced by NifV, the homocitrate synthase, from acetyl-coenzyme A and

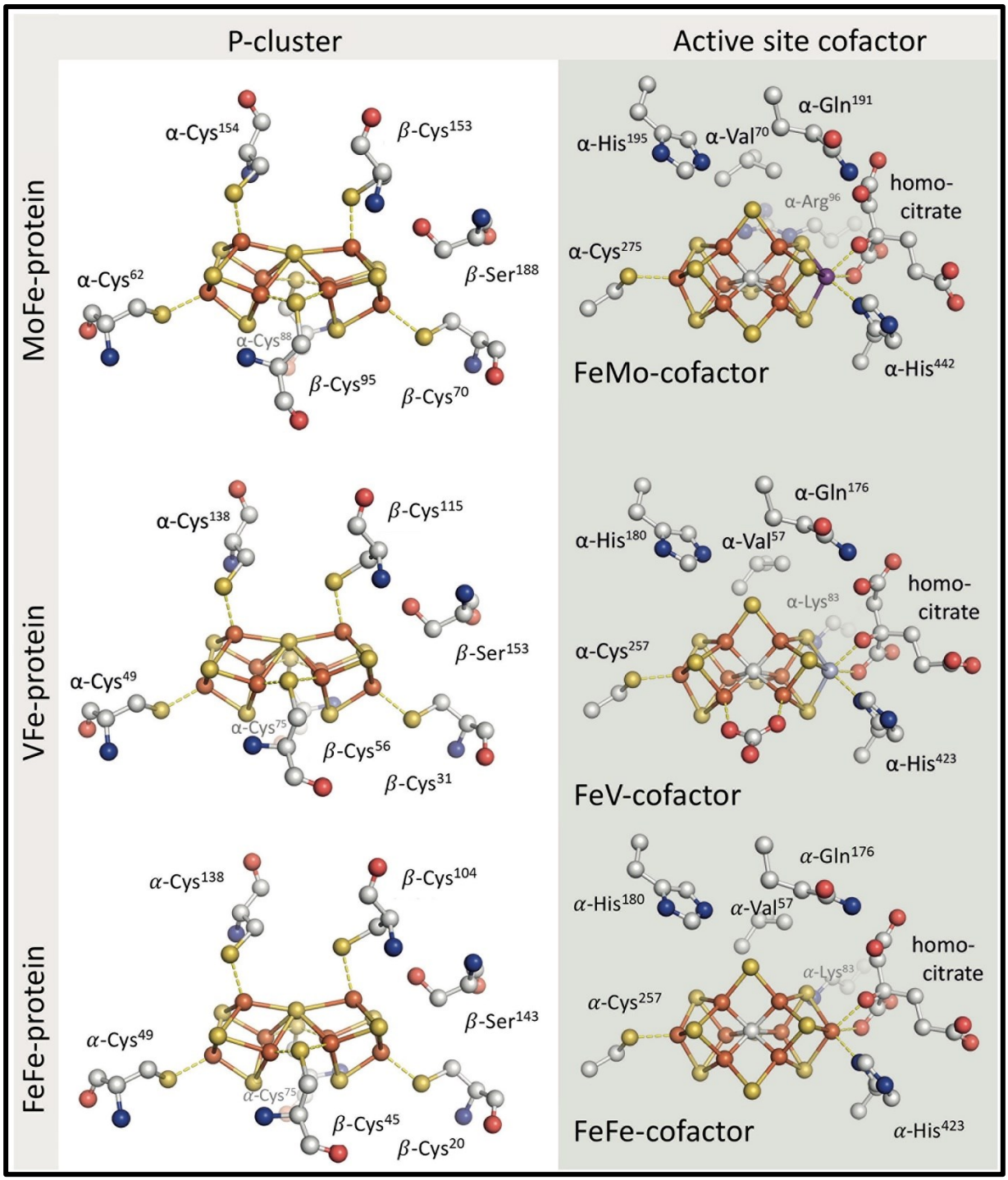


Figure 7. Metallocofactor structures and environment in the three nitrogenase isoforms from *A. vinelandii*. Shown P clusters are in the reduced state P^N. Colour code: rust (Fe), yellow (S), purple (Mo), red (O), grey (C), lavender (V). / This figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*.

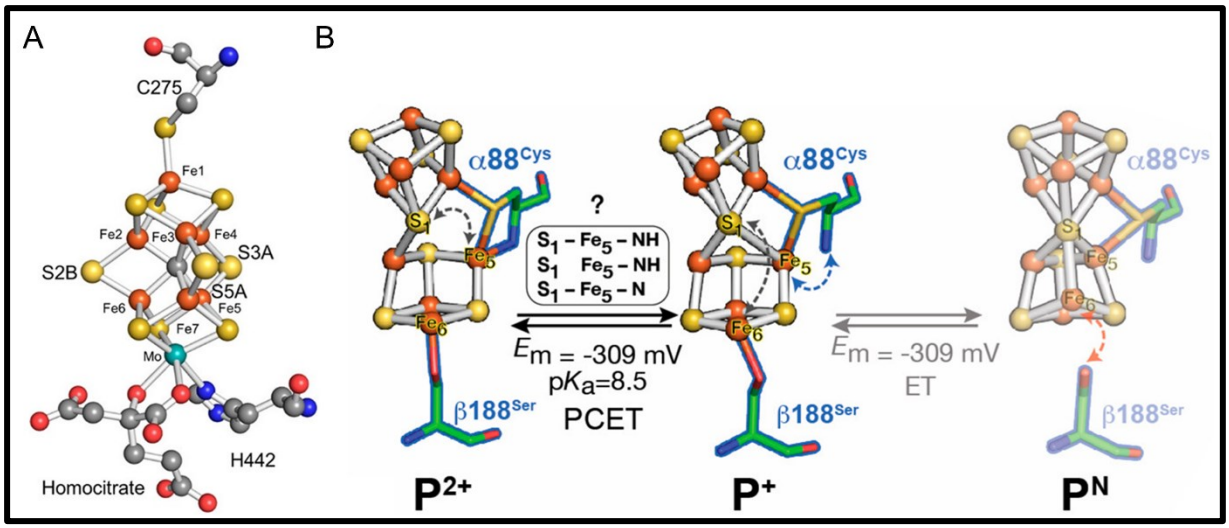


Figure 8. MoFe protein Metallocofactor structures and P cluster oxidation states (P^{2+}, P^{+}, and P^{N} states). **A.** The structure of FeMoco from *A. vinelandii*. **B.** Redox-coupled coordination changes of P cluster during the $P^{2+}\leftrightarrow P^{+}$ and $P^{+}\leftrightarrow P^{N}$ transition. / Panel A of this figure is reprinted with permission from Henthorn *et al.* (2019) (86). Copyright 2019 *American Chemical Society*. Panel B of this figure is reprinted with permission from Chica *et al.* (2022) (78). Copyright 2022 *American Chemical Society*.

2-oxoglutarate (87). The FeMoco is found in two states: dithionite-stabilized M^{N} and one-electron oxidised M^{ox}, with a mid-potential of -42 mV (65). All nitrogenase isoenzymes contain identical P clusters, but their catalytic cofactors differ from one another (40). They all have a conserved metallocofactor architecture containing the μ6-carbide (88) with almost identical bond distances and the key difference between them is the presence of either vanadium or iron as a substitute for the molybdenum atom in vanadium nitrogenases (VFe-co) and iron-only nitrogenases (FeFe-co), respectively (Figure 7) (68, 69, 89). Another major difference present in the VFeco is the substitution of one of the bridging sulfides (S3A) by a carbonate (CO_3^{2-}) ligand (Figure 7). However, its function remains unclear so far (68).

While their architecture is largely conserved, nitrogenase isoforms greatly differ in their abilities to reduce alternative substrates (Figure 9) (90-94). Namely, in addition to its natural substrates: dinitrogen and protons, nitrogenases have a plethora of alternative substrates, often small molecules with terminal double or triple bonds, which can be grouped into (52): *(i)* N_xH_x substrates: hydrazine (N_2H_4), diazene (N_2H_2) and diazene homologues (methyldiazene (CH_3-N=NH), dimethyldiazene (CH_3-N=N-CH_3) and diazirine (CH_2N_2)); *(ii)* NO_x substrates: nitrite (NO_2^-), nitrous oxide (N_2O) and hydroxylamine (NH_2OH); *(iii)* C-based substrates: alkynes, most notably acetylene (C_2H_2), ethylene (C_2H_2), carbon monoxide (CO), carbon dioxide (CO_2) and its analogues isocyanic acid/cyanate ($HNCO/OCN^-$), thiocyanate (SCN^-), and carbon disulfide (CS_2).

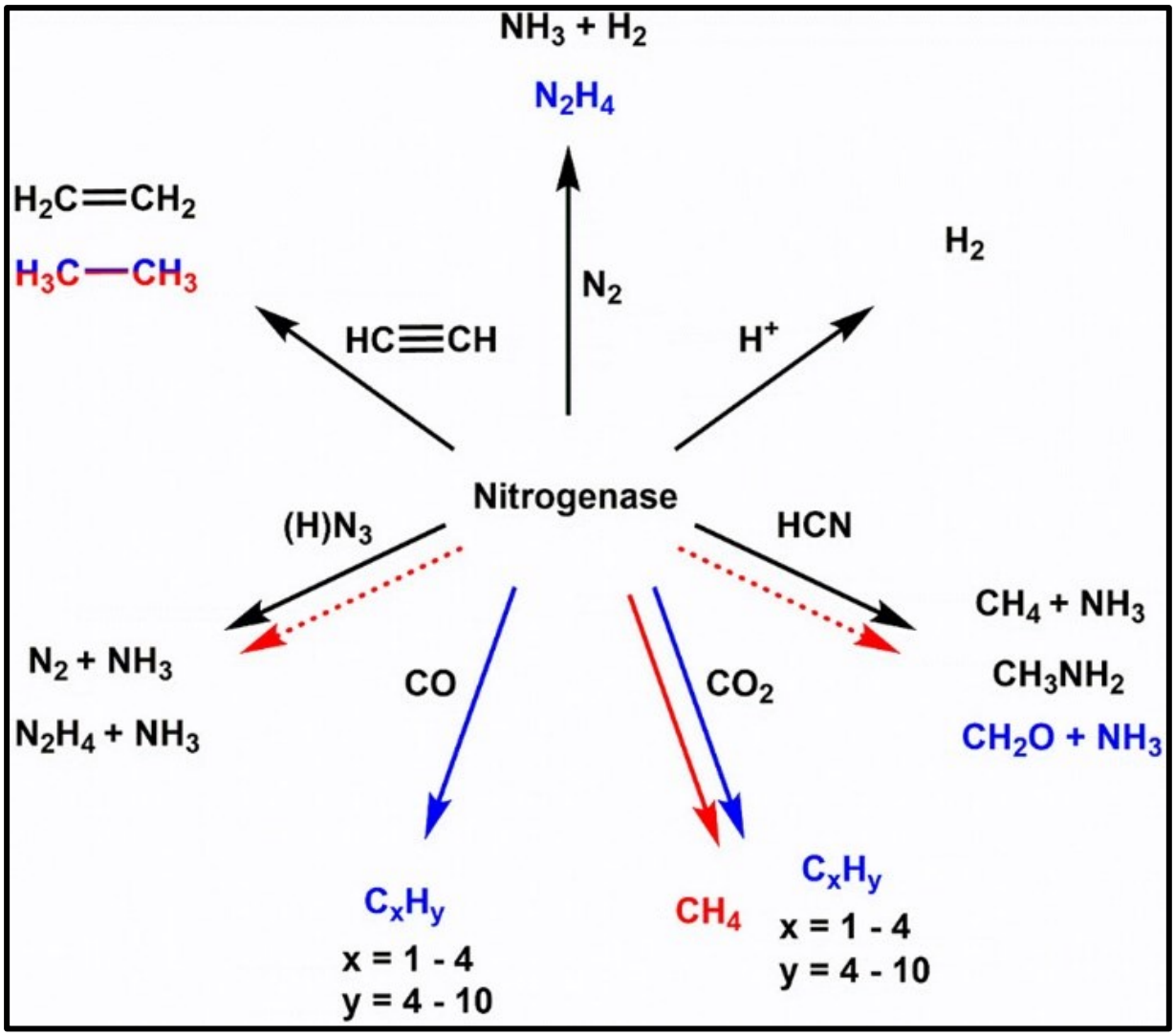


Figure 9. Reactivity of the nitrogenase isoforms. An overview of nitrogenase isoforms reactivity toward different substrates. Colour code: black (all nitrogenases), blue (V-nitrogenase), red (Fe-only nitrogenase), dots (not empirically proven). / This figure is reprinted with permission from Jasniewski AJ, Lee CC, Ribbe MW, Ribbe MW, Hu Y. 2020. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. Chemical Reviews 120:5107-5157 (67). Copyright 2020 *American Chemical Society*.

Finally, NifH and VnfH can function independently as reductases, converting CO_2 to CO and even further to hydrocarbons (95, 96). While many of these reactions may not be physiologically relevant, they are valuable in studying the mechanism and for future biotechnological applications. Although the three nitrogenase variants may be expected to behave similarly in catalysing the same nitrogen reduction reaction, their efficiency varies.

However, all three variants utilise the same overall eight-electron reaction mechanism (58, 89, 97-99).

1.3. The dynamic dance of nitrogenases: Thorneley-Lowe kinetic model. The overall stoichiometry of this eight-electron reaction mechanism can be represented as follows (100):

Mo-nitrogenase: $N_2 + 8\ H^+ + 16\ MgATP + 8\ e^- \rightarrow 2\ NH_3 + H_2 + 16\ MgADP + 16\ P_i$

V-nitrogenase: $N_2 + 12\ H^+ + 24\ MgATP + 12\ e^- \rightarrow 2\ NH_3 + 3\ H_2 + 24\ MgADP + 24\ P_i$

Fe-only nitrogenase: $N_2 + 20\ H^+ + 40\ MgATP + 20\ e^- \rightarrow 2\ NH_3 + 7\ H_2 + 40\ MgADP + 40\ P_i$

The "alternative" nitrogenases are less efficient in producing NH_3 in comparison to the Mo-nitrogenase as they require more ATP and evolve more H_2. Indeed, it has been observed that alternative nitrogenases exhibit lower activity compared to the molybdenum nitrogenase from the same organism when acting on the same set of substrates. For example, vanadium and iron nitrogenases from *A. vinelandii* show ~75% and 35% of Mo-nitrogenase N_2-reduction activity (67). These catalytic variations between isoforms are likely a result of the cofactor composition and protein environment (91, 94, 101-103). Similar reactivity was observed among the same isoforms of nitrogenase from different organisms in general, but due to the diverse experimental conditions, making direct comparisons is not straightforward (67).

It became evident that hydrogen evolution accompanying N_2-fixation is obligatory, and with that, a minimum requirement of 16 molecules of ATP per N_2 transformed to NH_3.

The overall reaction, following the Thorneley-Lowe (TL) model, can be represented by two cycles (Figures 10 and 11): *(i)* The Fe protein cycle, during which nucleotide exchange, reduction by flavodoxin/ferredoxin and electron transfer to the dinitrogenase occur; *(ii)* Dinitrogenase cycle, in which the metallocofactor cycles through 8 different increasingly reduced kinetic states (E_0-E_7) each representing a different step of the eight-electron transfer (Figure 10) (41, 45). Despite the incorporation of more detailed insights by many research groups, no significant conceptual revisions have been made to the TL model. However, it is important to note that the TL model was postulated using *Klebsiella pneumoniae* (now called *Klebsiella oxytoca*, NCBI:txid571) (104) as a model at 23° C and it is rarely used in quantitative applications (60). Furthermore, in practice, the stoichiometry of H_2:N_2 = 1 and MgATP:$2e^-$ = 4 is rarely observed (60). These discrepancies in assumed theoretical models and practical observations result from multiple factors, which are summarised and discussed in section 1.4. Nevertheless, the TL model remains the most developed kinetic and mechanistic outline for N_2 reduction by molybdenum-nitrogenase and can be expanded to the two other types.

The most detailed version of TL, obtained by pulsed EPR, revealed that N_2 binds to MoFe protein in the E_4 state. The E_4 state has been referred to as the *Janus* state, named after the Roman god with two faces. The name reflects the two possibilities of the reaction progression

at this critical stage, as it can proceed "forward" toward NH_3 formation or "backward" to the resting state by releasing H_2 (Figure 10A) (42). The E_4 state contains two bridging hydrides existing in the reductive elimination/oxidative addition (*re*/*oa*) equilibrium (Figure 10B) (105-108):

$E_4(4H) + N_2 \leftrightarrow E_4(2N2H) + H_2$

This state drives the reduction of N_2 to diazene (2N2H) intermediate, the most difficult step of N_2 reduction. Two forms of E_4 intermediate with bound N_2 have been observed: with two bridging hydrides and with bound N_2 after hydride protonation and H_2 displacement (109). Such reversible reductive elimination, documented in various transition metal complexes, has been demonstrated in all three nitrogenase isoforms. However, it was not conclusive whether or not the N-N bond was reduced and the intermediate with a reduced N≡N bond as well as the exact steps to reduce it remains uncharacterised.

While other substrates with the exception of H_2 are binding to more oxidised forms of the MoFe-protein (60). H_2 evolution is happening in the states E_2–E_4, while NH_3 is proposed to be released from some combination of E_5, E_6, and/or E_7 states. The dinitrogenase can release products or bind the substrates only when it is not complexed with the Fe protein and has to be reduced to the E_2-E_4 state before doing so (60).

The ATP hydrolysis cycle occurring within the Fe protein orchestrates the dynamic association/dissociation to the dinitrogenase. Conformational changes dictated by ATP/ADP+P_i/ADP bound states ensure the correct interprotein lock, which is critical for electron transfer (Figure 11) (65). The electrons flow from the donors (flavodoxins/ferredoxins) through the [4Fe-4S] Fe protein's cluster to the dinitrogenase's P cluster and FeMoco, finally reaching the substrate. However, since the P cluster exists in an all-ferrous P^N form (which cannot be further reduced, can be obtained by dithionite addition *in vitro*), it has been inferred that the intraprotein electron transfer in the dinitrogenase is likely conformationally gated (110). In this scenario, the docking of a reduced ATP-loaded Fe protein (111, 112) lowers P cluster's reduction potential, which allows the single electron transfer from the P cluster to the FeMoco, creating the "electron deficit" and giving it the name "deficit spending" mechanism (110, 111, 113). Since the distance between the P cluster and FeMoco remains unchanged, rearrangement of residues between the two metallocofactors is probably responsible for "the gating mechanism". Such repositioning would help the electronic coupling or influence cluster coordination (65, 114). The initial electron transfer from the P cluster to the FeMoco is followed by an electron transfer from the [4Fe-4S] cluster

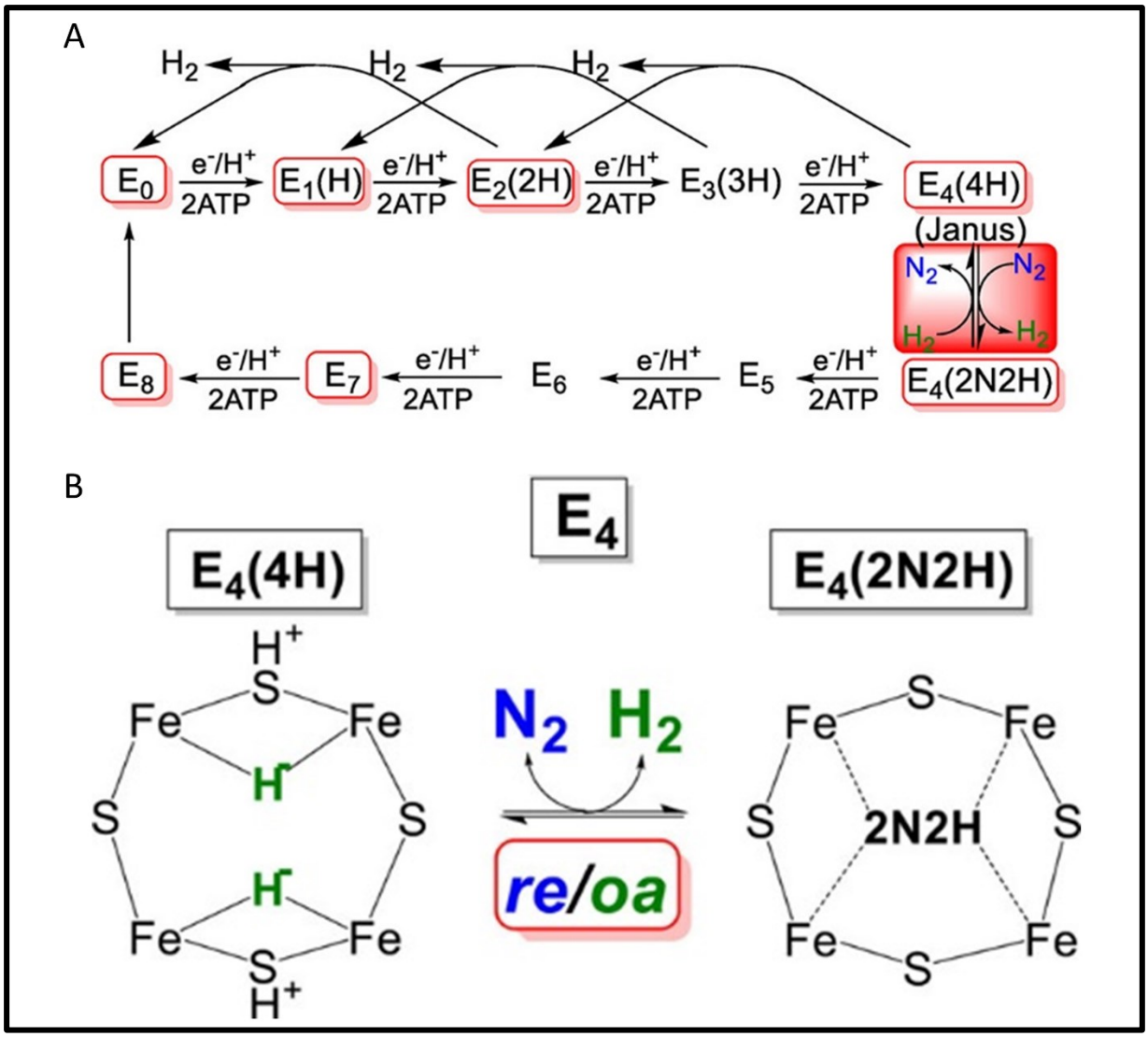


Figure 10. The nitrogenase kinetic cycle according to the Thorneley-Lowe model. **A.** Simplified scheme of the eight different states occurring during the catalytic cycle. n = number of e^-/H^+ added to nitrogenase, the stoichiometry of H/N bound to nitrogenase at each step is shown in brackets. The *re*/*oa* equilibrium is represented in the red box, while the E_n states, characterised by spectroscopy, have been shown with a red outline. **B.** Mechanistic coupling of N_2 binding to the FeMoco concomitant with the reductive elimination of H_2. / This figure is reprinted with permission from Seefeldt LC, Yang ZY, Lukoyanov DA, Harris DF, Dean DR, Raugei S, Hoffman BM. 2020. Reduction of Substrates by Nitrogenases. Chemical Reviews 120:5082-5106 (52). Copyright 2020 *American Chemical Society*.

to the P cluster (making up for the "deficit").

While it is still debated whether ATP hydrolysis occurs before or after the electron transfer to the P cluster (111, 115, 116), there is an agreement that the rate-limiting step of the cycle is the P_i release from the Fe protein (16-22 s^{-1}) and not the complex dissociation as thought previously (rate constant 759 s^{-1}) (117). Once the MgADP-bound oxidised state of the Fe protein disengaged from the dinitrogenase, the cycle can start again.

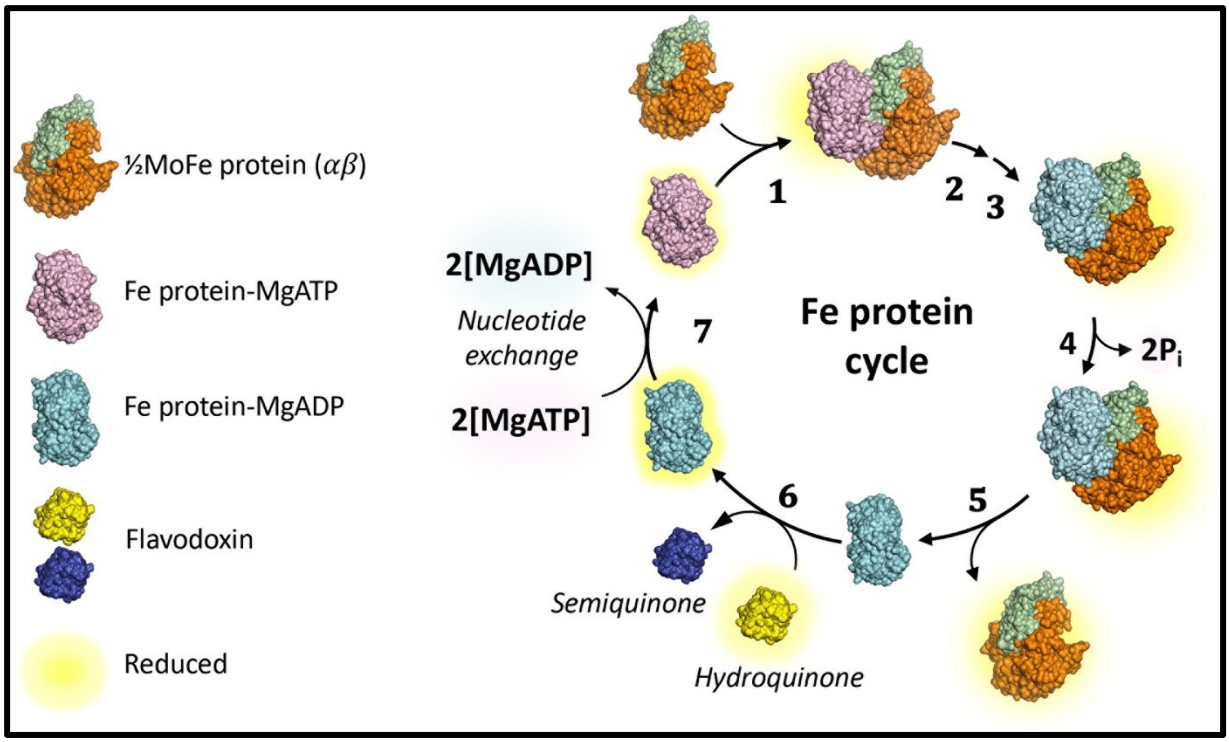


Figure 11. Thorneley-Lowe model of Fe-protein kinetic cycle. Reduced Fe-protein loaded with MgATP binds to the MoFe protein (Step 1). The [4Fe-4S] cluster of the Fe-protein is oxidised to [4Fe-P cluster4S]$^{1+}$ state after transferring one electron to the P cluster in a deficit-spending mechanism (Step 2). This is followed by MgATP hydrolysis (Step 3), the release of P_i (Step 4), and nitrogenase complex dissociation (Step 5). Oxidised Fe-protein is reduced by a flavodoxin/ferredoxin (Step 6) and can bind MgATP again (Step 7), entering a new cycle./ This figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*.

The cellular energy consumption during the reaction is very high. However, since the [4Fe-4S] cluster and the P cluster can exist as two different redox couples, it opens the possibility that the electron transfer is happening in pairs. Furthermore, NifH in the dithionite -reduced [4Fe-4S]$^{1+}$ state could form a productive complex with NifDK, as it undergoes the same nucleotide-dependent conformational changes, as in the [4Fe-4S]^{0} state. Multiple studies support this possibility (118, 119), but there is no definitive evidence for a two-electron transfer. Recent research suggests that electron transfer does not occur in pairs, however, while i*n vitro* experiments typically use dithionite (DT) as the reductant (which can only reduce the [4Fe-4S] cluster to the $^{1+}$ state), there is no DT involved in *in vivo* N_2 reduction. The experiments with *A. vinelandii* lysate also suggest that the electron transfer mechanism can be different in the presence of other physiological components *in vivo* (119). This would be highly advantageous in terms of energy efficiency, reducing it by 50% and facilitating the rapid achievement of the E_4 state of dinitrogenase, necessary for binding of N_2 (60, 65).

Nitrogenase does not have an exclusive electron donor *in vivo,* and different ferredoxins (120)

or flavodoxins (121) can fulfil this role, while for *in vitro* studies, sodium dithionite ($Na_2S_2O_4$) is traditionally used (122). Out of five different flavodoxins identified in the genomes of diazotrophs, only one can deliver electrons to nitrogenase *in vitro* (NifF) (123, 124). The ferredoxins found in diazotrophs proved to be more robust, with many of the forty-seven identified so far proven to have a role in electron transfer to nitrogenase (65). Although these ferredoxins and flavodoxins usually have a reduction potential comparable to NifH's, they cannot replace it as the electron donor to NifDK (65).

The key features of nitrogenase catalytic cycle can be summarised as follows: *(i)* NifH-NifDK form dissociable complex, *(ii)* this complex dissociates after every cycle of the electron transfer, *(iii)* the electron flow and ATP consumption rates are both constant irrespective of the substrate being reduced or the number of already transferred electrons (125-128), *(iv)* however, increasing the molar ratio of NifH to NifDK can provide a higher electron flux and increase the activity (the optimal ratio is isoform and substrate dependent, and can range from 20:1 or 30:1 (considered high) to 1:1 (low))(60), *(v)* the hydrolysis of two ATP molecules by NifH (per cycle) is coupled to the electron transfer and only happens in the presence of NifDK, *(vi)* NifDK active sites are independent, but might be cooperative, *(vii)* different substrates and inhibitors bind to different NifDK E_n states, *(vii)* H_2 production is obligatory during N_2 reduction, with the R-homocitrate, cluster sulfurs and the carbonate ligand in the V-nitrogenase being identified as the potential ionizable sites. The V- and Fe-only nitrogenase have been less characterised than the Mo-nitrogenase. Still, the available kinetic analyses suggest that they share a common mechanism with respect to N_2-reduction (67, 69, 100).

Most of the biochemical, biophysical, and spectroscopic insights previously described have been obtained from the work on the model organism *A. vinelandii*.

1.4. *Azotobacter vinelandii*: a historical model for understanding biological nitrogen fixation. During the early days of genetic studies of nitrogen fixation, *K. oxytoca* was a preferred model organism due to its genetic similarity to *Escherichia coli*, which had many available genetic tools already. However, in the 1980s, *K. oxytoca* was gradually replaced by other genetically tractable Gram-negative bacteria that were easier to grow and provided higher nitrogenase yields (up to 10% of the total cellular protein pool), such as *A. vinelandii* (40). *A. vinelandii* cemented its position as the principal model organism for studying nitrogen fixation after the first X-ray structures of both NifH (57) and NifDK (75) were solved, consolidating all the available information up to that point.

A. vinelandii was isolated from soil in 1903 in Vineland, New Jersey (129), but has since

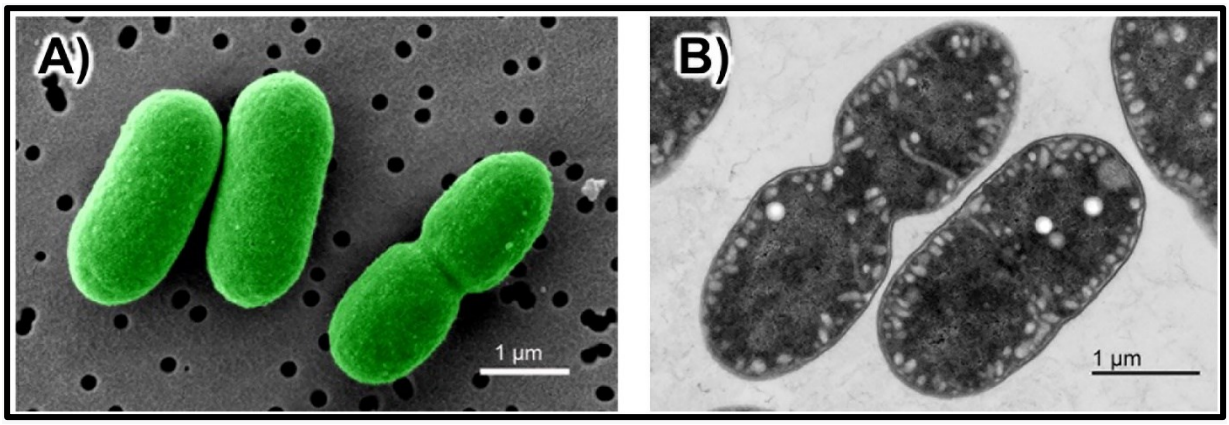


Figure 12. A. Scanning- and B. -transmission-electron micrographs of *A. vinelandii* cells. / This figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*.

been identified in alkaline soils, freshwater, and plant rhizospheres worldwide. This free-living, mesophilic (optimum growth temperature ~15-37 °C, depending on the strain and growth conditions), obligate aerobe has motile cells and varying cell morphology (from rods to cocci, Figure 12), depending on growth conditions (130). It is known for its ability to use a wide range of carbon substrates, including alcohols, organic acids, and sugars, as well as various fixed nitrogen sources such as ammonium, nitrate, nitrite, amino acids, adenine, and urea. *A. vinelandii* can also produce multiple polymers, such as polyhydroxyalkanoates (PHAs) and alginate, as well as multiple siderophores (130). Another notable feature of *A. vinelandii* is the formation of heterocysts resistant to drought and other detrimental environmental factors in response to carbon availability.

Although multiple important principles were postulated over the years using *A. vinelandii* as a model organism, such as the Lineweaver-Burk kinetic models (131) and the genetic code (132), its greatest contribution remains to the studies related to nitrogen fixation. Many fundamental principles, such as the general nitrogenase composition, structure, catalytic activity, biosynthesis, and genetics, have been discovered using *A. vinelandii* thanks to its valuable natural competency, which, when combined with established methods for genetic manipulation and engineering, makes it an excellent organism for research purposes (40, 130).

1.5. The blueprint of biological nitrogen fixation: genetic organisation of the *nif* operon. Biological nitrogen fixation is a challenging biochemical task requiring many gene products to operate optimally *in vivo*. This is why the genes encoding the nitrogenase enzymes are accompanied by a plethora of accessory proteins that play important roles in regulation, energy and reducing equivalent supply, and the biosynthesis and maturation of metal

cofactors (40). Every nitrogenase isoenzyme is encoded by a separate operon (40): Mo-nitrogenase is encoded by *nif* genes (*ni*trogen *f*ixation) (124), V-nitrogenase is encoded by *vnf* (*v*anadium *n*itrogen *f*ixation) (133), while the Fe-only nitrogenase is encoded by *anf* (*a*lternative *n*itrogen *f*ixation) genes (134, 135).b

Depending on the organism, the *nif* genes encoding the nitrogenase can be arranged as one or several operons or regulons. Sometimes, *nif* gene operons can also be found on plasmids (136). The composition of these operons varies from the minimal set of genes required for functional Mo-nitrogenase (*nifHDKENB*) *in vitro* to over more than 20 genes in some species (40, 66, 67). So far, this minimal set of genes used as a criterium for diazotroph identification is found in 19 bacterial (Bacteroidetes, Cyanobacteria, Firmicutes, Proteobacteria, Acidobacteria, Actinobacteria, Aquificae, Chlorobi, Chloroflexi, Chrysiogenetes, Deferribacteres, Elusimicrobia, Fusobacteria, Lentisphaerae, Candidatus Margulisbacteria, Nitrospirae, Planctomycetes, Spirochetes, and Verrucomicrobia) and only one archaeal phylum (Euryarchaeota) (137). This diversity is reflected in the variety of environments they inhabit, such as agricultural, marine, and freshwater sediments, wastewaters, and intestinal tracts of termites, some of which are human pathogens (39, 67). To the best of our knowledge, so far, there are no eukaryotic cells that can fix nitrogen.

The core nitrogenase genes in *Bacteria* are homologous to those in *Archaea*, but, as mentioned before, they are often part of more complex operons.

As an example, *A. vinelandii* has been found to have 103 genes involved in some aspect of nitrogen fixation (40). The *nif* genes that encode Mo-nitrogenase in this bacterium are split into two clusters located adjacent to the chromosome origin site (Figure 13): a major and minor cluster, which are accompanied by clusters of *naf* (*n*itrogenase *a*ssociated *f*actors), *nfa* (*n*itrogen *f*ixation *a*ssociated), *rnf* (*R*hodobacter *n*itrogen *f*ixation, orthologous to the genes originally described in *Rhodobacter capsulatus*) and *fix* (N_2 *fix*ation) genes (40). Major *nif* cluster contains 35 genes organised in 10 transcriptional units, which encode the nitrogenase (*nifDKH*) and proteins involved in: nitrogenase maturation (*nifMWZ*, *nafH*), the cofactor biosynthesis (*nifEN*), biosynthesis of [4Fe-4S] cluster (*nifUS*, *nafFG*), electron transfer (*nifF*), oxidation protection (*nafO*) and homocitrate synthase (*nifV*) (Figure 13) (40, 66). Minor *nif* cluster contains 17 genes split into 3 transcriptional units. It includes genes whose products are responsible for FeMoco biosynthesis (*nifB*), Fe-Mo insertion (*nafY*), Mo-trafficking (*nifOQ*), transcriptional regulation (*nifAL*), and nitrogenase expression (*rnf*) (40, 66).

While *vnf* genes are also separated into two distinct clusters, *anf* genes are all grouped

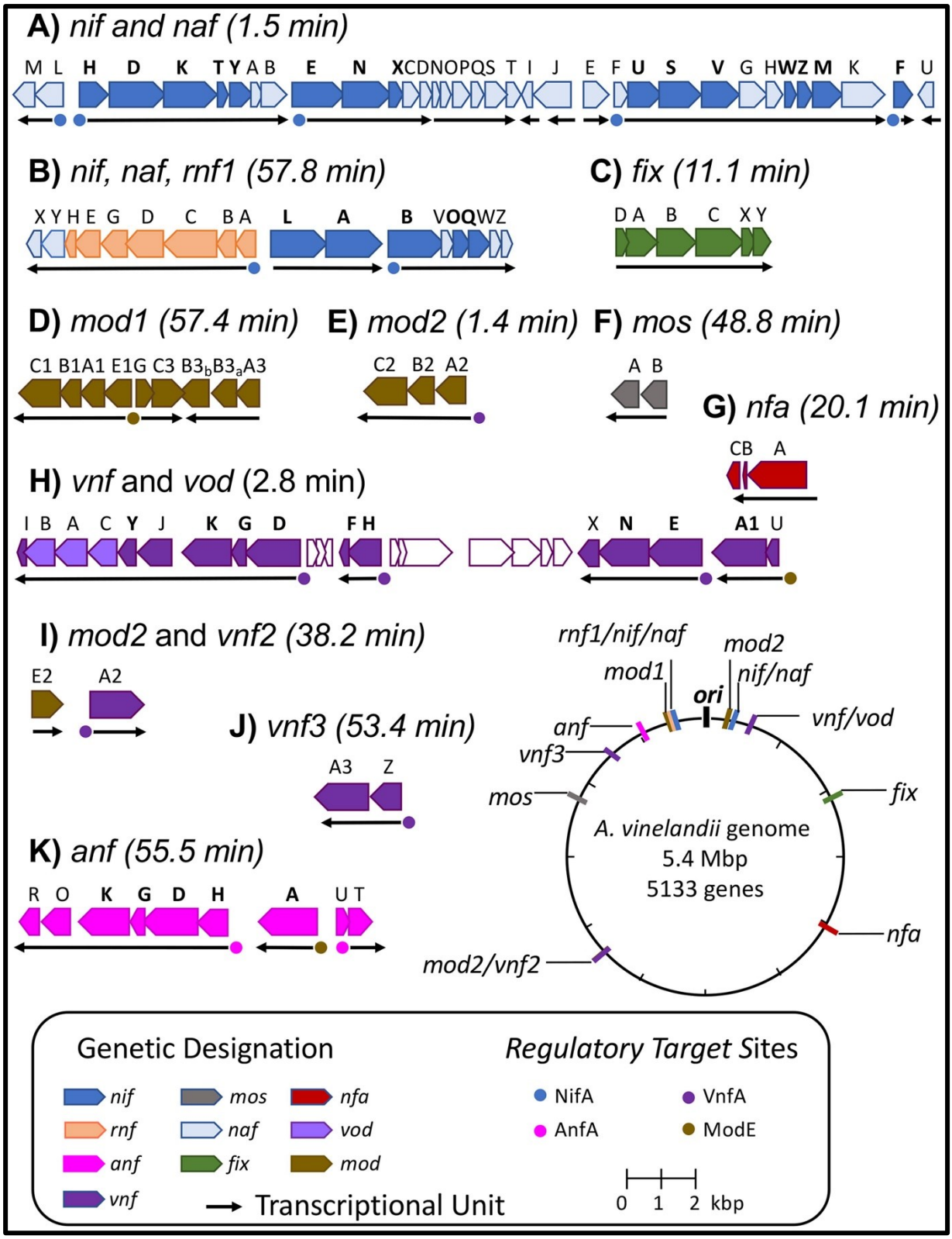


Figure 13. The genomic environment of *nif, vnf,* and *anf* operons and associated genes in *A. vinelandii*. The genomic location of each cluster is shown on the circular genome map and noted in the brackets. Arrows represent transcriptional units, while the transcription start sites are represented by dots. The summary of the exact genomic locations, functions, and available structures of the shown genes and their products is summarised in del Campo *et al.* (2023). / This figure is reprinted with permission from del Campo *et al.* (2023) (40). Copyright 2023 *Taylor & Francis*.

together. The availability of heterometals predominantly regulates the expression of alternative nitrogenases, and as a result, gene clusters involved in heterometal import and storage such as *mod* (*mo*lyb*d*enum transport), *mos* (*mo*lybdenum *s*torage), and *vod* (*v*anadium transport) play important roles (Figure 13) (40). Table 1 summarises the genes of *nif*, *vnf,* and *anf* operons in *A. vinelandii* and their functions. The full overview of all 103 genes is summarised elsewhere (40). Many of these 103 genes have yet unclear roles, and it is unknown if they directly impact the phenotype. However, bioinformatics and transcriptomic studies have confirmed their involvement in the nitrogen fixation. The transcriptomic profiling of *A. vinelandii* showed significant overlap in the expression of biosynthetic *nif* genes with *vnf* and *anf* operons. (138), indicating their interdependency and supporting the currently proposed evolutionary scenario.

The complexity of nitrogen fixation-related operons in *A. vinelandii*, which is notably higher compared to most other diazotrophs, is thought to be due to its strictly aerobic lifestyle (40, 66) and is primarily attributed to the optimisation of nitrogenase synthesis and regulation to accommodate the heightened aerobic productivity (139).

Table 1. Relevant *nif*, *vnf,* and *anf* genes from *A. vinelandii*. / This table is reprinted with permission from Jasniewski AJ, Lee CC, Ribbe MW, Hu Y. 2020. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. Chemical Reviews 120:5107-5157 (67). Copyright 2020 *American Chemical Society*.

Gene	Gene product label	Known Function
nifH	NifH, dinitrogenase reductase, component 2, reductase component, γ subunit, Fe protein	Mediation of ATP-dependent electron transfer during catalytic turnover, facilitates the formation of a P cluster on NifDK and facilitates the conversion of L- to M-cluster on NifEN via insertion of Mo and R-homocitrate.
nifD,K	NifDK, dinitrogenase, component 1, catalytic component, MoFe protein	Facilitates chemical transformation of N_2 to NH_3 at the active cofactor site
nifL		Serves as a negative transcription regulator of *nif* genes
nifA		Serves as a positive transcription regulator of *nif* genes
nifF	NifF, flavodoxin	Involved in the transfer of electrons to nitrogenase during catalysis or assembly
nifS	NifS, cysteine desulfurase	Transfers sulfur to NifU for the assembly of small FeS clusters
nifU	NifU	Serves as a scaffold for the assembly of small FeS clusters, which can then be used for M- and P cluster assembly
nifB	NifB	Mediates the radical SAM-dependent insertion of carbon concomitant with the formation of a Mo/homocitrate-free precursor of the M-cluster
nifE,N	NifEN	Serves as a scaffold for the maturation of the M-cluster, a structural/functional homologue of NifDK.
nifV	NifV, homocitrate synthase	Synthesizes *R*-homocitrate for M-cluster assembly
nifZ	NifZ	Serves as a key factor in the stepwise maturation of P clusters in NifDK, possibly through a chaperone-like function
nifM	NifM, peptidyl-prolyl cis-trans isomerase	Involved in the maturation of Fe protein
nifX	proposed carrier	In M-cluster assembly
nifY	proposed carrier	In M-cluster assembly
nafY	proposed carrier	In M-cluster assembly
nifW	NifW	Putative role in protecting nitrogenase from oxygen damage
nifO	NifO	Function unknown, although it resembles thioredoxin
nifT	NifT	Unknown function
vnfH	VnfH, dinitrogenase reductase, component 2, reductase component, γ	Mediation of ATP-dependent electron transfer during catalytic turnover, likely participates in the biosynthesis of VnfDGK cofactors by analogy to the *nif* system, but this has not been shown.

	subunit, Fe protein	
vnfD,G, K	VnfDGK, dinitrogenase, component 1, catalytic component, VFe protein	Facilitates chemical transformation of N_2 to NH_3 and CO/CO_2 into hydrocarbons at the active cofactor catalytic component, VFe protein site
vnfE,N	VnfEN	Putative involvement in V-cluster biosynthesis by analogy to NifEN
vnfA		Serves as a positive transcription regulator of *vnf* genes, also correlated to the expression of *anf* genes
vnfX	VnfX	Unknown function
vnfY	VnfY	Unknown function
vnfO	VnfO	Unknown function
vnfU	VnfU	Unknown function
anfH	AnfH, dinitrogenase reductase, component 2, reductase component, γ subunit, Fe protein	Mediation of ATP-dependent electron transfer during catalytic turnover, may participate in the biosynthesis of AnfDGK cofactors but has not been shown.
AnfD,G, K	AnfDGK, dinitrogenase, component 1, catalytic component, FeFe protein	Facilitates chemical transformation of N_2 to NH_3 at the active cofactor site
anfA		Serves as a positive transcription regulator of *anf* genes
anfO	AnfO	Unknown function
anfR	AnfR	Unknown function
anfU	AnfU	Unknown function

1.6. Challenges of studying the nitrogenase: an ultimate testing ground for biochemistry. With remarkable advancements made in understanding the structure of nitrogenase and its metallocofactors, as well as its assembly and activity, driven by technological breakthroughs in crystallography, spectroscopy, molecular biology, and biochemistry, it might be tempting to assume that we have already discovered everything there is to know about nitrogenases. However, it is essential to recognise that the quest for knowledge about this enzyme is far from complete. Many questions, particularly regarding the exact sequence of events occurring during the transition ATP hydrolysis/electron transfer and the exact pathway of N_2 reduction, remain open.

The study of nitrogenase kinetics is a significant challenge due to a multitude of factors that complicate the analysis, including high oxygen sensitivity, the dynamic nature of the complex, the existence of multiple intermediate states, and the heterogeneity in the nitrogenase population due to the presence of partially assembled and damaged clusters (60). On the same line, the purification of nitrogenase proteins is non-trivial. The nitrogenase is not

naturally highly overexpressed, and it's purification is a time-intensive process that has to be carried out under anaerobic conditions. Thus far, molybdenum nitrogenase has been natively isolated from *A. vinelandii* (140), *Azotobacter chroococcum* (141), *K. oxytoca* (142), *Clostridium pasteurianum* (143), *Gluconacetobacter diazotrophicus* (144) and *R. capsulatus* (55). Vanadium nitrogenase has been isolated only from two organisms: *A. vinelandii* (145) and *A. chroococcum* (146). In contrast, the iron-only nitrogenase has been isolated from *A. vinelandii* (147), *R. capsulatus* (148), *R. rubrum* (149) and *R. palustris* (150).

Structural insights are limited to an even smaller subset of organisms: the structures of dinitrogenase reductase are available from the bacteria *A. vinelandii* (all three isoforms, *Av*NifH (57), *Av*VnfH (58), *Av*AnfH (69)) and *C. pasteurianum* (*Cp*NifH, (151)) and the archaeon *M. acetivorans* (*Ma*NifH, (152)), while the structure of nitrogenase is available from *A. vinelandii* (all three isoforms, *Av*NifDK (75), *Av*VnfDKG (68), *Av*AnfDKG (69)), *K. oxytoca* (*Ko*NifDK, (153)), *C. pasteurianum* (*Cp*NifDK, (154)) and *G. diazotrophicus* (*Ga*NifDK, (155)).

Throughout history, structural information has proven vital in offering valuable context for the interpretation of data obtained from spectroscopy and different biochemical methods. As a result, future efforts should, among others, focus on broadening the scope of accessible structural information from phylogenetically distant microbes to explore the natural repertoire of nitrogenases with possible different catalytic properties.

Moreover, the demand for effective biocatalysts for industrial ammonia production (31) has led to a quest for microbial enzymes that can function under the harsh conditions of industrial processes, including high pressure/temperature/pH stability. The discovery of diazotrophic thermophilic diazotrophs has presented new options (156). Among them, (hyper)thermophilic methanogens represent a group with unique metabolism and energy conservation strategies (157), which stands out by their biotechnological potential (158).

1.7. Methanogens: the diversity and metabolism. Methanogens are a very diverse group of obligate anaerobes often called "biocatalysts" due to their ability to produce the storable energy carrier, methane (158). Their historical impact on methane release and shaping our planet´s atmosphere is indisputable (159). Nowadays, they still actively contribute to the carbon cycle by releasing half of the methane released in our atmosphere (160, 161) as they have colonised most anaerobic ecological niches worldwide (158, 162, 163). They can be found in various anoxic habitats, ranging from extreme niches of hydrothermal deep sea vents and hypersaline lakes, anaerobic sediments of rivers, lakes, marshes, and rice fields,

wastewaters to the plant tissues and digestive tracts of insects and animals, including humans (164). In these habitats, anaerobic bacteria, protozoa, and fungi hydrolyse biopolymers to monomers, lipids to glycerol and long-chain fatty acids, and ferment them to acetic acid, CO_2, and H_2, which are then used as substrates for methanogenesis. In other words, by living in syntrophic association with these decomposers, methanogens are operating the final step in organic matter degradation.

Currently, seven orders of methanogens are recognised: *Methanococcales*, *Methanobacteriales*, *Methanosarcinales*, *Methanomicrobiales*, *Methanopyrales*, *Methanocellales,* and *Methanomassilicoccales*, all belonging to *Euryarchaeota* (158), and some of them have already found application in the different industrial process: production of methane from crude oils, production of gas from organic matter, wastewater treatment, treatment of solid waste and development of micro-biogas plants and methanogenic bioelectrochemical systems (158, 165, 166). One of these technologies relies on the biogas conversion of methanogens using hydrogenotrophic methanogenesis, in which H_2 or formate are used as electron donors for CO_2 reduction and methane production ($4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$). Hydrogenotrophic methanogenesis is the most widespread out of the three described so far (hydrogenotrophic, methylotrophic, and acetoclastic) (167). However, due to the very low H_2 partial pressures (pH_2 1-10 Pa) in most of the natural habitats of methanogens, the energy yields from this metabolism are very low, at the limit of supporting active metabolism. Under these conditions, the free energy change of CO_2 reduction to methane is not enough for the synthesis of even 1 mole of ATP per mole of methane produced (157). Despite being energy extremophiles, methanogens perform a plethora of energetically demanding metabolic reactions, including N_2-fixation.

Nearly forty years ago, it was discovered that methanogens also have the ability to fix N_2. Simultaneously, it was shown that *Methanosarcina barkeri* (168) and the marine archaeon *Methanothermococcus thermolithotrophicus* could grow on N_2 as the sole nitrogen source (169). Such striking discovery opened up questions about the evolution of nitrogenase (discussed in detail in section 1.8.) and a reconsideration of the ecological role of these microbes in the nitrogen cycle. Alongside anaerobic methanotrophic archaea (ANME), they are the only archaeal group with a genomic predisposition for diazotrophy known so far (170). In both groups, N_2-fixation occurs only in the absence of other nitrogen sources. Thus, the ability to fix N_2 may be of special importance for the survival of methanogens in niches low on readily available nitrogen sources, such as sea floor, terrestrial hot springs, and lakes. The *nifH* gene, used as a marker to identify nitrogen-fixing microbes, has so far been

identified in multiple diverse anoxic habitats, such as deep seawater and hydrothermal vent fluids (171), oligotrophic open seas (172), deep-sea methane seep sediments (173), and N_2-limited soils of salt marshes (174). Because of their simple anaerobic metabolism and ancient origin (175), methanogens are attractive models for describing ancient (if not primordial) Life forms. Interestingly, the genomes of hyperthermophilic methanogens harbor the most reduced *nif* operons, which is especially interesting for consideration of evolutionary scenarios.

1.8. Nitrogenase evolution: tracing the origins of methanogenic nitrogen fixation. Nitrogenases are an ancient protein family that evolved around 1.5-2.2 billion years ago, although there are isotopic traces of potential nitrogenase activity dating as early as 3.2 billion years ago (176). Certainly, the nitrogenase evolved only once, but its subsequent evolutionary history is filled with gene duplication, fusion, loss events, and horizontal gene transfers (177).

There are two different proposed scenarios for the origin of the nitrogenase. According to the first scenario, nitrogen-fixing Last Common Universal Ancestor (LUCA) hypothesis, nitrogenase had already evolved in the LUCA, and gene loss, along with horizontal gene transfer, has been a dominant factor dictating the distribution of nitrogenase. This aligns with the fact that nitrogenase is found neither in eukaryotes nor in many entire phyla of prokaryotes (177). The second scenario, currently accepted as more plausible and with more evidence supporting it (137, 139, 178, 179), is the methanogen origin hypothesis. In this scenario, the nitrogenase evolved in methanogenic archaea and was transferred by horizontal gene transfer to the bacterial domain, most likely a member of Firmicutes (177). Narrow ecological niches in which different organisms could co-exist, so-called "evolutionary hotspots," were probably conducive to intensified horizontal gene transfer (177). For the nitrogenase, anaerobic, sulfidic basins in the redox-stratified ocean, with the occasional influx of oxygenated surface waters containing molybdenum (abundantly present during the early Proterozoic in marine environments (180, 181)), could have been these hot spots (178). The absence of nitrogenase in the early-branching bacterial lineages and eukaryotes altogether also supports this scenario.

Since the estimated concentrations of Fe in the oceans until ~2.2-2.5 years ago were higher than molybdenum (182, 183), it was assumed that Fe-only and V-nitrogenases evolved first (182, 184) and predated the Mo-nitrogenase (177). However, more recent phylogenetic analyses relying on ancestral sequence reconstruction point that Mo-nitrogenase is the ancestral one instead. This phylogeny suggested by Garcia *et al*. (Figure 14) (137) is consistent with previous analyses that suggested it (137, 139, 178). According to it, there are

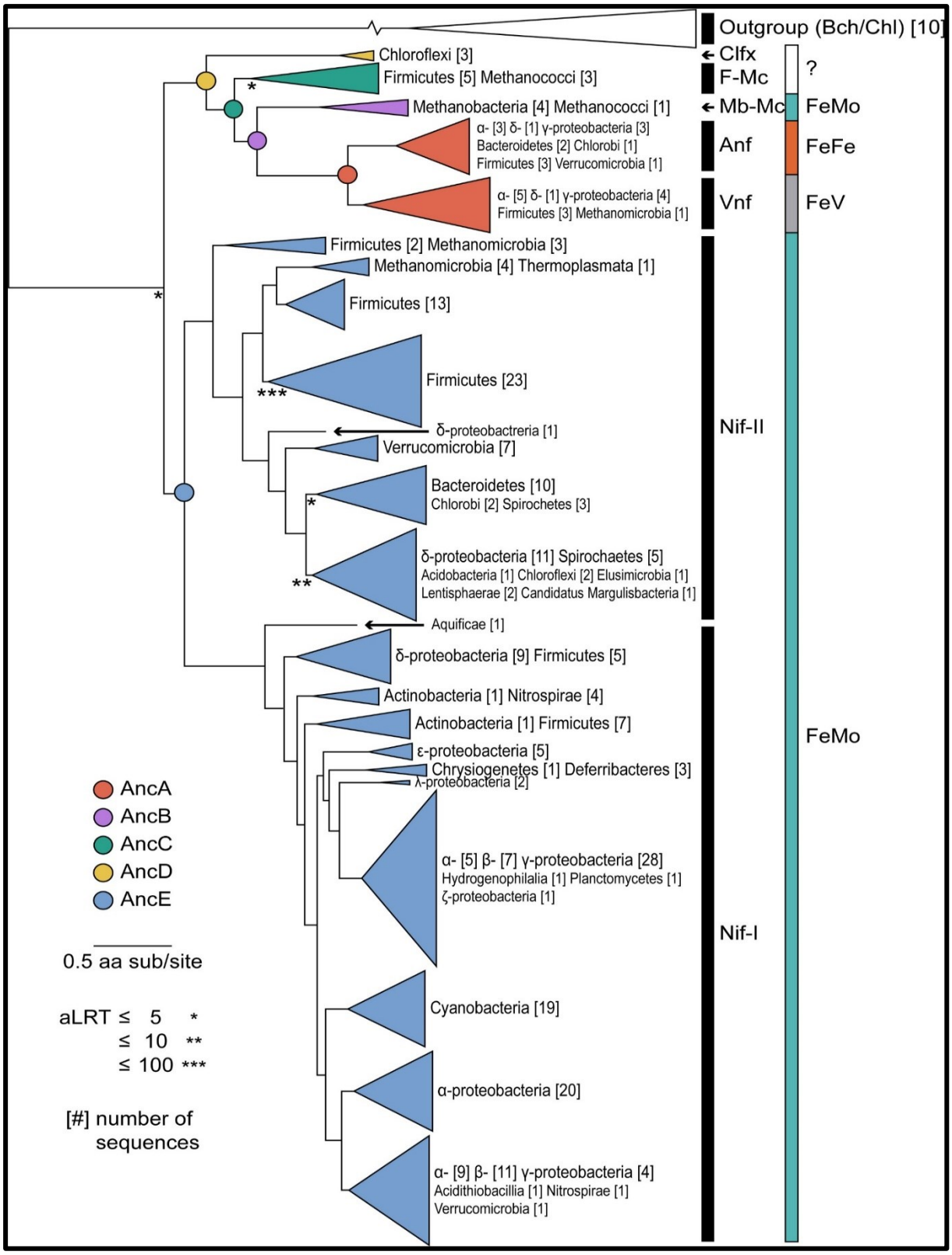


Figure 14. Maximum-likelihood protein tree of concatenated Nif/Anf/VnfHDK nitrogenases. Bch/ChlLNB was used as an outgroup. AncA–AncE = Ancestral nodes. An approximate likelihood-ratio test (aLRT) was used to derive branch support and branch length scale in units of amino acid substitutions per site. True branch length of the outgroup = 5,578 substitutions per site. Colour code: yellow = Chloroflexi, green = Firmicutes and Methanococci, purple = Methanobacteria and Methanococci, red = Vnf/And, blue = Nif**.** / This figure is reprinted with permission from Garcia *et al.* (2020) (137). Copyright 2020 *John Wiley & Sons Ltd*.

two major lineages of nitrogenases: *(i)* Mo-nitrogenases and *(ii)* V-, Fe-only nitrogenases including the three clades of uncharacterised nitrogenases. The model corroborates the observation that V- and Fe-only nitrogenases are only present in the genomes of organisms containing the Mo-nitrogenase isoform. Moreover, the V-/Fe- nitrogenase assembly relies on *nif* biosynthetic genes (138).

The FeMoco synthetic machinery encoded by *nifEN* arose from the gene duplication of *nifDK* (137). Parallel evolution of nitrogenase sequences, changes in metal specificity, and speciation in methanogens (137) resulting in the presence of all three nitrogenase types in this clade is another piece of evidence supporting the scenario that nitrogenase originated in this ancient group of archaea.

Interestingly, it has been proposed that the growing biomass of methanogens during the Archean itself created selective pressure for the evolution of nitrogenase (185). Namely, at this time, methanogens were dominating the biosphere (186), and high atmospheric methane concentrations were preventing abiotic nitrogen fixation through lightning discharge.

The diversification of diazotrophs into oxic niches and the evolution of strategies to protect the highly oxidative-damage-sensitive nitrogenase were driven by the competitive advantage conferred by N_2-fixation capabilities (139). These strategies include temporal (*Synechococcus*-like cyanobacteria), spatial (*Anabaena*-like cyanobacteria), and metabolic (*Azotobacte*r-like Proteobacteria) decoupling of nitrogen fixation and respiration (139). The anoxic to oxic transition was accompanied by a progressive evolution of accessory proteins involved in nitrogenase maturation and regulation. The dominant regulation strategy changed from the post-translational in anaerobes, to transcriptional regulation in obligate aerobes and facultative anaerobes (40, 139).

1.9. Balancing energy and nitrogen: the regulation of nitrogen fixation. Given the high metabolic expense of synthesizing, using, and maintaining such a complicated enzyme, it is not surprising that nitrogenase activity is carefully regulated. While the regulatory mechanisms differ among different taxa, different nodes of the nitrogen metabolism are universally controlled by a conserved and ancient superfamily of sensory proteins called P_{II} (187). Being found in all domains of Life, the P_{II} represents one of the most widely distributed superfamilies of signal transduction proteins known in nature (187, 188). The signal transduction occurs via direct interactions with the targets via the so-called T-loops, a typical structural feature conserved in the P_{II} proteins (187, 188). They have homotrimeric architecture, with each monomer harbouring a binding site for 2-oxoglutarate (2-OG) and

MgATP. The high cellular concentration of 2-OG signals nitrogen deficiency, whereas high cellular glutamine levels signal nitrogen sufficiency. This unique feature of detecting and interpreting the availability of cellular energy charge (ATP/ADP ratio), nitrogen (2-OG/glutamine), and carbon (2-OG) makes them a remarkable sensory tool to control central metabolism (187). In addition to the central role in nitrogen and carbon metabolism, they regulate the production of signaling metabolites (c-di-GMP), cofactors (NAD), and a range of transporters (187). P_{II} superfamily encompasses two deep-branching families that evolved most likely through a duplication: *(i)* GlnB-K family consisting of GlnB, GlnK, and similar proteins, and *(ii)* NifI family, consisting of $NifI_1$ and $NifI_2$ (189).GlnB-K and NifI have a role in the regulation of nitrogenase activity on different levels.

N_2-fixation is regulated on transcriptional and posttranslational levels, with notable differences between *Bacteria* and *Archaea* in the regulation mechanisms. In Gammaproteobacteria, including *A. vinelandii*, transcriptional regulation of *nif* genes is the dominant regulation strategy and is controlled by the transcriptional activator NifA and its antagonist NifL. In *A. vinelandii*, NifA and NifL are constitutively co-transcribed, and their expression is not dependent on nitrogen availability (40). NifA promotes the expression of *nif* genes in a typical enhancer-transcription factor fashion by binding to activator sequences upstream of *nif* genes and promoting the formation of the open σ^{54}-RNA polymerase complex (40). It can also sense the cellular nitrogen levels through 2-OG binding. NifA and NifL are a noncanonical two-component system since NifL inhibits NifA activity via a direct protein-protein interaction instead of phosphorylation. The P_{II} family protein, GlnK, further regulates the formation of the NifAL complex. GlnK ensures that that the formation of NifAL complex is coordinated with the availability of the nitrogen in the cell and it is accordingly reversibly uridylylated (Figure 15). If the cellular nitrogen levels are high, GlnK is not uridylylated and promotes a ternary complex formation with NifAL, inhibiting NifA activity (Figure 15). If the cellular nitrogen levels are low, GlnK is uridylylated, and 2-OG is bound to NifA, which causes the breakdown of the inhibitory GlnK-NifA-NifL complex (Figure 15).

In addition to nitrogen availability, the NifAL system responds to carbon availability (also signalled by 2-OG) and O_2 levels, each prioritised under different conditions. Namely, if the O_2 levels are high (the oxidised FAD bound to NifL promotes NifAL complex formation), the signals about the unfavourable redox state override the signals of favourable nitrogen and carbon levels, which would promote nitrogen fixation. Similarly, under low O_2 levels (the reduced $FADH_2$ bound to NifL would cause NifAL dissociation) and favourable carbon status, the signals of nitrogen excess will be decisive in preventing nitrogen fixation. Finally,

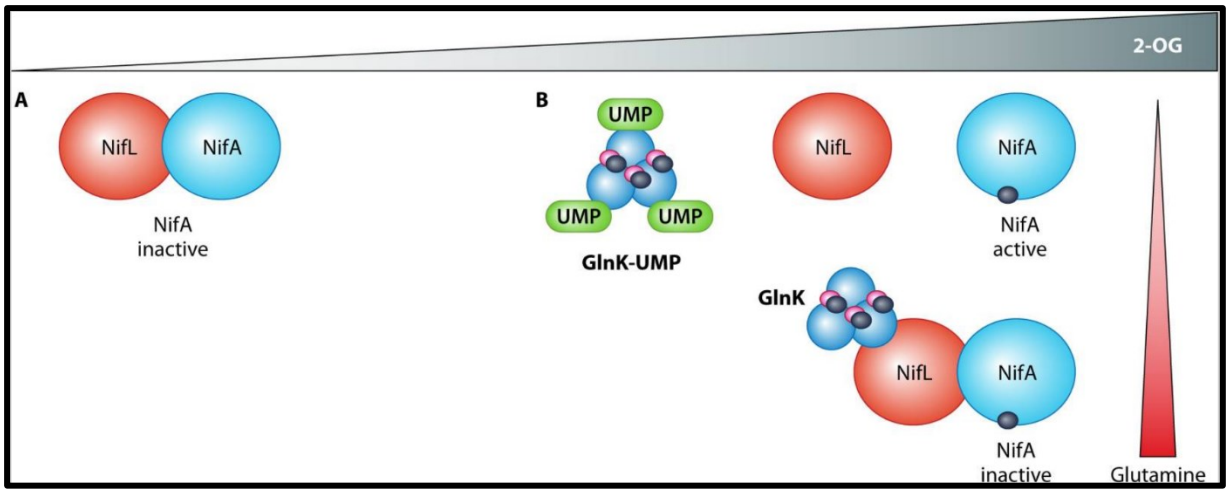


Figure 15. The mechanism of transcriptional regulation of *nif* genes by NifA-NifL-GlnK complex in *A. vinelandii* cells. The unconventional two-component system NifA-NifL controls the expression of *nif* genes in *A. vinelandii* in response to cellular oxygen, nitrogen, and carbon availability. 2-OG is the main signaling molecule of nitrogen availability in the cell. When the levels of 2-OG are high and signal low nitrogen levels, NifA is freely promotes the expression of *nif* genes. In contrast, when the 2-OG levels are low, signaling a high nitrogen level, deuridylylated GlnK promotes the formation of the NifA-NifL complex, inhibiting NifA. / This figure is reprinted with permission granted under the licence ID 1364204-1 from Huergo *et al.* (2015) (190). Copyright 2015 *American Society for Microbiology*.

nitrogen fixation will be inhibited under unfavourable carbon conditions, even with low O_2 and ammonia levels. It is important to note that this mechanism of NifAL action varies even among different Gammaproteobacteria. For example, in *K. oxytoca*, the expression of NifAL is under the control of the transcriptional factor NtrC and both uridylylated and non-uridylylated GlnK are preventing the NifAL complex formation (40).

The inhibition of nitrogenase activity on the post-translational level in bacteria is best studied and described in α-Proteobacteria such as *Rhodospirillum rubrum, Rhodobacter capsulatus*, and *Azospirillum brasilense*. This "switch-off" mechanism is mediated by the post-translational modification of nitrogenase reductase through reversible covalent ADP-ribosylation (189). The addition and removal of ADP-ribosyl moiety are catalysed by DRAT (dinitrogenase reductase ADP-ribosyltransferase) and DRAG (dinitrogenase reductase-activating glycohydrolase), respectively, whose activity is regulated by P_{II} family proteins GlnB, GlnK, and possibly GlnJ (189).

In *Archaea*, these regulatory networks are different and are best studied in the different species of *Methanosarcina* and *Methanococcus maripaludis* (158, 191), due to the availability of genetic tools. These studies uncovered different aspects of archaeal nitrogen fixation, notably their dependence on particular trace elements (192), presence of different nitrogenase

isoforms (193), organisation of *nif* operons (194-196), regulation of nitrogenase activity (197-204) and transcriptome of diazotrophic cells. The core organisation of *nif* operons present in methanogens, initially characterised in *M. maripaludis* (belonging to the *Methanococcales* (195)), has been shown to be conserved across different methanogenic species. In *M. maripaludis*, the *nif* operon contains seven genes: *nifHI$_1$I$_2$DKEN*, with *nifHDK* coding for the nitrogenase, *nifEN* for the MoFe-cofactor biosynthesis machinery, *nifI$_{1,2}$* for the regulatory proteins, while *nifX* still has an unknown role (199).

The transcriptional regulation of *nif* genes differs from that in Bacteria in two important manners: *(i)* it is governed by a distinct family of transcription factors, notably the NrpR repressor and NrpA activator, and *(ii)* it does not involve P_{II} proteins. NrpR was discovered in *M. maripaludis* and it is considered to be a global transcriptional regulator (master regulator) of genes involved in nitrogen metabolism in methanogens (205, 206). NrpR suppresses the transcription of *nif*, *nrpA,* and *glnA* genes (encoding the glutamine synthetase for NH_3-assimilation) by binding to the operators in their promoter regions (Figure 16A). 2-OG binding to NrpR releases its inhibition. Interestingly, similar to the NifAL bacterial system, the expression of NrpR is not regulated by nitrogen levels, and instead of forming an inhibitory complex with NrpA, NrpR binds to its operator sequence and prevents its transcription (Figure 16A).

In addition to the NrpA/R system, *Methanosarcinales* use small regulatory RNAs to regulate the transcriptional control of *nif* genes (204, 207). The best characterised sRNA$_{154}$ from *M. mazei* Gö1 enhances the translational initiation and stabilises the transcripts of *nifH*, *nifD*, *nifK*, *glnA,* and *nrpA* (203). The involvement of the sRNAs in the transcriptional control of *nif* genes or the homologues of sRNA$_{154}$ has not been reported in *Methanococcales.*

Like the α-Proteobacteria, archaea employ the nitrogenase switch-off mechanism via a post-translational regulation, with three notable differences: *(i)* the target of the switch off is the dinitrogenase instead of the nitrogenase reductase, *(ii)* the inhibition is mediated by a different P_{II} family protein, the NifI$_1$/NifI$_2$ complex (NifI$_{1,2}$), *(iii)* there are no covalent modifications of the target involved, and the inhibition is achieved through direct protein-protein contact instead (189). A study which aimed to correlate the different metabolic strategies with overall *nif* gene content and the phylogenetic distribution of individual *nif* genes in different taxa revealed that the distributions of *nifA* and *nifL* genes are positively correlated with the ability to use O_2 in metabolism, while the distributions of *nifI$_1$* and *nifI$_2$* genes were inversely correlated with the ability to use O_2 in metabolism, suggesting a shift in the primary strategies

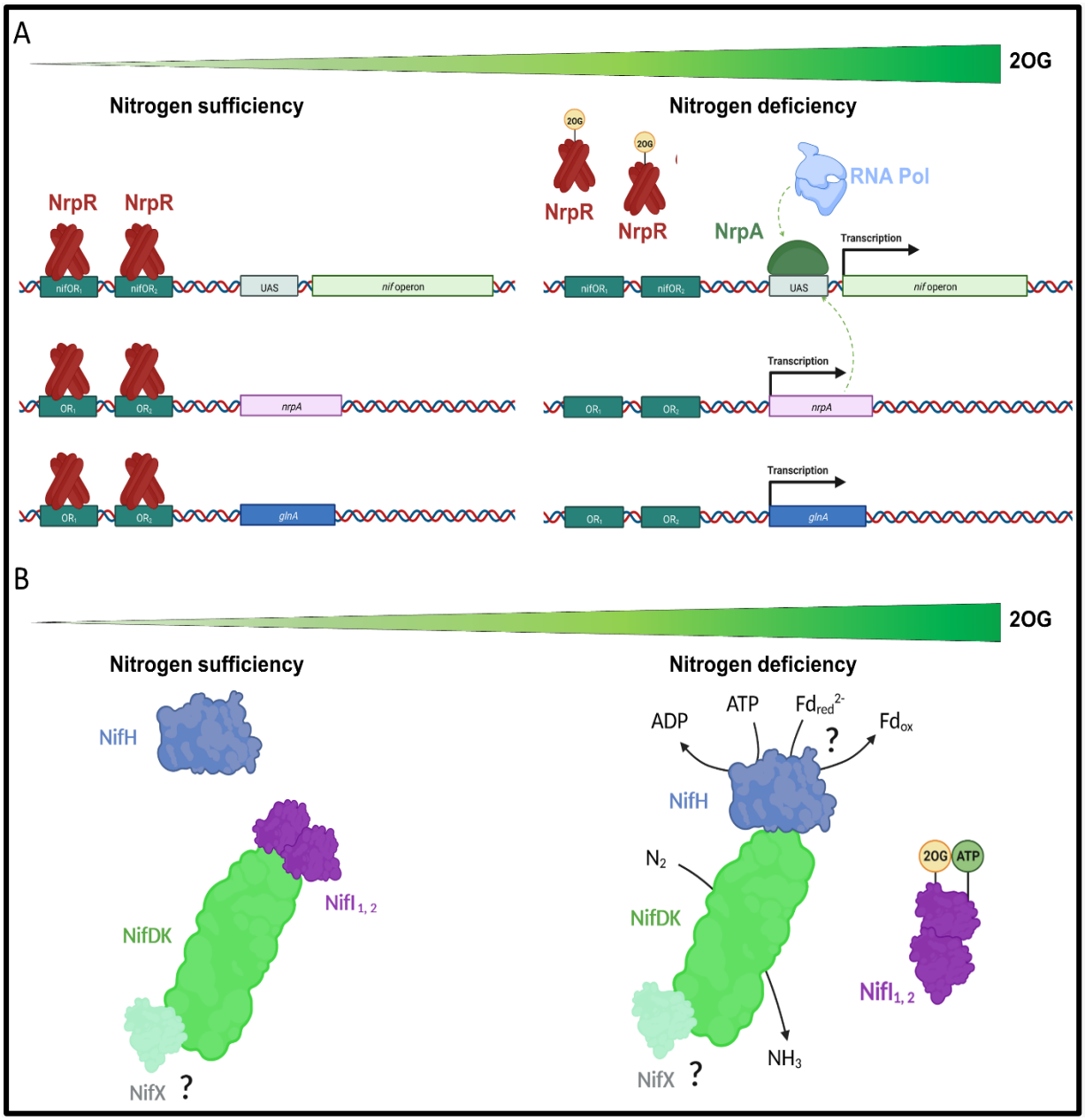


Figure 16. The mechanism of transcriptional and posttranscriptional regulation of nitrogenase in *Methanococcales*. A. Model of transcriptional regulation of *nif* genes expression by NrpR/NrpA in *Methanococcales*. Under low nitrogen conditions, NrpR prevents the expression of *nif, nrpA,* and *glnA* genes. This inhibition is released under high 2-OG levels signaling nitrogen deficiency. NrpA promotes the expression of the *nif* genes. **B.** Mechanism of the post-translational regulation of nitrogenase activity in *Methanococcales*. $NifI_{1,2}$ inhibitory complex regulates the access of the NifH to NifDK depending on the energetic cellular charge and nitrogen status. / This figure was created with Biorender (Biorender.com)

by which Nif is regulated during the transition from anaerobic ($NifI_{1,2}$) to aerobic (NifAL) metabolism (139).

This theory is further supported by the observation that nearly all organisms that lack *nifI1I2* have *nifA*. However, there are examples of taxa in which both

(anaerobic *Chlorobi*) or neither (aerobic cyanobacteria and anaerobic *Geobacter* spp.) *nifI*$_1$*I*$_2$ and *nifA* are found (139).

Like all P_{II} family proteins, $NifI_{1,2}$ responds to 2-OG and ATP levels to associate/dissociate with the target. While the T-loops are necessary for the binding of $NifI_{1,2}$ to the nitrogenase (198), they are not necessary for the $NifI_{1,2}$ complex formation. If enough nitrogen is available and a cellular energy deficiency, $NifI_{1,2}$ binds to the nitrogenase and prevents NifH association (Figure 16B). On the contrary, if the nitrogen becomes limiting and the energy is sufficient to sustain N_2-fixation, 2-OG and ATP bind to $NifI_{1,2}$ and release the nitrogenase inhibition (Figure 16B) (208).

$NifI_{1,2}$ complex also behaves differently compared to other P_{II} proteins (all homotrimers) in terms of oligomerization: although it is assumed to form a heterohexamer, $NifI_{1,2}$ can self-oligomerize in the presence of 2-OG to a higher-molecular-weight state (198). In the absence of the 2-OG, when it is complexed with nitrogenase, it also seems to form a higher-order molecular weight complex consistent with multiple dinitrogenases and $NifI_{1,2}$ molecules (198). The reason behind such oligomerization (199, 209) is unknown, and so is the mechanism of $NifI_{1,2}$ inhibition, which could physically block NifH binding or provoke conformational changes on the nitrogenase that would interfere with electron flow or catalysis.

1.10. (Hyper)thermophilic methanogens as the next generation of model organisms for studying biological nitrogen fixation. We have selected hyperthermophilic methanogens of order *Methanococcales* as our model organisms for this thesis due to a range of intriguing differences compared to the traditional diazotrophic model organisms. First, their nitrogenases belong to a phylogenetically distinct clade, suggesting the possibility of novel properties and regulatory mechanisms. Additionally, (hyper)thermophiles are known to possess robust enzymes and thrive in extreme conditions. These presumed unique adaptations enable nitrogen fixation at extreme temperatures (210), very high H_2-partial pressures (168, 169, 210) and WO_4^{2+} concentrations that would normally inhibit nitrogenases. These extraordinary environmental conditions under which they operate make them exceptionally well-suited for investigating the intricacies of nitrogenase regulation as well. As a result, we have chosen to focus our attention on two specific model organisms within this group.

The two model organisms used in this thesis, thermophilic *Methanothermococcus thermolithotrophicus* SN-1 (211) and hyperthermophilic *Methanocaldoccoccus infernus* ME (212), belong to the order of *Methanococcales*, which have only marine species. The

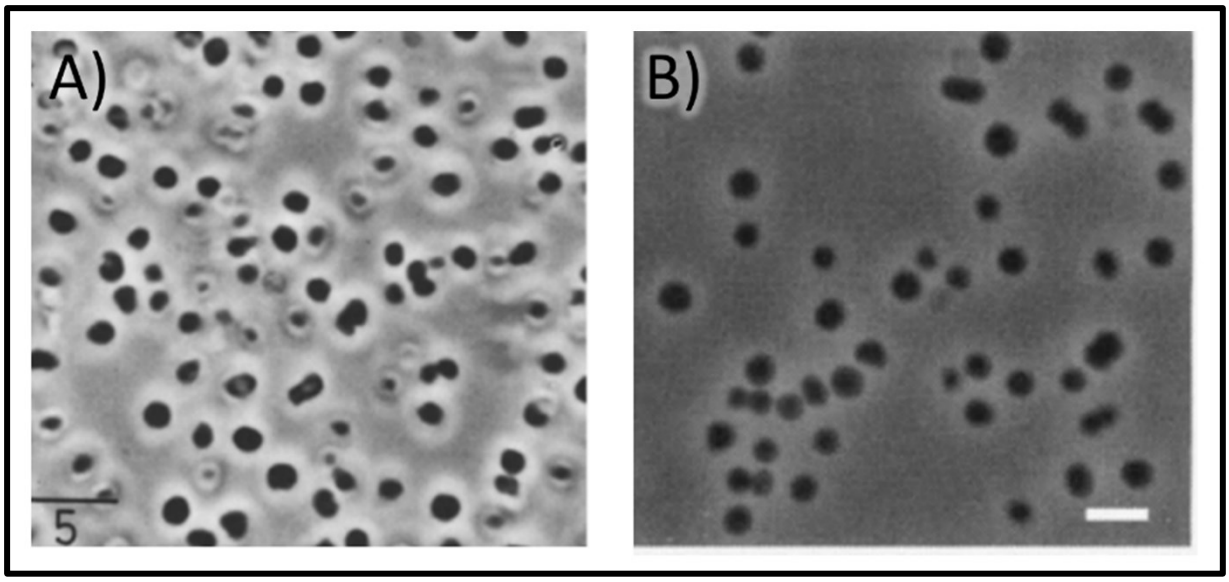


Figure 17. A. A phase contrast light micrograph of *Methanococcus thermolithotrophicus*. B. Phase contrast light micrograph of *Methanocaldococcus infernus*. / The Panel A of this figure is reprinted with permission granted under the licence ID 5553600732933 from Huber *et al.* (1982) (211). The Panel B of this figure is reprinted with permission granted under the licence ID 1356994-1 from Jeanthon *et al.* (1999) (212).

thermophile *M. thermolithotrophicus* has an optimal growth temperature of 65 °C and was isolated from shallow (0.5 m) geothermally heated sea sediments of Stufe di Nerone (Naples, Italy) (211). Coccoid cells have ~20 flagella and can occur as single cells or in pairs (Figure 17A). *M. thermolithotrophicus* can grow between 30 and 70 °C with an optimal pH between 6.5 and 7.5 (211). *M. infernus*, isolated from a deep-sea (3 000 m) white smoker hydrothermal vent at the Mid-Atlantic Ridge, grows at the optimal temperature of 85 °C (with a range between 50 and 93 °C and pH between 5.25 and 7.0) (212). Cells are also characteristically coccoid and heavily flagellated with several flagella tufts (Figure 17B) (212).

They both contain only one nitrogenase copy in their respective genomes, which is thought to be molybdenum-dependent and consequently annotated as *nif*. At the moment, genetic manipulations of these two organisms remain out of reach, but the tools developed for the closely related *M. jannaschii* have the potential to be accordingly adapted.

A native approach, in which the protein of interest is characterised directly from the native organism, leads to an unbiased view of the enzyme properly folded by dedicated chaperones, in complex with their natural partners, and loaded with the correct metallo-cofactor and post-translational modification, that will warranty an optimal enzymatic activity.

1.11. Research hypotheses and questions addressed in the thesis. This thesis was undertaken with the aim to characterize the diazotrophic metabolism and explore the nitrogen

assimilation system and its regulation in (hyper)thermophilic methanogens on the molecular level.

It was driven by four main objectives: *(i)* unravelling the molecular basis and identifying the protein partners involved in nitrogen fixation within (hyper)thermophilic methanogenic archaea through transcriptomic and physiological experiments **(Chapter 2)**; *(ii)* obtaining the native structure of archaeal dinitrogenase reductase via X-ray crystallography **(Chapter 3)**; *(iii)* pioneering the acquisition of the very first structure of archaeal nitrogenase via X-ray crystallography **(Chapter 4)**; *(iv)* unveiling the molecular mechanism behind nitrogenase post-translational regulation by P_{II} family proteins in *Archaea* **(Chapter 4)**. Furthermore, an established pipeline for anaerobic protein crystallisation was utilised to contribute to the accomplishment of two additional objectives: *(v)* determination of the functional and structural role of the Mo-nitrogenase MMB site **(Chapter 5)** and *(vi)* investigation of the cooperativity in the nitrogenase system **(Chapter 6)**. The two last chapters rely on *A. vinelandii* as a model organism. Based on these objectives, the following hypotheses were postulated:

H1: Diazotrophic growth causes metabolic remodelling in marine thermophilic methanogens.

Chapter 2 of the thesis presents a comprehensive transcriptomic analysis that answers the following questions: ***how nitrogen fixation is integrated into the methanogenic metabolism of M. thermolithotrophicus and what potential electron donors are involved in supporting nitrogenase activity.***

H2: The distinct evolutionary positions and thermostability of nitrogenase from Methanococcales can be explained by their unique structural features.

Chapter 3 provides the answers regarding ***the structures of dinitrogenase reductase from meso-, thermo-, and hyperthermophile marine methanogens and their compatibility with the phylogenetically distinct bacterial system*** from *A. vinelandii*. **Chapter 4** describes ***the first structural insights of the archaeal nitrogenase***.

H3: Nitrogenases from hyperthermophiles are an ideal model to study the mechanism of N_2-fixation catalysis.

Chapter 4 details ***resting and turn-over states trapped in the as-isolated enzyme from M. infernus, highlighting the physiological P^{1+} state of the P cluster***.

H4: The higher oligomerization of $NifI_{1,2}$ is required for its physiological role in regulating

the nitrogenase activity in (hyper)thermophilic archaea.

In **Chapter 4**, the thesis addresses two questions: the physiological and mechanistic ***reasons for the oligomerization of NifI$_{1,2}$ and the mechanism of nitrogenase inhibition at the molecular level.***

H5: The absence of the 16th Fe atom at the MMB site is not impacting Mo-nitrogenase functionality.

Chapter 5 answers ***whether the 16th Fe atom at the nitrogenase MMB site has any implications for the enzyme's electron transfer, catalytic activity, or structural stability*.**

H6: The nitrogenase exhibits negative cooperativity during NH_3 and H_2 reduction activity.

Chapter 6 gives an answer to the question of ***which mode of cooperativity nitrogenase exhibits during the reduction of physiological substrates*.**

1.12. An overview of enclosed manuscripts. Contributions to manuscripts.

Chapter 2: "Differential transcriptomics sheds light on the remodelling of gene expression during diazotrophy in the thermophilic methanogen *Methanothermococcus thermolithotrophicus*"

Nevena Maslać, Chandni Sidhu, Hanno Teeling, Tristan Wagner

Published in: *mBio Volume 13 • Number 6 • 22 November 2022*

DOI: https://doi.org/10.1128/mbio.02443-22

Author contributions: N.M. and T.W. designed the research. N.M. performed all culture experiments. C.S. processed the transcriptomic data. C.S., N.M., and T.W. interpreted the data. All authors wrote the paper.

In this manuscript, the growth of *M. thermolithotrophicus* under diazotrophic conditions is investigated using cultivation techniques and transcriptomic analysis. *M. thermolithotrophicus* was adapted to diazotrophy, and its culture medium was optimized. The physiological experiments revealed that diazotrophy depends on the presence of molybdenum, while the transcriptomic analysis highlighted that a substantial portion of the genome exhibited differential expression patterns during different stages of growth.

Chapter 3: "Structural comparison of nitrogenase reductase from three marine *Methanococcales*"

Provisional author list: Nevena Maslać, Cécile Cadoux, Pauline Bolte, Fenja Murken, Ross Dean Milton, Tristan Wagner

Manuscript in preparation (all presented models are fully refined and ready for validation)

N.M. and T.W. designed the research. C.C. and N.M. purified *Av*NifDK and *Av*NifH. F.M. optimised the protocol for the purification of recombinant *Mi*NifH. P.B. cultivated *M. infernus* cells and purified native *Mi*NifH. N.M. cultivated *M. thermolithotrophicus* cells and purified native *Mt*NifH. N.M. produced recombinant *Mm*/*Mt*/*Mi*NifH. N.M. performed the phylogenetic analysis, activity assays, and hrCN PAGE. X-ray data collection was performed by N.M. and T.W. Data processing, model building, structure refinement, validation and deposition were performed by N.M. and T.W. Structures were analysed by N.M. and T.W. R.D.M. and T.W. acquired funding to realise the project. The paper was written by N.M. and T.W. with contributions and final approval of all co-authors.

This study describes the first native purification of dinitrogenase reductase from thermophilic and hyperthermophilic marine methanogens. The enzymes were structurally characterised in the presence and absence of the MgADP. Furthermore, the study also includes the recombinant structure of dinitrogenase reductase from the mesophilic *M. maripaludis.* Despite exhibiting a high degree of conservation and retaining the defining structural features of P-loop NTPases, they do not demonstrate cross-reactivity in activity assays with Mo-nitrogenase from *A. vinelandii*, indicating the existence of specific functional differences.

Chapter 4: "Molecular basis of nitrogenase regulation by a P_{II} family protein in a hyperthermophilic methanogen"

Provisional author list: Nevena Maslać, Marie-Caroline Müller, Pauline Bolte, Tristan Wagner

Manuscript in preparation (presented models are not fully refined)

Author contributions: N.M. and T.W. designed the research. N.M. and P. B. cultivated *M. infernus,* purified and crystallized the proteins. N.M. cultivated *M. thermolithotrophicus*, purified and crystallized the proteins. X-ray data were collected by N.M. and T.W. Data processing, model building, and structure refinement of *Mt*Nif$I_{1,2}$ structure were performed by M-C.M. Data processing, model building, structure refinement, and analysis of all other presented structures were performed by N.M. and T.W. T.W. acquired funding to realize the

project. The paper was written by N.M. and T.W. with contributions and final approval of all co-authors. N.M. and T.W. are working on the manuscript.

This manuscript presents the groundbreaking achievement of elucidating the first crystal structure of an archaeal nitrogenase from the hyperthermophilic methanogen *Methanocaldococcus infernus*. The atomic resolution structures of the dinitrogenase alone describes new states of the metallocofactor, which have never been observed in an as-isolated Mo-nitrogenase. Furthermore, the study offers a complete understanding of the regulation system by providing the structure of the NifHDK complex and the inhibited $NifI_{1,2}$NifDK super complex of 0.86 MDa.

Chapter 5: "The 16th Fe of Mo-nitrogenase is not required for activity"

Cécile Cadoux*, Nevena Maslać*, Léa Di Luzio, Daniel Ratcliff, Wenyu Gu, Tristan Wagner, Ross D. Milton (*Equally contributed to the study)

Manuscript submitted to Angewandte Chemie

C.C., W.G. and R.D.M. designed the research. C.C., L.D.L., W.G. and R.D.M. constructed and produced the ß-D353G/D357G MoFe protein mutant. C.C. performed limited proteolysis; and C.C., L.D.L. and D.R. performed activity assays. N.M. performed the crystallisation. X-ray data collection was performed by N.M. and T.W. Data processing, model building, structure refinement, validation and deposition were performed by N.M. and T.W. Structures were analysed by N.M. and T.W. N.M. performed the sequence comparison analyses. R.D.M. and T.W. acquired funding to realize the project. The paper was written by all authors.

This study employed a combination of X-ray crystallography, bioinformatics, and biochemical and biophysical methods to investigate the influence of the MMB site (usually loaded with Fe) on the catalytic capabilities and stability of nitrogenase. A β-D353G/D357G variant (representing the mutated axial ligands of the 16th iron) of *A. vinelandii* NifDK was created to serve as a model system. Structural analysis revealed the loss of the MMB site and the absence of an anomalous signal. At the same time, further biochemical characterization indicated no significant alterations in the enzymatic properties or stability of the variant.

Chapter 6: "Nitrogen fixation and hydrogen evolution by sterically encumbered Mo-nitrogenase"

Cécile Cadoux, Daniel Ratcliff, Nevena Maslać, Wenyu Gu, Ioannis Tsakoumagkos, Sascha

Hoogendoorn, Tristan Wagner, Ross D. Milton

Published in: *JACS Au Volume 3 • Number 5 • 09 May 2023*

DOI: https://doi.org/10.1021/jacsau.3c00165

Author contributions: The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript. Credit: Cecile Cadoux conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing-original draft; Daniel Ratcliff conceptualization, data curation, formal analysis, investigation, methodology, writing-review & editing; Nevena Maslać conceptualization, data curation, formal analysis, investigation, methodology, writing-review & editing; Wenyu Gu conceptualization, formal analysis, investigation, methodology, writing-review & editing; Ioannis Tsakoumagkos conceptualization, formal analysis, investigation, methodology, writing-review & editing; Sascha Hoogendoorn conceptualization, data curation, formal analysis, investigation, methodology, resources, supervision, visualization, writing-review & editing; Tristan Wagner conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, resources, software, supervision, validation, visualization, writing-review & editing; Ross D. Milton conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing-original draft, writing-review & editing.

In this manuscript, the mode of cooperativity employed by nitrogenase was studied on sterically inhibited α-C45A/L158C variant of *A. vinelandii* Mo-nitrogenase created for this purpose. While the crystal structure revealed an almost identical conformation to the wild-type, biochemical characterization revealed an N_2-fixation activity with a maximum velocity supporting the negative cooperativity.

This thesis was funded by the Max Planck Society.

CHAPTER 2: Differential transcriptomics sheds light on the remodeling of gene expression during diazotrophy in the thermophilic methanogen *Methanothermococcus thermolithotrophicus*

Nevena Maslać[1], Chandni Sidhu[1], Hanno Teeling[1], Tristan Wagner[1]

[1] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

Published in:

mBio Volume 13 • Number 6 • 22nd November 2022

DOI: https://doi.org/10.1128/mbio.02443-22

Recieved: 31st of August 2022

Accepted: 1st of November 2022

Published online: 22nd of November 2022

Correspondence to Tristan Wagner (*twagner@mpi-bremen.de*)

The publication can be accesed at:

https://journals.asm.org/doi/10.1128/mbio.02443-22

DOI: *https://doi.org/10.1128/mbio.02443-22*

2.1. Abstract

Some marine thermophilic methanogens are able to perform energy-consuming nitrogen fixation despite deriving only little energy from hydrogenotrophic methanogenesis. We studied this process in *Methanothermococcus thermolithotrophicus* DSM 2095, a methanogenic archaeon of the order *Methanococcales*, that contributes to the nitrogen pool in some marine environments. We successfully grew this archaeon under diazotrophic conditions in both batch and fermenter cultures, reaching the highest cell density reported so far. Diazotrophic growth depended strictly on molybdenum and, in contrast to other diazotrophs, was not inhibited by tungstate or vanadate. This suggests an elaborate control of metal uptake and a specific metal recognition system for the insertion into the nitrogenase cofactor. Differential transcriptomics of *M. thermolithotrophicus* grown under diazotrophic conditions with ammonium-fed cultures as controls revealed upregulation of the nitrogenase machinery including chaperones, regulators, and molybdate-importers, as well as simultaneous upregulation of an ammonium-transporter and a putative pathway for nitrate/nitrite utilization. The organism thus employs multiple synergistic strategies for uptake of nitrogen nutrients during the early exponential growth phase without altering transcription levels for genes involved in methanogenesis. As a counterpart, genes coding for transcription and translation processes were downregulated, highlighting the maintenance of an intricate metabolic balance to deal with energy constraints and nutrient limitations imposed by diazotrophy. This switch in the metabolic balance included unexpected processes, such as upregulation of the CRISPR-Cas system, probably caused by drastic changes in transcription levels of putative mobile and virus-like elements.

2.2. Importance

The thermophilic anaerobic archaeon, *M. thermolithotrophicus*, is a particularly suitable model organism to study the coupling of methanogenesis to diazotrophy. Likewise, its capability to simultaneously reduce N_2/CO_2 into NH_3/CH_4 with H_2 makes it a viable target for biofuel production. We optimized *M. thermolithotrophicus* cultivation, resulting in considerably higher cell yields and enabling the successful establishment of N_2-fixing bioreactors. Improved understanding of the N_2-fixation process would provide novel insights into metabolic adaptations that allow this energy-limited extremophile to thrive under diazotrophy, for instance by investigating its physiology and uncharacterized nitrogenase. We demonstrate that diazotrophic growth of *M. thermolithotrophicus* is exclusively dependent on molybdenum, and complementary transcriptomics corroborated the expression of the

molybdenum nitrogenase system. Further analyses of differentially expressed genes during diazotrophy across three cultivation time points revealed insights into the response to nitrogen limitation and the coordination of core metabolic processes.

2.3. Keywords

Methanogenic archaea, N_2-fixation, thermophile, nitrogenase, nitrogen-starvation, differential transcriptomics, energy limitation, molybdenum.

2.4. Introduction

Methanogenic archaea generate about one gigaton of methane per year (1). This amounts to about half of the greenhouse gas methane in our atmosphere (2, 3). Methanogenesis is a strictly anaerobic process occurring in habitats in which electron acceptors other than CO_2 are depleted (4). It is accepted that under natural conditions methanogenesis provides an extremely low energy yield, causing methanogens to thrive close to the thermodynamic limits of life (1, 5, 6) Therefore, it came as a surprise when in 1984 two studies proved that *Methanothermococcus thermolithotrophicus* (7) and *Methanosarcina barkeri* (8), can perform nitrogen fixation, a very energy-demanding metabolic process that requires the hydrolysis of at least 16 ATP per molecule of fixed N_2 (9).

The *nifH* gene is widely used as a marker to identify nitrogen-fixing organisms. Multiple environmental studies detected *nifH* genes and transcripts from methanogens in diverse anoxic habitats, such as deep seawater and hydrothermal vent fluids (10), oligotrophic open seas (11), deep-sea methane seep sediments (12), and N_2-limited soils of salt marshes (13). Collectively these findings show that methanogens are indeed actively fixing N_2 in nature and thereby contribute considerably to global nitrogen cycling. It also implies that methanogens and anaerobic methanotrophs (ANMEs) (14) can overcome the largest activation barrier in biology of +251 $kJ.mol^{-1}$ (15) to break the N_2 triple bond.

The reduction of N_2 to NH_3 is catalyzed by the nitrogenase enzyme complex. The overall organization of this complex is highly conserved among *Bacteria* and *Archaea*. It is composed of a dinitrogenase reductase (iron protein, NifH) and dinitrogenase (iron-molybdenum protein, NifDK) containing one [$MoFe_7S_9C$-(R)-homocitrate] iron-molybdenum cofactor (FeMo-co) and a [8Fe-7S] P cluster (15). The nitrogenase-encoding *nif* genes and a plethora of accessory proteins with roles in regulation, biosynthesis and maturation of metal cofactors, can be clustered within one or in several operons or regulons depending on the organism (15-18). In addition to Nif, representing the most widespread and well-studied

system, at least two alternative nitrogenases are known: vanadium nitrogenase (VFe protein: Vnf) and iron-only nitrogenase (FeFe protein: Anf) (15). Both are considered to be evolutionary related to the Nif system (19).

Phylogenetic analyses suggest that the ancestral nitrogenase originated in anaerobic methanogenic archaea, from where it was subsequently transferred into the bacterial domain, most probably initially to the *Firmicutes* (15, 19-22). Despite the shared origin and a minimal conserved set of genes required for functionality, archaeal nitrogenases are very distinct from their bacterial homologs. This is reflected in the genetic organization of nitrogenase operon(s), regulation of their expression by the repressor NrpR and activator NrpA, and the unique mode of post-translational activity regulation through direct protein-protein interactions (23-25). These differences have been extensively studied in the mesophilic methanogen *Methanococcus maripaludis* (16, 17, 26-30) and different *Methanosarcina* strains (31-36) using well-established genetic systems for both species.

In this context it is noteworthy that *M. maripaludis* features a single operon consisting exclusively of *nif* genes (16, 28) (Fig. S1), whereas some *Methanosarcinales* harbor additional *vnf* and *anf* operons coding for V- and Fe-only nitrogenases, respectively. However, these operons all share the core organization of the *nif* operon in which *nifH*, *nifD* and *nifK* genes code for the two structural subunits of the nitrogenase system, while *nifE* and *nifN* play a putative role in the synthesis of FeMo-co. Genes *nifI*$_1$ and *nifI*$_2$ encode P_{II}-family regulators that are conserved in archaea, while *nifX* has no homology with the homonymous bacterial gene (Fig. S1). NifX, absent in *Methanosarcinales*, has an unknown function. It is not required for nitrogen fixation in *M. maripaludis*, since *nifX* in-frame deletion mutants are still capable of diazotrophic growth (17).

Only a few studies exist on the diazotrophic physiology of (hyper)-thermophilic *Methanococcales*. The N_2-fixation capabilities of these hydrogenotrophic methanogens are particularly interesting for several reasons: *(i*) their nitrogenases form a separate evolutionary branch (19) (Fig. S2, Fig. S3), *(ii*) they can perform N_2-fixation under very high H_2-partial pressure usually inhibiting nitrogenases (7, 8, 37), and *(iii*) they can fix nitrogen at the highest temperature ever described so far (i.e. up to 92 °C) (37). Together this suggests a unique adaptation that allows N_2-fixation to operate at both high temperatures (7) and extreme energy limitations.

We investigated the integration of N_2-fixation with other metabolic processes in thermophilic *M. thermolithotrophicus* by using a combined approach of physiological tests and differential

transcriptomics. Our results provide new insights into the adaptive strategies of *M. thermolithotrophicus* to energy and nutrient limitation stress inflicted by Mo-dependent diazotrophy.

2.5. Results

2.5.1. Nitrogen acquisition by *M. thermolithotrophicus*. We selected *M. thermolithotrophicus* DSM 2095 due to its remarkable chemolithoautotrophic capabilities and fast growth at 65 °C in mineral medium (38, 39). Series of incubations with different NH_4Cl concentrations, ranging from 0.1 to 16.8 mM, showed that *M. thermolithotrophicus* required a minimum of 10 mM NH_4Cl for best growth (Fig. 1A) in our optimized medium (see Materials and Methods for composition) (39, 40). Higher NH_4Cl concentrations of 25-200 mM NH_4Cl did not result in higher cell yields, and a 500 mM excess of NH_4Cl led to a decrease (Fig. S4). Continuous monitoring of ammonia consumption during growth on 10 mM NH_4Cl over a span of 26 hours confirmed depletion proportional to the observed increase in biomass (Fig. 1B). The culture reached stationary phase after 18 h during which most of the ammonia was consumed beyond the reliable detection limit.

M. thermolithotrophicus was adapted to diazotrophic conditions after three successive transfers to NH_4Cl-free medium. Under this condition, dissolved N_2 was the sole available nitrogen source. The cell density of the diazotrophic culture was six times higher than reported by Belay and colleagues in their original study on the discovery of diazotrophy in methanogens in 1984 (e.g., 1.85 achieved in 48 h as compared to 0.4 achieved in ~20 h (7)). Ammonia possibly released in the medium was measured during diazotrophic growth (Fig. 1C), and traces were detected during the mid-exponential phase reaching a maximum of 40 µM ammonia in the stationary phase (Fig. 1C). However, we cannot exclude that detected ammonia also originated from cell lysis rather than from active excretion.

In addition to batch diazotrophic cultures, *M. thermolithotrophicus* was successfully grown in a 10 L fermenter continuously supplied with H_2/CO_2 and N_2. To maintain the sulfur source in the medium, we replaced Na_2S used in batch cultures (which would be flushed out as H_2S) with Na_2SO_3 or Na_2SO_4 (see Materials and Methods)(39). Diazotrophic growth was not affected by HSO_3^-, a known inhibitor of methanogenesis (Fig. 1D) (41). While the final cell yield was similar to the one observed in batch cultures, division times were shorter (Fig. 1D). Fermenter-grown cultures supplemented with NH_4Cl had higher final yields than diazotrophic fermenter-grown cultures. However, in comparison to NH_4^+-grown batch cultures, the

difference was not very pronounced, as final yields were only slightly higher and the division times were similar.

2.5.2. N_2-fixation is molybdenum-dependent. We then investigated the nitrogenase type used for diazotrophy, taking into account that the *M. thermolithotrophicus* genome features only one *nifDK*. It must be noted that *M. thermolithotrophicus* resembles *M. maripaludis* in terms of physiology, and the latter was shown to require molybdenum (Mo) for its diazotrophic growth (28). Therefore, we first tested for molybdenum (Mo) dependency. Growth of cultures lacking either Mo, vanadium (V) or both metals in the medium were monitored simultaneously. Depletion of Mo and V from the media was achieved by three successive culture transfers to the same media without Mo and V. The highest optical density ($OD_{600\ nm}$) was reached when *M. thermolithotrophicus* was grown in the control medium containing both Mo and V. Growth was similar in the diazotrophic culture incubated without V, but with a lower final $OD_{600\ nm}$ (Fig. 2A). This trend was reproducible and suggests that V acts as a potential growth stimulator under such conditions. No growth was observed in the absence of Mo, with or without V (Fig. 2A). The minimal required Mo concentration of 0.1 µM was determined in a separate series of incubations with Mo concentrations ranging from 0.01 to 100 µM (Fig. 2B). To check, if Mo is essential for methanogenesis (42) or metabolic processes other than nitrogen fixation, we supplemented the culture grown in absence of Mo with NH_4Cl, which restored growth after an overnight incubation (Fig. 2A).

An inhibitory effect of tungstate (WO_4^{2-}) on diazotrophic growth has already been observed in *A. vinelandii* OP (43)*, Methanosarcina barkeri* 227 (32) and *M. maripaludis* strain S2 (28). In this case, WO_4^{2-} can inhibit MoO_4^{2-} uptake (44) and it can be incorporated into the nitrogenase, thereby rendering the protein inactive (43, 45, 46). These inhibition mechanisms are competitive and thus dependent on the MoO_4^{2-}/WO_4^{2-} ratio. Surprisingly, there was no observable inhibitory effect in *M. thermolithotrophicus* (Fig. 2C), even at a 1:1 MoO_4^{2-}/WO_4^{2-} ratio. In contrast, diazotrophic growth of *M. maripaludis* was inhibited at a 10:1 MoO_4^{2-}/WO_4^{2-} ratio (28). This observation might be explained by different affinities of the WO_4^{2-} and MoO_4^{2-} transport systems in both species. The ModABC transporter (47, 48) is present in both organisms and should transport MoO_4^{2-} and WO_4^{2-} (49). In addition, *M. thermolithotrophicus* has a highly specific WO_4^{2-} transporter TupABC (50) that is lacking in *M. maripaludis*. Instead, *M. maripaludis* has a third type of tungstate transporter WtpABC (49, 51), which can transport WO_4^{2-} and MoO_4^{2-} but has a higher affinity for WO_4^{2-} than has been shown for Mod and Tup in *Pyrococcus furiosus* (51). Furthermore, it is also known that Mod transporters can have different affinities for both oxyanions in different organisms (49, 52, 53), suggesting that

both transporter specificities and the specificity of the FeMo-co insertion machinery contribute to the W tolerance during diazotrophy.

2.5.3. Diazotrophy shifts expression of a large number of genes. Transcriptomics and DNA-microarrays have been used to investigate the complex metabolic and regulatory networks that control N_2-fixation in two model organisms*: Azotobacter vinelandii* (54) and *Methanosarcina mazei* Gö1 (36, 55, 56). These studies led to the discovery of three gene clusters (*rnf1*, *rnf2* and *fix*) coding for electron transfer systems that provide reducing equivalents to the nitrogenase in *A. vinelandii*, as well as multiple potential transcriptional and sRNA regulators in *Methanosarcina mazei* Gö1. Here, we used comparative transcriptome profiling of N_2-fixing *M. thermolithotrophicus* cultures versus NH_4^+-grown cultures to investigate the metabolic adaptations induced by nitrogen fixation. The experiment was conducted in biological triplicates at three different time points: 3 h (early exponential phase), 21 h (stationary phase-starvation due to the exhaustion of the gas phase), and 25 h post-inoculation (see Materials and Methods, Fig. 3A, and Supplementary Table 1). The gas phase was exchanged after sampling at the 21 h time point, since the cultures consumed H_2/CO_2. The robustness of our sample data was reflected in a corresponding principal component analysis (PCA), in which triplicates from different conditions and time points are clustered (Fig. 3B). More importantly, it revealed that 92% of the variation could be explained by the different sample treatments.

Transcriptome profiling across the three time-points revealed prominent changes (Fig. 3C). Out of the total 1,751 predicted genes (open reading frames, ORFs), 1,737 genes were found to have non-zero read counts. At 3 h, 12.3% (214) were differentially expressed with most (122/214) being upregulated under diazotrophic conditions. More than half of these genes (70/122) were of an unknown function without any homologs in the databases (their probable role is discussed further below). The difference was smaller at 21 h, with only 3.3% (57) of the total expressed genes having differential expressions, however at 25 h, 17.2% (302) were differentially expressed, with most of them downregulated (257/302) under diazotrophic conditions (Supplementary Table 2).

To check the overall expression of genes irrespective of differential expression, we mapped reads to predicted genes in terms of transcripts per million (TPM). Details of this analysis are summarized in the sections below.

2.5.4. Nitrogen- and molybdate-acquisition genes are mostly affected during the early exponential phase. As anticipated, upregulated genes at 3 h included the *nif* operon

(*nifHI$_1$I$_2$DKXEN)* and molybdate-acquisition genes (*modABC*). This illustrates a swift response of *M. thermolithotrophicus* to diazotrophic conditions. The highest log_2-fold change (FC) was observed for *nifI$_1$I$_2$HDKX* genes with values ranging from 8.8 to 7.0 (Supplementary Table 2). The proteins of the FeMo-co biosynthetic machinery, *nifE and nifN*, were also among the upregulated genes with log_2FC of 5.3 and 4.6 respectively (Fig. 4A-B). An essential gene for FeMo-co biosynthesis, *nifB*, (57, 58), was transcribed at a lower level with a log_2FC of 1.8.

The *mod* operon, coding for the three subunits of a molybdate ABC transporter (*modABC*), was also upregulated (Fig. 4A), with *modA* being strongly expressed with a log_2FC of 5.5 (Supplementary Table 2). This agrees with our physiological data that show a clear Mo-dependency under diazotrophic conditions. Although three *modA* instances are present in the genome, only the *modA* gene in vicinity of *nif* operon that is part of the complete *mod* operon was highly expressed. The *tupA* gene coding for the tungstate transporter was also upregulated (log_2FC: 1.7), strengthening the hypothesis that both transporter specificities and adaptation of the cofactor biosynthetic machinery contribute to the W tolerance during diazotrophy.

The other genes upregulated at 3 h included the ammonium transporter *amtB$_2$* and its P_{II}-family regulatory protein *glnK$_2$* with the log_2FC of 6.2 and 5.8, respectively (Supplementary Table 2). The genome of *M. thermolithotrophicus* harbors two different *amtB* genes with their associated *glnK* regulators (40). Interestingly, *amtB$_1$* and *glnK$_1$* were not differentially expressed (Supplementary Table 2, Figure 4A-B), which might point to the existence of an internal regulator or an additional promotor. In addition to the *nif* and *amtB/glnK* operons, which are known to be upregulated during diazotrophy from previous studies (54, 55, 59), the ammonium assimilating glutamine synthetase (*glnA*) gene was also upregulated (Supplementary Table 2). However, the transcript level of the second enzyme involved in the N-assimilation, the glutamate synthase, remained unchanged, suggesting that GlnA is the rate limiting step in ammonium assimilation.

Unexpectedly, *narK* coding for a putative nitrate transporter and *narB* coding for a molybdopterin-dependent nitrate reductase (60), were also highly expressed during diazotrophy in the early exponential stage (Fig. 4A-B, Supplementary Table 2). If nitrate would be imported and reduced in the cell, the oxidant nitrite would be generated and could damage the highly oxidation-sensitive methanogenic machinery (61). A gene coding for a F_{420}-sulfite reductase isoform (also belonging to Fsr group 1 but different to the one naturally

expressed under sulfite condition described in (39)), which co-occurs with *narK* and *narB*, would be a plausible candidate for nitrite detoxification, since Fsr has been recognized to catalyze nitrite reduction (39). *M. thermolithotrophicus* has been reported to be able to grow on nitrate as the only source of nitrogen (62), which might involve these genes.

In addition, three genes coding for the enzymes of the molybdopterin biosynthetic pathway (*moaE*, *mobB* and *moeB*) were upregulated. They might be involved in supplying the putative nitrate reductase with molybdopterin, while the formylmethanofuran dehydrogenase enzyme used for methanogenesis harbors a tungstopterin (42). The pathway of molybdopterin biosynthesis has been extensively studied and it is highly conserved (63). The pathway of tungstopterin biosynthesis is thought to be homologous up to the step of metal insertion (64). It is proposed that the biosynthetic machinery is able to distinguish between the two metals and insert the correct metal into the respective enzymes by employing MoeA isoenzymes selective for either molybdate or tungstate (64, 65). All archaeal genomes sequenced so far encode two *moeA* isoforms sharing around 40% identity (64). This might explain how organisms are able to correctly express different proteins with molybdopterin and tungstopterin simultaneously (64). However, this hypothesis still lacks experimental validation.

As described in Fig. 4C-D, the expression of genes partaking in N_2-fixation and molybdenum acquisition remained unchanged at 21 h and 25 h and therefore, no differential expression was observed. The overall expression of genes in terms of TPM followed a similar pattern. An established marker for metabolic activity of methanogens, the transcription level of *mcrA* (66), was used as a reference point for comparison (Fig. 4B-D).

None of the methanogenesis pathway genes showed any difference in transcription levels at the early exponential phase. However, after 25 h nearly all genes involved in methanogenesis were downregulated. The expression of F_{420}-reducing hydrogenase was maintained at the same level at all time-points to supply reduced F_{420} from H_2-oxidation. Hydrogenotrophic *Methanococcales* can alternatively use formate as electron source (67, 68), and in this case a putative formate transporter (*fdhC*) and formate dehydrogenase (*fdhF*) were found to be downregulated at 3 h. In addition to methanogenesis, numerous anabolic processes were shut down at 25 h, such as carbon-assimilation (e.g. pyruvate:ferredoxin oxidoreductase), amino-acid metabolism (e.g. ketol-acid reductoisomerase), lipid biosynthesis (e.g. hydroxymethylglutaryl-CoA synthase), ATP synthesis (i.e. ATP-synthase) and vitamins and coenzymes biosynthesis (e.g. *hemE*). Such a decrease in catabolic and anabolic processes

combined with the downregulation of genes involved in S-layer formation and cellular division (e.g. *ftsZ*) suggests a fine-tuned metabolic mode of energy saving to prioritize nitrogen fixation.

2.5.5. Transcriptional and translational machineries are considerably downregulated. Both the transcriptional and translational machineries responded negatively to diazotrophic conditions after 3 h. This included downregulation of RNA polymerase subunits (*rpoA₂H*), the sigma factor 70 (*rpoD*, controlling the transcription of housekeeping genes), and 35 ribosomal proteins (Supplementary Table 2). Our ribosomal RNA expression data has the limitation that rRNA removal treatment was performed before sequencing. Still, we could detect downregulation of rRNA expression, which was corroborated by downregulation of ribosomal proteins. Some genes with a putative function in tRNA/ribosome biogenesis and biosynthesis of nucleotide and amino acid precursors were also negatively affected, corroborating a deep impact on the overall translation process under N_2-fixing conditions. For example, two key enzymes (transketolase and transaldolase) of the non-oxidative branch of the pentose phosphate pathway (69) for synthesis of nucleotides and histidine (from ribose 5-phosphate), as well as aromatic amino acids (from erythrose 4-phosphate) precursors were downregulated (Supplementary Table 2). This is another notable difference from *Methanosarcina*, in which the upregulation of genes involved in the synthesis of aromatic amino acids was observed (55). Notably, $tRNA^{Thr}$ and $tRNA^{Ser}$ were upregulated while $tRNA^{Ala}$ and $tRNA^{Pro}$ were downregulated at 3 h. Taken together, changes in the expression of all the mentioned genes contribute to the restriction of the entire translation process. The DNA replication system seems not to be impacted, since we could not observe any changes in the transcript levels of DNA polymerase encoding genes. This is in line with the results of Belay and coworkers, who showed that during diazotrophic growth the cellular protein content of *M. thermolithotrophicus* was significantly reduced (7).

2.5.6. High expression of CRISPR-Cas genes and putative viral genes. Components of the CRISPR-Cas virus defense system were also upregulated under diazotrophy at 3 h, including Csm1-5 of the Csm effector complex and CRISPR-associated proteins (Cas proteins) Cas4-6 (Supplementary Table 2). This might be explained by presence of a prophage that is expressed when the cells are energy depleted or otherwise stressed. The Phaster server, a tool for finding prophages in bacterial genomes (70), was unsuccessful in detecting any complete prophages in the *M. thermolithotrophicus* genome. However, we identified a locus of 28 co-occurring ORFs that were highly expressed at all-time points (Fig. 5). Twenty out of these genes had no characterized homologs, while eight encoded putative virus-like, replication,

and mobile genetic elements. We used models generated by Alphafold2 (71) and a membrane prediction tool to gain further information on these sequences and examined 18 confident models (Fig. 5). Five of these proteins were predicted to be secreted or embedded in the membrane, but structural homologs were scarce based on the predicted models (Supplementary Table 3). The protein encoded by the third gene of the locus (Fig. 5) has already been structurally characterized in *Thermococcales* (72) and is believed to be a virus-like element. While *Methanococcales* share some of these genes, *Methanosarcina mazei*, a model organism for the *Methanosarcinales*, contains only two of them in its genome (Supplementary Table 3). These 28 co-occurring genes might be derived from a small plasmid transferred by conjugation, or represent a yet unknown prophage, although we could not identify a protein that could be responsible for independent insertion or replication of this element. This region could therefore be considered part of the dark matter in archaeal genomes (73). However, active transcription of all the genes of this locus might be a plausible explanation for the observed upregulation of the CRISPR-Cas system.

Another locus of 26 co-occurring ORFs was downregulated under diazotrophic conditions at all-time points (Fig. S5). Again, most of these genes had unknown functions, with the exception of a putative transcriptional regulator, a mini chromosome maintenance protein, and a recombinase (Supplementary Table 4). Four of these genes were predicted to encode secreted proteins, and eight contained transmembrane segments. Based on these observations, we assume that these 26 co-occurring open reading frames are also derived from mobile elements or prophage-associated genes.

2.6. Discussion

Diazotrophy allows microbes to survive when nitrogen becomes limited in their natural habitats. This is also the case for (hyper)thermophilic methanogens, such as *M. thermolithotrophicus*, which has been shown to rely on diazotrophy in different environments (10-11, 13). A notable exception among *Methanococcales* is *Methanocaldococcus jannaschii*, a non-diazotrophic methanogen isolated from a white smoker on the East Pacific Rise. While ammonium concentrations at the *M. jannaschii* isolation site were not measured (74), some parts of the East Pacific Rise, such as Guaymas Basin, are known to feature notable ammonium concentrations (e.g. 15.3 mM (75)). Availability of this inorganic nitrogen source might have resulted in a complete loss of nitrogen fixing abilities in *M. jannaschii*.

Due to its active N_2-fixation, *M. thermolithotrophicus* represents one of the contributors to the available nitrogen pool in specific environments (10-11, 13). In our laboratory cultures, *M.*

thermolithotrophicus released only minute amounts of up to 40 µM ammonia to the medium, which might have just resulted from passive diffusion or cell lysis. It therefore seems that N_2-fixation *in M. thermolithotrophicus* is precisely controlled to avoid any losses, likely by the $NifI_{1,2}$ regulation system (Fig. 6). The excessive energy cost inflicted by nitrogenase activity was noticeable in our physiology experiments (Fig. 1), as it reduced final yields by a factor of two in both batch and fermenter grown cultures.

Transcriptome profiling under diazotrophic conditions revealed that *M. thermolithotrophicus* relies on multiple synergistic strategies that ensure both sufficient N_2-fixation and energy preservation to support cellular growth (Fig. 6). In addition, *M. thermolithotrophicus* enhances nitrogen acquisition by increasing ammonia uptake via its *amtB*$_2$ transporter and glutamine synthetase overall activity - a strategy that has been described before based on proteomics of nitrogen-starved *M. maripaludis* cultures (59). While upregulation of the *nif* and *amtB/glnK* operons under nitrogen limitation have been previously reported in other diazotrophs (54, 55, 59), the upregulation of a putative nitrate transporter (*narK*), a nitrate reductase (*narB*), and a new isoform of F_{420}-dependent sulfite reductase (*fsr*) reported in this study is so far unique to *M. thermolithotrophicus*.

Like *M. thermolithotrophicus*, the hyperthermophile *Methanocaldococcus infernus* has also been reported to grow on nitrate as sole nitrogen source (76). *M. infernus* was isolated from a smoking crater of the Logatchev hydrothermal vent field (76, 77) from which other nitrate-reducing organisms have been isolated as well (78). Although hydrothermal fluids have been reported to be depleted in nitrate and nitrite (79), bottom seawater can contain nitrate for use as electron acceptor and nitrogen source. The enzyme which can possibly be used for intracellular nitrate reduction by these methanogens is the nitrate reductase NarB. NarB is expected to harbor a tungstopterin or molybdopterin cofactor, which might explain upregulation of molybdopterin biosynthesis genes (*moaE*, *mobB*, *moeB*) under diazotrophic conditions. In this context, the molybdate transporter Mod would play a dual role, supplying Mo for both the nitrogenase and NarB metallo-cofactor.

Strict dependence on Mo for diazotrophy has also been described for *Methanococcus maripaludis* (28). Likewise, our results support that the single nitrogenase operon of *M. thermolithotrophicus* encodes a molybdenum nitrogenase. This is corroborated by the *M. thermolithotrophicus* phylogenetic position (19), as well as high transcription levels of the *mod* operon (encoding a molybdate ABC transporter) adjacent to the nitrogenase operon under diazotrophy. In addition to *modA* within the *mod* operon, *M. thermolithotrophicus*

features two additional *modA* instances. These are not co-localized with *modBC* genes and were not differentially expressed in our experiments. Although multiple *modA* genes are present in different diazotrophic organisms, including *A. vinelandii* and the methanogens *M. maripaludis* and *M. mazei*, the benefit of multiple *modA* genes is not understood yet. Interestingly, in *M. mazei* (55), no upregulation of Mod transporters was observed. In *Methanosarcinales* expression of *modABC* is under the control of the transcriptional regulator ModE. This regulator also affects transcriptional regulation of different nitrogenases types, if these are present (80, 81). *Methanococcales* are devoid of *modE* homologs, which might explain the observed upregulation of the *mod* operon under diazotrophic growth in *M. thermolithotrophicus*.

The *nif* operon exhibited the highest $\log_2$FC values between both conditions. However, changes in transcript levels of the *nifE* and *nifN* FeMo-co biosynthesis genes were less prominent than those of the structural *nif* genes. The former are possibly required in lower amounts, as was the case for the FeMo-co biosynthesis gene *nifB*. Levels of *nifX* transcription were similar to those of the *nifDK* nitrogenase genes, suggesting potential association with the nitrogenase complex. The *nifX* gene is not essential for nitrogen fixation in *M. maripaludis*, and has not been detected in the nitrogenase complex of this archaeon. Instead, it has been proposed to take on the role of the FeMo-co biosynthesis genes (17). However, since *M. maripaludis nifX* is not homologous to *M. thermolithotrophicus nifX*, future experimental evidence is required to clarify its putative function and association to NifDK in (hyper)-thermophilic methanogens.

The enormous ATP investment required for diazotrophy did not affect the transcription of methanogenesis genes at the onset of diazotrophic growth. This observation corroborates previous studies (55, 59) stating that the proportion of the overall energy required for maintenance increases when switching to diazotrophy (82). The high energy demand for nitrogen fixation is counterbalanced by drastic downregulation of the transcription and translation machineries. The question is, whether this is an *a priori* and thus targeted metabolic adaptation, or a mere general consequence of energy starvation. It has been described that transcription and translation are reduced upon nutrient limitations. After all, 40-70% of the cellular ATP pool in growing bacteria is attributed to protein synthesis (83). It is also known that bacteria can balance tRNA abundances when under stress to selectively regulate the translation of stress-induced proteins. These proteins contribute to the response and adaptation to different types of stress, including nutrient limitation (84). It has recently been shown that *M. maripaludis* adopts a resource relocation strategy upon energy depletion

(85): instead of reducing ribosome numbers, the cells rather redistribute available energy and decrease catabolic and ribosomal activities; a strategy also described in bacteria (86). Therefore, our results suggest a targeted response associated with switching to a diazotrophic lifestyle. In comparison, no such changes in the transcript levels of transcriptional and translational genes were detected in either *Azotobacter vinelandii* (54) or *M. mazei* (55) when switching to diazotrophic growth, even though in the case of the former 30% of the genes is differentially expressed (as observed by differential transcriptomics), and in case of the latter 5% (as observed by microarray analysis).

Studies in *M. mazei* and *M. maripaludis* confirmed the prominent regulation of nitrogen metabolism at the transcriptional level, particularly the importance of the global nitrogen regulatory repressor NrpR of nitrogen assimilation genes (24, 25, 35, 87). NrpR represses transcription by binding to the *nif* and *glnA* promoter regions in a 2-oxoglutarate-dependent manner. NrpR from *M. thermolithotrophicus* shares 38.2% identity with the one from *M. mazei.* Interestingly, NrpR expression levels are not regulated by nitrogen supply (24, 25). Thus, although a homolog of NrpR does exist in *M. thermolithotrophicus*, its transcription level remained unchanged during diazotrophic growth. In contrast to NrpR, NrpA, the *nif* promotor-specific activator, sharing 32% identity with the one from *M. mazei*, is known to be upregulated upon nitrogen limitation (88). We did not observe this in our experiments suggesting that NrpA could be constitutively expressed in *M. thermolithotrophicus.*

One major difference between *M. mazei* and *M. thermolithotrophicus*, however is the role of regulatory sRNAs. *Methanosarcina mazei* Gö1 expresses multiple sRNA genes in response to nitrogen limitation (56), with $sRNA_{154}$ being the best characterized (36). In *M. mazei* $sRNA_{154}$ stabilizes the nitrogenase, *glnA1* and *nrpA* mRNAs (36). While $sRNA_{154}$ is highly conserved within *Methanosarcinales*, there are no $sRNA_{154}$ homologs in *Methanococcales* (89). Consequently we did not detect such a pattern in *M. thermolithotrophicus*.

Transcriptome analysis can provide valuable information on proteins involved in nitrogen-fixation. For instance, transcriptome analysis allowed the discovery that in *A. vinelandii* electron bifurcation is coupled to a flavodoxin that fuels the nitrogenase (54). However, we could not detect any putative novel candidates that could shuttle electrons to the nitrogenase, and we propose reduced ferredoxin as a candidate for electron delivery, a known electron carrier for anabolic reactions. *M. thermolithotrophicus* would use H_2-oxidation for ferredoxin reduction, an endergonic process that requires the coupling of the influx of sodium ions by the

Ech complex. The reduced ferredoxin would then drive the reduction of NifH for N_2-fixation (Figure 6).

Here we present novel insights into the metabolic rebalancing that methanogens employ to accommodate Mo-dependent diazotrophy. Further studies at the protein level are required to decipher the mechanistic of this complex adaptation in greater detail. In particular molecular investigations of the regulatory functions of $NifI_{1,2}$, the role of NifX, and the intrinsic properties of the thermostable NifHDK complex will unveil the secrets of the astonishing diazotrophic capabilities of *M. thermolithotrophicus*.

2.7. Material and methods

2.7.1. Growth conditions. *Methanothermococcus thermolithotrophicus* DSM 2095 (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) was grown under anoxic conditions in minimal mineral media with 1 bar overpressure of either H_2:CO_2 (80%:20%) for non-diazotrophic or 1.2 bar overpressure H_2:CO_2:N_2 (58.2%:14.5%:27.3%) for diazotrophic cultures. Cultures were grown in 250 mL serum flasks (Glasgerätebau Ochs, Bovenden, Germany) sealed with rubber stoppers and aluminium crimps in final volumes of 10 mL with a 1:10 inoculum. Serum flasks and media were made anoxic prior to inoculation by sparging with N_2 and two final gas exchanges with H_2:CO_2 (80%:20%). Incubation was done at 65 °C, in the dark, without shaking.

2.7.2. Medium composition. The used minimal mineral medium was prepared as described in Müller *et al*. 2021 (40), but with replacement at equal final concentration of $Fe(NH_4)_2(SO_4)_2$ x $12H_2O$ by $FeCl_2$ x $4H_2O$, and of Na_2SeO_3 x $5H_2O$ by Na_2SeO_4. The used trace metal solution 100 fold concentrated contained 7.1 mM nitrilotriacetic acid, 0.45 mM $MnCl_2$ x $2H_2O$, 0.68 mM $FeCl_3$ x $6H_2O$, 0.41 mM $CaCl_2$, 0.76 mM $CoCl_2$, 0.66 mM $ZnSO_4$ x $6H_2O$, 0.28 mM $CuSO_4$, 0.19 mM Na_2MoO_4 x $2H_2O$, 0.38 mM $NiCl_2$ x $6H_2O$ and 0.19 mM VCl_3. The final pH was adjusted to 6.0 by addition of NaOH pellets. The final media were subsequently made anoxic by several degassing and N_2-addition cycles (minimum 25 cycles). The same medium, but without NH_4Cl, was used for diazotrophic cultures. Na_2S was used as both a reductant and sulfur source at a final concentration of 1.5 mM in all cases.

2.7.3. Adaptation to diazotrophic conditions. *M. thermolithotrophicus* was adapted to growth under diazotrophic conditions after NH_4Cl depletion from the media by three successive transfers to the same media without NH_4Cl with a headspace containing 1.2 bar

H_2:CO_2:N_2 (58.2%:14.5%:27.3%) as described above. Cultures grown in media without NH_4Cl with a gas phase of 1 bar H_2:CO_2 (80%:20%) were used as negative controls.

2.7.4. Influence of trace metal availability on diazotrophic growth and tungstate inhibition. To determine influences of Mo and V on diazotrophic growth, we depleted media of Mo and V by three successive transfers of diazotrophic *M. thermolithotrophicus* cultures to media prepared as described above, but without Na_2MoO_4 x $2H_2O$ or VCl_3 or without both. The minimal concentration of Mo for diazotrophic growth was determined in a series of incubations, in which an already Mo-depleted diazotrophic culture was supplemented with Na_2MoO_4 x $2H_2O$ concentrations ranging from 0.01 to 100 µM. W inhibition was tested by supplementing the already W-depleted diazotrophic culture with 0.001 to 10 µM Na_2WO_4 x $2H_2O$. A supplementation with the concentration of 100 µM gave unreproducible results (data not shown). W depletion was done as described above, with transfers to the media prepared without Na_2WO_4 x $2H_2O$.

2.7.5. Cultivation in a fermenter. *M. thermolithotrophicus* was continuously grown in a 10 L fermenter (BIOSTAT® B plus, Sartorius, Göttingen, Germany) under diazotrophic conditions with either 10 mM Na_2SO_3 or Na_2SO_4 as sulfur source instead of Na_2S. The final culture volume was 7 L for the culture grown with Na_2SO_3 and 6 L for the culture grown with Na_2SO_4. The culture was continuously sparged with H_2:CO_2 (80%:20%) and N_2 in the ratio of 1:1 and stirred with the speed of 500 rotation per minute, at 65 °C. As an inoculum, cultures cultivated in the same media were used in a ratio of 1:10.

2.7.6. Ammonia measurement. Ammonia concentrations were measured in the culture supernatant: 0.5 mL culture aliquots were subsampled aerobically at each time point, cells were pelleted by centrifugation for 5 min at 15,700 x *g* using 5415R Microcentrifuge (Eppendorf, Hamburg, Germany) and the supernatant was frozen at -20° C until further use. Ammonia was measured by the salicylate-nitroprussidine method (90) in ROTILABO® F-profile microtitration plates (Carl Roth GmbH, Karlsruhe, Germany). Standards in the range from 0 to 600 µM NH_4Cl were prepared in the same medium used for cultivation. 80 µM of salicylate reagent (424.7 mM sodium salicylate, 193.8 mM tri-sodium citrate dihydrate, 193.8 mM di-sodium tartrate dihydrate and 0.95 mM sodium nitroprusside dihydrate) and 80 µM of hypochlorite reagent (10% sodium hypochlorite and 1.5 M NaOH mixed in the ratio of 1:36) were added to 40 µl of each sample. The plate was additionally mixed for 5 minutes on a shaker and incubated in the dark at the room temperature for 45 min. The absorbance was

measured at 650 nm using an Infinite 200 PRO plate reader (Tecan, Männedorf, Switzerland) at room temperature.

2.7.7. Evolutionary analyses. NifHDK, VnfHDK and AnfHDK sequences of 35 selected species were aligned using MUSCLE (91) (default parameters) in MEGA11 (92, 93). Afterwards, ambiguous positions were removed for each possible pairing (pairwise deletion option). The final alignment contained a total of 485 NifH positions, 729 NifD positions and 563 NifK positions. The sequence's evolutionary history was inferred using the Neighbor-Joining method (94) with the JTT matrix model for multiple substitutions (95). The analysis was conducted in MEGA11 with ChlLNB from *Chlorobium limicola* as outgroup.

2.7.8. Transcriptomics set-up. The cultures for transcriptomic profiling were grown in triplicates in batch as described above, but scaled up to culture volumes of 60 mL in 1 L pressure resistant Duran bottles. Inocula used to start the cultures were adapted to the respective conditions prior to inoculation as described above. Subsamples were taken anaerobically on ice after 3 h, 21 h and 25 h, and transferred to an anaerobic chamber, where they were pelleted and subsequently frozen in liquid nitrogen immediately after being taken out of the anaerobic chamber. Samples were stored at -80 °C until further use. Samples were shipped on dry ice to the Max Planck Genome Centre in Cologne, where they were processed and sequenced. RNA extraction and quality control, including rRNA removal, were also done at the Max Planck Genome Centre in Cologne.

2.7.9. Transcriptome sequencing and analysis. Transcriptome sequencing was performed on an Illumina HiSeq 3000 (San Diego, CA, USA). Information on the raw reads are summarized in Supplementary Table 1. Raw RNA reads were quality trimmed and repaired using the *bbduk* and *repair.sh* scripts of the BBMap v35.14 suite (https://sourceforge.net/projects/bbmap/). Reads with a minimum length of 70 bp and a quality score of 20 were filtered for ribosomal RNAs using SortMeRNA v3.0.3 (96). Remaining mRNA reads were mapped against the *M. thermolithotrophicus* reference genome using Bowtie2 (97) as part of the SqueezeMeta v1.3.1 pipeline (98). The *DESeq2* R package (99) was subsequently used to calculate $\log_2$-fold changes, standard errors, test statistics and adjusted p-values. Changes in expression levels with adjusted p-values (padj) <0.05 and a minimum twofold change ratio ($\log_2$-fold change of 1 or higher) were considered significant (54).

The high-quality genome of *M. thermolithotrophicus* was used as a reference for mapping RNA reads. The genomic DNA of *M. thermolithotrophicus* was isolated using the protocol

from Platero *et al.* (100) and sequenced on PacBio Sequel II platform using a single SMRT cell at the Max Planck Genome Center in Cologne. The genome was assembled using Flye v2.7 (101) and annotated as part of part of SqueezeMeta pipeline. Briefly, the ORFs were predicted using Prodigal (102) and similarity searches for GenBank (103), eggNOG (104), KEGG (105), were done using Diamond (106). HMM homology searches were done by HMMER3 (107) for the Pfam database (108).

2.7.10. Sequence submission. The *M. thermolithotrophicus* genome sequence and the transcriptome raw reads are available under the ENA project number PRJEB53446.

2.7.11. Structural modeling and bioinformatic analysis. Alphafold2 was run with default parameters for all generated models. Predictions of membrane regions and overall topology were run on the DeepTMHMM server (https://dtu.biolib.com/DeepTMHMM/)(109) as of June 20th, 2022).

2.8. Acknowledgements

We thank the Max Planck Institute for Marine Microbiology for continuous support and the Max Planck Genome Centre Cologne for RNA library preparation and sequencing. We also thank Dr. Mark Schweizer for setup the Alphafold2 pipeline and Dr. Susanne Erdmann for her helpful discussion regarding the virus-like/mobile elements encoding regions.

2.9. Funding

This study was funded by the Max Planck Society.

2.10. Author contributions

N.M. and T.W. designed the research. N.M. performed all culture experiments. C.S. processed the transcriptomic data. C.S., N.M. and T.W. interpreted the data and all authors wrote the paper.

2.11. References

1. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nat Rev Microbiol 6:579-591.

2. Shima S, Huang G, Wagner T, Ermler U. 2020. Structural Basis of Hydrogenotrophic Methanogenesis. Annu Rev Microbiol 74:713-733.

3. Offre P, Spang A, Schleper C. 2013. Archaea in Biogeochemical Cycles. Annu Rev Microbiol 67:437-457.

4. Conrad R. 2020. Importance of hydrogenotrophic, aceticlastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: A mini review. Pedosphere 30:25-39.

5. Buckel W, Thauer RK. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na(+) translocating ferredoxin oxidation. Biochim Biophys Acta 1827:94-113.

6. Lemaire ON, Jespersen M, Wagner T. 2020. CO_2-Fixation Strategies in Energy Extremophiles: What Can We Learn From Acetogens? Front Microbiol 11:486.

7. Belay N, Sparling R, Daniels L. 1984. Dinitrogen fixation by a thermophilic methanogenic bacterium. Nature 312:286-288.

8. Murray PA, Zinder SH. 1984. Nitrogen fixation by a methanogenic archaebacterium. Nature 312:284-286.

9. Harris DF, Lukoyanov DA, Kallas H, Trncik C, Yang Z-Y, Compton P, Kelleher N, Einsle O, Dean DR, Hoffman BM, Seefeldt LC. 2019. Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism To Achieve N_2 Reduction. Biochemistry 58:3293-3301.

10. Mehta MP, Butterfield DA, Baross JA. 2003. Phylogenetic Diversity of Nitrogenase (*nifH*) Genes in Deep-Sea and Hydrothermal Vent Environments of the Juan de Fuca Ridge. Appl Environ Microbiol 69:960-970.

11. Man-Aharonovich D, Kress N, Zeev EB, Berman-Frank I, Béjà O. 2007. Molecular ecology of *nifH* genes and transcripts in the eastern Mediterranean Sea. Environ Microbiol 9:2354-2363.

12. Dang H, Luan X, Zhao J, Li J. 2009. Diverse and Novel *nifH* and *nifH*-Like Gene Sequences in the Deep-Sea Methane Seep Sediments of the Okhotsk Sea. Appl Environ Microbiol 75:2238-2245.

13. Bae HS, Morrison E, Chanton JP, Ogram A. 2018. Methanogens Are Major Contributors to Nitrogen Fixation in Soils of the Florida Everglades. Appl Environ Microbiol 84:1-16.

14. Dekas AE, Poretsky RS, Orphan VJ. 2009. Deep-Sea Archaea Fix and Share Nitrogen in Methane-Consuming Microbial Consortia. Science 326:422-426.

15. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. 2018. Exploring the alternatives of biological nitrogen fixation. Metallomics 10:523-538.

16. Kessler PS, Blank C, Leigh JA. 1998. The *nif* Gene Operon of the Methanogenic Archaeon *Methanococcus maripaludis*. J Bacteriol 180:1504-1511.

17. Kessler PS, Leigh JA. 1999. Genetics of Nitrogen Regulation in *Methanococcus maripaludis*. Genetics 152:1343-1351.

18. Halbleib CM, Ludden PW. 2000. Regulation of Biological Nitrogen Fixation. J Nutr 130:1081-1084.

19. Boyd ES, Hamilton TL, Peters JW. 2011. An alternative path for the evolution of biological nitrogen fixation. Front Microbiol 2:1-11.

20. Boyd ES, Anbar AD, Miller S, Hamilton TL, Lavin M, Peters JW. 2011. A late methanogen origin for molybdenum-dependent nitrogenase. Geobiology 9:221-232.

21. Boyd ES, Peters JW. 2013. New insights into the evolutionary history of biological nitrogen fixation. Front Microbiol 4:1-12.

22. Boyd ES, Garcia Costas AM, Hamilton TL, Mus F, Peters JW. 2015. Evolution of Molybdenum Nitrogenase during the Transition from Anaerobic to Aerobic metabolism. J Bacteriol 197:1690-1699.

23. Leigh JA, Dodsworth JA. 2007. Nitrogen Regulation in Bacteria and Archaea. Annu Rev Microbiol 61:349-377.

24. Lie TJ, Leigh JA. 2003. A novel repressor of *nif* and *glnA* expression in the methanogenic archaeon *Methanococcus maripaludis*. Mol Microbiol 47:235-246.

25. Lie TJ, Wood GE, Leigh JA. 2005. Regulation of *nif* Expression in *Methanococcus maripaludis*: ROLES OF THE EURYARCHAEAL REPRESSOR NrpR, 2-OXOGLUTARATE, AND TWO OPERATORS. J Biol Chem 280:5236-5241.

26. Blank CE, Kessler PS, Leigh JA. 1995. Genetics in Methanogens: Transposon Insertion Mutagenesis of a *Methanococcus maripaludis nifH* Gene. J Bacteriol 177:5773-5777.

27. Kessler PS, Daniel C, Leigh JA. 2001. Ammonia Switch-Off of Nitrogen Fixation in the Methanogenic Archaeon *Methanococcus maripaludis*: Mechanistic Features and Requirement for the novel GlnB homologues, $NifI_1$ and $NifI_2$. J Bacteriol 183:882-889.

28. Kessler PS, McLarnan J, Leigh JA. 1997. Nitrogenase Phylogeny and the Molybdenum Dependence of Nitrogen Fixation in *Methanococcus maripaludis*. J Bacteriol 179:541-543.

29. Dodsworth JA, Cady NC, Leigh JA. 2005. 2-Oxoglutarate and the PII homologues $NifI_1$ and $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. Mol Microbiol 56:1527-1538.

30. Dodsworth JA, Leigh JA. 2006. Regulation of nitrogenase by 2-oxoglutarate-reversible, direct binding of a PII-like nitrogen sensor protein to dinitrogenase. Proc Natl Acad Sci USA 103:9779-9784.

31. Lobo AL, Zinder SH. 1990. Nitrogenase in the Archaebacterium *Methanosarcina barkeri* 227. J Bacteriol 172:6789-6796.

32. Lobo AL, Zinder SH. 1988. Diazotrophy and Nitrogenase Activity in the Archaebacterium *Methanosarcina barkeri* 227. Appl Environ Microbiol 54:1656-1661.

33. Chien Y-T, Zinder SH. 1996. Cloning, Functional Organization, Transcript Studies, and Phylogenetic Analysis of the Complete Nitrogenase Structural Genes (*nifHDK2*) and Associated Genes in the Archaeon *Methanosarcina barkeri* 227. J Bacteriol 178:143-148.

34. Ehlers C, Veit K, Gottschalk G, Schmitz RA. 2002. Functional organization of a single *nif* cluster in the mesophilic archaeon *Methanosarcina mazei* strain Gö1. Archaea 1:143-150.

35. Weidenbach K, Ehlers C, Kock J, Ehrenreich A, Schmitz RA. 2008. Insights into the NrpR regulon in *Methanosarcina mazei* Gö1. Arch Microbiol 190:319-332.

36. Prasse D, Förstner KU, Jäger D, Backofen R, Schmitz RA. 2017. $sRNA_{154}$ a newly identified regulator of nitrogen fixation in *Methanosarcina mazei* strain Gö1. RNA Biol 14:1544-1558.

37. Mehta MP, Baross JA. 2006. Nitrogen Fixation at 92 °C by a Hydrothermal Vent Archaeon. Science 314:1783-1786.

38. Huber H, Thomm M, König H, Thies G, Stetter KO. 1982. *Methanococcus thermolithotrophicus*, a Novel Thermophilic Lithotrophic Methanogen. Arch Microbiol 132:47-50.

39. Jespersen M, Pierik AJ, Wagner T. 2022. The structure of the F_{420}-dependent sulfite-detoxifying enzyme from *Methanococcales* reveals a prototypical sulfite-reductase with assimilatory traits. bioRxiv doi:10.1101/2022.04.07.487323:2022.04.07.487323.

40. Müller MC, Wagner T. 2021. The Oxoglutarate Binding Site and Regulatory Mechanism Are Conserved in Ammonium Transporter Inhibitors GlnKs from *Methanococcales*. Int J Mol Sci 22:8631.

41. Balderston WL, Payne WJ. 1976. Inhibition of Methanogenesis in Salt Marsh Sediments and Whole-Cell Suspensions of Methanogenic Bacteria by Nitrogen Oxides. Appl Environ Microbiol 32:264-269.

42. Wagner T, Ermler U, Shima S. 2018. Formyl-Methanofuran Dehydrogenase, p 1-18. *In* Scott RA (ed), Encyclopedia of Inorganic and Bioinorganic Chemistry. John Wiley & Sons, Ltd.

43. Benemann JR, Smith GM, Kostel PJ, McKenna CE. 1973. TUNGSTEN INCORPORATION INTO *AZOTOBACTER VINELANDII* NITROGENASE. FEBS Lett 29:219-221.

44. Bulen WA. 1961. EFFECT OF TUNGSTATE ON THE UPTAKE AND FUNCTION OF MOLYBDATE IN *AZOTOBACTER AGILIS*. J Bacteriol 82:130-134.

45. Nagatani HH, Brill WJ. 1974. NITROGENASE V. THE EFFECT OF Mo, W and V ON THE SYNTHESIS OF NITROGENASE COMPONENTS IN *AZOTOBACTER VINELANDII*. Biochim Biophys Acta 362:160-166.

46. Siemann S, Schneider K, Oley M, Müller A. 2003. Characterization of a Tungsten-Substituted Nitrogenase Isolated from *Rhodobacter capsulatus*. Biochemistry 42:3846-3857.

47. Grunden AM, Shanmugam KT. 1997. Molybdate transport and regulation in bacteria. Arch Microbiol 168:345-354.

48. Self WT, Grunden AM, Hasona A, Shanmugam KT. 2001. Molybdate transport. Res Microbiol 152:311-321.

49. Aguilar-Barajas E, Díaz-Pérez C, Ramírez-Díaz MI, Riveros-Rosas H, Cervantes C. 2011. Bacterial transport of sulfate, molybdate, and related oxyanions. Biometals 24:687-707.

50. Makdessi K, Andreesen JR, Pich A. 2001. Tungstate Uptake by a Highly Specific ABC Transporter in *Eubacterium acidaminophilum*. J Biol Chem 276:24557-24564.

51. Bevers LE, Hagedoorn PL, Krijger GC, Hagen WR. 2006. Tungsten Transport Protein A (WtpA) in *Pyrococcus furiosus*: the First Member of a New Class of Tungstate and Molybdate Transporters. J Bacteriol 188:6498-6505.

52. Balan A, Santacruz-Pérez C, Moutran A, Ferreira LC, Neshich G, Gonçalves Barbosa JA. 2008. Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. Biochim Biophys Acta 1784:393-399.

53. Imperial J, Hadi M, Amy NK. 1998. Molybdate binding by ModA, the periplasmic component of the *Escherichia coli mod* molybdate transport system. Biochim Biophys Acta 1370:337-346.

54. Hamilton TL, Ludwig M, Dixon R, Boyd ES, Dos Santos PC, Setubal JC, Bryant DA, Dean DR, Peters JW. 2011. Transcriptional Profiling of Nitrogen Fixation in *Azotobacter vinelandii*. J Bacteriol 193:4477-4486.

55. Veit K, Ehlers C, Ehrenreich A, Salmon K, Hovey R, Gunsalus RP, Deppenmeier U, Schmitz RA. 2006. Global transcriptional analysis of *Methanosarcina mazei* strain Gö1 under different nitrogen availabilities. Mol Gen Genomics 276:41-55.

56. Jäger D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA. 2009. Deep sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. Proc Natl Acad Sci USA 106:21878-21882.

57. Jenner LP, Cherrier MV, Amara P, Rubio LM, Nicolet Y. 2021. An unexpected P-cluster like intermediate *en route* to the nitrogenase FeMo-co. Chem Sci 12:5269-5274.

58. Kang W, Rettberg LA, Stiebritz MT, Jasniewski AJ, Tanifuji K, Lee CC, Ribbe MW, Hu Y. 2021. X-Ray Crystallographic Analysis of NifB with a Full Complement of Clusters: Structural Insights into the Radical SAM-Dependent Carbide Insertion During Nitrogenase Cofactor Assembly. Angew Chem Int Ed Engl 60:2364-2370.

59. Xia Q, Wang T, Hendrickson EL, Lie TJ, Hackett M, Leigh JA. 2009. Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen *Methanococcus maripaludis*. BMC Microbiol 9:149.

60. Moreno-Vivián C, Cabello P, Martínez-Luque M, Blasco R, Castillo F. 1999. Prokaryotic Nitrate Reduction: Molecular Properties and Functional Distinction among Bacterial Nitrate Reductases. J Bacteriol 181:6573-6584.

61. Duin EC, Wagner T, Shima S, Prakash D, Cronin B, Yanez-Ruiz DR, Duval S, Rumbeli R, Stemmler RT, Thauer RK, Kindermann M. 2016. Mode of action uncovered for the specific reduction of methane emissions from ruminants by the small molecule 3-nitrooxypropanol. Proc Natl Acad Sci USA 113:6172-6177.

62. Belay N, Jung KY, Rajagopal BS, Kremer JD, Daniels L. 1990. Nitrate as a Sole Nitrogen Source for *Methanococcus thermolithotrophicus* and Its Effect on Growth of Several Methanogenic Bacteria. Curr Microbiol 21:193-198.

63. Schwarz G. 2005. Molybdenum cofactor biosynthesis and deficiency. Cell Mol Life Sci 62:2792-2810.

64. Bevers LE, Hagedoorn P-L, Hagen WR. 2009. The bioinorganic chemistry of tungsten. Coord Chem Rev 253:269-290.

65. Seelmann CS, Willistein M, Heider J, Boll M. 2020. Tungstoenzymes: Occurrence, Catalytic Diversity and Cofactor Synthesis. Inorganics 8:44.

66. Alvarado A, Montañez-Hernández LE, Palacio-Molina SL, Oropeza-Navarro R, Luévanos-Escareño MP, Balagurusamy N. 2014. Microbial trophic interactions and *mcrA* gene expression in monitoring of anaerobic digesters. Front Microbiol 5:597.

67. Lupa B, Hendrickson EL, Leigh JA, Whitman WB. 2008. Formate-Dependent H_2 Production by the Mesophilic Methanogen *Methanococcus maripaludis*. Appl Environ Microbiol 74:6584-6590.

68. Wood GE, Haydock AK, Leigh JA. 2003. Function and Regulation of the Formate Dehydrogenase Genes of the Methanogenic Archaeon *Methanococcus maripaludis*. J Bacteriol 185:2548-2554.

69. Bräsen C, Esser D, Rauch B, Siebers B. 2014. Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation. Microbiol Mol Biol Rev 78:89-175.

70. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16-21.

71. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596:583-589.

72. Keller J, Leulliot N, Soler N, Collinet B, Vincentelli R, Forterre P, van Tilbeurgh H. 2009. A protein encoded by a new family of mobile elements from Euryarchaea exhibits three domains with novel folds. Protein Sci 18:850-855.

73. Makarova KS, Wolf YI, Forterre P, Prangishvili D, Krupovic M, Koonin EV. 2014. Dark matter in archaeal genomes: a rich source of novel mobile elements, defense systems and secretory complexes. Extremophiles 18:877-893.

74. Jones WJ, Leigh JA, Mayer F, Woese CR, Wolfe RS. 2004. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. Arch Microbiol 136:254-261.

75. Wankel SD, Bourbonnais A, Charoenpong C. 2017. 8. Microbial nitrogen cycling processes at submarine hydrothermal vents, p 179-222. *In* Jens K (ed), Life at Vents and Seeps, vol 2. De Gruyter, Berlin, Boston.

76. Jeanthon C, L'Haridon S, Reysenbach AL, Vernet M, Messner P, Sleytr UB, Prieur D. 1998. *Methanococcus infernu*s sp. nov., a novel hyperthermophilic lithotrophic

methanogen isolated from a deep-sea hydrothermal vent. Int J Syst Bacteriol 48 Pt 3:913-919.

77. Petersen S, Kuhn K, Kuhn T, Augustin N, Hékinian R, Franz L, Borowski C. 2009. The geological setting of the ultramafic-hosted Logatchev hydrothermal field (14°45′N, Mid-Atlantic Ridge) and its influence on massive sulfide formation. Lithos 112:40-56.

78. Miroshnichenko ML, Kostrikina NA, Chernyh NA, Pimenov NV, Tourova TP, Antipov AN, Spring S, Stackebrandt E, Bonch-Osmolovskaya EA. 2003. *Caldithrix abyssi* gen. nov., sp. nov., a nitrate-reducing, thermophilic, anaerobic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent, represents a novel bacterial lineage. Int J Syst Evol Microbiol 53:323-329.

79. Vetriani C, Voordeckers JW, Crespo-Medina M, O'Brien CE, Giovannelli D, Lutz RA. 2014. Deep-sea hydrothermal vent *Epsilonproteobacteria* encode a conserved and widespread nitrate reduction pathway (Nap). ISME J 8:1510-1521.

80. Studholme DJ, Pau RN. 2003. A DNA element recognised by the molybdenum-responsive transcription factor ModE is conserved in Proteobacteria, green sulphur bacteria and Archaea. BMC Microbiol 3:24.

81. Chanderban M, Hill CA, Dhamad AE, Lessner DJ. 2021. *Methanosarcina acetivorans* simultaneously produces molybdenum, vanadium, and iron-only nitrogenases in response to fixed nitrogen and molybdenum depletion. bioRxiv doi:10.1101/2021.06.03.447018.

82. Fardeau ML, Peillex JP, Belaïch JP. 1987. Energetics of the growth of *Methanobacterium thermoautotrophicum* and *Methanococcus thermolithotrophicus* on ammonium chloride and dinitrogen. Arch Microbiol 148:128-131.

83. Pontes MH, Sevostyanova A, Groisman EA. 2015. When Too Much ATP Is Bad for Protein Synthesis. J Mol Biol 427:2586-2594.

84. Torrent M, Chalancon G, de Groot NS, Wuster A, Madan Babu M. 2018. Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. Sci Signal 11:eaat6409.

85. Müller AL, Gu W, Patsalo V, Deutzmann JS, Williamson JR, Spormann AM. 2021. An alternative resource allocation strategy in the chemolithoautotrophic archaeon *Methanococcus maripaludis*. Proc Natl Acad Sci USA 118.

86. Khaova EA, Kashevarova NM, Tkachenko AG. 2022. Ribosome Hibernation: Molecular Strategy of Bacterial Survival (Review). Appl Biochem Microbiol 58:213-231.

87. Lie TJ, Dodsworth JA, Nickle DC, Leigh JA. 2007. Diverse homologues of the archaeal repressor NrpR function similarly in nitrogen regulation. FEMS Microbiol Lett 271:281-288.

88. Weidenbach K, Ehlers C, Schmitz RA. 2014. The transcriptional activator NrpA is crucial for inducing nitrogen fixation in *Methanosarcina mazei* Gö1 under nitrogen-limited conditions. FEBS J 281:3507-3522.

89. Prasse D, Schmitz RA. 2018. Small RNAs Involved in Regulation of Nitrogen Metabolism. Microbiol Spectr 6.

90. Bower CE, Holm-Hansen T. 1980. A Salicylate–Hypochlorite Method for Determining Ammonia in Seawater. Can J Fish Aquat Sci 37:794-798.

91. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

92. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547-1549.

93. Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38:3022-3027.

94. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425.

95. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275-282.

96. Kopylova E, Noé L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28:3211-3217.

97. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-359.

98. Tamames J, Puente-Sánchez F. 2019. SqueezeMeta, A Highly Portable, Fully Automatic Metagenomic Analysis Pipeline. Front Microbiol 9:3349.

99. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

100. Martín-Platero AM, Valdivia E, Maqueda M, Martínez-Bueno M. 2007. Fast, convenient, and economical method for isolating genomic DNA from lactic acid bacteria using a modification of the protein "salting-out" procedure. Anal Biochem 366:102-104.

101. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103-1110.

102. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119.

103. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2016. GenBank. Nucleic Acids Res 44:D67-72.

104. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286-293.

105. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28:27-30.

106. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59-60.

107. Eddy SR. 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform 23:205-211.

108. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam

protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279-285.

109. Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. bioRxiv doi:10.1101/2022.04.08.487609:2022.04.08.487609.

2.12. Figures and legends

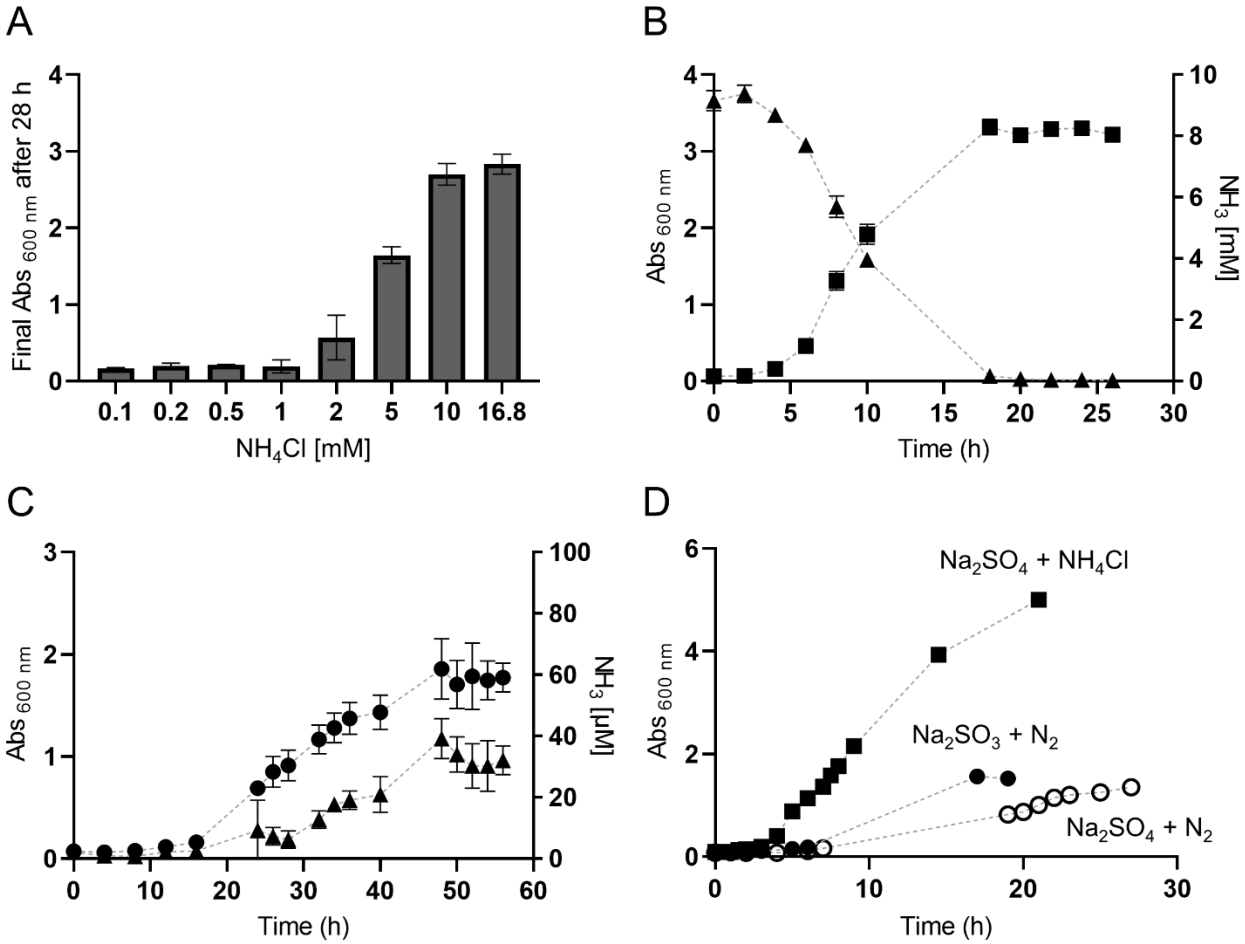


FIG. 1 Growth of *Methanothermococcus thermolithotrophicus* on different nitrogen sources. (A) Final Abs $_{600\ nm}$ of *M. thermolithotrophicus* cultures after 28 h of incubation with different NH_4Cl concentrations. (B) Growth curve of *M. thermolithotrophicus* cultures grown on 10 mM NH_4Cl (squares) and NH_4Cl consumption during growth (triangles). (C) Growth curve of diazotrophic *M. thermolithotrophicus* cultures grown on N_2 as the sole nitrogen source (circles) and NH_3 release during the growth (triangles). (D) Growth curves of non-diazotrophic *M. thermolithotrophicus* on 16.8 mM NH_4Cl (squares), a diazotrophic culture grown on Na_2SO_3 (full circles) and a diazotrophic culture grown on Na_2SO_4 (empty circles) in a fermenter. Measurements for A have been performed in duplicates. Measurements for B and C have been performed in triplicates.

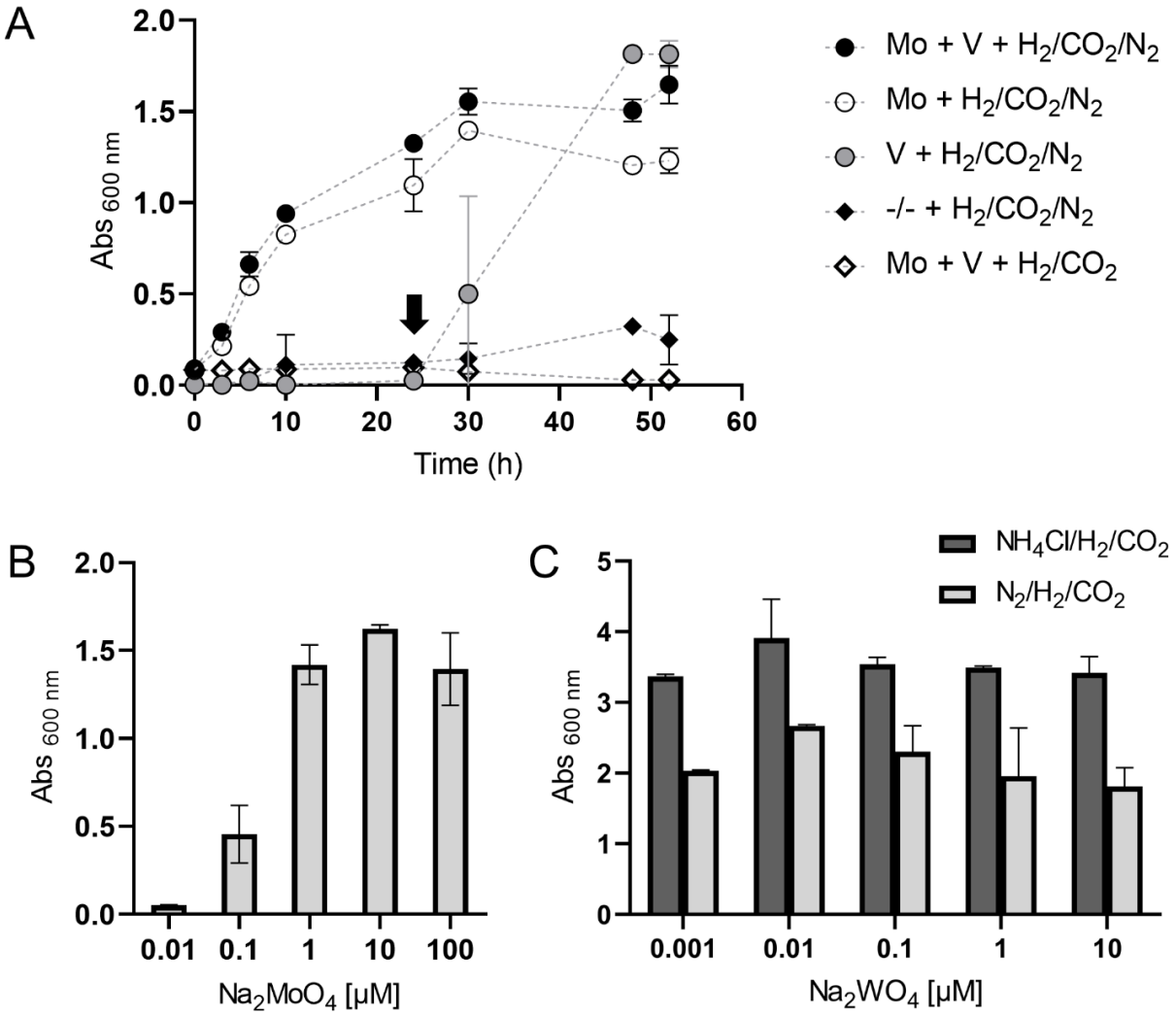


FIG. 2 Influence of trace metal availability on diazotrophic growth of *M. thermolithotrophicus.* (A) Growth curves of diazotrophic *M. thermolithotrophicus* grown in the medium with both Mo and V (full circles), without V (empty circles), without Mo (gray circles), without both Mo and V (full rhomboids) and negative control (empty rhomboids). The arrow indicates the supplementation of the culture without Mo with 16.8 mM NH_4Cl. (B) Final Abs $_{600\ nm}$ of *M. thermolithotrophicus* as a function of $[MoO_4^{2+}]$ concentration. (C) Final Abs $_{600\ nm}$ of non-diazotrophic (dark gray bars) and diazotrophic (light gray bars) *M. thermolithotrophicus* as a function of $[WO_4^{2+}]$ concentration. 10 µM of Na_2MoO_4 was used for this experiment. Measurements for A have been performed in triplicates. Measurements for B and C have been performed in duplicates.

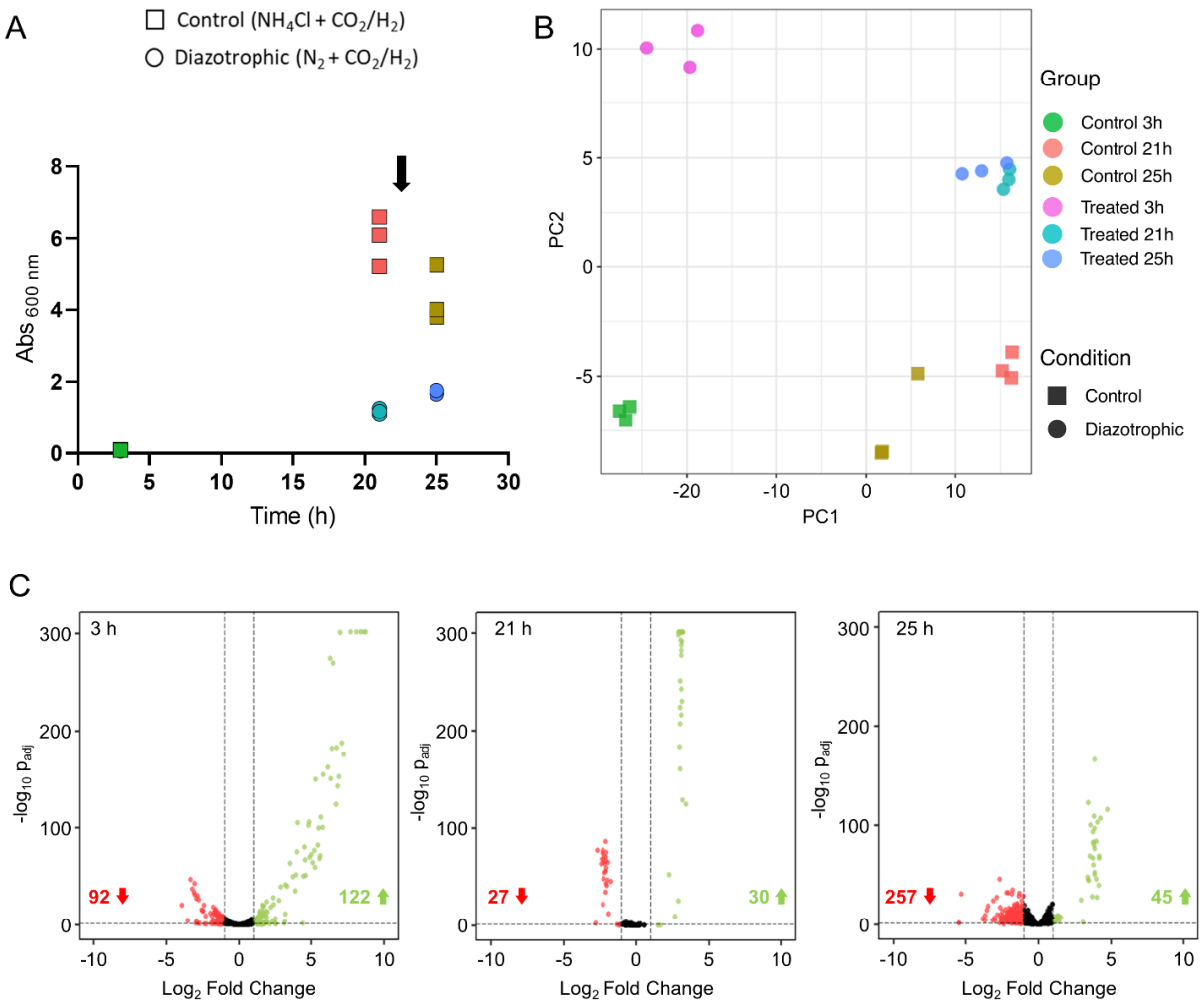


FIG. 3 Differential gene expression between NH_4Cl-grown and diazotrophic cultures. (A) Points during growth of the control (squares) and diazotrophic (circles) cultures at which samples for transcriptomic profiling were taken (each of the triplicates is shown separately). The arrow indicates the complete gas phase exchange of the cultures after the sample at 21 h was taken. (B) PCA plot showing the first two principal components that explain the variability in the data using the regularized log count data. Control corresponds to the NH_4Cl grown cultures. (C) Volcano plots of differentially expressed genes in the diazotrophic culture versus control at different time points. Negative and positive $\log_2$ fold changes >1 with adjusted $p < 0.05$ are shown in red and green, respectively. Values in red and green indicate the number of down- and up-regulated genes, respectively.

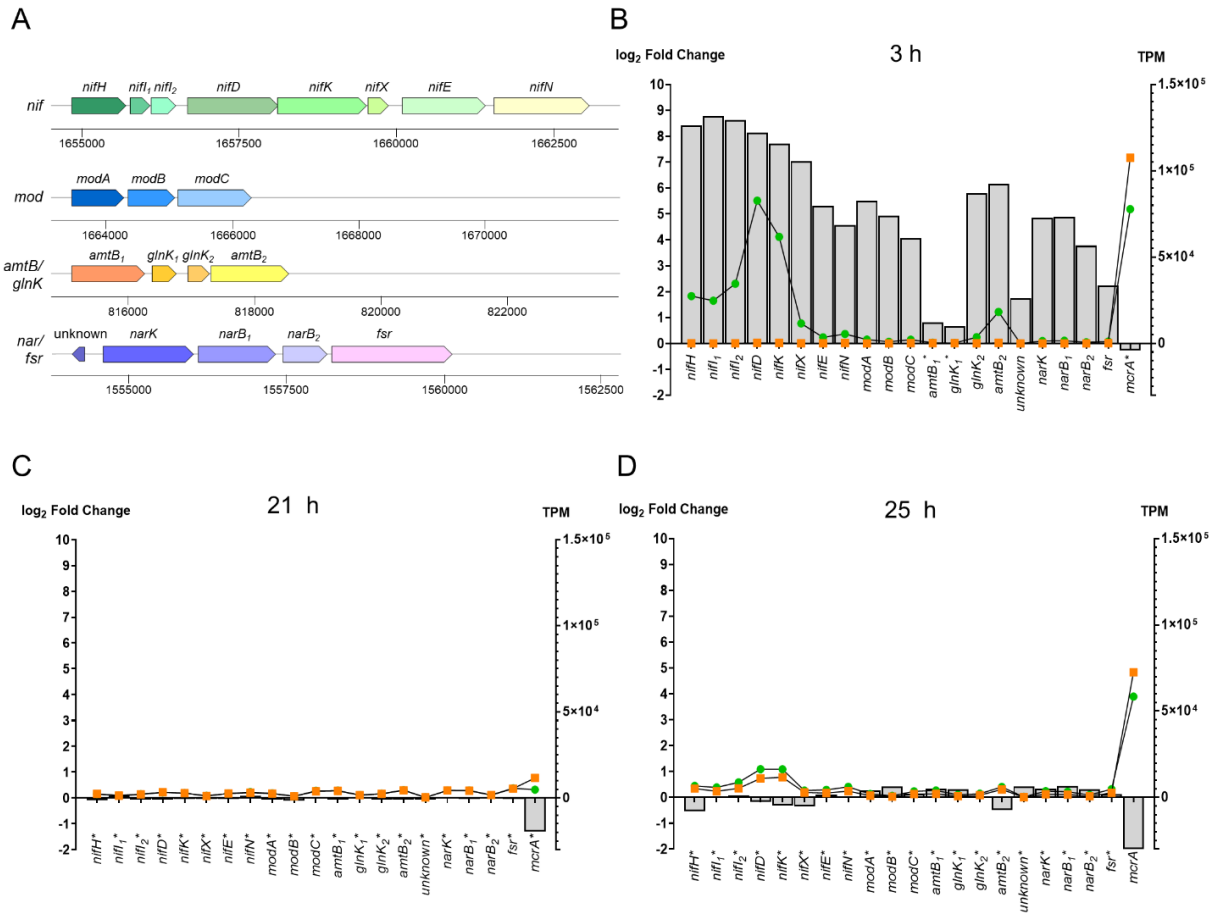


FIG. 4 Changes in the expression over time of selected genes involved in nitrogen acquisition. (A) Arrangement of the *nif, mod, amtB*/*glnK* and *nar/fsr* operons in *M. thermolithotrophicus.* (B-D) Log_2-fold changes (light gray bars) and transcript numbers per million (TPM) of the controls (orange squares) and diazotrophic cultures (green circles) of nitrogen acquisition genes compared tothe *mcrA* gene as reference after 3 h (B), 21 h (C) and 25 h (D). Genes that were not differentially expressed at a given time point are marked by an asterisk.

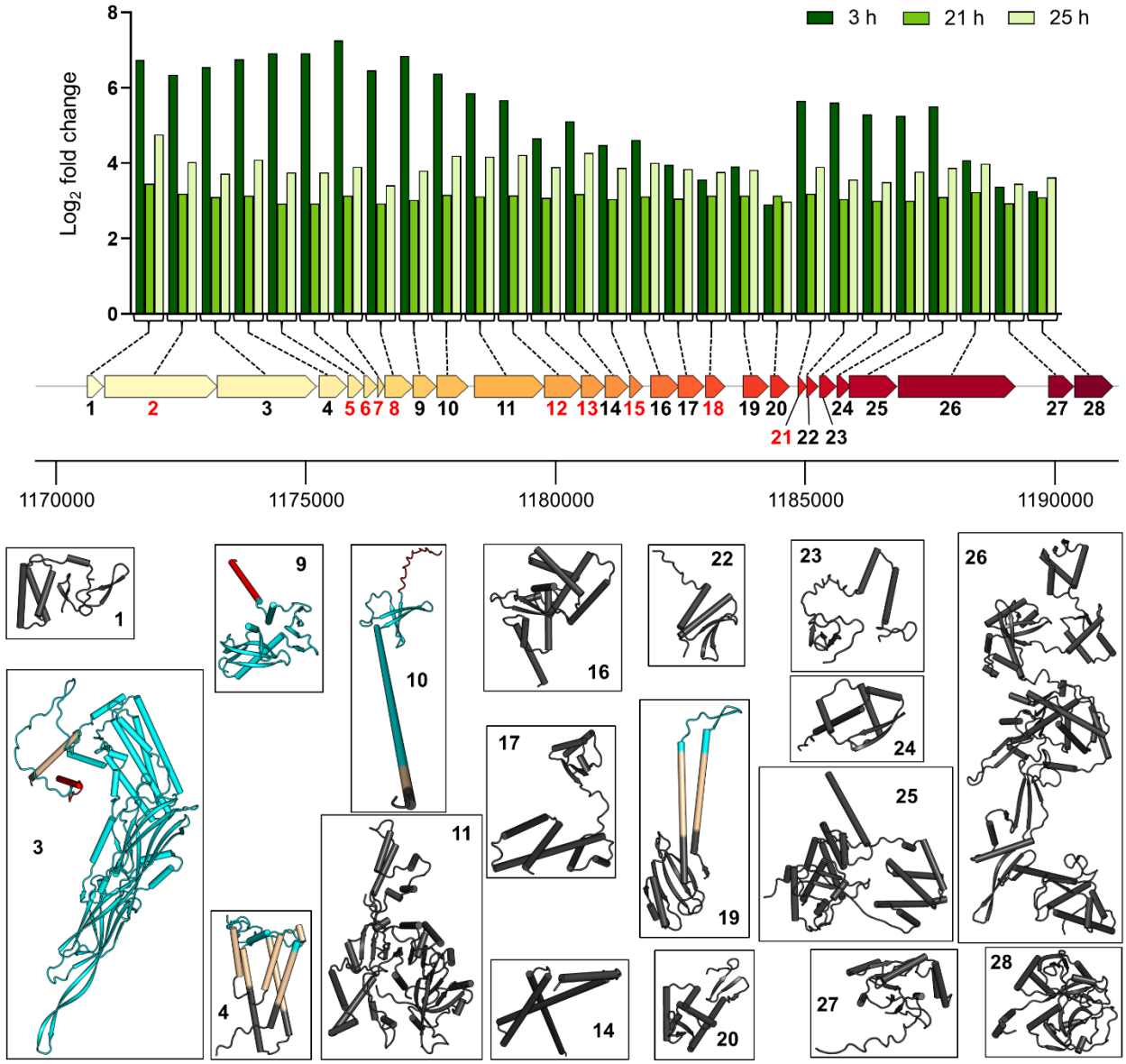


FIG. 5 Log_2 fold changes over time for 28 co-occurring genes of unknown function that are of putative viral origin. AlphaFold 2 (72) models are represented by cartoons and color-coded as follows: black for intracellular segments, cyan for predicted extracellular segments, ocher for transmembrane segments and red for predicted signal sequences. The topology prediction was made via the DeepTMHMM server (106). Genes with numbered red labels highlight models with an overall per-residue confidence score (pLDDT) below 75. These models are not presented due to their low confidence scores.

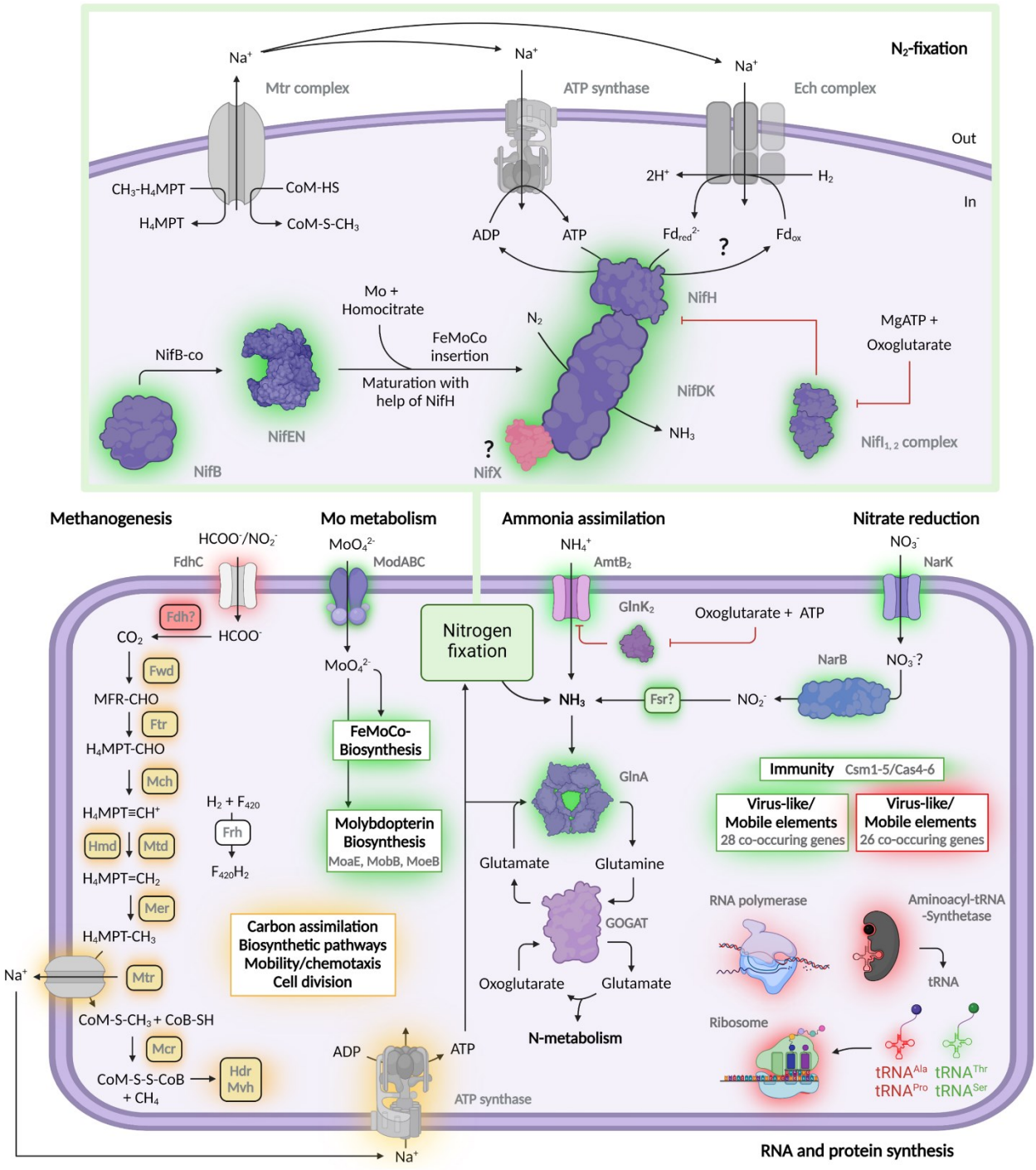


FIG. 6 Influence of nitrogen fixation on the metabolism of *M. thermolithotrophicus*. In the scheme, differentially expressed genes involved in methanogenesis, nitrogen fixation, ammonia assimilation, nitrate reduction, molybdate import, molybdopterin biosynthesis, immunity, mobile elements, transcription and translation, as well the tRNAs are shown. Differentially upregulated and downregulated transcripts at 3 h post-inoculation are highlighted with a green and red glow, respectively. An orange glow highlights differentially downregulated gene transcripts at 25 h post-inoculation. Full names of enzymes can be found in Supplementary Table 2. This picture has been created with Biorender.

2.13. Supplemental Material

Due to their large size, the supplementary tables were not included in the printed thesis and can be accessed through the links listed bellow.

Supplementary Table 1. Mapping statistics**.**

The table can be accessed under the following link:
https://journals.asm.org/doi/suppl/10.1128/mbio.02443-22/suppl_file/mbio.02443-22-s0001.xlsx

Supplementary Table 2. Summary of up/downregulated genes pairwise.

The table can be accessed under the following link:
https://journals.asm.org/doi/suppl/10.1128/mbio.02443-22/suppl_file/mbio.02443-22-s0002.xlsx

Supplementary Table 3. Upregulated genes coding for putative virus-like elements in region 1 and their conservation among selected methanogens.

The table can be accessed under the following link:
https://journals.asm.org/doi/suppl/10.1128/mbio.02443-22/suppl_file/mbio.02443-22-s0003.xlsx

Supplementary Table 4. Downregulated genes coding for putative virus-like elements in region 2 and their conservation among selected methanogens.

The table can be accessed under the following link:
https://journals.asm.org/doi/suppl/10.1128/mbio.02443-22/suppl_file/mbio.02443-22-s0004.xlsx

Supplementary Table 5. Accession numbers for sequences used in phylogenetic reconstruction.

The table can be accessed under the following link:
https://journals.asm.org/doi/suppl/10.1128/mbio.02443-22/suppl_file/mbio.02443-22-s0005.xlsx

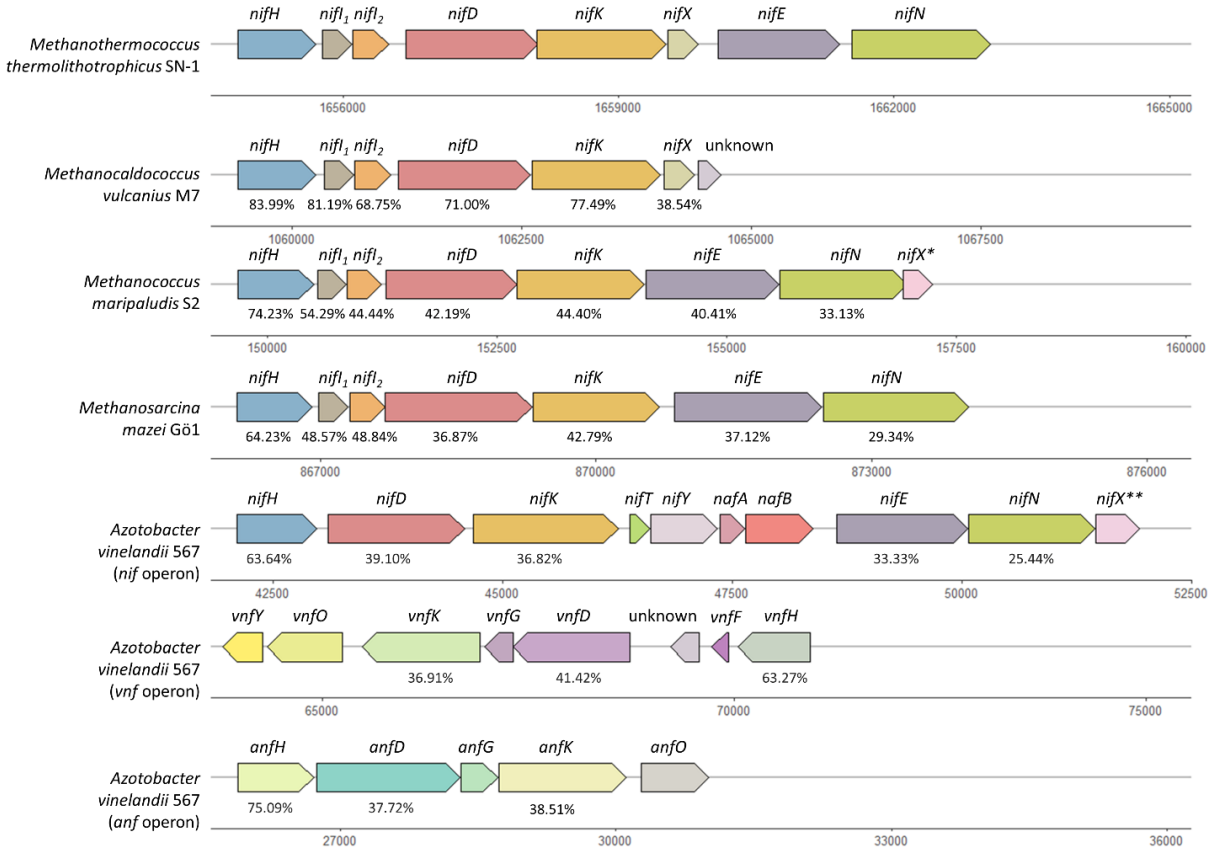


FIG. S1 Genomic environment of *nif*, *vnf* and *anf* genes from selected diazotrophic methanogens and *A. vinelandii*. The gene *nifX* from *M. thermolithotrophicus* SN-1 (also referred as strain DSM 2095 in the main text), *nifX* from *M. maripaludis* S2 and *nifX* from *A. vinelandii* 567 do not share sequence homologies.

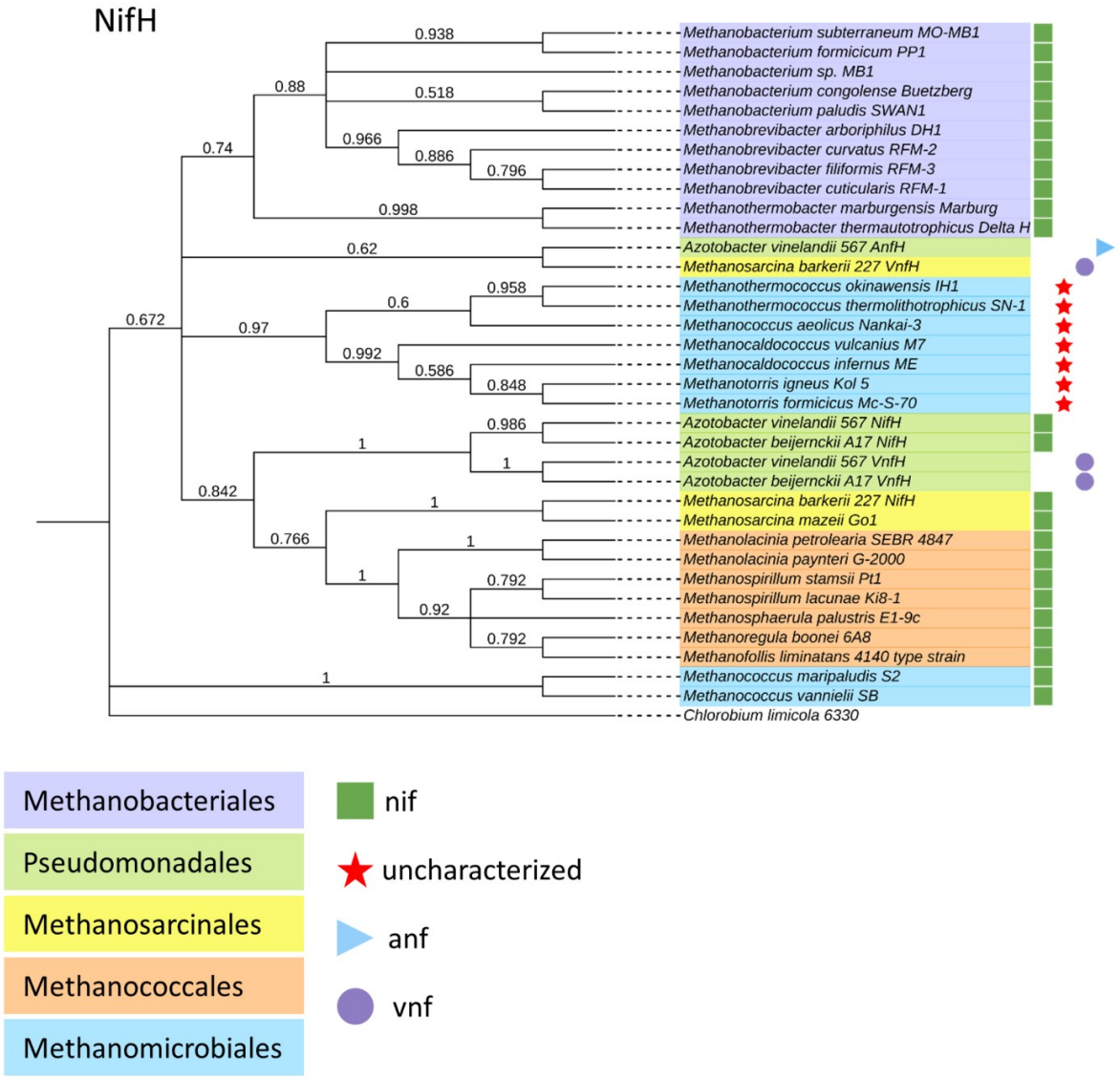


FIG. S2 Evolutionary analysis of 35 NifH sequences. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (1). Evolutionary distances (2) are in the units of the number of amino acid substitutions per site. ChlL (light-independent protochlorophyllide reductase) from *Chlorobium limicola* was used as an outgroup. Accession numbers for sequences used in phylogenetic reconstruction can be found in Supplementary Table 5.

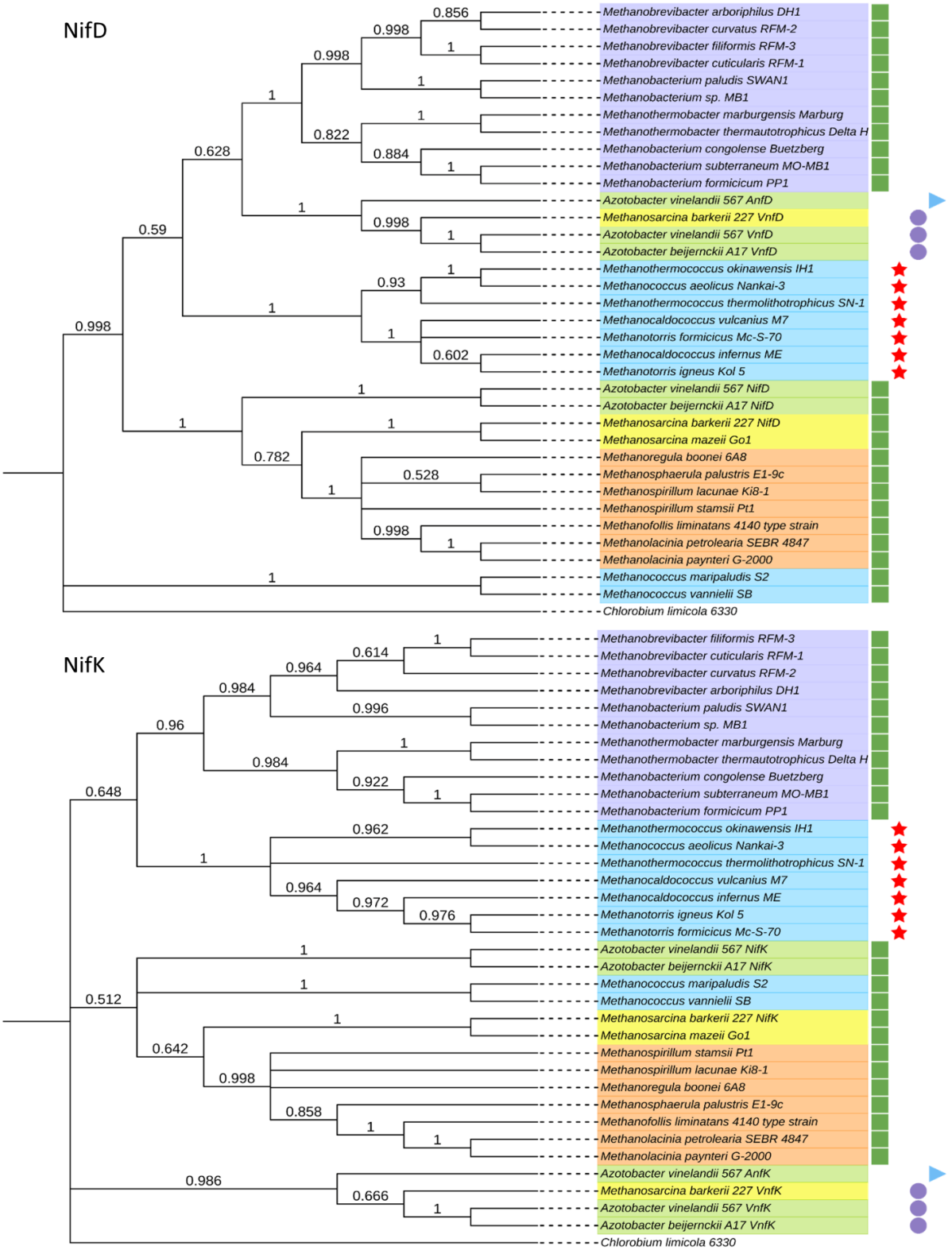


FIG. S3 Evolutionary analysis of 35 NifD and NifK sequences. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (1). Evolutionary distances (2) are in the units of the number of amino acid substitutions per site. ChlNB from *Chlorobium limicola* were used as an outgroup. The same color-coding and symbols are used as in Fig. S2. Accession numbers for sequences used in phylogenetic reconstruction are provided in Supplementary Table 5.

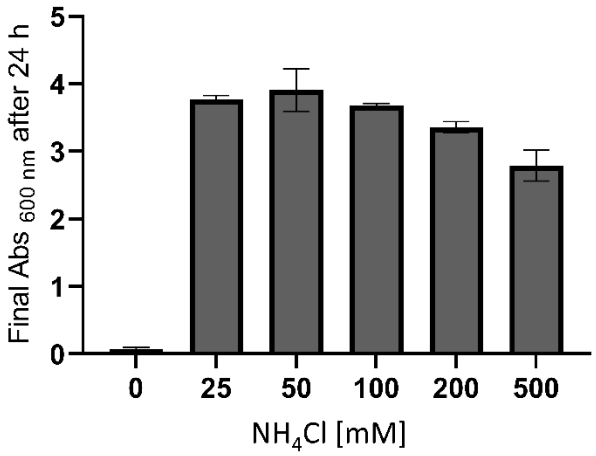


FIG. S4 Final Abs $_{600nm}$ of *M. thermolithotrophicus* after 24 h of incubation with different NH_4Cl concentrations. All measurements were done in triplicates.

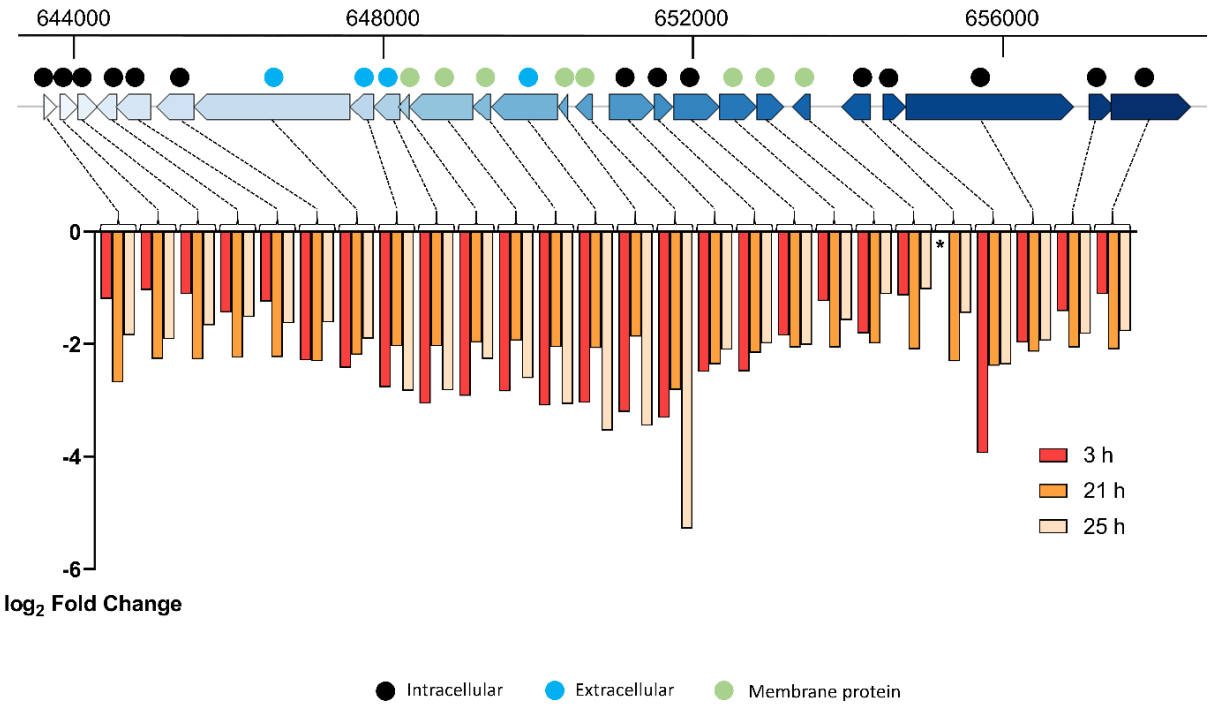


FIG. S5 Log_2 fold changes overtime for the 26 co-occurring genes of unknown function. The downregulation of this genomic region in coordination with the upregulation of the other detected putative region of viral origin (Figure 5), might imply the competition of the two for the availability of transcriptional and translational machineries or is the result of a general cellular stress. The topology prediction was made via the DeepTMHMM server and is color-coded as follows: black for intracellular proteins, cyan for secreted proteins, and green for proteins containing transmembrane segments. WP numbers and predicted functions can be found in Supplementary Table 4. Gene that was not differentially expressed at 3 h is marked by an asterisk.

References

1. Felsenstein J. 1985. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. Evolution 39:783-791.

2. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275-282.

CHAPTER 3: Structural comparison of nitrogenase reductase from three marine *Methanococcales*

Nevena Maslać[1], Cécile Cadoux[2,3], Pauline Bolte[1], Fenja Murken[1], Ross Milton[2,3], Tristan Wagner[1*]

[1] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

[2] Department of Inorganic and Analytical Chemistry, Faculty of Sciences, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

[3] National Centre of Competence in Research (NCCR) Catalysis, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

Manuscript in preparation (*all presented models are fully refined and ready for validation*)

Correspondence to Tristan Wagner (*twagner@mpi-bremen.de*)

3.1. Abstract

NifH catalyses the ATP-dependent electron delivery to the Mo-dependent dinitrogenase, a critical step in biological N_2-fixation. While NifHs have been extensively studied in Bacteria, structural information is limited on their archaeal counterparts. Archaeal NifHs are considered more ancient, particularly the ones from *Methanococcales*, a group of marine hydrogenotrophic methanogens, which includes N_2-fixers growing at temperatures near 92 °C. Here, we investigate NifH from three *Methanococcales* by structural and biochemical means, offering the first structures from meso-, thermo-, and hyperthermophile methanogens. While the protein from *Methanococcus maripaludis* (37 °C) was obtained through heterologous recombinant expression, *Methanothermococcus thermolithotrophicus* (65 °C) and *Methanocaldococcus infernus* (85 °C), adapted to diazotrophic conditions, were utilised to native purify the enzyme. The structures from *M. thermolithotrophicus* crystallised as-isolated present high flexibility, while the MgADP complexes obtained for NifH from the three methanogens are superposable and more rigid. All newly described archaeal NifHs are highly conserved with structural homologues. They retain key structural features of P-loop NTPases and share similar electrostatic profiles with the Mo-dependent counterpart from the bacterial representative *Azotobacter vinelandii*. Interestingly, compared to the NifH from the phylogenetically distant *Methanosarcina acetivorans*, these reductases do not cross-react with Mo-containing nitrogenase from *A. vinelandii*. However, they associate with bacterial nitrogenase when ADP-AlF_4 is added to mimic a transient reactive state. Accordingly, detailed surface analyses suggests that subtle substitutions would affect optimal binding during the catalytic cycle between the NifH from *Methanococcales* and the bacterial dinitrogenase, implying differences in the N_2-machinery from these ancient archaea.

3.2. Keywords

Nitrogen fixation, NifH, Fe protein, nitrogenase reductase, thermostable proteins, X-ray crystallography, [4Fe-4S]-cluster.

3.3. Introduction

Nitrogen, a key element for Life, is ubiquitous in the atmosphere in the form of dinitrogen (N_2) but is not readily bioavailable and must be converted to its most reduced state: ammonia (NH_3) [1]. The N_2 can be reduced either by the industrial Haber-Bosch process or by biological N_2-fixation via nitrogenases. The Haber-Bosch process plays a major role in supporting global agriculture, with 80-85% of the produced NH_3 used as fertiliser [2, 3] for

crops feeding half the world's population [4]. While this artificial process is essential to sustain our modern society, it has significant drawbacks. This is underlined by its energy consumption, fossil fuels dependency [5, 6] and its association with environmental problems ranging from water and soil pollution and acidification to the resulting loss of biodiversity [7-9]. The high-energy requirements and harsh conditions of this reaction are due to the extreme stability of the N_2 triple bond (with a bond energy of 941 $kJ.mol^{-1}$) [10]. In the search for sustainable, robust, and effective alternative catalysts [2, 11], a researchers turned to studying the biological N_2-fixation. Nitrogenases are remarkable enzymes performing the same process under ambient conditions, and considerable efforts over the past decades have been made to obtain insights into their catalytic cycle as they serve as the blueprints to design new catalysts through biomimicry.

Nitrogenases are metalloenzymes composed of two subunits: the MoFe protein (referred to as the *nifDK* gene product: NifDK) and the Fe protein or nitrogenase reductase (referred to as the *nifH* gene product NifH). The heterotetrameric NifDK ($\alpha_2\beta_2$) contains two metalloclusters: a [8Fe-7S] P cluster and a [Mo:7Fe:9S:C] FeMo-cofactor, the latter being the catalyst and representing the most complex metallocofactor in biology [12-14]. Two other systems exist in which the molybdenum is exchanged for a vanadium or an iron atom instead [10, 15]. These two other isoforms, considered to have evolved from the Mo-containing nitrogenase, are encoded in different operons, containing a different set of Fe proteins (VnfH and AnfH for the V and Fe system, respectively) and chaperones involved in cofactor biosynthesis and partners with still elusive functions [16-18].

All known Fe proteins belong to the P-loop NTPases and form homodimers, with each monomer folded as a single α/β-type domain family containing a nucleotide-binding site [19, 20]. A cubane [4Fe-4S] cluster bridges the homodimer [14, 21]. The main biological role of NifH is to act as an electron donor for N_2-fixation. The elegant ballet of electron transfer and nucleotide exchange is schematised in Figure 1. The binding of a reduced NifH loaded with MgATP to the NifDK lowers the reduction potential of the P cluster, triggering the electron transfer to the FeMoco in a so-called 'deficit spending mechanism' [22-24]. The lost electron is then replenished by the electron carried by the [4Fe-4S] cluster of NifH and depends on the hydrolysis of two ATP molecules [22]. The subsequent phosphate release has been shown to be the rate-limiting step in the nitrogen reduction cycle [25]. NifH then dissociates from NifDK, proceeds to the nucleotide exchange and is reduced back by either ferredoxin [26] or a flavodoxin [27]. Following the Thorneley-Lowe model [28], a minimum of 8 cycles is

required to accomplish the complete reduction of one N_2 molecule by the Mo-nitrogenase, which leads to the hydrolysis of 16 ATP (Fig. 1) [15, 29, 30].

Electron transfer to the substrate and ATP hydrolysis have constant rates, irrespective of the molecule that has to be reduced. Indeed, in addition to dinitrogen, nitrogenase can reduce a variety of other substrates, including acetylene, hydrogen cyanide, thiocyanate, azide ions, methyl isocyanide and carbon monoxide. The respective affinity for these substrates fluctuates depending on the isoform [31, 32]. NifH can also act as a reductase on its own, catalysing the reduction of CO_2 to CO [33] and even further reducing CO to hydrocarbons, albeit at low rates [34, 35]. This feature is particularly important for the development of CO_2 remediation strategies.

Despite of being of great interest for biotechnology, the structural information gathered for the Fe protein is limited to the bacteria *Azotobacter vinelandii* [21, 36, 37] and *Clostridium pasteurianum* [38] and the archaeon *Methanosarcina acetivorans* [39]. In contrast, the NifH homologues from *Methanococcales*, considered more ancient and phylogenetically distant from *Methanosarcina* [40, 41], present new interesting perspectives for the field.

The *Methanococcales* order gathers hydrogenotrophic methanogens (feeding on H_2 and CO_2) exclusively marine, with species showing a growth optimum ranging from 20 °C to 98 °C [42-44]. Metagenomic and metatranscriptomics studies concluded that *Methanococcales* are the main N_2-fixers in some ecological niches [45-48]. They exhibit impressive diazotrophic abilities under laboratory conditions [41, 49, 50]. For instance, *Methanocaldococcus vulcanius* holds the world record for the highest temperature for N_2-fixation and *Methanothermococcus thermolithotrophicus* presents high resistance to tungstate under diazotrophic conditions [41, 51]. These strict anaerobes must rely on temperature- and pressure-stable nitrogenases which represent candidates for novel bioinspired catalysts capable of withstanding harsh industrial conditions [52].

Here, we report the first structural insights of NifH from methanogens belonging to the *Methanococcales* order: the mesophilic *Methanococcus maripaludis* (*Mm*NifH) [53], the thermophilic *Methanothermococcus thermolithotrophicus* (*Mt*NifH) [54] and the hyperthermophilic *Methanocaldococcus infernus* (*Mi*NifH) [55]. Combining biochemical and structural analyses, we unveiled their specific properties and examined their cross-reactivity with the distant bacterial system from *A. vinelandii*.

3.4. Results

3.4.1. Phylogeny of nitrogenase reductases from *Methanococcales*. The neighbour joined protein tree of selected NifH, VnfH and AnfH sequences branches in two main NifH lineages: the base-branching NifH of some hydrogenotrophic methanogens (including the protein from *M. maripaludis*) and the more diversified and recent [40], group containing sequences found in bacteria and archaea, with most of Anf and Vnf gathering in a single monophyletic branch (Fig. 2A). This clustering is consistent both with the previously described concatenated NifDKH protein trees [18, 40] and individually observed subunits [41]. It is also one of the main arguments in favour of the currently accepted theory stating that the Nif system evolved in hydrogenotrophic methanogens [16] and Nif predating Vnf and Anf [18, 40]. Hence, these Nif systems found in methanogens such as *M. maripaludis* are particularly intriguing as they are closer to the roots of the phylogenetic tree, suggesting a more ancient origin and, maybe, ancestral features. In accordance with the phylogenetic tree, *Mt*NifH and *Mi*NifH are highly similar and share 85.16 % of sequence identity, while *Mm*NifH shares roughly 75 % identity with them (Fig. 2B). The three sequences are closer to *Av*AnfH (in the range of 74 % sequence identity) compared to *Av*NifH or *Av*VnfH.
*Mm*NifH, *Mt*NifH and *Mi*NifH were further biochemically investigated due to their distinct and unique phylogenetic positions, but also to study the basis of their thermo-adaptation.

3.4.2. NifHs from *Methanococcales* share the typical P-loop NTPase fold. Following our improved protocol, *M. thermolithotrophicus* and *M. infernus* adapted to a diazotrophic lifestyle were cultivated on large scales (see Material and Methods). *Mt*NifH and *Mi*NifH were successfully natively purified under anaerobic conditions through three chromatographic steps, yielding satisfying enrichment for structural attempts (Supplementary Fig. 1). Because of the cell yield limitation, the NifH homologue from *M. maripaludis* was recombinantly expressed using *Escherichia coli* as a host (see Material and Methods). Since the required quantities for activity assays were not compatible with the yield obtained from native purification, *Mt*NifH/*Mi*NifH were alternatively recombinantly produced in *E. coli* (Supplementary Fig. 1).

The structure of nucleotide-free *Mt*NifH (Fig. 1, state 1) was obtained in a monoclinic $P2_1$ (form 1, Supplementary Fig. 2A-B) and a tetragonal $P4_32_12$ form (form 2, Fig. 3B and Supplementary Fig. 2C), refined to 2.30-Å and 1.90-Å, respectively (Table 1). *Mm*NifH and *Mi*NifH were tentatively crystallised as-isolated but were not yielding exploitable crystals. However, we did obtain exploitable crystals in the presence of MgADP for all three proteins.

The MgADP-bound structures (Fig. 1, state 6) of *Mm*NifH, *Mt*NifH and *Mi*NifH were refined to 1.70-Å, 2.74-Å and 2.80-Å, respectively (Table 1).

All structural features previously identified as essential for NifH functionality are conserved in all three methanogenic systems. Figure 3A displays a sequence alignment from ESPript (https://espript.ibcp.fr, [56]) using the model of *Mt*NifH as input, highlighting the 11 α-loops and 8 β-sheets, which organise in the Rossmann fold characteristic for the P-loop NTPases (Fig. 3B). The P-loop, switch I and switch II, hallmarks of P-loop NTPases, are perfectly conserved (Fig. 3A) reflecting the critical importance of these motifs for functionality. The P-loop (residues 17-24 in *Mt/Mi*NifH and 9-16 in *Mm*NifH) plays a major role in the coordination of the bound nucleotides and Mg^{2+}, and goes through conformational changes upon ligand binding together with switch I (residues 46-52 in *Mt*/*Mi*NifH and 38-44 in *Mm*NifH, also participating in Mg^{2+} coordination) and switch II (residues 133-140 in *Mt*/*Mi*NifH and 125-132 in *Mm*NifH) regions. Switch II is directly involved in [4Fe-4S] cluster coordination and stabilises the P-loop through two anchoring points: S16-D125/G11-D129 in *Mm*NifH and S24-D133/G19-D137 in *Mt*/*Mi*NifH.

Although very well conserved, the structures of *Mt*NifH in the nucleotide-free state are highly flexible, especially at the dimeric interface (Fig. 4A, C and D). The ligand binding "locks" NifH in the conformation ready for binding to nitrogenase, making it less flexible. The nucleotide-binding restraining the natural flexibility of the Fe protein has already been mentioned in previous works [57, 58] and is illustrated in Figure 4.

In an attempt to identify structural features contributing to the remarkable thermostability of *Mt*NifH and *Mi*NifH homodimers, we used the PISA server [59] to quantify the free energy of assembly dissociation (ΔG^{diss}), which is expected to increase proportionally to the temperature. The analysis, performed on all MgADP-bound states, surprisingly show that the hyperthermophilic *Mi*NifH has the lowest average ΔG^{diss} (21.55 kcal/mol) (Table 2), while *Av*NifH has the highest (37 kcal/mol). In this regard, *Av*NifH also has the highest number of salt bridges (14), while *Mt*NifH and *Mi*NifH have fewer salt bridges than *Mm*NifH. Such absence of correlation of stable dimeric interface for proteins coming from (hyper)thermophiles suggests the involvement of other cellular factors (e.g. osmolytes, high potassium phosphate concentration) in thermostability.

3.4.3. The [4Fe-4S] cluster and ligand binding sites are conserved in NifHs from *Methanococcales*. The [4Fe-4S]-cluster at the interface of the NifH homodimers is coordinated by Cys97/Cys132 in *Mm*NifH and Cys105/Cys140 *Mt*NifH/*Mi*NifH (Fig. 6A).

The proximity of the [4Fe-4S] cluster is perfectly conserved, including the homologous residues to Arg100 of *A. vinelandii*, which is participating in CO_2 reduction.

Based on the electron density map, the MgADP fully occupied the ligand binding pocket in the three structures (Fig. 5). The P-loop, switch I and II adopt a similar, nearly identical, position. Due to their perfect sequence conservation, it was expected that these structural motifs and the residues involved in ligands binding would exhibit similar tri-dimensional conservation. Our three models corroborate this hypothesis, and Figure 6 details the identical ligand binding network across the bacterial and archaeal NifHs. The only exception was observed in the recombinantly produced *Mm*NifH, in which a C-terminal extension was introduced to add a His-tag. Here, the Arg213, which should stack the adenine ring together with Val217, has been displaced by the artificially introduced Arg277 and stabilised by the Asp275 and Ser186. While the overall geometry of the ADP is unaffected (Fig 6B), subtle changes in the nucleotide exchange dynamic cannot be excluded.

The comparative analysis of the electrostatic surfaces between the NifHs from *Methanococcales* and the three Fe proteins from *A. vinelandii* presents a similar overall charge distribution pattern. The expected interaction area between the Fe protein and the dinitrogenase displayed a positive patch around the [4Fe-4S] cluster, which becomes negatively charged in the outer ring (Fig. 7, top view). This general pattern is common among all studied Fe proteins, including Vnf and Anf systems. While *Mm*NifH and *Mi*NifH have two distinct positive patches on the outer ring, we presume that the electrostatic surface charges should still be complementary enough to that of *Av*NifH and would therefore support cross-reactivity with the Mo-nitrogenase from *A. vinelandii* as previously shown for *Av*VnfH and *Av*AnfH [36].

3.4.4. Methanogenic NifHs are not active *in vitro* when combined with *Av*NifDK. Cross-reactivity experiments have been conducted previously, e.g., N_2-fixation activity between a Fe protein and a dinitrogenase from a different system or organism. These works confirmed that the two other Fe protein isoforms of *A. vinelandii* (*Av*VnfH and *Av*AnfH) and the two isoforms from *Methanosarcina acetivorans* (*Ma*NifH and *Ma*VnfH) can cross-react with NifDK albeit exhibiting lower rates. To monitor the cross-reactivity between the Fe proteins from the *Methanococcales* and the bacterial NifDK, we tested the physiological reaction of N_2-reduction by quantifying the final reaction product, ammonia. The NifH proteins used in the assay were produced recombinantly to circumvent the yield limitations of native

purification. NifDK and NifH from *A. vinelandii* were purified as previously described [60] and served as a positive control for the reaction.

Despite our attempts, we could not detect any reliable activity for the enzymes from *Methanococcales* even when these proteins were used immediately after their purification without any freezing steps (Fig. 8). Since the affinity of these Fe proteins for *Av*NifDK might be relatively low, we increased the molar ratio of NifDK:NifH up to 1:50 instead of the classic 1:16. Here again, no reasonable activities could be detected when compared to the negative controls. *Mt*NifH:*Av*NifDK at a moalr ratio of 1:50 resulted in 8.2% of the activity measured with *Av*NifH. However, this relative activity has to be cautiously taken since the negative controls also exhibit a slight background (Fig. 8).

The apparent incompatibility between the methanogenic NifHs and *Av*NifDK prompted the question of whether, despite the seemingly overall conserved interaction interface, there are indeed subtle differences in singular residues causing it and whether the complex between the two can be formed.

3.4.5. NifHs from *Methanococcales* bind *Av*NifDK upon MgADP-AlF_4^- addition. To check if *Mm*/*Mt*/*Mi*NifHs can form a complex with *Av*NifDK, we incubated the different proteins coupled with MgADP-AlF_4^- and observed the complex formation through migration shifts on native PAGE as a proxy. MgADP-AlF_4^- is traditionally used as an analogue in nucleotide switch proteins and nitrogenase specifically to mimic the transitory state of the NifH-NifDK complex occurring before the ATP hydrolysis (Fig.1, state 4) [61]. Once formed, the *Av*NifDKH-MgADP-AlF_4^- is considered indefinitely stable [62]. Here again, we used *Av*NifDK and *Av*NifH as positive control. A shift in the migration of *Av*NifDK is observable upon the addition of all components, which is caused by the complex formation (Fig. 9). While the *Av*NifDK shift only occurs when the Fe protein, ADP, Mg^{2+} and AlF_4^- is added, *Mt*NifH presents a different profile in which the shift happens in the absence of $MgCl_2$ addition, and a further shift is present in the complete mix. While the effect of the second shift appearing has already been seen in the native *Av*NifDKH complex in our previous work [60] and might come from an excessive binding of *Mt*NifH on *Av*NifDK, the binding in the absence of Mg^{2+} is more elusive. It is probable that *Mt*NifH would have partially retained Mg^{2+} during purification or can bind to *Av*NifDK when the ADP-AlF_4^- only is present. The Fe protein from *Methanococcales* locked on the bacterial dinitrogenase does not reflect a physiological state that would explain the conflicting absence of activity. Even if NifHs from *Methanococcales* can bind *Av*NifDK, it does not mean that the formed complex can allow for

successful electron transfer. Indeed, as shown in Figure 1, during the turnover the Fe protein dynamically changes the conformation, while NifDK remains relatively static, implying the involvement of subtle rearrangement which could be missing in this case.

3.4.6. Mapping of the residues involved in the NifDK-interaction during catalysis. To deepen our investigation of the Fe protein–dinitrogenase interaction, we analysed the conservation of the surface residues engaging in protein-protein contacts in the four different *Av*NifH-NifDK complexes described by Tezcan *et al.* [57]. The four snapshots represent the nucleotide-free, MgADP-bound, ADP-AlF_4^-bound, and MgAMPPCP-bound (mimicking the MgATP-bound *Av*NifDKH complex) states of NifH [61]. The three central residues, Arg101, Arg141 and Lys171, interacting with NifDK in the four complexes, are all conserved in *Mm*/*Mt*/*Mi*NifHs. Gly66, making contacts with NifDK in all complexes except the nucleotide-free state, is also conserved in NifHs from *Methanococcales*. In contrast, Asn174, establishing contacts in all states except the MgAMPPCP-bound, is substituted by a Glutamate. Some other charged residues involved in contacts in some states are also not conserved in the three NifHs from *Methanococcales*: Glu113, Glu69, Asp70, Glu112, and Glu142. An exhaustive list of these substitutions found across *Mm*/*Mt*/*Mi*NifHs can be found in Supplementary Table 2.

Some of these residues could impact the interaction by obstructing the optimal contact for docking and electron transfer or generating a charge repulsion that would disrupt the process.

3.5. Discussion

The importance of understanding the molecular details of nitrogenase catalysis remains of uttermost interest for agriculture [2-4], ecology [6, 63], biofuel production [52] and evolution [18, 64] alike. So far, *A. vinelandii* was the paradigm model organism for studying biological nitrogen fixation, but shifting the focus to understudied models may provide unexpectedly valuable insights. Among other candidates, methanogens stand out due to the distant phylogeny of their nitrogenases, their unique physiology and their metabolic features, which are already applied [11, 52, 65]. The *Methanococcales* group contains (hyper)thermophilic species, which are particularly good models to investigate nitrogen fixation under high temperatures. These archaea play important roles in marine environments where they contribute to the nitrogen input for the microbial community [51]. On the way to in depth understanding of the nitrogen fixation in these peculiar microorganisms, we elucidated the first component of their machinery: the nitrogenase reductase. We obtained the first NifHs from a thermophile and hyperthermophile by directly purifying the enzymes from the

organisms. Their structural characterisation, together with NifH from *M. maripaludis* recombinantly expressed in *E. coli,* confirms that all Fe proteins characterised so far share the same architecture and harbour the conserved key elements of the functional P-loop NTPases, the P-loop, the *switch I* and *switch II*.

Strikingly, despite their extremely high tri-dimensional conservation with *Av*NifH and our efforts, we could not detect a relevant NH_3-production when *Mm*/*Mt*/*Mi*NifHs are combined with *Av*NifDK in a cross-reactivity assay. The N_2-reduction reaction between cross-kingdom proteins is possible [35], as it has been shown that the recombinant *Ma*NifH and *Ma*VnfH can electronically fuel *Av*NifDK, albeit with lower rates compared to the native system (11% and 36% relative activity, respectively). One reason for the lack of observed activities might be the slow electron transfer rate, favouring protonolysis and rendering N_2 fixation inefficient. This should be verified in the future by quantification of the produced H_2 along with the NH_3 quantification. The temperature at which the assay was carried out and the discrepancy between the expected optimum of 65 and 85 °C for *Mt*NifH and *Mi*NifH might also drastically influence the dynamic and electron transfer rate. Still, it does not explain the lack of activity for *Mm*NifH. Another plausible explanation is the differences in terms of contacts. Available data on cross-reactivity of the three Fe proteins from *A. vinelandii* [37], the one from *M. acetivorans* [35] and this study shows a lack of correlation between the sequence identity and the production of NH_3 between the hybrid systems. This is exemplified when considering the crossover between the three Fe protein isoforms and the three different dinitrogenase isoforms. Since genetic systems exist in the mesophile *M. maripaludis,* it would be interesting to check if diazotrophic growth can physiologically occur with hybrid systems.

Experimental results of cross-reactivity assays provide additional insights for understanding and modeling binding mechanistic, energetics, electron transfer efficiencies and engineering of nitrogenases to capture targeted substrates and intermediates. The substitutions listed in our analysis, which are part of the binding surface with the partner, also suggest the existence of an adaptation on the surface of the nitrogenase from *Methanococcales*. While it is expected that the similar binding area should, in principle, reflect the similar binding mechanism, some rearrangement might have occurred due to additional partners present in *Methanococcales*. For instance, Leigh´s work highlighted how the $NifI_{1,2}$ couple, belonging to the P_{II} superfamily of regulatory proteins, interferes with NifH binding efficiency. Such a post-translational regulation system might require changes of the dinitrogenase surface, for example the observed substitutions in *Mm*/*Mt*/*Mi*NifHs. Structural insights into the whole nitrogenase complex from these ancient microorganisms will unveil its unique features, but

more importantly, it could be used as a new model to deepen our knowledge of biological N_2-fixation.

3.6. Material and methods

3.6.1. *M. thermolithotrophicus* growth conditions. Both *Mt*NifH without ligands and with bound MgADP were obtained from derepressed *M. thermolithotrophicus* cells. *M. thermolithotrophicus* strain DSM 2095 (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) was cultivated on 10 mM NH_4Cl with a gas phase of 1 bar H_2/CO_2 in 5 l pressure-resistant Duran bottles, standing at 37 ºC. Each bottle contained 300 ml of medium, with a 1:10 inoculum. As the inoculum, diazotrophic cells adapted to the diazotrophic condition and grown as described in Maslać *et al.* 2022 [41]. After 48 h of incubation, cells were transferred in an anaerobic tent (atmosphere containing a mixture of N_2:CO_2 90:10) and harvested anaerobically at 6 000 x *g* for 30 min at room temperature. The pellet was washed with the same medium without NH_4Cl. The centrifugation step was repeated as described. Finally, the pellet was resuspended in the final volume of 600 ml NH_4Cl free medium and transferred to 2 l pressure-resistant Duran bottles, each containing 150 ml of culture with a gas phase of 0.6 bar H_2/CO_2 and 0.6 bar N_2. Resuspended cells were incubated for 24 h, standing at 37 ºC. Harvesting was done by centrifugation at 6 000 x *g* for 30 min at room temperature. In both cases, the pellets were flash frozen in liquid nitrogen and stored at -80 ºC before use. The medium was prepared as described in Maslać *et al.* 2022 [41], with the addition of 10 mM NH_4Cl in the case of NH_4Cl-containing medium.

3.6.2. Purification of native *Mt*NifH from *M. thermolithotrophicus* cells. Cell pellets of *M. thermolithotrophicus* were (~ 4 g of wet cells from the 1.5 l of total culture in the first batch from which NifH without ligands was obtained, ~ 16 g of wet cells from the 6.5 l of total culture in the second batch from which NifH with bound MgADP was obtained) thawed at room temperature and sonicated inside an anaerobic chamber (N_2/CO_2 (90:10%) atmosphere) after being diluted with ion exchange buffer (IEX; 50 mM Tris/HCl pH 8.0, 2 mM dithiothreitol (DTT)). Sonication was done in the same anaerobic chamber (KE76 probe Bandelin SONOPULS, Sigma, Berlin, Germany) at 70-75% intensity (10 times 10 seconds). The lysate was centrifuged at 35 000 x *g* for 30 minutes at 18 ºC in the case of the first batch and at 75 600 x *g* for 1 h at 18 ºC in the case of the second batch and transferred to an anaerobic Coy tent (N_2/H_2 (97:3%) atmosphere). The supernatant was filtered through a 0.2 µm filter (Sartorius, Göttingen, Germany) and loaded on a 20 ml HiTrap DEAE

Sepharose Fast Flow column (Cytiva, Freiburg, Germany) previously equilibrated with IEX Buffer. Elution was done with a 0-500 mM NaCl linear gradient for 18 column volumes (CVs) at a 3 ml/min flow rate. The fractions containing *Mt*NifH were obtained between 300.5 and 349.5 mM NaCl. The fractions of interest were pooled together and loaded on a 10 ml HiTrap Q Sepharose High-Performance column (GE Healthcare Life Sciences, Munich, Germany) after being filtered. Elution from Q Sepharose was performed by applying a 400-600 mM NaCl linear gradient (with the same IEX buffer above with 50 mM Tris/HCl at pH 9.0) for 9 CVs at a 1.5 ml/min flow rate. *Mt*NifH eluted in fractions between 418.2-450 mM NaCl. Fractions were pooled, diluted 1:3 with hydrophobic exchange buffer (HIC; 25 mM Tris/HCl pH 7.6, 2 M $(NH_4)_2SO_4$, 2 mM DTT) and filtered before being injected on a 5 ml Phenyl Sepharose High-Performance column (GE Healthcare Life Sciences, Munich, Germany). Under this condition, *Mt*NifH was not binding to the column and was collected in the flow through. The final yield of the purification was 0.66 mg of *Mt*NifH from the first batch and 8.19 mg *Mt*NifH from the second batch.

3.6.3. *M. infernus* growth conditions. *M. infernus* strain DSM 11812 (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany). The cells were adapted to the diazotrophic condition and grown in a medium similar to the one described in Maslać *et al.* 2022 [41] with the following modification of the growth medium: instead of Na_2SeO_4, Na_2SeO_3 x $5H_2O$ was used as a final concentration of 2 µM; the concentration of Na_2WO_4 x H_2O was increased to 100 µM, and the medium contained Na_2SO_4 at the final concentration of 10 mM. The final pH of the medium was adjusted to 6.5 aerobically. The cells were cultivated in 1 l pressure-resistant Duran bottles, containing 60 ml of medium with a gas phase of 0.5 bar H_2/CO_2 and 0.5 bar N_2, standing at 75 ºC. Cultures were inoculated 1:20 using diazotrophic *M. infernus* cells as an inoculum. The cells were harvested and stored as described for *M. thermolithotrophicus* cells above.

3.6.4. Purification of native *Mi*NifH. *M. infernus* cells (~ 25.62 g wet cells obtained from 5.5 l of culture) were thawed and sonicated as described above, using the same IEX buffer. A soluble fraction of the cell lysate was prepared anaerobically by centrifugation at 45 000 x *g* for 1 hour at room temperature. Purification was done inside an anaerobic chamber with N_2/H_2 (97:3%) atmosphere. The filtered supernatant was loaded on a 10 ml HiTrap Q Sepharose High-Performance column (GE Healthcare Life Sciences, Munich, Germany) and eluted with a 250-700 mM NaCl linear gradient for 4.5 CVs at a 1.5 ml/min flow rate.

The fractions containing *Mi*NifH obtained between 313.7-406.5 mM NaCl were pooled and diluted 1:2.5 with HIC buffer, filtered as described above and loaded on 5 ml Phenyl Sepharose column (GE Healthcare Life Sciences, Munich, Germany). Elution was done with a 1.9-0.6 M $(NH_4)_2SO_4$ linear gradient for 13 CVs at a 1 ml/min flow rate. *Mi*NifH was obtained between 0.943 and 0.726 M $(NH_4)_2SO_4$. The final yield of the purification was 3.12 mg for the *Mi*NifH.

3.6.5. *Escherichia coli* strains and plasmids are used to produce recombinant NifH proteins. In this study, three different *E.coli* strains were used (Supplementary Table 3): *(i) E.coli* DH5α carrying WG35 vector with *M. maripaludis* NifH (original construct from James Swartz, Stanford Chemical Engineering (Stanford University, Stanford, USA)) *(ii) E. coli* BL21 Δ*iscR* (original construct from Wenyu Gu, Department of Chemical Engineering and Civil and Environmental Engineering (Stanford University, Stanford, USA)) and *(iii) E.coli* DH5α carrying pDB1282 plasmid (original construct from Dennis Dean, Department of Biochemistry, Virginia Tech, Blacksburg, USA)).

3.6.6. Synthetic gene constructs. *Mt*NifH and *Mi*NifH sequences (Supplementary material) were codon optimised for *E. coli*, synthesised and cloned into pET-28(+) vectors using the restriction sites NcoI and NdeI, with a stop codon (TGA) incorporated before NdeI (Supplementary Table 3). All these steps were performed by GenScript (GenScript Corp, Piscataway, NJ, USA).

3.6.7. Growth conditions of *E. coli* BL21 Δ*iscR* WG *Mm*NifH and cell harvesting. *E. coli* BL21 Δ*iscR* WG *Mm*NifH were grown in an 8 l fermenter in Terrific Broth medium (12 g/l tryptone, 24 g/l yeast extract, 5 g/l glycerol, 1:10 diluted TB salts (0.17 M KH_2PO_4, 0.72 M K_2HPO_4)) containing 10 mM MOPS buffer pH 7.4, 2 mM Ferric ammonium citrate and glucose, kanamycin and ampicillin all at a final concentration of 0.5% w/v. The cells were incubated at 37 ºC until they reached an OD_{600nm} of 1.3. Here, the temperature was decreased to room temperature, and the fermenter switched to anaerobic condition by sparging with 100% N_2. The cells were induced immediately after equilibration by adding 0.5 mM IPTG, 25 mM fumarate and 2 mM cysteine final. After the overnight expression, the cells were harvested anaerobically by centrifugation for 15 minutes at 17 700 x *g* at 10 ºC. The cell pellet was washed with anaerobic phosphate-buffered saline (PBS, 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) inside an anaerobic chamber (N_2/CO_2 (90:10%) atmosphere) before being pelleted again as described above. The pellets were flash frozen in liquid nitrogen and stored anaerobically at -80°C until use.

3.6.8. Growth conditions of *E. coli* BL21 pDB 1282 pET-28(+) *Mt*/*Mi*NifH and cell harvesting. *E. coli* BL21 pDB 1282 pET-28(+) *Mt*/*Mi*NifH cells were grown in 1 l Duran flasks containing 0.33 l medium. The medium composition and the growth conditions were the same as described above for *E. coli* BL21 Δ*iscR* WG *Mm*NifH cells. The culture was grown aerobically at 37 ºC, shaking at 180 rpm until an $OD_{550\ nm}$ of 0.5 was reached. Upon reaching this OD, the headspace was exchanged to 100% N_2 by aseptic flushing using the Hungate technique and the bottles were sealed with sterile gas-tight rubber stoppers and caps. To induce the cells, anaerobic IPTG (250 µM final), fumarate (25 mM final), cysteine (2 mM final) and 0.2% arabinose were added by sterile filtration. The cultures were further incubated at 25 ºC, shaking at 180 rpm overnight. The cells were harvested anaerobically by centrifugation for 15 minutes at 17 700 x *g* at 10 ºC, washed with anaerobic PBS, pelleted again, flash frozen in liquid nitrogen and stored anaerobically at -80°C until use.

3.6.9. Purification of recombinant NifH proteins. In all cases, the respective cell pellets of transformed cells were thawed on ice and resuspended in Lysis Buffer (50 mM $Na_2HPO_4{\cdot}2H_2O$, 500 mM NaCl, 20 mM imidazole, 5% glycerol, pH 8.0) inside of an anaerobic chamber (N_2/CO_2 (90:10%) atmosphere). Sonication was done in the same anaerobic chamber (KE76 probe Bandelin SONOPULS, Sigma, Berlin, Germany) at 75% intensity (30 times 10 seconds). The cell lysate was centrifuged at 20 442 x *g* for 45 minutes at 4°C to remove cell debris and unbroken cells. The supernatant (soluble fraction) was transferred to an anaerobic tent (N_2/H_2 (97:3%) atmosphere), filtered through a 0.2 µm filter (Sartorius, Göttingen, Germany), and loaded on 10 ml His-Trap nitrilotriacetic acid column loaded with Nickel (Ni-NTA, GE Healthcare Life Sciences, Munich, Germany). The column was previously washed and equilibrated with the Lysis Buffer. Elution was done with a 20-320 mM imidazole linear gradient for 6 CVs at a 1 ml/min flow rate in all cases. *Mm*NifH eluted between 169.56 and 257.48 mM imidazole, *Mt*NifH eluted between 245.56 mM and 283.24 mM imidazole, while *Mi*NifH was eluted between 205.44 mM and 257.20 mM imidazole. Based on SDS PAGE and brown colour, purest fractions were pooled together and diluted with 3 volumes of IEX buffer, filtered, and loaded on a 10 ml HiTrap Q Sepharose High-Performance column (GE Healthcare Life Sciences, Munich, Germany). Elution was done with a 0-800 mM NaCl linear gradient for 6 CVs at a 1 ml/min flow rate. The fractions containing *Mm*NifH were obtained between 506.2 and 614.2 mM NaCl, *Mt*NifH was eluted between 454.5 mM and 560.8 mM NaCl, while *Mi*NifH was eluted between 401.1 mM and 546.8 mM NaCl. The final yield of the purification was 6 mg for the *Mm*NifH, 13.4 mg for the *Mt*NifH and 32.4 mg for the *Mi*NifH,

3.6.10. Crystallisation. All samples were concentrated with a 10-kDa cutoff with a VivaSpin amicons (Sartorius, Göttingen, Germany) to the desired concentration. All samples were crystallised inside an anaerobic chamber (N_2/H_2 (97:3%) atmosphere, 20 °C). Prior to crystallisation, samples were centrifuged at 13 000 x *g* for 3 min to remove macro-aggregates and dust. When the proteins were co-crystallised with ligands, the ligands were mixed with the sample before the centrifugation step. The crystallisation was done in 96-Well MRC 2-Drop polystyrene plates (SWISSCI) containing 90 µl of crystallisation solution in the reservoir in all cases. *Mm*NifH was crystallised at a final concentration of 13 mg/ml with 2 mM ATP and 2 mM $MgCl_2$ by mixing 0.7 µl of crystallisation solution with 0.7 µl of protein sample. Crystals were obtained in 20 % w/v polyethylene glycol 3 000, 100 mM Tris pH 7.0 and 200 mM calcium acetate. *Mt*NifH nucleotide-free was crystallised at a final concentration of 11 mg/ml by spotting 0.55 µl of crystallisation solution with 0.55 µl of protein sample. The crystallisation solution contained 30 % v/v 2-methyl-2,4-pentanediol, 100 mM sodium acetate, pH 4.6 and 20 mM calcium chloride. *Mt*NifH with ligands was crystallised at a final concentration of 32.7 mg/ml, with 10 mM ATP and 10 mM $MgCl_2$ by spotting 0.5 µl of crystallisation solution (20 % w/v polyethylene glycol 8 000, 100 mM TRIS, pH 8.5 and 200 mM magnesium chloride) with 0.5 µl of protein sample. *Mi*NifH was crystallised at a final concentration of 18.3 mg/ml, with 2 mM ATP and 2 mM $MgCl_2$, by spotting 0.7 µl of crystallisation solution with 0.7 µl of protein sample. In this case, the crystallisation solution contained 20 % w/v polyethylene glycol 3 350, 100 mM Bis-Tris propane pH 8.5 and 200 mM sodium nitrate. Sealed plates were stored inside the same anaerobic chamber where the crystallisation was performed.

3.6.11. Data collection, integration, scaling, model building and refinement. Data were collected at Synchrotron SOLEIL French National Synchrotron Facility (Paris, France, beamline Proxima I) and The Swiss Light Source Synchrotron (Villigen, Switzerland, beamlines PXI and PXIII) (see Table 1) and were integrated with *autoPROC* [66]. All datasets were treated with Staraniso [67]. *Mt*NifH nucleotide-free form 2 was solved by molecular replacement with the nucleotide-free model from *Av*NifH (PDB code 2NIP). All *Mt*NifH models were solved by molecular replacement with *Mt*NifH nucleotide-free form 2 as a template. *Mm*NifH and *Mi*NifH structures were solved by molecular replacement with the MgADP complex from *Av*NifH (PDB code 6N4L). All molecular replacements were done with Phaser [68] from the Phenix suite [69]. All models were manually optimised with *COOT* [70]. Refinement was performed with Phenix without applying non-crystallography symmetry and using a translation-libration screw. Models were refined with riding hydrogens omitted in

the final deposited models. All structures were validated by the Molprobity server [71] (on the 15th of April, 2023).

3.6.12. *A. vinelandii* growth conditions. Medium composition. *A. vinelandii* RS1 (derived from strain DJ) [60, 72] preculture was grown in a 250 ml baffled culture flask (glass) with a vented cap containing 120 ml of sterile liquid media (modified Burke's medium). Each preculture was started by inoculation from a fresh agar plate and grown aerobically at 200 rpm at 30 °C until $OD_{600\ nm}$ > 2.0 (usually overnight). Under sterile conditions, the preculture was used at a 1:10 ratio to inoculate the 12 l of liquid media dispatched in 4 x 5 l baffled flasks (Corning CLS431684, 3 l of culture per bottle) with vented caps. Cells were grown at 200 rpm at 30 °C overnight until an $OD_{600\ nm}$ > 1.5 was reached. Cells were collected by serial centrifugation runs at 4 500 x *g*, 15 min at room temperature. The supernatant was discarded, and the cells were gently resuspended in a fresh growth medium devoid of NH_4OAc or another source of fixed nitrogen (prepared fresh but not under sterile conditions) in the incubator shaker (~400 ml media per bottle). After resuspension, the culture was divided evenly between the remaining media/bottles and incubated at 200 rpm and 30 °C for another 3 hours to facilitate derepression of *the nif* operon. Cells were harvested by serial centrifugation collections as above. Finally, the cells were scooped into a plastic bag and stored at -80 °C until further use. Modified Burke's medium was prepared by adding 120 ml of an autoclaved 100 x phosphate buffer solution (0.46 M K_2HPO_4, 0.15 M KH_2PO_4), 40 ml of a filter-sterilised 3 M NH_4OAc solution, 12 ml of a filter-sterilised 0.1 M Fe solution ($FeCl_3.6H_2O$) and 12 ml of a filter-sterilised 10 mM Mo solution ($Na_2MoO_4.2H_2O$) to 12 l of autoclaved salt medium (60 mM sucrose, 8.1 mM $MgSO_4 \cdot 7H_2O$, 6.1 mM $CaCl_2$), all at room temperature, under sterile conditions. All buffers and solutions were prepared with MilliQ water (18.2 MΩ cm). Generally, large volumes (> 50 ml) were sterilised by autoclaving; smaller solutions were sterilised by filtration (0.45 µm syringe filters).

3.6.13. *Av*NifDK and *Av*NifH purification. Cell pellet (~ 77 g) was anaerobically thawed and resuspended for 45 minutes in lysis buffer (50 mM Tris/HCl pH 8, 5 mM dithionite, 37% v/v glycerol) with a 2:1 ratio (buffer: cells v/w) inside a Coy anaerobic chamber (N_2:H_2 95:5, Coy Laboratory Products, USA). Treated cells were collected by centrifugation at 12 000 x g for 25 min at 4 °C in deoxygenated and anoxically sealed poly(propylene) bottles. The cell pellet was resuspended in 200 ml of glycerol-free lysis buffer (containing a few µg of DNase) and lysed by osmotic shock/shaking. Lysed cells were incubated on ice for 20 min and pelleted by anaerobic centrifugation at 26 000 x g at 4 °C for 1 hour. The obtained supernatant was mixed with 20 ml of post-lysis buffer (2 M NaCl, 234 mM Tris, pH 8.0) and loaded on a

20 ml His Prep FF 16/10 column equilibrated with chelating equilibrating buffer (50 mM Tris, 0.3 M NaCl, pH 8.0). The flow through containing the *Av*NifH was collected. The elution was done with chelating elution buffer (50 mM Tris, 0.3 M NaCl, 300 mM imidazole, pH 8.0) in two steps. The first step containing 20 mM imidazole was used to remove weakly bound proteins, and the second at 300 mM imidazole step was to elute *Av*NifDK. Collected *Av*NifDK was diluted 1:3 (v/v) with 50 mM Tris/HCl pH 8.0 and 5 mM EDTA-disodium salt and loaded on a 20 ml Q Sepharose HiPrep HP 16/10 column equilibrated with ion exchange buffer A (50 mM Tris/HCl pH 8.0). The elution was done with ion exchange buffer B (50 mM Tris/HCl, 1 M NaCl, pH 8.0) in two steps: 100 mM NaCl washing step and 1 M NaCl step to elute *Av*NifDK. *Av*NifH was diluted and loaded on a 20 ml Q Sepharose HiPrep HP 16/10 column the same way as described above for *Av*NifDK and eluted with 200-650 mM linear NaCl gradient over 7 column volumes. Fractions containing *Av*NifH were determined based on SDS PAGE, pooled together and concentrated before loading on a Sephacryl S-200-HR HiPrep 26/60 320 ml column. Both samples were concentrated using a Merck Millipore stirred concentrator cell equipped with a 100 (for *Av*NifDK) or 30 kDa (for *Av*NifH) molecular weight cutoff membrane (fed with ultra-high-purity N_2 5.0) and flash frozen in liquid nitrogen as 20 µL pellets and stored in liquid nitrogen until further use. The final yield of the purification was 109.25 mg for the *Av*NifH and 370.37 mg for the *Av*NifDK.

3.6.14. Activity assays. Activity assays were performed in triplicate and contained a final volume of 1 ml per reaction. The reactions were performed in 13 ml septum-sealed glass vials (Wheaton) in deoxygenated buffer containing an ATP-regenerating system (5 mM ATP, 30 mM phosphocreatine, 0.6 mg/ml bovine serum albumin, 200 µg/ml creatine phosphokinase (from Rabbit muscle, Merck/Roche, Darmstadt, Germany), 10 mM sodium dithionite and 100 mM MOPS/NaOH at pH 7.0). All reactions contained 0.1 mg MoFe protein and 16.6 or 50 molar equivalents of Fe protein (corresponding to 0.5 mg and 1.56 mg, respectively). All reaction vials were prepared and sealed inside an anaerobic chamber under a 100% N_2 atmosphere and vented to atmospheric pressure before starting. Reactions were performed within a standing water bath (30 °C) and initiated by adding $MgCl_2$ (from a 1 M stock, 10 mM final concentration). After 8 minutes, the reactions were quenched by adding 300 µL of 400 mM EDTA (pH 8.0). Ammonia was quantified by the *ortho*-phthalaldehyde method using NH_4Cl as the standard, as reported previously [73-75].

3.6.15. High-resolution Clear Native PAGE (hrCN PAGE) and MgADP-AlF_4^- stabilisation. To test the formation of MgADP-AlF_4^--stabilised NifDK-NifH complexes, all samples were incubated at 20 °C inside an anaerobic chamber (N_2/H_2 (97:3%)) for 1 h in 100

mM Tris/HCl pH 7.6, with 20 mM NaF, 1 mM $AlCl_3$, 1 mM ADP and 2 mM $MgCl_2$ before loading on hrCN PAGE. The samples were diluted to 0.6 mg/ml in the same buffer. In all cases, NifDK and NifH were mixed in a molar ratio of 1:5 (corresponding to 45 µg of NifDK and 67.5 µg of NifH) and the mixture was diluted to a final concentration of 0.6 mg/ml prior to the loading on the gel, in the corresponding buffer. The 3-15% linear polyacrylamide gradient hrCN PAGE was run in the same anaerobic chamber and prepared according to Lemaire *et al.* [76]. Cathode buffer contained 50 mM Tricine, 15 mM Bis-Tris/HCl, pH 7, 0.05% (w/v) sodium deoxycholate, 2 mM DTT and 0.01% (w/v) dodecyl maltoside, while the anode buffer contained 50 mM Bis-Tris/HCl buffer pH 7 and 2 mM DTT. Electrophoresis was run with a constant 40 mA current (PowerPacTM Basic Power Supply, Bio-Rad). Protein bands were visualised with Ready BlueTM Protein Gel stain (Sigma Aldrich, Hamburg, Germany). NativeMarkTM Unstained Protein Standard (ThermoFischer Scientific, Dreieich, Germany) was used as the protein ladder.

3.6.16. Phylogenetic analysis. Selected NifH, VnfH and AnfH sequences were aligned using MUSCLE [77] (default parameters) in MEGA11 [78, 79]. Ambiguous positions were removed for each possible pairing (pairwise deletion option). The final alignment contained a total of 332 NifH positions. The phylogenetic tree was constructed using the Neighbour-Joining method [80] with the JTT matrix model for multiple substitutions [81]. The analysis was conducted in MEGA11 with ChlLNB from *Chlorobium limicola* as an outgroup.

3.7. Acknowledgements

We thank the Max Planck Society and Max Planck Institute for Marine Microbiology for their continuous support. We also thank Christina Probian and Ramona Appel for their technical assistance in the Microbial Metabolism laboratory. We thank Wenyu Gu for sharing WG35 vector with *M. maripaludis* NifH, Dennis Dean for sharing *E.coli* DH5α pDB1282 strain and James Swartz for sharing *E. coli* BL21 Δ*iscR strain.* We thank the Swiss Light Source and SOLEIL Synchrotrons, especially the staff of beamline PXI from SLS and PROXIMA-1 from SOLEIL, for their advice during data collection.

3.8. Funding

This study was funded by the Max Planck Society.

3.9. Author contributions

N.M. and T.W. designed the research. C.C. and N.M. purified *Av*NifDK and *Av*NifH. F.M. optimised the protocol for the purification of recombinant *Mi*NifH. P.B. cultivated *M. infernus* cells and purified native *Mi*NifH. N.M. cultivated *M. thermolithotrophicus* cells and purified native *Mt*NifH. N.M. produced recombinant *Mm*/*Mt*/*Mi*NifH. N.M. performed the phylogenetic analysis, activity assays, and hrCN PAGE. X-ray data collection was performed by N.M. and T.W. Data processing, model building, structure refinement, validation and deposition were performed by N.M. and T.W. Structures were analysed by N.M. and T.W. R.D.M. and T.W. acquired funding to realise the project. The paper was written by N.M. and T.W. with contributions and final approval of all co-authors.

3.10. References

1. Zhang, X., Ward, B. B. & Sigman, D. M. (2020) Global Nitrogen Cycle: Critical Enzymes, Organisms, and Processes for Nitrogen Budgets and Dynamics, *Chemical Reviews.* **120**, 5308-5351.

2. Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z. & Winiwarter, W. (2008) How a century of ammonia synthesis changed the world, *Nature Geoscience.* **1**, 636-639.

3. Nayak-Luke, R., Bañares-Alcántara, R. & Wilkinson, I. (2018) "Green" Ammonia: Impact of Renewable Energy Intermittency on Plant Sizing and Levelized Cost of Ammonia, *Industrial & Engineering Chemistry Research.* **57**, 14607-14616.

4. Smil, V. (2002) Nitrogen and food production: proteins for human diets, *AMBIO: A Journal of the Human Environment.* **31**, 126-131.

5. Philibert, C. (2017) Renewable energy for industry, *Paris: International Energy Agency.* **65**.

6. Ritter, S. K. (2008) The Haber-Bosch reaction: an early chemical impact on sustainability, *Chemical & Engineering News.* **86**.

7. Van Grinsven, H. J., Ward, M. H., Benjamin, N. & De Kok, T. M. (2006) Does the evidence about health risks associated with nitrate ingestion warrant an increase of the nitrate standard for drinking water?, *Environmental Health.* **5**, 1-6.

8. Purcell, J., Breitburg, D., Decker, M., Graham, W., Youngbluth, M., Raskoff, K., Rabalais, N. & Turner, R. (2001) Coastal Hypoxia: Consequences for Living Resources and Ecosystems, *American Geophysical Union*, 77-100.

9. Van Breemen, N., Burrough, P., Velthorst, E., Van Dobben, H., de Wit, T., Ridder, T. & Reijnders, H. (1982) Soil acidification from atmospheric ammonium sulphate in forest canopy throughfall, *Nature.* **299**, 548-550.

10. Mus, F., Alleman, A. B., Pence, N., Seefeldt, L. C. & Peters, J. W. (2018) Exploring the alternatives of biological nitrogen fixation, *Metallomics.* **10**, 523-538.

11. Atalah, J., Cáceres-Moreno, P., Espina, G. & Blamey, J. M. (2019) Thermophiles and the applications of their enzymes as new biocatalysts, *Bioresource Technology.* **280**, 478-488.

12. Kirn, J. & Rees, D. (1992) Crystallographic structure and functional implications of the nitrogenase molybdenum–iron protein from *Azotobacter vinelandii*, *Nature.* **360**, 553-560.

13. Kim, J. & Rees, D. (1992) Structural models for the metal centers in the nitrogenase molybdenum-iron protein, *Science.* **257**, 1677-1682.

14. Einsle, O. & Rees, D. C. (2020) Structural Enzymology of Nitrogenase Enzymes, *Chemical Reviews.* **120**, 4969-5004.

15. Eady, R. R. (1996) Structure-function relationships of alternative nitrogenases, *Chemical Reviews.* **96**, 3013-3030.

16. Boyd, E. S., Anbar, A. D., Miller, S., Hamilton, T. L., Lavin, M. & Peters, J. W. (2011) A late methanogen origin for molybdenum-dependent nitrogenase, *Geobiology.* **9**, 221-232.

17. Boyd, E. S., Garcia Costas, A. M., Hamilton, T. L., Mus, F. & Peters, J. W. (2015) Evolution of Molybdenum Nitrogenase during the Transition from Anaerobic to Aerobic metabolism, *Journal of Bacteriology.* **197**, 1690-1699.

18. Garcia, A. K., McShea, H., Kolaczkowski, B. & Kaçar, B. (2020) Reconstructing the evolutionary history of nitrogenases: Evidence for ancestral molybdenum-cofactor utilization, *Geobiology.* **18**, 394-411.

19. Koonin, E. V. (1993) A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif, *Journal of Molecular Biology.* **229**, 1165-1174.

20. Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases, *Journal of molecular biology.* **317**, 41-72.

21. Georgiadis, M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. & Rees, D. (1992) Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*, *Science.* **257**, 1653-1659.

22. Howard, J. B. & Rees, D. C. (1994) Nitrogenase: a nucleotide-dependent molecular switch, *Annual Review of Biochemistry.* **63**, 235-264.

23. Danyal, K., Dean, D. R., Hoffman, B. M. & Seefeldt, L. C. (2011) Electron transfer within nitrogenase: evidence for a deficit spending mechanism, *Biochemistry.* **50**, 9255-9263.

24. Lanzilotta, W. N. & Seefeldt, L. C. (1997) Changes in the midpoint potentials of the nitrogenase metal centers as a result of iron protein-molybdenum-iron protein complex formation, *Biochemistry.* **36**, 12976-12983.

25. Yang, Z.-Y., Ledbetter, R., Shaw, S., Pence, N., Tokmina-Lukaszewska, M., Eilers, B., Guo, Q., Pokhrel, N., Cash, V. L. & Dean, D. R. (2016) Evidence that the P_i release event is the rate-limiting step in the nitrogenase catalytic cycle, *Biochemistry.* **55**, 3625-3635.

26. Mortenson, L. E. (1964) Ferredoxin and ATP, requirements for nitrogen fixation in cell-free extracts of *Clostridium pasteurianum*, *Proceedings of the National Academy of Sciences.* **52**, 272-279.

27. Martin, A., Burgess, B., Iismaa, S., Smartt, C., Jacobson, M. & Dean, D. (1989) Construction and characterisation of an *Azotobacter vinelandii* strain with mutations in the genes encoding flavodoxin and ferredoxin I, *Journal of bacteriology.* **171**, 3162-3167.

28. Burgess, B. K. & Lowe, D. J. (1996) Mechanism of molybdenum nitrogenase, *Chemical Reviews.* **96**, 2983-3012.

29. Schneider, K. & Müller, A. (2004) 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) pp. 281-307, Springer Netherlands.

30. Harris, D. F., Lukoyanov, D. A., Kallas, H., Trncik, C., Yang, Z.-Y., Compton, P., Kelleher, N., Einsle, O., Dean, D. R., Hoffman, B. M. & Seefeldt, L. C. (2019) Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism To Achieve N_2 Reduction, *Biochemistry.* **58**, 3293-3301.

31. Seefeldt, L. C., Yang, Z.-Y., Duval, S. & Dean, D. R. (2013) Nitrogenase reduction of carbon-containing compounds, *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* **1827**, 1102-1111.

32. Hu, Y. & Ribbe, M. W. (2015) Nitrogenase and homologues, *J Biol Inorg Chem.* **20**, 435-45.

33. Rebelein, J. G., Stiebritz, M. T., Lee, C. C. & Hu, Y. (2017) Activation and reduction of carbon dioxide by nitrogenase iron proteins, *Nature Chemical Biology.* **13**, 147-149.

34. Stiebritz, M. T., Hiller, C. J., Sickerman, N. S., Lee, C. C., Tanifuji, K., Ohki, Y. & Hu, Y. (2018) Ambient conversion of CO_2 to hydrocarbons by biogenic and synthetic [Fe_4S_4] clusters, *Nature Catalysis.* **1**, 444-451.

35. Hiller, C. J., Stiebritz, M. T., Lee, C. C., Liedtke, J. & Hu, Y. (2017) Tuning Electron Flux through Nitrogenase with Methanogen Iron Protein Homologues, *Chemistry.* **23**, 16152-16156.

36. Rohde, M., Trncik, C., Sippel, D., Gerhardt, S. & Einsle, O. (2018) Crystal structure of VnfH, the iron protein component of vanadium nitrogenase, *Journal of Biological Inorganic Chemistry.* **23**, 1049-1056.

37. Trncik, C., Müller, T., Franke, P. & Einsle, O. (2022) Structural analysis of the reductase component AnfH of iron-only nitrogenase from *Azotobacter vinelandii*, *Journal of Inorganic Biochemistry.* **227**, 111690.

38. Schlessman, J. L., Woo, D., Joshua-Tor, L., Howard, J. B. & Rees, D. C. (1998) Conformational variability in structures of the nitrogenase iron proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum*, *Journal of Molecular Biology.* **280**, 669-85.

39. Rettberg, L. A., Kang, W., Stiebritz, M. T., Hiller, C. J., Lee, C. C., Liedtke, J., Ribbe, M. W. & Hu, Y. (2019) Structural Analysis of a Nitrogenase Iron Protein from *Methanosarcina acetivorans*: Implications for CO_2 Capture by a Surface-Exposed [Fe_4S_4] Cluster, *mBio.* **10**, e01497-19.

40. Boyd, E. S., Hamilton, T. L. & Peters, J. W. (2011) An alternative path for the evolution of biological nitrogen fixation, *Frontiers in Microbiology.* **2**, 1-11.

41. Maslać, N., Sidhu, C., Teeling, H. & Wagner, T. (2022) Comparative Transcriptomics Sheds Light on Remodeling of Gene Expression during Diazotrophy in the Thermophilic Methanogen *Methanothermococcus thermolithotrophicus*, *mBio.* **13**, e02443-22.

42. Angelidaki, I., Karakashev, D., Batstone, D. J., Plugge, C. M. & Stams, A. J. M. (2011) Chapter sixteen - Biomethanation and Its Potential in *Methods in Enzymology* (Rosenzweig, A. C. & Ragsdale, S. W., eds) pp. 327-351, Academic Press.

43. Liu, Y. & Whitman, W. B. (2008) Metabolic, Phylogenetic, and Ecological Diversity of the Methanogenic Archaea, *Annals of the New York Academy of Sciences.* **1125**, 171-189.

44. Miller, J. F., Shah, N. N., Nelson, C. M., Ludlow, J. M. & Clark, D. S. (1988) Pressure and Temperature Effects on Growth and Methane Production of the Extreme Thermophile *Methanococcus jannaschii*, *Applied and Environmental Microbiology.* **54**, 3039-3042.

45. Mehta, M. P., Butterfield, D. A. & Baross, J. A. (2003) Phylogenetic Diversity of Nitrogenase (*nifH*) Genes in Deep-Sea and Hydrothermal Vent Environments of the Juan de Fuca Ridge, *Applied and Environmental Microbiology.* **69**, 960-970.

46. Man-Aharonovich, D., Kress, N., Zeev, E. B., Berman-Frank, I. & Béjà, O. (2007) Molecular ecology of *nifH* genes and transcripts in the eastern Mediterranean Sea, *Environmental Microbiology.* **9**, 2354-2363.

47. Dang, H., Luan, X., Zhao, J. & Li, J. (2009) Diverse and Novel *nifH* and *nifH*-Like Gene Sequences in the Deep-Sea Methane Seep Sediments of the Okhotsk Sea, *Applied and Environmental Microbiology.* **75**, 2238-2245.

48. Bae, H. S., Morrison, E., Chanton, J. P. & Ogram, A. (2018) Methanogens Are Major Contributors to Nitrogen Fixation in Soils of the Florida Everglades, *Applied and Environmental Microbiology.* **84**, 1-16.

49. Belay, N., Sparling, R. & Daniels, L. (1984) Dinitrogen fixation by a thermophilic methanogenic bacterium, *Nature.* **312**, 286-288.

50. Kessler, P. S., McLarnan, J. & Leigh, J. A. (1997) Nitrogenase Phylogeny and the Molybdenum Dependence of Nitrogen Fixation in *Methanococcus maripaludis*, *Journal of Bacteriology.* **179**, 541-543.

51. Mehta, M. P. & Baross, J. A. (2006) Nitrogen Fixation at 92 °C by a Hydrothermal Vent Archaeon, *Science.* **314**, 1783-1786.

52. Enzmann, F., Mayer, F., Rother, M. & Holtmann, D. (2018) Methanogens: biochemical background and biotechnological applications, *AMB Express.* **8**, 1.

53. Jones, W. J., Paynter, M. J. B. & Gupta, R. (1983) Characterisation of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt marsh sediment, *Archives of Microbiology.* **135**, 91-97.

54. Huber, H., Thomm, M., König, H., Thies, G. & Stetter, K. O. (1982) *Methanococcus thermolithotrophicus*, a Novel Thermophilic Lithotrophic Methanogen, *Archives of Microbiology.* **132**, 47-50.

55. Jeanthon, C., L'Haridon, S., Reysenbach, A. L., Vernet, M., Messner, P., Sleytr, U. B. & Prieur, D. (1998) *Methanococcus infernu*s sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent, *International Journal of Systematic and Evolutionary Microbiology.* **48 Pt 3**, 913-919.

56. Robert, X. & Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.* **42**, W320-324.

57. Tezcan, F. A., Kaiser, J. T., Mustafi, D., Walton, M. Y., Howard, J. B. & Rees, D. C. (2005) Nitrogenase complexes: multiple docking sites for a nucleotide switch protein, *Science.* **309**, 1377-1380.

58. Rutledge, H. L. & Tezcan, F. A. (2020) Electron Transfer in Nitrogenase, *Chemical Reviews.* **120**, 5158-5193.

59. Krissinel, E. & Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol.* **372**, 774-797.

60. Cadoux, C., Ratcliff, D., Maslać, N., Gu, W., Tsakoumagkos, I., Hoogendoorn, S., Wagner, T. & Milton, R. D. (2023) Nitrogen Fixation and Hydrogen Evolution by Sterically Encumbered Mo-Nitrogenase, *JACS Au.* **3**, 1521-1533.

61. Schindelin, H., Kisker, C., Schlessman, J. L., Howard, J. B. & Rees, D. C. (1997) Structure of ADP·AlF_4^--stabilised nitrogenase complex and its implications for signal transduction, *Nature.* **387**, 370-376.

62. Renner, K. A. & Howard, J. B. (1996) Aluminum Fluoride Inhibition of Nitrogenase: Stabilisation of a Nucleotide·Fe protein·MoFe protein Complex, *Biochemistry.* **35**, 5353-5358.

63. Udvardi, M., Brodie, E. L., Riley, W., Kaeppler, S. & Lynch, J. (2015) Impacts of Agricultural Nitrogen on the Environment and Strategies to Reduce these Impacts, *Procedia Environmental Sciences.* **29**, 303-303.

64. Garcia, A. K., Harris, D. F., Rivier, A. J., Carruthers, B. M., Pinochet-Barros, A., Seefeldt, L. C. & Kaçar, B. (2023) Nitrogenase resurrection and the evolution of a singular enzymatic mechanism, *eLife.* **12**, e85003.

65. Mehta, R., Singhal, P., Singh, H., Damle, D. & Sharma, A. K. (2016) Insight into thermophiles and their wide-spectrum applications, *3 Biotech.* **6**, 81.

66. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. & Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox, *Acta Crystallographica Section D.* **67**, 293-302.

67. Tickle, I., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C. & Bricogne, G. (2018) Staraniso, *Global Phasing Ltd, Cambridge, UK*.

68. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007) Phaser crystallographic software, *Journal of applied crystallography.* **40**, 658-674.

69. 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. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr D Struct Biol.* **75**, 861-877.

70. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D: Biological Crystallography.* **66**, 486-501.

71. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, D. C. (2018) MolProbity: More and better reference data for improved all-atom structure validation, *Protein Sci.* **27**, 293-315.

72. Milton, R. D., Cai, R., Abdellaoui, S., Leech, D., De Lacey, A. L., Pita, M. & Minteer, S. D. (2017) Bioelectrochemical Haber–Bosch Process: An Ammonia-Producing H_2/N_2 Fuel Cell, *Angewandte Chemie International Edition.* **56**, 2680-2683.

73. Corbin, J. L. (1984) Liquid chromatographic-fluorescence determination of ammonia from nitrogenase reactions: A 2-min assay, *Applied and Environmental Microbiology.* **47**, 1027-1030.

74. Brown, K. A., Harris, D. F., Wilker, M. B., Rasmussen, A., Khadka, N., Hamby, H., Keable, S., Dukovic, G., Peters, J. W. & Seefeldt, L. C. (2016) Light-driven dinitrogen reduction catalyzed by a CdS: nitrogenase MoFe protein biohybrid, *Science.* **352**, 448-450.

75. Milton, R. D., Abdellaoui, S., Khadka, N., Dean, D. R., Leech, D., Seefeldt, L. C. & Minteer, S. D. (2016) Nitrogenase bioelectrocatalysis: Heterogeneous ammonia and hydrogen production by MoFe protein, *Energy and Environmental Science.* **9**, 2550-2554.

76. Lemaire, O. N., Infossi, P., Ali Chaouche, A., Espinosa, L., Leimkühler, S., Giudici-Orticoni, M.-T., Méjean, V. & Iobbi-Nivol, C. (2018) Small membranous proteins of the TorE/NapE family, crutches for cognate respiratory systems in Proteobacteria, *Scientific reports.* **8**, 13576.

77. Edgar, R. C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity, *BMC Bioinformatics.* **5**, 113.

78. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, *Mol Biol Evol.* **35**, 1547-1549.

79. Tamura, K., Stecher, G. & Kumar, S. (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11, *Mol Biol Evol.* **38**, 3022-3027.

80. Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol Biol Evol.* **4**, 406-425.

81. Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences, *Comput Appl Biosci.* **8**, 275-282.

82. Rees, D. C., Akif Tezcan, F., Haynes, C. A., Walton, M. Y., Andrade, S., Einsle, O. & Howard, J. B. (2005) Structural basis of biological nitrogen fixation, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* **363**, 971-984.

83. Seefeldt, L. C., Peters, J. W., Beratan, D. N., Bothner, B., Minteer, S. D., Raugei, S. & Hoffman, B. M. (2018) Control of electron transfer in nitrogenase, *Current Opinion in Chemical Biology.* **47**, 54-59.

3.11. Figures and legends

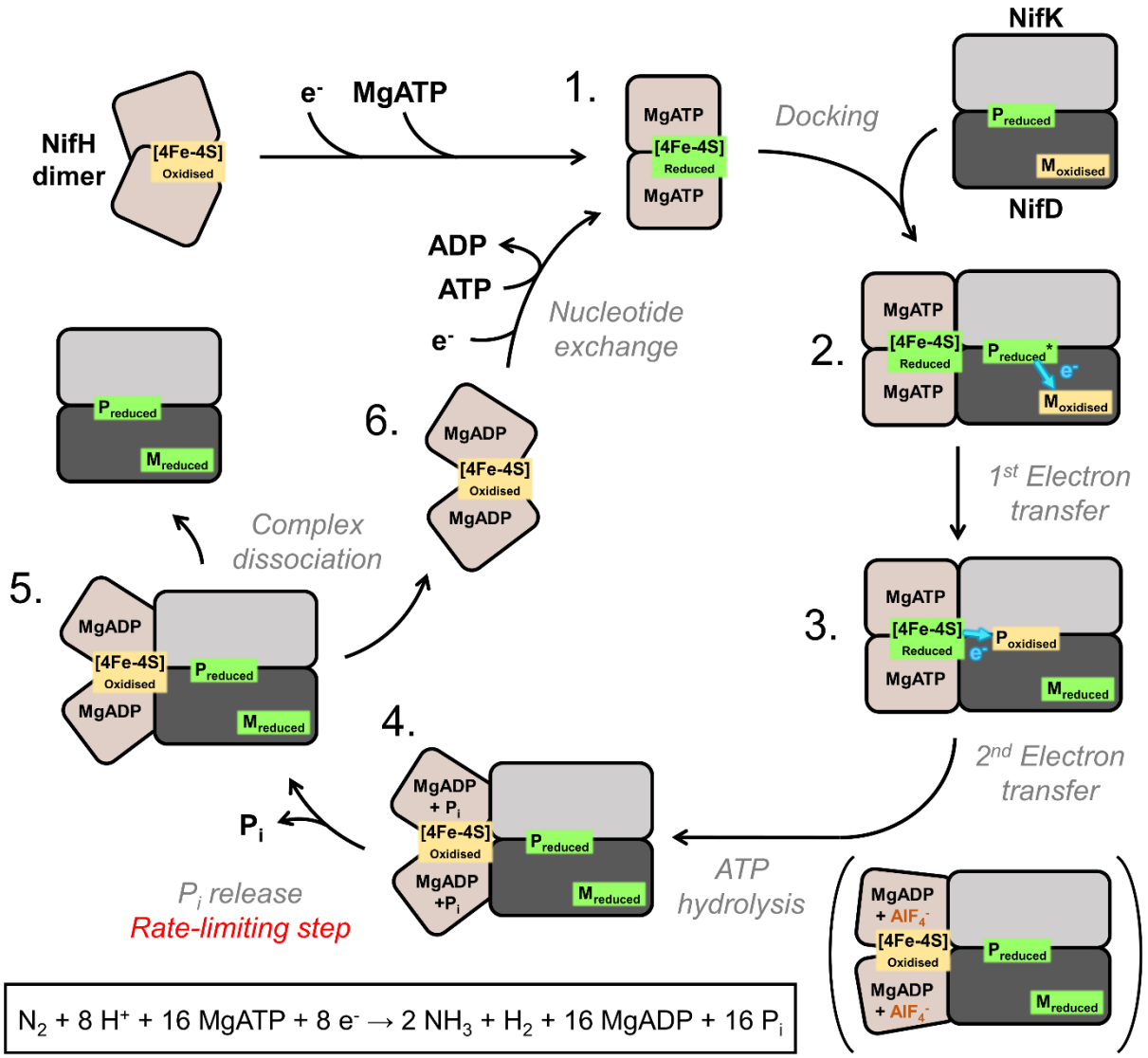
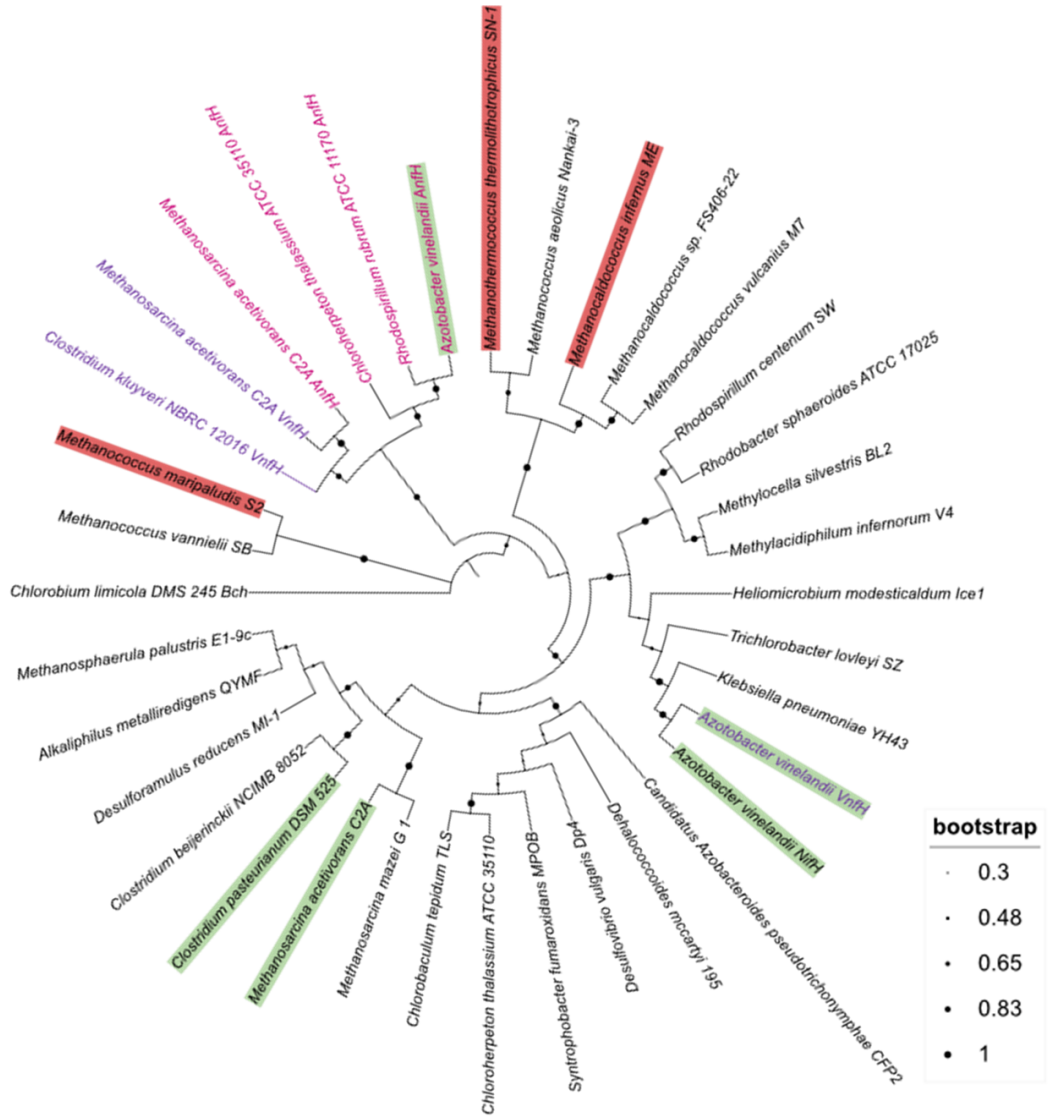


Fig. 1. Order of events for the nitrogenase catalytic cycle during a single turnover and according to the "deficit spending model" [23, 82] and based on [83]. The scheme depicts the six different states (labelled as 1-6) of the NifH cycle. NifH is shown as a homodimer loaded with a reduced [4Fe-4S] cluster and MgATP in its active state (state 1). Only one-half of the NifDK heterotetramer is displayed for clarity, with whom NifH associates (state 2). NifDK-NifH association triggers an electron transfer (state 2 to 3) from the activated P cluster (labelled as P*) to the FeMoco (labelled as M). A second electron transfer follows from the reduced [4Fe-4S] cluster of NifH to the oxidised P cluster (states 3 to 4). In the next step, the ATP is hydrolysed (state 4 to 5), leading to the NifH homodimer containing 2MgADP/2P_i and the oxidised [4Fe-4S] cluster (state 4). The artificial complex of NifDK-NifH with NifH in complex with MgADP and AlF_4^- mimics this intermediate transition in which the [4Fe-4S]-cluster is adequately positioned for electron transfer to the P cluster [57, 58]. After P_i release (state 4 to 5), NifH-MgADP disengaged from NifDK (state 6) and can be reloaded with an electron and ATP to start a new cycle.

A



B

% identity	*Av*NifH	*Av*VnfH	*Av*AnfH	*Mm*NifH	*Mt*NifH	*Mi*NifH	*Cp*NifH
*Av*NifH	100.00	91.00	63.87	66.18	63.27	63.50	69.74
*Av*VnfH		100.00	64.96	65.82	63.27	63.87	91.35
*Av*AnfH			100.00	73.72	75.09	73.72	64.23
*Mm*NifH				100.00	74.73	75.09	73.82
*Mt*NifH					100.00	85.16	73.72
*Mi*NifH						100.00	75.09
*Cp*NifH							100.00

Fig. 2. Phylogenetic analyses and sequence conservation of NifH from *Methanococcales* and homologues. **A.** Evolutionary analysis of 37 selected NifH/VnfH/AnfH sequences. Node statistics are represented by black dots whose diameter depends on the node score (see bootstrap legend). Structures currently available are highlighted with green background. Structures presented in this work are highlighted with red background. . Labels corresponding to NifH, VnfH and AnfH are colored in black, purple and pink text, respectively. ChlL (light-independent protochlorophyllide reductase) from *Chlorobium limicola* was used as an outgroup. Accession numbers for sequences used in phylogenetic reconstruction can be found in Supplementary Table 1. **B.** Sequence identity matrix between selected NifH/VnfH/AnfH proteins. Colour code: red-green from the lowest to the highest percentual identity value.

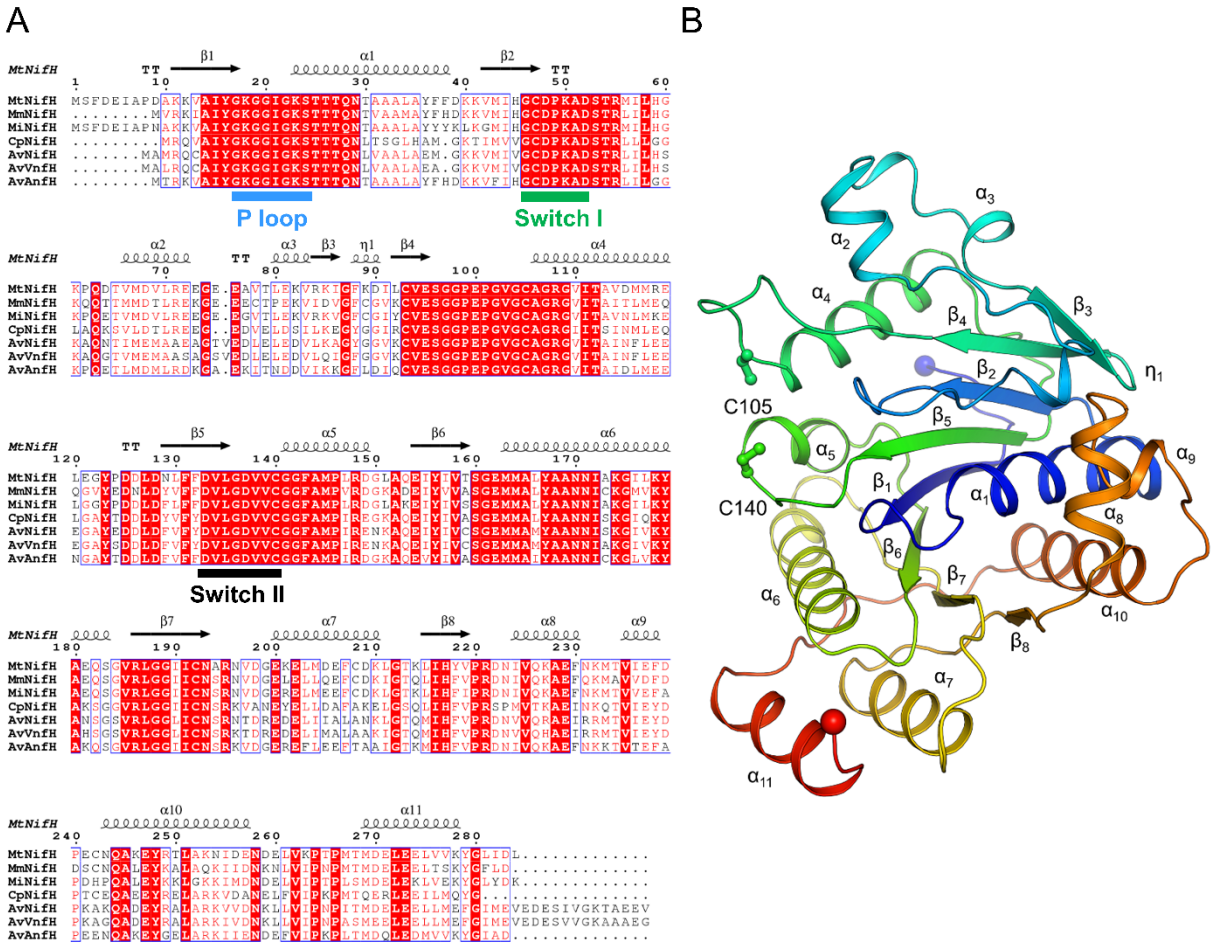


Fig. 3. A. Residues conservation in selected NifH, VnfH and AnfH proteins. Sequence alignment of *Mt*NifH, *Mi*NifH, *Mm*NifH, *Ma*NifH (PDB code 6NZJ), *Av*NifH (PDB code 2NIP), *Av*Vnf (PDB code 6Q93), *Ma*Vnf (GenBank access code AAM04632.1) and *Av*Anf (PDB code 7QQA). The alignment was done by MUSCLE [76], while the conserved residues and the corresponding secondary structures were analysed by ESPript (*https://espript.ibcp.fr*, [56]) on the structure of *Mt*NifH. **B.** Secondary structures of *Mt*NifH monomer labelled based on ESPript prediction. The chain is coloured as a rainbow, starting from the C terminus (blue) to the N terminus (red).

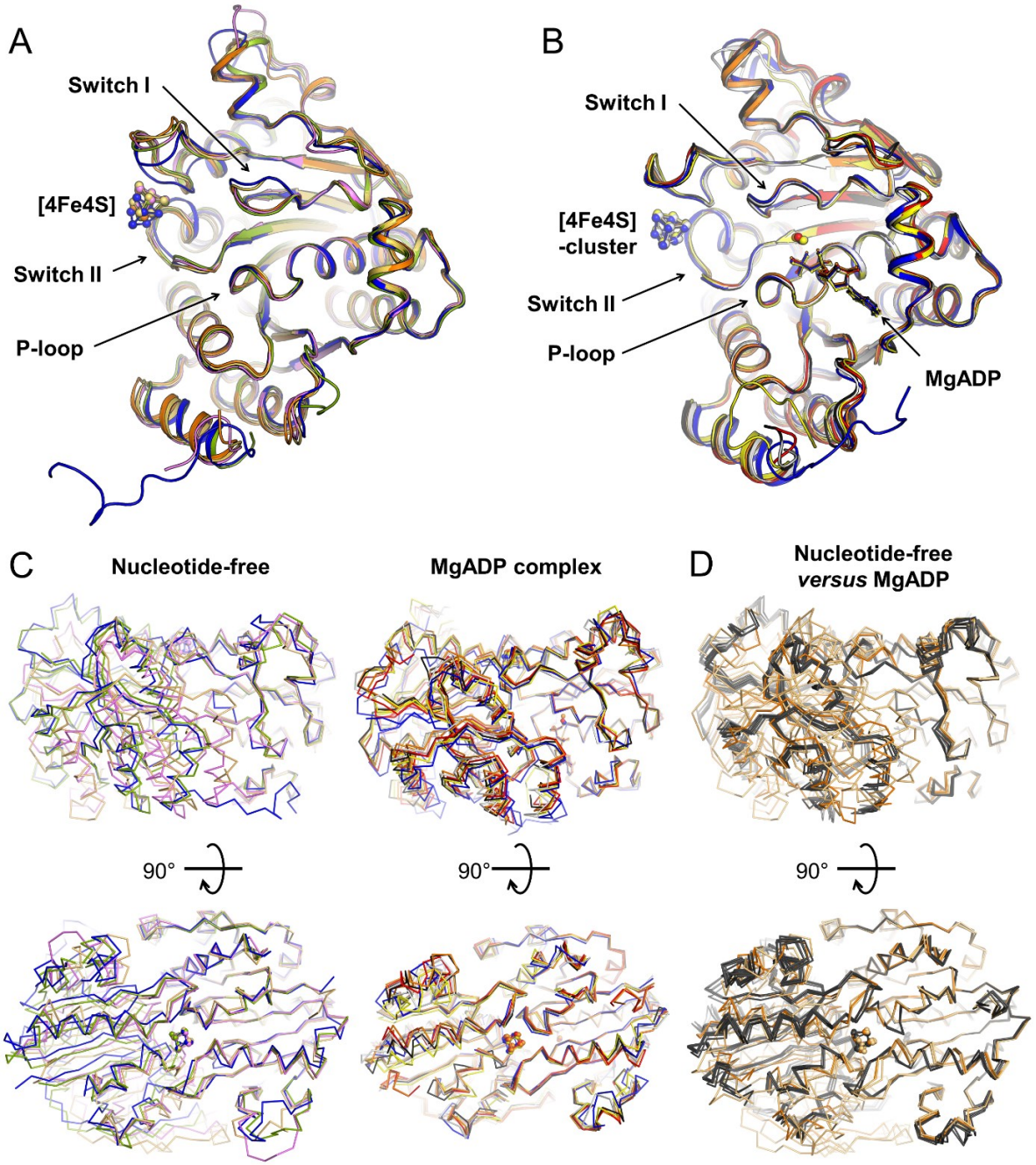


Fig. 4. A. Superposition of nucleotide-free monomer from *Mt*NifH form 1 (yellow-orange) and form 2 (orange), *Ma*NifH (PDB: 6NZJ, violet), *Cp*NifH (PDB: 1CP2, split-pea) and *Av*NifH (PDB: 2NIP, sky blue). **B**. Superposition of MgADP-bound monomer from *Mm*NifH (yellow), *Mt*NifH (orange), *Mi*NifH (red), *Av*NifH (PDB: 6N4L, blue), *Av*VnfH (PDB: 6Q93, white) and *Av*AnfH (PDB: 7QQA, black), in the same pose as panel A. All models are pictured in cartoons with ligands in balls and sticks. **C.** Left: side and top views of NifH homodimers nucleotide-free with the same models and colour coding as in panel A. Right: side and top views of MgADP-bound NifH homodimers with the same models and colour-coding as in panel B. **D**. Superposition of nucleotide-free *Mt*NifH form 2 (orange), with the two homodimers occupying the asymmetric unit of form 1 (yellow-orange) and the six MgADP-bound *Mt*NifH homodimers occupying the asymmetric unit (black). For C and D, all models displayed as ribbons were superposed on the right *Mt*NifH monomer (nucleotide-free, form 2).

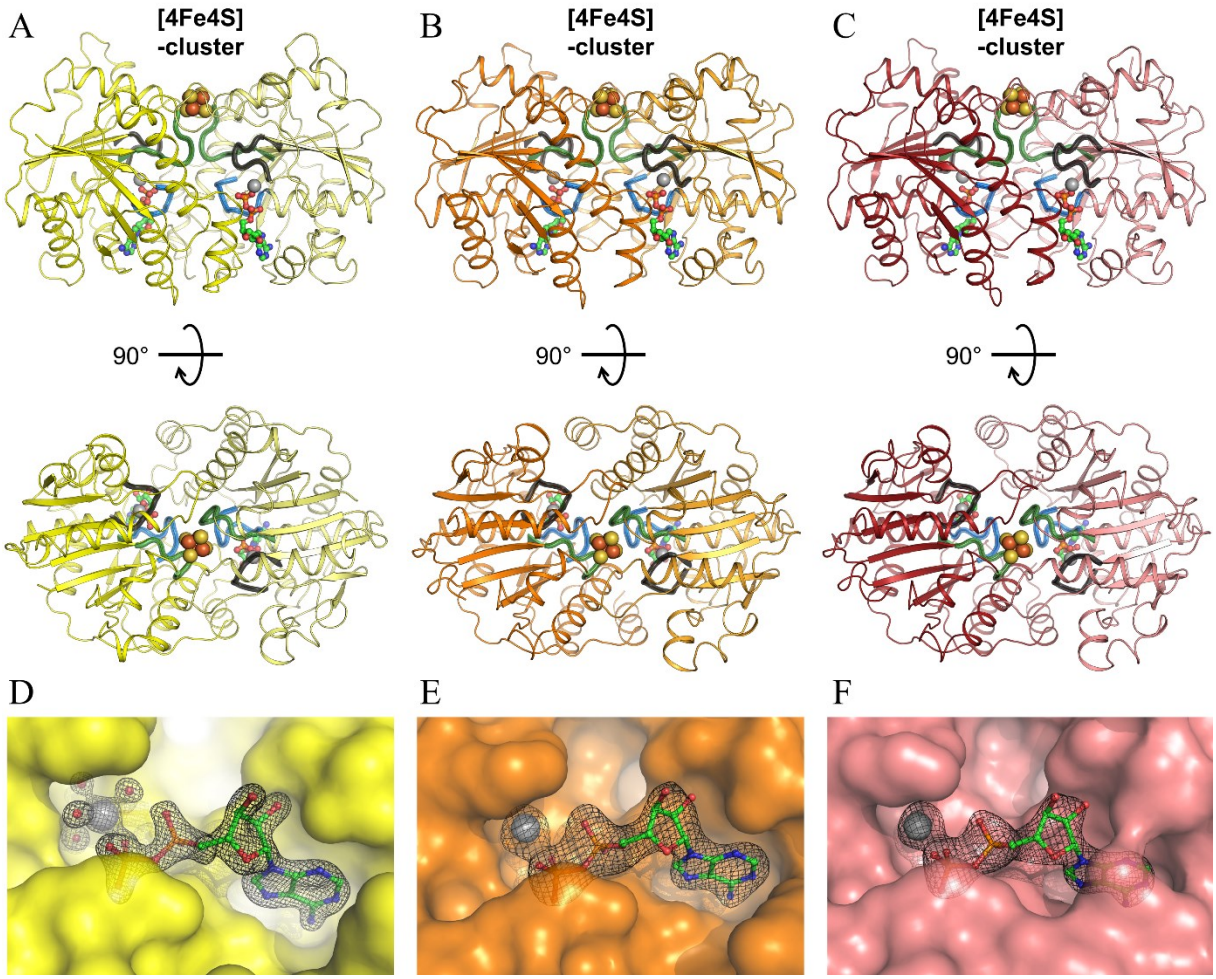


Fig. 5. A-C. Side and top views of MgADP-bound *Mm*NifH (A, yellow), *Mt*NifH (B, orange) and *Mi*NifH (C, red). Homodimers are displayed in cartoons with one monomer in a fainter colour than the other. Important structural features are colour-coded as follows: Switch I – forest green, Switch II – black, and P-loop – marine blue. **D-F.** Electron density of MgADP within the ligand binding site in *Mm*NifH ligand binding site of *Mm*NifH (D, yellow), *Mt*NifH (E, orange) and *Mi*NifH (F, red). A surface represents the proteins. All electron density maps ($2F_o$-F_c) are contoured to 2-σ. Due to the lower resolution of the models, waters were not modelled in E and F panels. All ligands are depicted in ball and stick models with carbon, nitrogen, oxygen, phosphorus, magnesium, sulfur and iron coloured in green, blue, red, grey, orange, yellow and dark orange.

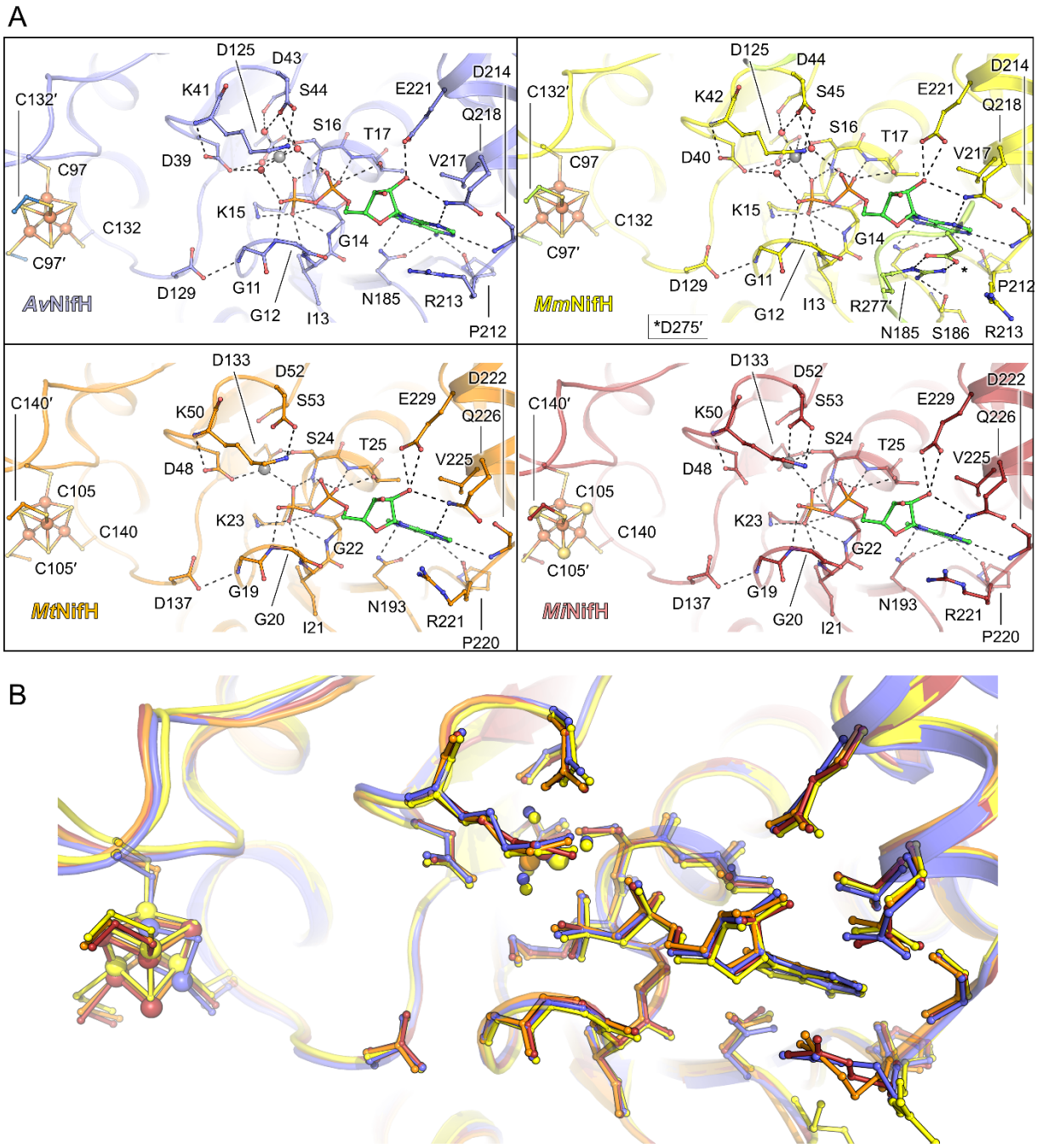


Fig. 6. A. Conservation of the [4Fe-4S]-cluster and MgADP binding site in *Av*NifH (PDB: 6N4L, blue), *Mm*NifH (pale yellow), *Mt*NifH (orange), *Mi*NifH (firebrick red). The protein is pictured in transparent cartoons on all panels, with the residues coordinating the [4Fe-4S] and MgADP shown as balls and sticks. Carbon is coloured appropriately with the identity of the protein. Nitrogen, oxygen, sulphur, phosphorus, magnesium and iron are coloured blue, red, yellow, orange, grey and light orange, respectively. **B**. Compilation of all four ligand binding sites from panel A with *Av*NifH (PDB: 6N4L), *Mm*NifH, *Mt*NifH, and *Mi*NifH coloured in blue, yellow, orange, and firebrick red, respectively. For clarity, the C-terminal artificial extension in *Mm*NifH, shown as lemon colour in A, has been excluded in panel B. The primed numbers indicate the residues carried by the second monomer.

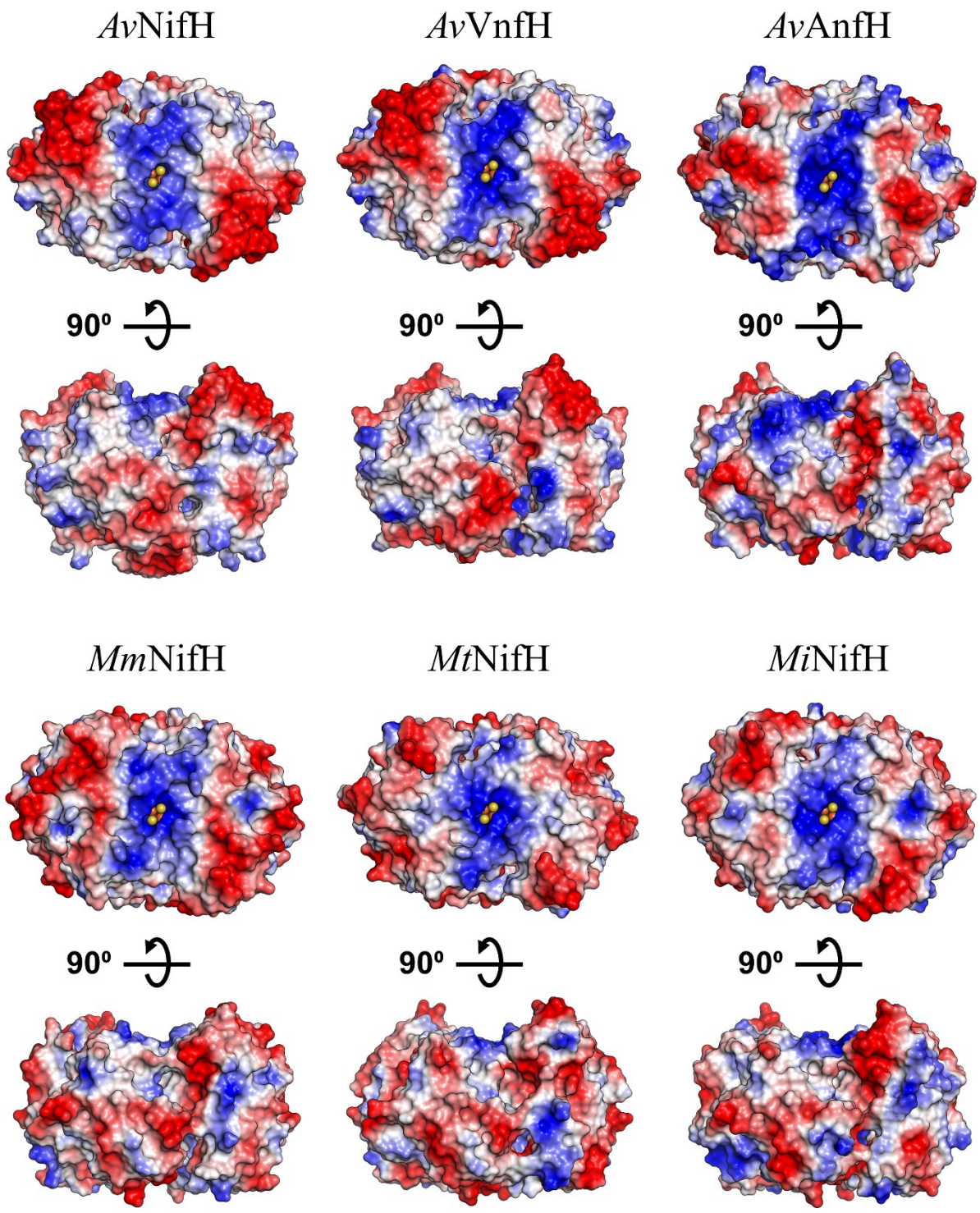


Fig. 7. Electrostatic surface properties of selected NifH, VnfH and AnfH proteins, all in MgADP-bound state. For all models, the colour code reflects the charges on the surface from negative (red) to positive (blue). PDB codes for *Av*NifH, *Av*VnfH and *Av*AnfH are 6N4L, 6Q93, and 7QQA, respectively. The [4Fe-4S]-cluster is depicted as balls with sulphur and iron coloured in yellow and orange, respectively.

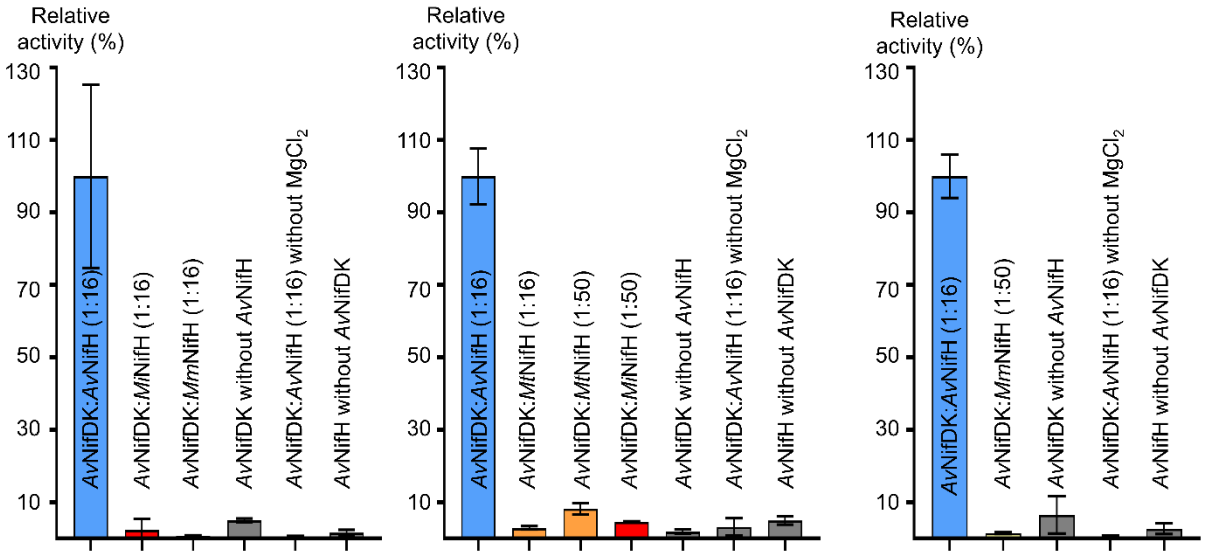


Fig. 8. Cross-reactivity of *Mm*/*Mt*/*Mi*NifH proteins with *Av*NifDK is shown as a relative activity for NH_3-production from N_2-reduction compared to *Av*NifDKH. Relative activities for *Av*NifH, *Mm*NifH, *Mt*NifH, *Mi*NifH and negative controls are highlighted in blue, yellow, orange, red and grey, respectively. *Mm*/*Mt*/*Mi*NifH does not show activity when combined with *Av*NifDK in a 1:16 ratio *in vitro*. Only *Mt*NifH shows some activity at a ratio of 1:50. The three panels represent three different sets of experiments. All experiments have been performed in triplicates.

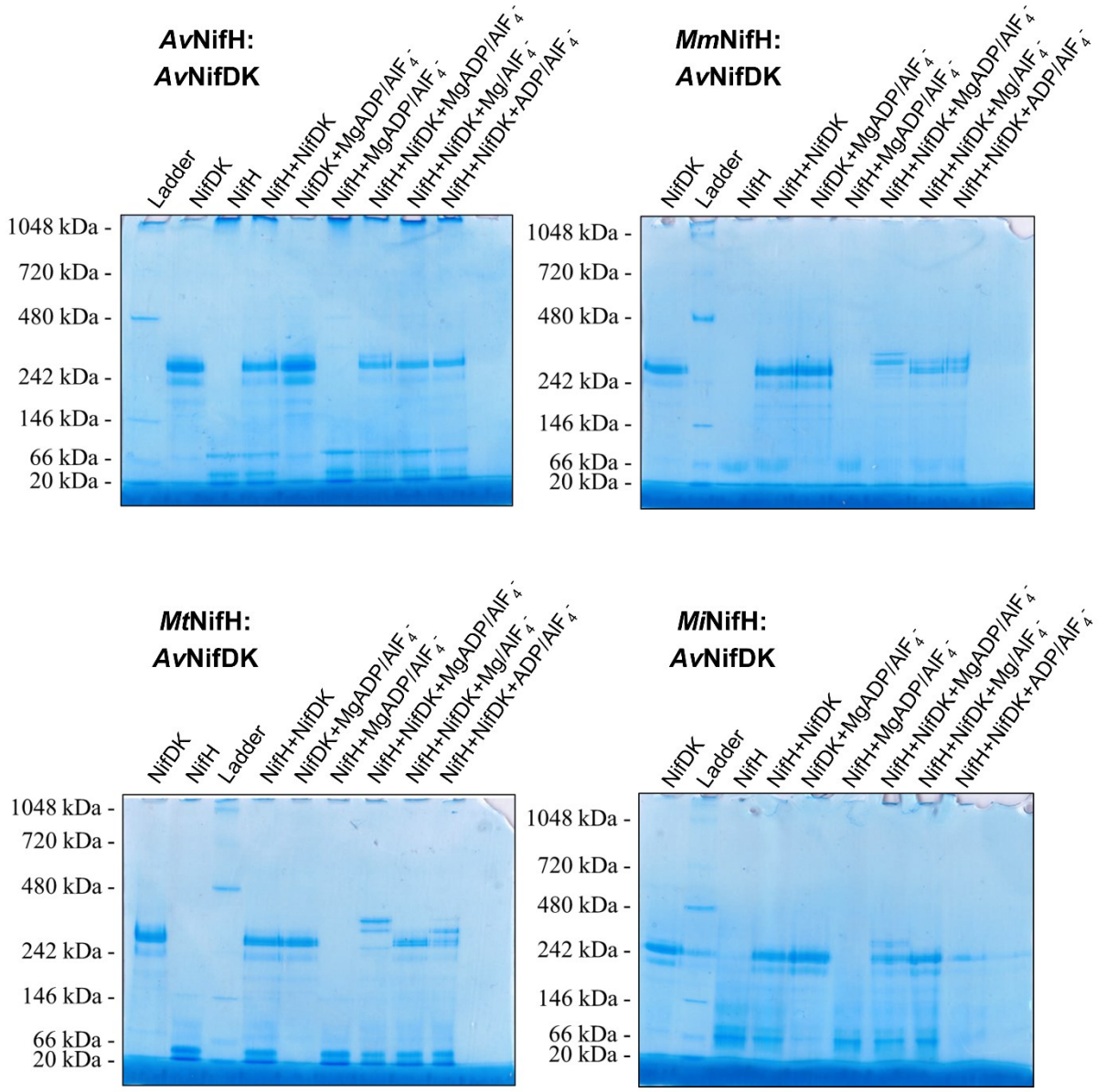


Fig. 9. Native PAGE analysis of the recombinantly expressed NifHs from methanogens in complex with *Av*NifDK upon MgADP/AlF_4^- addition. All samples were incubated at 20 °C inside an anaerobic chamber (N_2/H_2 (97:3%)) for 1 h in 100 mM Tris/HCl pH 7.6, with 20 mM NaF, 1 mM $AlCl_3$, 1 mM ADP and 2 mM $MgCl_2$ (when described) before loading on native PAGE. NifDK:NifH molar ratio was 1:5 in all cases. The gels were run in the same anaerobic chamber on a 3-15% gradient gels. 3 µg of protein total per well in all wells (see Material and Methods).

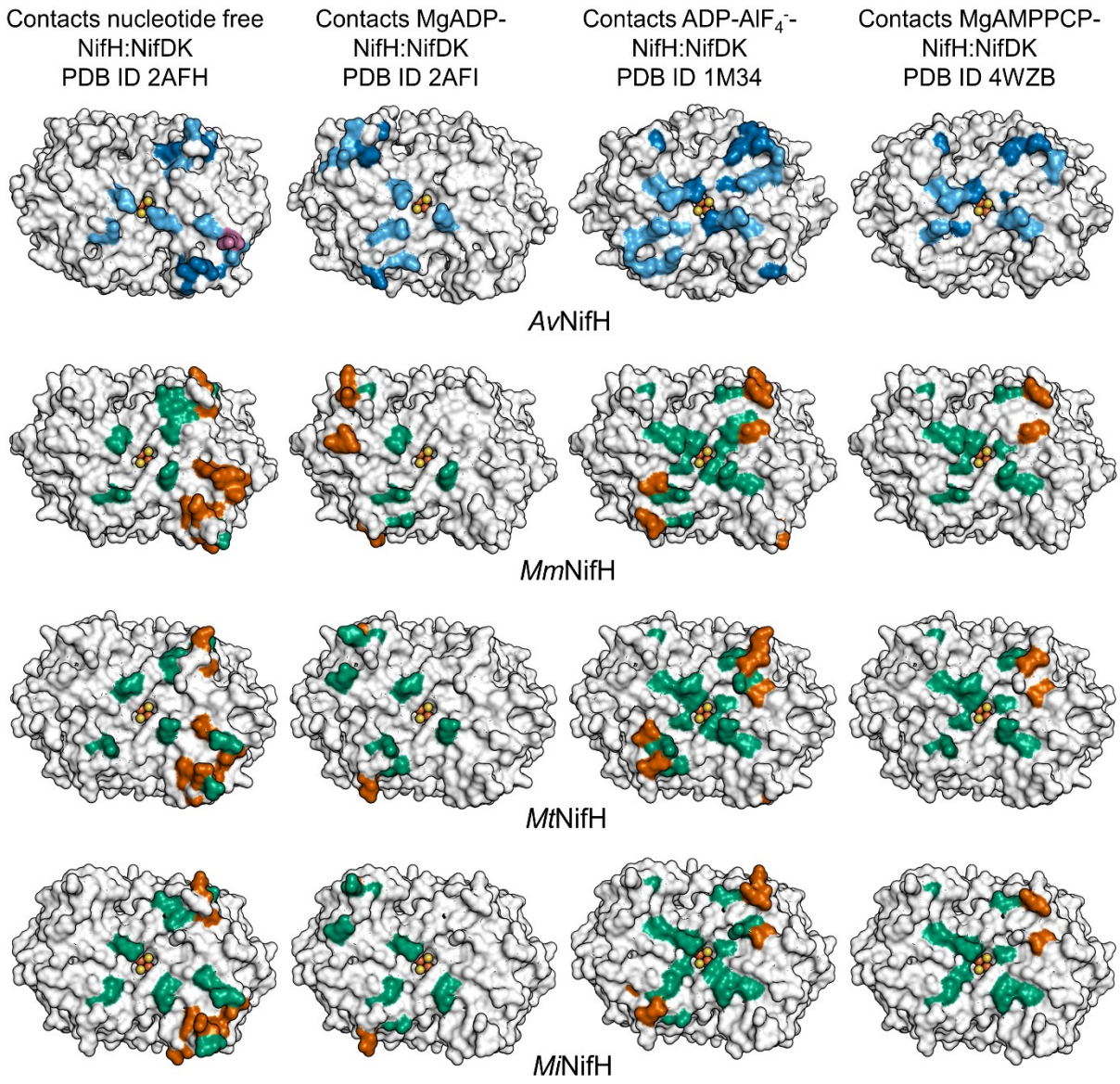


Fig. 10. Mapping of the *Av*NifH residues interacting with *Av*NifDK in four different nitrogenase complexes as described in the Supplementary Table S4 from Tezcan *et al.* [57]. All structures are displayed from the top view. Residues making contacts are colour coded as follows: dark blue- the contact involves the main chain atoms of a specified residue; sky blue- side chain groups mediate the contacts; magenta-the contact involves both the main and the chain. For NifH from methanogens, the equivalent residues making the contacts with the dinitrogenase are colour coded as follows: red-the labelled residue is not conserved between *Av*NifH and *Mm*/*Mt*/*Mi*NifH, green-the labelled residue is conserved between *Av*NifH and *Mm*/*Mt*/*Mi*NifH. The summary of all labelled residues can be found in Supplementary Table 2. The three methanogenic NifH represented models are the MgADP-bound state.

Table 1. Data collection and refinement statistics.

	***Mt*NifH as isolated form 1**	***Mt*NifH as isolated form 2**	***Mm*NifH-MgADP**	***Mt*NifH-MgADP**	***Mi*NifH-MgADP**
Data collection					
Synchrotron source	SOLEIL, Proxima I	SOLEIL Proxima I	SLS, PXIII	SLS, PXIII	SLS, PXI
Wavelength (Å)	1.74013	1.74013	0.99999	1.00004	0.99999
Space group	$P2_1$	$P4_32_12$	$P2_1$	$P1$	$P2_12_12_1$
Resolution (Å)	68.15 – 2.31 (2.35– 2.31)	52.00-1.91 (1.94-1.91)	78.82– 1.70 (1.86– 1.70)	71.00 – 2.74 (2.90 – 2.74)	120.26 – 2.49 (2.83 – 2.49)
Cell dimensions					
a, b, c (Å)	52.82, 83.98, 116.60	98.90, 98.90, 61.14	68.55, 111.09, 78.82	61.48, 87.59, 182.05	147.95, 156.73, 206.45
α, β, γ (°)	90, 91.14, 90	90, 90, 90	90, 90, 90	93.978, 98.58, 109.65	90, 90, 90
R_{merge}(%)[a]	24.0 (134.2)	15.9 (312.0)	21.9 (123.2)	28.4 (122.4)	17.6 (231.4)
R_{pim} (%)[a]	8.5 (56.6)	2.5 (64.4)	9.0 (51.1)	11.7 (50.9)	4.0 (52.9)
$CC_{1/2}$[a]	0.990 (0.517)	0.997 (0.616)	0.992 (0.468)	0.983 (0.526)	0.998 (0.680)
I/σ_I[a]	5.2 (1.5)	15.9 (1.4)	7.5 (1.6)	4.5 (1.6)	15.5 (2.0)
Spherical completeness[a]	100 (100)	100 (100)	73.5 (15.9)	72.1 (22.8)	54.4 (8.6)
Ellipsoidal completeness[a]	/	/	94.2 (57.6)	90.5 (63.1)	91.2 (77.5)
Redundancy[a]	8.7 6.5)	41.2 (23.9)	6.9 (6.6)	6.7 (6.7)	20.8 (19.3)
Nr. unique reflections[a]	45,087 (2,223)	24,200 (1,183)	94,344 (4,718)	66,142 (3,154)	91,224 (4,487)
Refinement					
Resolution (Å)	48.47 – 2.31	52.01 – 1.91	46.89 – 1.70	39.93 – 2.74	49.26-2.49
Number of reflections	45,074	24,197	94,335	66,105	91,144
R_{work}/R_{free}[b] (%)	19.63/23.49	20.32/23.05	16.18/18.59	21.90/25.43	22.20/24.20
Number of molecules/AU					
Number of atoms	8,977	2,278	10,146	25,980	26,165
Protein	8,553	2,123	8,598	25,495	25,660
Ligands/ions	46	29	205	462	464
Solvent	378	126	1,343	23	41
Mean B-value ($Å^2$)	37.92	46.31	20.79	60.62	79.27
Molprobity clash score, all atoms	4.08	1.86	1.66	2.56	2.61
Ramachandran plot					
Favoured regions (%)	97.92	97.09	98.38	97.45	98.41
Outlier regions (%)	0	0	0	0.06	0.03
rmsd[c] bond lengths (Å)	0.006	0.013	0.009	0.009	0.009
rmsd[c] bond angles (°)	0.874	1.431	0.913	1.213	1.258
PDB ID code					

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Table 2. Analysis of the interface stability in different NifH homologues. ΔG^{diss} = the free energy of assembly dissociation, N_{HB} = number of hydrogen bonds, N_{SB} = number of salt bridges, T °C = optimal temperature for the host growth.

	***Av*NifH**	***Av*VnfH**	***Av*AnfH**	***Mm*NifH**	***Mt*NifH**	***Mi*NifH**
ΔG^{diss} [kcal/mol]	37	25.83 ± 4.04	23.80 ± 0.42	24.25 ± 2.19	24.20 ± 1.31	21.55 ± 1.08
N_{HB}	14	12 ± 1	7 ± 4	18 ± 1	9 ± 1	9 ± 2
N_{SB}	14	7 ± 3	3 ± 1	10 ± 2	6 ± 2	8 ± 1
T °C	30 °C	30 °C	30 °C	37 °C	65 °C	85 °C

3.12. Supplemental Material

Supplementary Table 1. Accession Numbers for Sequences Used in Phylogenetic Reconstruction

Supplementary Table 2. Conservation of NifH Residues interacting with NifDK

Supplementary Table 3. Vectors, and strains used in this study.

***Mm/Mt/Mi*NifH construct sequences**

Supplementary Figure 1. SDS PAGE of purified *M. infernus* and *M. thermolithotrophicus* proteins.

Supplementary Figure 2. Growth of diazotrophic *M. infenus*.

Table S1. Accession Numbers for Sequences Used in Phylogenetic Reconstruction.

Taxon	NifH
Alkaliphilus metalliredigens QYMF	WP_012064614.1
Azotobacter vinelandii AvOP AnfH	WP_012703362.1
Azotobacter vinelandii AvOP NifH	WP_012698831.1
Azotobacter vinelandii AvOP VnfH	WP_012698955.1
Candidatus *Azobacteroides pseudotrichonymphae* CFP2	WP_012573569.1
Candidatus *Methanosphaerula palustris* E1-9c	WP_012617253.1
Chlorobium tepidum TLS	WP_010933198.1
Chloroherpeton thalassium ATCC 35110	WP_012500369.1
Chloroherpeton thalassium ATCC 35111	WP_012499583.1
Clostridium beijerinckii NCIMB 8052	ABR34169
Clostridium kluyveri NBRC 12016	WP_012102140.1
Clostridium pasterianum DSM 525	WP_003447877.1
Dehalococcoides mccartyi 195	WP_010936850.1
Desulfotomaculum reducens MI-1	WP_011879120.1
Desulfovibrio vulgaris DP4	WP_011176593.1
Geobacter lovleyi SZ	WP_012468732.1
Heliobacterium modesticaldum Ice1	WP_012282218.1
Klebsiella pneumoniae YH43	BAS34484.1
Methanocaldococcus infenus ME	WP_013099459.1
Methanocaldococcus sp. FS406-22	WP_012979647.1
Methanocaldococcus vulcanius M7	WP_015733159.1
Methanococcus aeolicus Nankai-3	WP_011974142.1
Methanococcus maripaludis S2	WP_011170797.1
Methanococcus vannielii SB	WP_011971883.1
Methanosarcina acetivorans CA2 AnfH	AAM04624.1
Methanosarcina acetivorans CA2 NifH	AAM07246.1
Methanosarcina acetivorans CA2 VnfH	AAM04632.1
Methanosarcina mazei Go1	WP_011032670.1
Methanothermococcus thermolithotrophicus SN-1	WP_018154785.1
Methylacidiphilum infernorum V4	WP_012464212.1
Methylocella silvestris BL2	WP_012592585.1
Nostoc sp. PCC 7120	WP_010995626.1
Rhodobacter sphaeroides ATCC 17025	WP_011908284.1
Rhodospirillum centenum SW	WP_012568805.1
Rhodospirillum rubrum ATCC 11170	WP_011389149.1
Syntrophobacter fumaroxidans MPOB	WP_011697884.1
Chlorobium limicola DSM 245	WP_012467084.1

Table S2. Conservation of NifH Residues interacting with NifDK.

	*Av*NifH		*Mm*NifH	*Mt*NifH	*Mi*NifH
	chain 1	**chain 2**			
nf-Av2:Av1 intermolecular contacts (PDB ID 2AFH)	Ala g2-62		Leu 62	Leu 70	Leu 70
	Ala g2-65		Lys 65	Glu 73	Glu 73
	Thr g2-67		Glu 67	Glu 75	Glu 75
	Glu g2-69		Glu 68	Glu 76	Glu 76
	Leu g2-71		Cys 70	Val 78	Val 78
		Arg g1-101	Arg 100	Arg 108	Arg 108
	Arg g2-101		Arg 100	Arg 108	Arg 108
	Asn g2-108		Thr 107	Asp 115	Asn 115
	Glu g2-112		Gln 111	Glu 119	Glu 119
	Glu g2-112		Gln 111	Glu 119	Glu 119
	Glu g2-113		His 112	Leu 120	Leu 120
		Arg g1-141	Arg 140	Arg 148	Arg 148
	Lys g2-171		Lys 170	Lys 178	Lys 178
	Tyr g2-172		Tyr 171	Tyr 179	Tyr 179
	Ala g2-173		Ala 172	Ala 180	Ala 180
	Asn g2-174		Glu 173	Glu 181	Glu 181
	Gly g2-176		Ser 175	Ser 183	Ser 183
	Arg g1-179		Arg 178	Arg 186	Arg 186
adp-Av2:Av1 intermolecular contacts (PDB ID 2AFI)					
		Gly 66	Gly 66	Gly 74	Gly 74
		Val 68	gap	gap	gap
		Glu 69	Glu 68	Glu 76	Glu 76
		Asp 70	Glu 69	Ala 77	Gly 77
	Arg 101		Arg 100	Arg 108	Arg 108
		Arg 101	Arg 100	Arg 108	Arg 108
		Glu 112	Gln 111	Glu 119	Glu 119
		Arg 141	Arg 140	Arg 148	Arg 148
		Lys 171	Lys 170	Lys 178	Lys 178
		Asn 174	Glu 173	Glu 181	Glu 181

alf-Av2:Av1 (PDB ID 1M34)		Gly 66	Gly 66	Gly 74	Gly 74
	Asp 70		Glu 69	Ala 77	Gly 77
	Cys 98		Cys 97	Cys 105	Cys 105
		Cys 98	Cys 97	Cys 105	Cys 105
	Arg 101		Arg 100	Arg 108	Arg 108
		Arg 101	Arg 100	Arg 108	Arg 108
	Thr 105		Thr 104	Thr 113	Thr 112
		Thr 105	Thr 104	Thr 113	Thr 112
	Gly 134		Gly 133	Gly 141	Gly 141
		Gly 134	Gly 133	Gly 141	Gly 141
	Arg 141		Arg 140	Arg 148	Arg 148
		Arg 141	Arg 140	Arg 148	Arg 148
	Glu 142		Asp 141	Asp 149	Asp 149
		Glu 142	Asp 141	Asp 149	Asp 149
	Lys 171		Lys 170	Lys 178	Lys 178
		Lys 171	Lys 170	Lys 178	Lys 178
	Asn 174		Glu 173	Glu 181	Glu 181
	Ser 175		Gln 174	Gln 182	Gln 182
		Ser 175	Gln 174	Gln 182	Gln 182
pcp-Av2:Av1 intermolecular contacts (PDB ID 4WZB)	Gly 66		Gly 66	Gly 74	Gly 74
	Cys 98		Cys 97	Cys 105	Cys 105
	Arg 101		Arg 100	Arg 108	Arg 108
		Arg 101	Arg 100	Arg 108	Arg 108
	Thr 105		Thr 104	Thr 113	Thr 112
		Thr 105	Thr 104	Thr 113	Thr 112
	Gly 134		Gly 133	Gly 141	Gly 141
		Gly 134	Gly 133	Gly 141	Gly 141
		Arg 141	Arg 140	Arg 148	Arg 148
	Glu 142		Asp 141	Asp 149	Asp 149
	Lys 171		Lys 170	Lys 178	Lys 178
	Ser 175		Gln 174	Gln 182	Gln 182

main chain
side chain
conserved
not conserved

Table S3. Vectors, and strains used in this study.

Vectors	
pDB1282 [1, 2]	with arabinose-inducible operon containing key components of the *isc* operon (*iscSUA*, *hscBA*, *fdx*, *iscX*)
WG35_*Mm*NifH	WG35 vector with *Mm*NifH
pET-28(+)_*Mt*NifH	with *Mt*NifH between NcoI and NdeI restriction sites
pET-28(+)_*Mi*NifH	with *Mi*NifH between NcoI and NdeI restriction sites
Strains	
E. coli DH5α WG35_*Mm*NifH	carrying WG35 vector with *M. maripaludis* NifH (original construct from Wenyu Gu, Department of Chemical Engineering and Civil and Environmental Engineering, Stanford University, Stanford, USA)
E.coli DH5α pDB1282	carrying pDB1282 plasmid (original construct from Dennis Dean, Department of Biochemistry, Virginia Tech, Blaksburg, USA)
E. coli BL21 Δ*iscR* [3]	*iscR* has been replaced with kanamycin resistance (on the genome) (original construct from James Swartz, Stanford Chemical Engineering, Stanford University, Stanford, USA)
A. vinelandii RS1	8 x HIS-tag between codons 2/3 of *nifD* [4, 5]. Comparable to strain DJ1141 [6]. Produced from *A. vinelandii* DJ (BTCC).

***Mm*NifH sequence:**

Green Bold = Start codon

Bold = Stop codon

Black=NifH gene sequence

Blue = Linker

Orange = His Tag

ATGGTGCGCAAGATCGCTATCTATGGCAAAGGTGGTATCGGCAAGTCTACTACGA
CGCAAAATACAGTCGCGGCAATGGCACACTTCCATGATAAGAAGGTATTCATTCA
TGGCTGTGACCCAAAGGCGGATAGTACCCGTCTTATCTTACATGGGAAACAGCAG
GTTACGATGATGGATACGCTGCGCGAGAAGGGTGAGGATGAATGCACGCCAGAC
AAGGTTATCGAGGTTGGTTTCGGGGGAGTGAAGTGCGTCGAGTCTGGTGGTCCTG
AACCCGGTGTTGGTTGTGCGGGTCGCGGTGTCATTACGGCTATTACGTTGATGGA
GCAGCATGGGGTCTACGAGGACGATTTGGATTTCGTATTCTTTGACGTTTTGGGC
GATGTAGTTTGTGGGGGTTTTGCGATGCCCGTCCGTGATGGGAAAGCTGACGAAA
TTTATGTAGTCGCCAGCGGAGAAATGATGGCGTTATACGCAGCTAACAACATCTG
CAAGGGCATGGTCAAGTATGCGGAGCAATCCGGGGTCCGTCTGGGAGGCATTATT
TGTAACTCACGCAATGTGGACGGCGAGCTGGATCTTCTGCAAGAATTCTGCGATA
AGATCGGCACCCAGCTGATCCATTTTGTACCGCGCGACAACATCGTACAAAAAGC
AGAATTTCAGAAAAAAGCAGTGGTCGACTACGATGATACCTGTAATCAGGCACT
GGAATATAAAGAATTGGCTCGTAAGATCATTGAGAACGAAAACCTGGTAATCCC
AACGCCAATGACTATGGACGAGTTAGAAGAGCTGACTTCGAAATATGGATTTCTG
GACGGACGTGCGATAGAGGGTCGTCACCACCACCACCACCAC**TGA**

Optimized *Mt*NifH sequence cloned into pET28a(+):

Green Underlined Bold = Restriction site (NcoI and NdeI)

Bold italic = Start codon

Bold = Stop codon

Black=NifH gene sequence

Blue = Linker

Orange = His Tag

Purple = Thrombin cleavage site

CC***ATG***GGAAGTTTTGATGAAATTGCACCAGATGCTAAAAAAGTAGCAATTTATGG
AAAAGGAGGTATTGGGAAATCAACCACAACCCAAAATACCGCAGCAGCACTTGC
ATACTTTTTCGATAAAAAAGTTATGATTCACGGATGTGACCCAAAAGCAGATTCA
ACAAGAATGATTCTCCACGGAAAACCACAGGATACCGTTATGGATGTACTTAGGG
AAGAGGGGGAAGAGGCGGTTACTCTTGAAAAAGTAAGAAAAATAGGATTTAAAG
ACATATTATGTGTAGAAAGTGGTGGTCCAGAACCTGGTGTAGGATGTGCAGGTAG
AGGGGTTATTACTGCAGTTGATATGATGAGAGAGCTCGAAGGATACCCTGATGAT
TTAGACAACTTGTTCTTCGATGTCCTTGGGGACGTTGTATGCGGTGGTTTCGCTAT
GCCGCTTAGAGACGGTCTTGCACAGGAAATCTACATTGTTACATCAGGGGAGATG
ATGGCATTATATGCCGCAAACAACATTGCAAAAGGTATCTTAAAGTACGCTGAAC
AGTCTGGAGTTAGACTTGGAGGTATAATCTGTAACGCAAGAAATGTCGATGGTGA
AAAAGAGTTAATGGATGAATTCTGCGATAAACTTGGAACCAAATTAATCCACTAC
GTTCCAAGGGACAACATTGTACAGAAAGCTGAATTCAACAAAATGACCGTTATTG
AGTTCGATCCAGAATGTAATCAAGCAAAAGAATACAGAACTCTGGCAAAAAACA
TTGATGAAAATGACGAACTTGTGAAACCAACTCCAATGACTATGGATGAATTGGA
AGAATTAGTTGTAAAATACGGATTAATTGACTTAAGCAGCGGCCTGGTGCCGCGC
GGCAGCAGCAGCGGCCATCATCATCATCATCAC**TGA**CATATG

Optimized *Mi*NifH sequence cloned into pET28a(+):

Green Underlined Bold = Restriction site (NcoI and NdeI)

Bold italic = Start codon

Bold = Stop codon

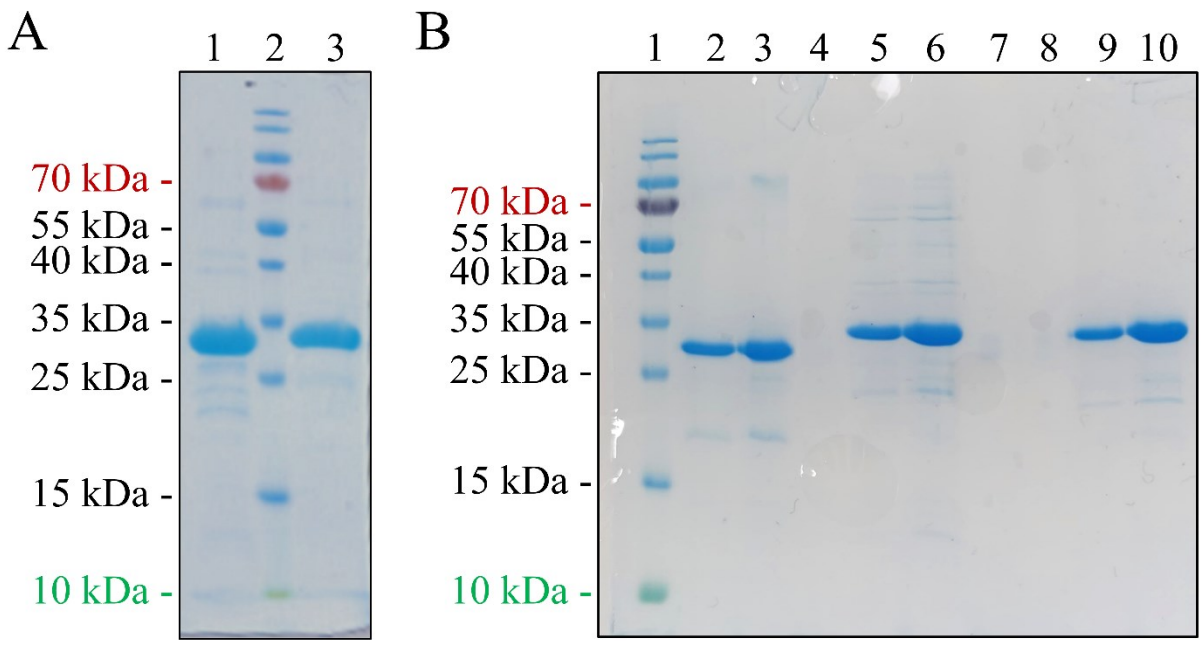
Black=NifH gene sequence

Blue = Linker

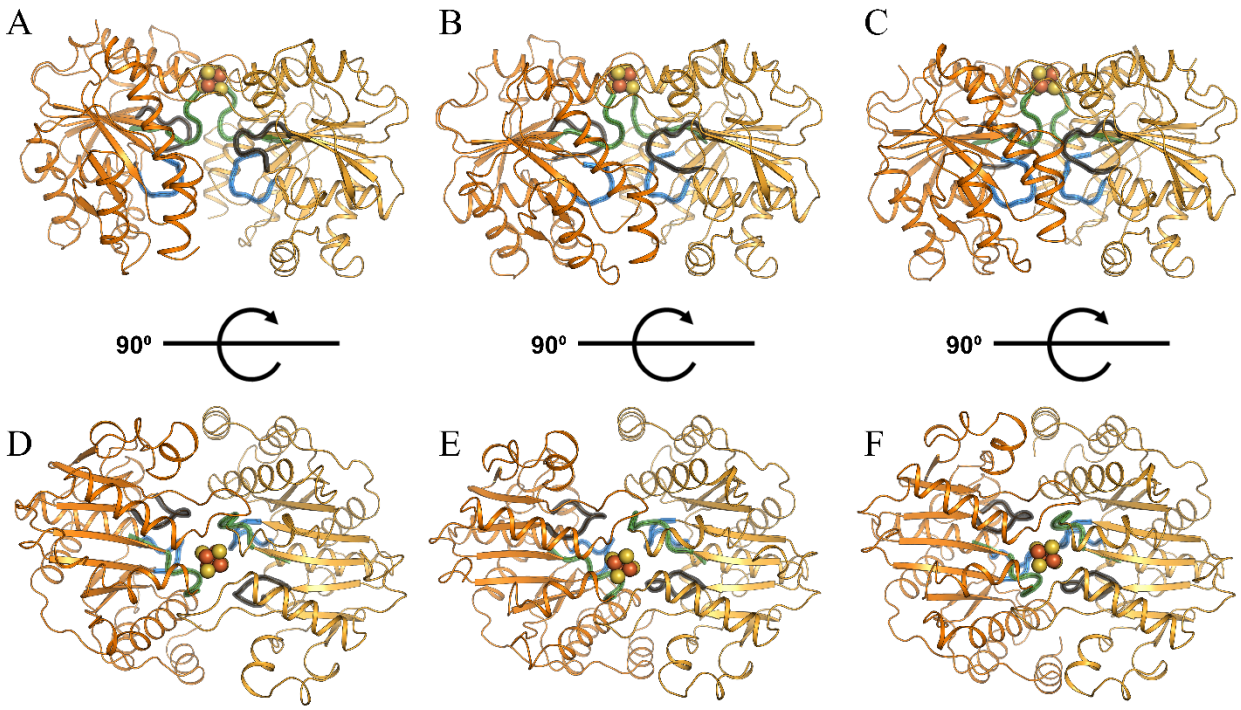
Orange = His Tag

Purple = Thrombin cleavage site

CC***ATG***GGAAGCTTTGACGAAATTGCCCCAAATGCTAAGAAAGTAGCAATTTACGG
AAAAGGAGGGATAGGGAAGAGTACAACTACACAAAATACAGCAGCAGCTTTAGC
ATACTACTACAAATTAAAGGGAATGATTCACGGCTGTGATCCTAAGGCGGATTCA
ACAAGAATGATATTACATGGAAAGCCCCAAGAGACAGTTATGGATGTTTTGAGA
GAAGAAGGAGAAGAAGGAGTTACCTTAGAAAAGGTTAGAAAAGTAGGATTCTGT
GGAATTTATTGTGTTGAATCTGGAGGACCAGAGCCAGGGGTTGGATGTGCAGGA
AGAGGGGTTATAACTGCTGTCAATCTAATGAAAGAGCTTGGTGGATATCCAGATG
ATTTAGACTTTCTCTTCTTTGACGTCCTTGGGGATGTTGTCTGTGGAGGTTTTGCC
ATGCCTCTAAGGGATGGGTTGGCTAAGGAAATTTACATAGTTTCCTCTGGAGAAA
TGATGGCACTTTACGCAGCTAACAACATTGCTAAGGGTATCTTAAAATATGCTGA
GCAGTCAGGAGTTAGATTAGGAGGAATAATTTGTAACTCAAGAAATGTTGATGG
AGAAAGAGAGTTAATGGAAGAGTTCTGTGATAAGTTAGGAACTAAGCTTATTCAC
TTCATTCCAAGAGACAATATAGTTCAGAAGGCTGAGTTTAACAAGATGACAGTAG
TTGAATTTGCTCCAGATCATCCACAGGCTTTAGAGTACAAGAAGTTAGGAAAGAA
GATTATGGACAATGATGAACTTGTAATTCCAACTCCATTAAGCATGGATGAGTTG
GAGAAGTTGGTTGAGAAGTATGGGTTATATGACAAAAGCAGCGGCCTGGTGCCG
CGCGGCAGCAGCAGCGGCCATCATCATCATCATCAC**TGA**CATATG



Supplementary Fig. 1. A. SDS PAGE profile of the purified native *Mt/Mi*NifH. Lanes: 1 = *Mt*NifH, 2 = molecular weight marker, 3 = *Mi*NifH, (5 µg of protein per well). **B.** SDS PAGE profiles of the purified recombinant *Mm*NifH, *Mt*NifH and *Mi*NifH. Lanes: 1 = molecular weight marker, 2 = *Mm*NifH, 2.5 µg of protein in the well, 3 = *Mm*NifH, 5 µg of protein in the well, 4 = empty, 5 = *Mt*NifH, 2.5 µg of protein in the well, 6 = *Mm*NifH, 5 µg of protein in the well, 7 = empty, 8 = empty, 9 = *Mi*NifH, 2.5 µg of protein in the well, 10 = *Mi*NifH, 5 µg of protein in the well.



Supplementary Fig. 2. A. Side view of *Mt*NifH in apo state (form 1, *P* 2_1 space group, dimer 1). **B.** Side view of *Mt*NifH in apo state (form 1, *P* 2_1 space group, dimer 2). **C.** Side view of *Mt*NifH in apo state (form 2, *P* 4_3 2_1 2 space group). **D.** Top view of *Mt*NifH in apo state (form 1, *P* 2_1 space group, dimer 1). **E.** Top view of *Mt*NifH in apo state (form 1, *P* 2_1 space group, dimer 2). **F.** Top view of *Mt*NifH in apo state (form 2, *P* 4_3 2_1 2 space group). Important structural features are colour coded as follows in all panels: Switch I – forest green, Switch II – black, P-loop – marine blue.

References

1. Zheng, L., Cash, V. L., Flint, D. H. & Dean, D. R. (1998) Assembly of iron-sulfur clusters. Identification of an *iscSUA-hscBA-fdx* gene cluster from *Azotobacter vinelandii*, *J Biol Chem.* **273**, 13264-72.

2. McCarty, R. M., Krebs, C. & Bandarian, V. (2013) Spectroscopic, steady-state kinetic, and mechanistic characterization of the radical SAM enzyme QueE, which catalyzes a complex cyclization reaction in the biosynthesis of 7-deazapurines, *Biochemistry.* **52**, 188-98.

3. Schwartz, C. J., Giel, J. L., Patschkowski, T., Luther, C., Ruzicka, F. J., Beinert, H. & Kiley, P. J. (2001) IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins, *Proceedings of the National Academy of Sciences.* **98**, 14895-14900.

4. Roth, L. E. & Tezcan, F. A. (2012) ATP-Uncoupled, Six-Electron Photoreduction of Hydrogen Cyanide to Methane by the Molybdenum–Iron Protein, *Journal of the American Chemical Society.* **134**, 8416-8419.

5. Milton, R. D., Cai, R., Abdellaoui, S., Leech, D., De Lacey, A. L., Pita, M. & Minteer, S. D. (2017) Bioelectrochemical Haber–Bosch Process: An Ammonia-Producing H_2/N_2 Fuel Cell, *Angewandte Chemie - International Edition.* **56**, 2680-2683.

6. Christiansen, J., Goodwin, P. J., Lanzilotta, W. N., Seefeldt, L. C. & Dean, D. R. (1998) Catalytic and Biophysical Properties of a Nitrogenase Apo-MoFe Protein Produced by a *nifB*-Deletion Mutant of *Azotobacter vinelandii*, *Biochemistry.* **37**, 12611-12623.

CHAPTER 4: Molecular basis of nitrogenase regulation by a P_{II} family protein in a hyperthermophilic methanogen

Nevena Maslać[1], Marie-Caroline Müller[1], Pauline Bolte[1], Tristan Wagner[1*]

[1] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

Manuscript in preparation (*presented models are not fully refined*)

Correspondence to Tristan Wagner (*twagner@mpi-bremen.de*)

4.1. Abstract

Nitrogenases are the only enzymes catalysing the N_2 reduction, holding significant potential for biotechnological advancements in NH_3-production. While mesophilic bacterial models were mainly used to decipher the intricacies of nitrogenase reaction, exploring phylogenetically distant organisms could shine a new light on this process. Here, we present near-atomic resolution structures of an archaeal nitrogenase isolated from a hyperthermophilic methanogen. The overall highly conserved protein harbours a molybdenum-containing cofactor. The positions of key residues and the displaced belt sulfide align with the turnover states identified in vanadium and iron-only bacterial nitrogenases, confirming a universal mechanism for N_2reduction. The P^{1+} state of the P cluster, modelled in two alternate conformations, suggests an electron-depleted form waiting to be reloaded by the Fe protein. The nitrogenase in complex with its reductase, stabilised by MgADP-AlF_4^-, displays a similar organisation seen in the bacterial homologue in which the [4Fe-4S] and P clusters are electronically connected. $NifI_{1,2}$ structures in a switch-off state bound to MgATP/2-oxoglutarate and a switch-on state bound to the nitrogenase reveal the regulation mechanism in archaea. The 2($NifI_1$)$NifI_2$ dimer engulfs and sequestrates the nitrogenase by its protruding T-loops, forming a closed-ring structure sterically blocking the Fe protein access. The thermostable and ancient nature of this nitrogenase provides invaluable insights into the mechanism, regulation, and history of the biological fixation apparatus.

4.2. Keywords

Archaeal nitrogenase • turnover state • P_{II} family proteins • post-translational regulation • N_2-fixation

4.3. Introduction

Nitrogen fixation, converting dinitrogen gas (N_2) to ammonia (NH_3), is critical to the biogeochemical nitrogen cycle. Fixed nitrogen forms are necessary to sustain Life on Earth (1). In our modern society, fixed nitrogen can enter the biosphere in three ways: (*i*) geochemically through lightning discharge amounting to ~ 3%; (*ii*) through biological nitrogen fixation by nitrogenase enzymes, contributing ~ 32%; and (*iii*) by the industrial Haber-Bosch process (HB), providing ~65% (1, 2). In more than a century since its invention, the HB has transformed our society by increasing crop yield after the wide use of ammonia-based fertilisers (3-5). While it supports food production for estimated 3.5 billion people (5), the HB has a very high ecological and economic cost (6): from the non-renewable energy consumption, greenhouse gases CO_2, CO, and N_2O emissions to eutrophication and usage of

expensive synthetic catalysts in a highly specialised ammonia production plants at high temperature and pressure (300-550 °C, 150-250 atm). Therefore, developing efficient and cost-effective catalysts inspired by nitrogenase operating under atmospheric conditions is essential, but it represents a major scientific challenge in the last 80 years when the nitrogenase was first isolated.

All nitrogenases are two-component systems that form transitory complexes when active and represent the only enzyme capable of reducing the N_2 triple bond to form NH_3 through an overall eight electron reaction mechanism (7). The first component is the homodimeric Fe protein (referred to as NifH, the *nifH* gene product) responsible for electron delivery through an ATP hydrolysis-controlled mechanism (8-10). The electrons are donated from the [4Fe-4S]-cluster coordinated at the dimer interface. The second catalytic subunit, MoFe protein, harbours both the electron-relaying [8Fe:7S] P cluster and the N_2-reducing active site. Based on the heterometal present at the active site cofactor, three nitrogenase isoforms have been described so far: *(i)* Mo-nitrogenase (referred to as NifDK, the *nifDK* gene product) present in all known diazotrophs, with [Mo:7Fe:9S:C]:homocitrate cluster (8, 10), *(ii)* V-nitrogenase with [V:7Fe:8S:C:CO_3]:homocitrate cluster (11) and *(iii)* Fe-only nitrogenase with [8Fe:9S:C]:homocitrate cluster (12). Although the three nitrogenase isozymes have the same overall mechanism (7, 12-14), they show different reactivity towards both physiological and alternative substrates (9, 15-18).

Thorneley and Lowe postulated that MoFe protein passes through eight different states, representing eight electron transfer steps to allow the complete N_2 reduction (states E_0 to E_7) (2, 8). Each electron transfer step depends on the Fe protein cycle that passes through different phases of nucleotide exchange and [4Fe-4S]-cluster oxidation/reduction (2, 8), adding complexity to the overall mechanism (10). Despite the great progress in understanding the nitrogenases´ mechanistic, supported by decades of structural biology, several proposed mechanistic models of N_2 reduction have been postulated. The key mechanistic feature observed in the so-called "turnover" models is the liability of the belt sulfides (S2B, S3A, and S5A) of the FeMoco. Particularly, the described CO-inhibited structure of the MoFe protein of *Azotobacter vinelandii* presents a displacement of the sulfide S2B (19). In the V- (20) and Fe-only (12) nitrogenases from the same organism, the "turnover" state showed the same displacement of the sulfide accompanied by a reposition of perfectly conserved glutamine and accommodation of the sulfide nearby. These snapshots fit the Thorneley and Lowe model in which the E_2 state involves a bridging hydride displacing the S2B.

Further mechanistic insights require capturing different E*n* states *in crystallo* or by spectroscopy. Structural studies have been limited exclusively to Bacterial representatives of only three different taxonomic classes: *A. vinelandii* (11, 12, 21) and *Klebsiella pneumoniae* (renamed *Klebsiella oxytoca*, NCBI:txid571 (22)) (23) (Gammaproteobacteria), *Clostridium pasteurianum* (Clostridia) (24) and *Gluconacetobacter diazotrophicus* (Alphaproteobacteria) (25). The recently reported structure of iron-only nitrogenase from *A. vinelandii* (12), completing the characterisation of all three known bacterial isoforms, made the lack of structural insights on archaeal nitrogenases even more apparent and marked it as a new emerging frontier in the field.

Despite the shared evolutionary origin of archaeal and bacterial nitrogenases (26-28), diazotrophs in *Archaea* are characterised by multiple unique features, particularly in aspects related to nitrogen metabolism regulation (29). Archaea rely on a post-translational "switch-off" mechanism by P_{II} family proteins $NifI_{1,2}$ (29-32), which is thought to be the ancient hallmark of anaerobic diazotrophs (33). P_{II} proteins are a superfamily of signal transducers regulating central metabolism in all three domains of Life. They integrate signals about cellular energy, nitrogen, and carbon status through sensing 2-oxoglutarate (2-OG) and ATP concentrations (34, 35). Levels of 2-OG, the precursor of glutamate and glutamine, are particularly well conveying the nitrogen availability in the cells. All described P_{II} proteins are homotrimeric and interact with their target proteins through long, flexible extensions named T-loops, which undergo conformational changes upon ligand binding (34-36). Under high-nitrogen availability and low-energy charge, the $NifI_{1,2}$ binding on the nitrogenase disrupts NifH docking, interrupting N_2-fixation and ATP consumption, while high 2-OG and ATP levels unlocked $NifI_{1,2}$ from the nitrogenase and allowed its turnover under nitrogen-limiting conditions. Such a regulatory mechanism is dramatically different from the one found in Bacteria, where the inhibition is achieved by covalent post-translational modification of NifH (29). Intriguingly, it has been reported that $NifI_{1,2}$ can form higher-order complexes both with itself and the nitrogenase (31). However, the exact oligomerization states, the details of its interaction with the nitrogenase provoking the Fe protein binding inhibition remain unknown.

While genetically tractable *Methanococcus maripaludis* and different species of *Methanosarcina* may seem like attractive choices for archaeal models (37, 38), (hyper)thermophilic methanogens with their robust, thermostable and barotolerant nitrogenases (39, 40) might be potentially superior models for a range of industrial applications. Moreover, phylogenetic studies proposed that nitrogenases from

(hyper)thermophilic *Methanococcales* represent an ancient (if not the progenitor) lineage that would have predated the diversification to the three known metal isoforms (26, 41).

By obtaining structural snapshots of an archaeal molybdenum nitrogenase isolated natively from a hyperthermophilic methanogen *Methanocaldococcus infernus* (*Mi*NifDK in further text), this present study offers a complete view of the archaeal system, unravelling different states of the metallocofactors and a comprehensive view of the post-translational regulation process.

4.4. Results

4.4.1. *Mi*NifDK has a compact architecture. *M. infernus* harbours only one nitrogenase isoform in its genome that has a similar identity to each of the three isoforms from *A. vinelandii*. The diazotrophic growth of the archaeon requires a minimum of 0.1 µM Mo for observable growth (Supplementary Figure 1), as previously observed for the thermophile *Methanothermococcus thermolithotrophicus* (42), suggesting that the nitrogenase should contain a FeMoco. $NifI_{1,2}$ was systematically coeluted with NifDK during all purification steps. Therefore, an anionic exchange chromatography in the presence of 2-OG and MgATP was performed to separate $NifI_{1,2}$ and obtain a pure NifDK sample for downstream crystallisation (see Material and Methods and Supplementary Figure 2). Interestingly, the pure NifDK sample contains an additional band on sodium dodecyl sulfate-polyacrylamide gel electrophoresis, which should correspond to NifX, product of *nifX* gene in the operon.

*Mi*NifDK crystallised in a monoclinic ($P2_1$) and a triclinic ($P1$) crystalline form that were refined to 1.37-Å and 1.21-Å, respectively (Supplementary Figure 3 and Table S1). Both crystalline forms contain one biological unit in the asymmetric unit: the typical α2β2 ($NifD_2K_2$) heterotetramer (Figure 1A), identical to the organisation of the bacterial molybdenum nitrogenase. The typical central architecture made of three successive Rossmann folds (βαβ) is conserved in both subunits. Overall, *Mi*NifDK does not contain the large N- and C- terminal extensions observed in *Av*NifDK and *Av*AnfH, making it more compact and comparable to *Av*VnfDK. Partial conservation or unique differences between *Mi*NifD/*Mi*NifK with the other three isoforms from *A. vinelandii* are reported in Supplementary Figure 4. The major differences between *Mi*NifD and its homologues are located in the helix-loop region (residues 187-204) and in a slight extension spanning residues 378-382. For *Mi*NifK, a slight restructuration is visible in segments 121-134 and 179-185, followed by two extensions in 201-205 and 224-232. Interestingly, *Mi*NifDK shares similar subtle structural features in

comparison the three other systems. For instance, some loop positions are identical to the loops of *Av*AnfDK, while others are positioned differently, and the other way around.

4.4.2. Differences in the metallocofactors. The two metal clusters occupy the same highly conserved positions as described for all three characterised bacterial isoforms so far (Figure 1A). While the FeMoco is encompassed by the three βαβ domains of NifD, the P cluster, positioned at the two-fold axis, is bridging the two subunits. We tentatively modelled an Mg^{2+} coordinated at the mononuclear metal binding site (MMB site) in both crystalline forms based on the electron density profile and the absence of any anomalous signal at 7160 eV and 12700 eV. An Mg^{2+} has also been modelled in *Av*VnfDKG and *Av*AnfDKG (11), in contrast to *Av*NifDK in which the MMB site was occupied with Ca^{2+} or Fe^{2+} (43). The two metal clusters were modelled based on the high-resolution electron density map, and the position of the iron atoms was confirmed based on the anomalous signal map obtained from the dataset collected at the Fe-K edge (7120/7160 eV) (Figure 1B, Supplementary Figure 5B, Supplementary Figure 6B, 6D).

In the monoclinic form, the [8Fe:7S] P cluster presents a perfect fit with the previously observed P^{2+} (oxidised state) in which the central sulfur (S1) atom bridging the two [4Fe-4S] is tetracoordinated. In this state, the P cluster is coordinated by two terminal cysteines and one bridging cysteine from each subunit: $Cys20^{K}$, $Cys45^{K}$, $Cys103^{K}$, $Cys46^{D}$, $Cys72^{D,}$ and $Cys134^{D}$ (Figure 1B, Supplementary Figure 5A). The sharp electron density profile of the P cluster in the triclinic form indicates the presence of a P^{1+} state in two alternative conformations with a pentacoordinated S1. In the more populated state (refined to 70% occupancy), the $Ser143^{K}$ binds the Fe6, while Fe5 binds S1. In the second state (refined to 30% occupancy), the $Ser143^{K}$ disengaged from Fe6, which bridges S1, while Fe5 binds the main chain of $Cys72^{D}$ (Figure 1B). Located at the NifD-NifK interface, the P cluster is the key relay site of interprotein electron transfer from the [4Fe-4S] cluster of NifH to the FeMoco. Since the distance between the two clusters does not allow for a direct electron transfer, it has been postulated that perfectly conserved tyrosine and alanine act as mediators (12). The same pair is conserved in *Mi*NifDK as well, namely $Tyr48^{K}$ and $Ala49^{D}$.

The metal signature in the catalytic cofactor, experimentally determined based on the difference map obtained from the anomalous signal map collected at 19900 eV (low remote) and 20100 eV (peak of molybdenum fluorescence), is unequivocally Mo (Supplementary Figure 6). It retains characteristic trigonal prismatic geometry (*D*3h) with a molybdenum atom at the center of a three-fold rotational axis (Figure 1C). Fe2-Fe7 forms the body of the

prism, a hallmark of all described nitrogenase catalytic cofactors. Fe1 is coordinated by a conserved cysteine residue, Cys260^{D} (Figure 1C), while Mo atom is anchored by the homocitrate moiety and the conserved His430^{D} (Figure 1C). The distances between the Mo atom and its axial ligands are well conserved between the two crystalline forms and similar to the atomic resolution structure of *Av*NifDK (12). The central ligand of the prism has been assigned to a carbide based on the electron density, and the two halves of the prism are finally joined through three bridging sulfides (S2B, S3A, S5A) (Figure 1C).

While the monoclinic form reflects the so-called resting state (44), the triclinic form exhibits a mixture of the resting state and the turnover state first described in the VFe-nitrogenase (20) and, recently, in the Fe-only nitrogenase (12). In the resting form, the Fe2 and Fe6 atoms of the FeMoco are bridged by the belt sulfide S2B, while the nearby Gln171^{D} residue points toward the homocitrate moiety, holding the S2B displacement site (Figure 1C, Supplementary Figure 6). However, in the alternative conformation, the side chain of Gln171^{D} rotated toward the metallocluster and displaced the sulfide that can be appropriately modelled (Figure 1C). We tentatively modelled a nitrogen atom as a bridging ligand in the same way as the VFe- and Fe-only nitrogenase models. It must be stated here that the true identity of this bridging ligand will have to be further investigated.

4.4.3. NifI$_{1,2}$ sequesters *Mi*NifDK by forming a supercomplex. The native NifDK- NifI$_{1,2}$ complex isolated from *M. infernus* was obtained by omitting the 2-OG and MgATP during the purification (Supplementary Figure 2). The crystal structure of the complex, refined to 2.76-Å resolution (Table S1), unveiled three dinitrogenases bound tightly together by NifI$_{1,2}$, which was surprising because of its oligomeric organisation. While all P$_{II}$ family proteins so far have been characterised as homotrimers, NifI$_{1,2}$ organizes as a heterotrimer formed by two NifI$_1$ and one NifI$_2$. As previously seen, the trimeric interface is stabilised by a β-sheets core surrounded by α-helices (Supplementary Figure 7). The heterotrimer further dimerises by NifI$_2$ through a helix bundle. Each heterotrimer can independently bind one NifDK, thus the biological unit of the NifI$_{1,2}$ heterohexamer trapped two nitrogenase molecules, similar to an antibody and its target. The final quaternary structure is a closed circular supercomplex (Figure 2A), which maximizes the inhibition efficiency.

NifI$_{1,2}$ heterotrimer binds NifDK with a total area of 2113-Å^2 (detailed in Figure 2B-C), with the majority of contacts between NifI$_2$ and NifDK (1418-Å^2). As already proposed by Dodsworth and Leigh in 2006 (31), the T-loop of NifI$_2$ (residues 36-70) plays a central role in the binding by establishing multiple contacts between NifD and NifK. However, despite its

importance, its residues involved in binding are not conserved, and the substitutions in other microorganisms might have a role in fine tuning the binding efficiency based on their physiological needs.

Most interactions are conserved and guided by charges (Supplementary Figure 8) with non-negligible Van der Waals contributions (Figure 2B). Nitrogenase tetramers interact with each other via a salt bridge beetween $E152^D$ and $K153^D$ and are placed in a way to avoid steric repulsion. None of the *A. vinelandii* dinitrogenase isoforms would clash under such a configuration when overlaid. In other words, neither the N- and C- terminal extensions nor the extra subunit G would prevent the circular arrangement.

4.4.4. The conformational changes of $NifI_{1,2}$ underlying the inhibition mechanism. In order to gain insight into the regulatory switch, we tentatively crystallised the fraction of *Mi*$NifI_{1,2}$ bound to 2-OG/MgATP retrieved after the separation from the dinitrogenase. Despite our efforts, the obtained crystals diffracted to a low resolution of 4-Å. The same strategy was used on the thermophile *M. thermolithotrophicus* grown diazotrophically (42). Here, the highly homologous $NifI_{1,2}$ (*Mt*$NifI_{1,2}$) yielded exploitable crystals in a $P3_221$ space group, leading to a structure refined to 2.73-Å. As previously observed in the nitrogenase complex, *Mt*$NifI_{1,2}$ holo structure is formed of one heterohexamer with the $NifI_2$-$NifI_2$ dimeric contacts, making this organisation a possible universal feature of $NifI_{1,2}$ proteins (Figure 3A).

$NifI_{1,2}$ contains three ligand binding sites located at the interface between the different subunits (Figure 3A). The residues involved in the ligand coordination are similar between $NifI_2$ and $NifI_1$, the adenosine is stabilised through main chain hydrogen bonds, the phosphate backbone is coordinated by ionic bonds and a hydrogen bond from the main chain. The Mg^{2+} cation is coordinated by 2-OG and the ATP and locked by a conserved Glutamine (Gln39 in *Mt*$NifI_1$ and *Mt*$NifI_2$) (Figure 3B-D). The position of the three-ligand binding sites at dimeric-interfaces underscores the importance of communication between them and allows for coordinated conformational change upon ligand binding. As expected, the interaction between 2-OG/MgATP stabilises a critical region at the base of the T-loops restraining flexibility and imposing its lock at a particular angle. When overlaid with the NifDK-$NifI_{1,2}$ complex, it appears obvious that the conformational constraints imposed by the molecular sensors 2-OG and ATP would provoke the dissociation from NifDK by sterical encumbrance (Figure 3E-F and Supplementary Figure 9). In its apo state, T-loops protruding from the protein core would be free to engage in contact with its target by adequately changing their positions and imposing the inhibition of the N_2-fixation.

4.4.5. The dynamic docking site of NifH overlays with NifI$_{1,2}$ on NifDK. The NifDK-NifI$_{1,2}$ structure establishes a rationale for the inhibition mode as it seems to result from a physical hindrance of the NifH binding. If it is the case, NifI$_{1,2}$ and NifH will share the same binding interface on NifDK. To further investigate this hypothesis and characterize the complete nitrogenase system, we reconstituted the native *Mi*NifDKH complex. Since the two nitrogenase subunits are only transiently associated during catalysis and have to be artificially stabilised to allow structural characterisation (45), MgADP-AlF_4^- was added during reconstitution. The latter mimics the transitory physiological state of the NifH-NifDK complex before the ATP hydrolysis and is known to rigidify the interaction between both partners (46) (see Material and Methods). The obtained crystals of the MgADP-AlF_4^- stabilised *Mi*NifDKH complex belongs to the $P2_1$ space group, and the structure was refined to a resolution of 2.44 Å. In this complex, *Mi*NifDK is in an identical conformation as its resting state model, as evidenced by the root mean square deviation of 0.299 Å for all atoms. The overall structure fits relatively well with the bacterial homologue (Supplementary Figure 10), including the two-fold symmetry axis of the Fe protein aligned almost to the pseudo-two-fold axis of NifDK, allowing an optimal docking position for electron transfer (Figure 4A). The B-factor profile indicates a high vibration of NifH on NifDK, which hampers precise modelling of one of the sites and corroborates the natural flexibility of the system. The distance between the [4Fe-4S]-cluster of NifH and the P cluster of NifDK is 17.6-Å, which is in the same range as observed in *A. vinelandii* (10, 45). In such a configuration, the interaction area between *Mi*NifH and *Mi*NifDK scores 1809-Å^2. This corresponds to half of the buried surface found in the *Av*NifDKH complex (*i.e.*, 3400-3600 Å^2) and could be a trait of the hyperthermophilic nature in *M. infernus*. A closer inspection of the contact area undoubtedly argues that NifI$_{1,2}$ and NifH compete for the same site (Figure 4B), hitherto explaining the molecular basis of the nitrogenase regulation in *Methanococcales*.

4.5. Discussion

Throughout evolution, the genetic makeup and regulatory mechanisms of biological N_2-fixation have changed substantially. Such changes were primarily influenced by environmental conditions such as the availability of metals and the expansion of organisms into oxic habitats (2, 33). With the most minimalistic N_2-fixation operon, phylogenetic traits, and structural features, *M. infernus* nitrogenase can be regarded as closely related to ancestral forms. Furthermore, the molybdenum nitrogenase (the primordial isoform) evolved within hydrogenotrophic methanogens (26, 28). With hydrogenotrophic methanogenesis being one of the lowest energy-yielding reactions in known metabolism (47) and nitrogen fixation one

of the energetically most expensive ones (13), it may seem paradoxical. However, the inherent ingenuity of this system lies in its simplicity and elegant mechanisms developed to precisely control every aspect of energy expenditure.

Our work confirms and clarifies the physiological role of the higher oligomerization states of $NifI_{1,2}$ alone and in complex with the dinitrogenase as observed previously (31). Instead of forming a heterodimer of $(NifI_1)3$-$(NifI_2)3$, heterotrimers that dimerize through $NifI_2$ are formed, offering a double-headed inhibitor able to build a closed circular supercomplex. Due to its dynamic nature and limited binding area on NifDK, NifH will not be capable of competing with $NifI_{1,2}$ for the docking site, therefore, avoiding an ATP consumption and loss of reducing the power that could be distributed to other metabolic networks. It is much more energetically efficient to regulate the formation of complex conducive to electron transfer through nucleotide recycling and utilise the existing NifH pool than through possible degradation and *ab initio* synthesis of NifH. The physical sequestration of the dinitrogenase might go beyond inhibition: it might also serve as a protection mechanism to extend NifDK lifetime and, most importantly, protect its precious metallocofactors that require a high amount of cellular energy to be produced.

Once the cellular nitrogen level becomes critical, but the energy charge is sufficient, 2-OG and ATP bind to $NifI_{1,2}$, promoting the switch-off (Figure 4C). We propose that the switch-off is translated by a steric repulsion provoked by a readjustment of the T-loops upon ligand binding and should act as observed in the case of the P_{II} family ammonia transporters (34, 36).

Once liberated, NifDK can engage in N_2-fixation together with its reductase. Three distinct NifH docking geometries reflecting different turnover stages have been previously identified, depending on the NifH nucleotide state (45) (10). Since the MgADP-AlF_4^- stabilised *Mi*NifDK-NifH complex fits the DG2 state observed in *A. vinelandii*, we suspect that the Fe protein would follow a similar trajectory (Supplementary Figure 10, Table S2, S3). A comprehensive description of the electron transfer between the two nitrogenase subunits is central to understanding the overall catalytic mechanism because of its multiple unique features. According to the proposed conformational gated electron transfer model (48), structural remodelling would be coupled to the differences in the P cluster state and coordination. The “deficit-spending” model (49-51) predicts that the initial electron transfer happens from the P cluster to the FeMo co because, functionally, P cluster exists in the P^N-reduced resting state (52), and no further reduction has been reported. P cluster consequentially has a “deficit” of one electron and now exists in a P^{1+} state. The second

electron transfer from the [4Fe-4S] cluster is repaying this deficit, bringing back the P cluster to the P^N state. While the P cluster can be further oxidised to P^{2+} (as observed in the monoclinic structure) and P^{3+} states (53, 54), it is accepted that the P^N/P^{1+} couple is the one occurring during the physiological catalytic cycle. Until recently, only P^N resting state and the P^{2+} oxidised state were observed in structural models (53), since the P^{1+} state is difficult to observe during turnover, as the rates of the second electron transfer can be greater than 1700 s^{-1}. Thus, the P^{1+} state depicted at the atomic level in the first characterised archaeal nitrogenase described here is particularly important. The second site of interest, the FeMoco site, presents analogous movements of the close-by Glutamine and S2B to the ones observed in *Av*VnfDK (20) and *Av*AnfDK (12). It confirms the structural flexibility of the bridging sulfide as the universal key characteristic of the nitrogenase catalytic cofactors and suggests a universal mechanism of reductive catalysis for all nitrogenases from the *Bacteria* and *Archaea* kingdoms. This new model from a hyperthermophile has the potential to open the route for the future characterisation of transient turnover states, and completing the puzzle of one of nature's most enigmatic catalytic cycles.

4.6. Material and methods

4.6.1. Growth conditions of *M. infernus* and *M. thermolithotrophicus*, medium composition, and adaptation to diazotrophic conditions. *Mi*NifDK, *Mi*NifDKI$_{1,2}$ complex, and *Mi*NifH were purified natively from *M. infernus* DSM 11812 (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany). *Mt*NifI$_{1,2}$ was purified natively from diazotrophic *M. thermolithotrophicus* DSM 2095 (Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany).

The growth medium was prepared according to Maslać *et al.* 2022, with the substitution of Na_2SeO_4 with Na_2SeO_3 x $5H_2O$ (2 µM, increase of the Na_2WO_4 x H_2O concentration to 100 µM, the addition of 10 mM Na_2SO_4 and a pH adjustment to 6.5 in case of *M. infernus* cultures. Adaptation to diazotrophic conditions, growth, harvesting, and storage of the cells were handled under an anaerobic atmosphere as described in Maslać *et al.* 2022.

4.6.2. Native purification of *Mi*NifDK, *Mi*NifH, and *Mi*NifDKI$_{1,2}$ complex. *M. infernus* cell pellets (in the case of batch 1 from which *Mi*NifDK in the resting state was obtained ~25 g of wet cells obtained from 5.5 l of culture, in the case of batch 2 from which *Mi*NifDKH and *Mi*NifDKI$_{1,2}$ complexes were obtained: 27.21 g of wet cells obtained from 5.5 l of culture) were thawed at room temperature and sonicated inside the anaerobic chamber (N_2/CO_2 with a 90:10% ratio) in Ion Exchange Buffer (IEX; 50 mM Tris/HCl pH 8.0, 2 mM

dithiothreitol (DTT)). Cells were sonicated with the KE76 probe (Bandelin SONOPULS, Sigma, Berlin, Germany) at 75% intensity in the same anaerobic chamber. The lysate was centrifuged at 45 000 x *g* for 60 minutes at 25 ºC and transferred to an anaerobic Coy tent (N_2/H_2 with a 97:3% ratio). The whole procedure was performed under yellow light at 20 °C. The supernatant was filtered through a 0.2 µm filter (Sartorius, Göttingen, Germany) and loaded on a 10 ml HiTrap Q-Sepharose High-Performance column (QSeph, Cytiva, Freiburg, Germany) previously equilibrated with IEX Buffer. The sample was eluted over a 190-550 mM NaCl linear gradient for 6 column volumes (CVs) at a 1.5 ml/min flow rate. *Mi*NifDK with bound *Mi*$NifI_{1,2}$ was obtained between 456.5 and 550 mM NaCl. *Mi*NifH pool was obtained between 346.4 and 440.7 mM NaCl.

Mi$NifDKI_{1,2}$ pool was diluted with three volumes of hydrophobic exchange buffer (HIC, 25 mM Tris/HCl pH 8.0, 2 M $(NH_4)_2SO_4$ and 2 mM DTT)) and loaded on a 1.7 ml Source 15 PHE (phenyl) column (Cytiva, Freiburg, Germany). *Mi*$NifDKI_{1,2}$ was eluted with 1-0.3 M $(NH_4)_2SO_4$ linear gradient for 23.5 column volumes (CV) at a 1 ml/min flow rate. Under these conditions the protein eluted between 0.79 and 0.51 mM $(NH_4)_2SO_4$.

One-third of this pool was loaded on the QSeph column and eluted with a pH gradient from 50 mM Tris/HCl pH 8.0 to 500 mM sodium malate pH 5.4 (buffed with KOH pellets, Merck, Darmstadt, Germany) in the presence of 2 mM DTT. The sample was diluted with three volumes of IEX buffer prior to loading. In this case, the pure *Mi*$NifDKI_{1,2}$ has eluted between 219.85 – 272.2 mM malate.

The other two-thirds of the *Mi*$NifDKI_{1,2}$ pool obtained from Source 15 PHE was diluted with three volumes of IEX buffer containing 2 mM ATP, 2 mM $MgCl_2$ and 10 mM 2-OG and loaded on the QSeph equilibrated with the same ligand-containing buffer after a minimum one hour incubation. The proteins were eluted with a linear gradient of 100-500 mM NaCl. *Mi*NifDK eluted between 121.2 and 194.8 mM NaCl, while $NifI_{1,2}$ eluted between 267.0 and 336.2 mM NaCl.

*Mi*NifH pool obtained from the initial QSeph was diluted with three volumes of HIC buffer and injected into a 5 mL Phenyl Sepharose High-Performance column (Cytiva, Freiburg, Germany). The elution was done with a linear gradient from 1.64-0.7 M $(NH_4)_2SO_4$ over 8 column volumes, with *Mi*NifH eluting between 0.92 and 0.7 M $(NH_4)_2SO_4$. Finally, the obtained pool was passed through the Qseph with a Tris-malate gradient as described above for *Mi*NifDK. *Mi*NifH eluted from 163.6 – 186.2 mM malate.

Isolated *Mi*NifH was incubated for 20 minutes with *Mi*NifDK in the following buffer: 25 mM Tris/HCl pH 8.0, 20 mM NaF, 1 mM $AlCl_3$, 1 mM ADP, and 2 mM $MgCl_2$. This sample was loaded on the QSeph equilibrated with the same buffer and eluted with a gradient from 100-600 mM NaCl. *Mi*NifDKH eluted from 109.10 – 377.20 mM NaCl.

4.6.3. Purification of native *Mt*NifI$_{1,2}$. Pellets of *M. thermolithotrophicus* grown under N_2-fixation were thawed and sonicated as described above. The lysate was centrifuged at 75 600 x *g* for 1 h at 18 ºC and transferred to an anaerobic Coy tent (N_2/H_2 with a 97:3% ratio). *Mt*NifI$_{1,2}$ was purified in four steps. After filtering through a 0.2 µm filter (Sartorius, Göttingen, Germany), the whole supernatant was loaded on a 20 ml HiTrap DEAE Sepharose FF column (Cytiva, Freiburg, Germany) equilibrated with IEX buffer. The sample was eluted in a gradient from 150-500 mM NaCl over 9 column volumes. *Mt*NifI$_{1,2}$ coeluted with *Mt*NifDK in the range of 276.6 – 352.5 mM NaCl. The obtained pool was again diluted with three volumes of IEX buffer and loaded on the QSeph in the presence of MgATP and 2-OG in the same conditions as described above (2 mM $MgCl_2$, 2 mM ATP, 10 mM 2-OG). The protein pool from which NifI$_{1,2}$ crystallized still contained NifDK based on the SDS-PAGE profile and was obtained between 387.6 and 431.8 mM NaCl from a 200-600 mM NaCl gradient over 9 column volumes. It was further diluted with three volumes of HIC buffer and loaded on Phenyl Sepharose High-Performance column, from which it eluted once the concentration of $(NH_4)_2SO_4$ reached 0 mM after 800 – 0 mM gradient done over 12 column volumes. Finally, the sample was injected on Superdex 200 10/300 GL, from which it eluted after 9 ml. Based on the SDS-PAGE profile the obtained pool was labeled *NifDKX*. However, transparent NifI$_{1,2}$ crystalls were obtained from it, most likely due to the elution overlap during the purification.

4.6.4. Crystallisation. Prior to crystallisation, all samples were concentrated by centrifugation in VivaSpin amicons (Sartorius, Göttingen, Germany). Concentrators with 3 kDa, 10 kDa, and 30 kDa cutoffs were used for NifI$_{1,2}$, NifH, and NifDK/NifDKI$_{1,2}$/NifDKH, respectively. Samples were further centrifuged at 13 000 x *g* for 3 min to remove macro-aggregates and dust and crystallised inside an anaerobic chamber (N_2/H_2 with a 97:3% ratio at 20°C). Crystallisation was done by sitting drop method in 96-Well MRC 2-Drop polystyrene Crystallisation Plates (SWISSCI) plate containing 90 µl of crystallisation solution in the reservoir in all cases. *Mi*NifDK monoclinic and triclinic forms were crystallised at a final concentration of 2.95 and 22.4 mg/ml, respectively. Both were obtained by mixing 0.7 µl of protein sample with 0.7 µl of 30 % v/v 2-Methyl-2,4-pentanediol, 100 M Tris pH 8.5, 500

mM sodium chloride, and 8 % w/v polyethylene glycol (PEG) 8 000. The monoclinic form has an elongated plate morphology, while the triclinic are rather thin orthorhombic rods.

*Mi*NifDKI$_{1,2}$ was crystallised at a final concentration of 26.7 mg/ml by spotting 0.6 µl of crystallisation solution with 0.6 µl of protein sample. The crystallisation solution contained 0.2 M sodium chloride, 2 M ammonium sulfate, and 0.1 M sodium cacodylate pH 6.5. Crystal shape was hexagonal rods. Prior to freezing to liquid nitrogen, the crystals were soaked in the crystallisation solution supplemented with 30 % v/v glycerol for a few seconds.

*Mt*NifI$_{1,2}$ holo was crystallised from a sample at a final concentration of 2.0 mg/ml by spotting 0.7 µl of crystallisation solution with 0.7 µl of protein sample. Crystals were obtained in the solution containing 50 % v/v Ethylene glycol, 100 mM Tris pH 8.5, and 200 mM Magnesium chloride. The morphology was bipyramids.

*Mi*NifDKH was crystallised at a final concentration of 17.18 mg/ml with 20 mM NaF, 1 mM $AlCl_3$, 1 mM ADP, and 2 mM $MgCl_2$ by spotting 0.52 µl of crystallisation solution with 0.52 µl of protein sample. Crystals were obtained in the solution containing 0.2 M magnesium acetate tetrahydrate and 20 % w/v PEG 3 350 (SG1 Screen, Molecular Dimensions, Rotherham, United Kingdom). Prior to freezing to liquid nitrogen, the crystals were soaked in the crystallisation solution supplemented with 20 % v/v glycerol for a few seconds.

4.6.5. Data processing and refinement. Data were collected at 100 K at the Swiss Light Source (beamlines PXI and PXIII, Villigen, Switzerland) and European Synchrotron Radiation Facility (beamline FIP2, Grenoble, France) (see Table S1) and were integrated with *autoPROC* (55). All models were solved by molecular replacement using *Phaser* from *PHENIX*. *Av*NifDK at atomic resolution (PDB 3U7Q) was used as a template to solve *Mi*NifDK. GlnK$_2$ holo from *M. thermolithotrophicus* (PDB 7P50) was used as a template to solve *Mi*NifI$_{1,2}$. For *Mi*NifDKI$_{1,2}$, the refined structures of *Mi*NifDK and *Mi*NifI$_{1,2}$ were used as templates. Finally, *Mi*NifDKH was solved by using *Mi*NifDK as a template and the structure of *Mi*NifH (described in **Chapter 3** of this thesis). All models were manually optimised with *COOT* (56). Refinement was performed with PHENIX.refine (57) or Buster (58) without applying non-crystallography symmetry and using a translation-libration screw for all except both *Mi*NifDK models. For the two structures of *Mi*NifDK, the refinement was performed with all atoms considered anisotropic. All models were refined with hydrogens in riding positions and were validated by the Molprobity tool integrated with *PHENIX*.

4.7. Acknowledgments. We thank the Max Planck Institute for Marine Microbiology for continuous support. We also thank Christina Probian and Ramona Appel for their technical assistance in the Microbial Metabolism laboratory. We thank the Swiss light Source Synchrotron and the staff of beamlines PXI and PXIII. We would like to give special thanks to the FIP2 beamline at the European Synchrotron Radiation Facility, and Dr. Sylvain Engilberge for his help during anomalous data collection.

4.8. Funding. This study was funded by the Max Planck Society.

4.9. Author contributions. N.M. and T.W. designed the research. N.M. and P. B. cultivated *M. infernus*, purified and crystallised the proteins. N.M. cultivated *M. thermolithotrophicus*, purified and crystallised the proteins. X-ray data were collected by N.M. and T.W. Data processing, model building, and structure refinement of *Mt*NifI$_{1,2}$ structure were performed by M-C.M. Data processing, model building, structure refinement, and analysis of all other presented structures were performed by N. M. and T.W. T.W. acquired funding to realize the project. The paper was written by N. M. and T.W. with contributions and final approval of all co-authors.

4.10. References

1. Stevens CJ. 2019. Nitrogen in the environment. Science 363:578-580.

2. Martin del Campo JS, Rigsbee J, Bueno Batista M, Mus F, Rubio LM, Einsle O, Peters JW, Dixon R, Dean DR, Dos Santos PC. 2023. Overview of physiological, biochemical, and regulatory aspects of nitrogen fixation in *Azotobacter vinelandii*. Critical Reviews in Biochemistry and Molecular Biology 57:492-538.

3. Erisman JW, Sutton MA, Galloway J, Klimont Z, Winiwarter W. 2008. How a century of ammonia synthesis changed the world. Nature Geoscience 1:636-639.

4. Galloway JN, Cowling EB. 2002. Reactive Nitrogen and The World: 200 Years of Change. AMBIO: A Journal of the Human Environment 31:64-71, 8.

5. Smil V. 2004. Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production. MIT press.

6. Sutton MA, Howard CM, Kanter DR, Lassaletta L, Móring A, Raghuram N, Read N. 2021. The nitrogen decade: mobilizing global action on nitrogen to 2030 and beyond. One Earth 4:10-14.

7. Jasniewski AJ, Lee CC, Ribbe MW, Ribbe MW, Hu Y. 2020. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. Chemical Reviews 120:5107-5157.

8. Einsle O, Rees DC. 2020. Structural Enzymology of Nitrogenase Enzymes. Chemical Reviews 120:4969-5004.

9. Seefeldt LC, Yang ZY, Lukoyanov DA, Harris DF, Dean DR, Raugei S, Hoffman BM. 2020. Reduction of Substrates by Nitrogenases. Chemical Reviews 120:5082-5106.

10. Rutledge HL, Tezcan FA. 2020. Electron Transfer in Nitrogenase. Chemical Reviews 120:5158-5193.

11. Sippel D, Einsle O. 2017. The structure of vanadium nitrogenase reveals an unusual bridging ligand. Nature Chemical Biology 13:956-960.

12. Trncik C, Detemple F, Einsle O. 2023. Iron-only Fe-nitrogenase underscores common catalytic principles in biological nitrogen fixation. Nature Catalysis doi:10.1038/s41929-023-00952-1.

13. Harris DF, Lukoyanov DA, Kallas H, Trncik C, Yang Z-Y, Compton P, Kelleher N, Einsle O, Dean DR, Hoffman BM, Seefeldt LC. 2019. Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism To Achieve N_2 Reduction. Biochemistry 58:3293-3301.

14. Eady RR. 1996. Structure-function relationships of alternative nitrogenases. Chemical Reviews 96:3013-3030.

15. Harris DF, Yang Z-Y, Dean DR, Seefeldt LC, Hoffman BM. 2018. Kinetic understanding of N_2 reduction versus H_2 evolution at the E_4 (4H) Janus state in the three nitrogenases. Biochemistry 57:5706-5714.

16. Zheng Y, Harris DF, Yu Z, Fu Y, Poudel S, Ledbetter RN, Fixen KR, Yang ZY, Boyd ES, Lidstrom ME, Seefeldt LC, Harwood CS. 2018. A pathway for biological methane production using bacterial iron-only nitrogenase. Nature Microbiology 3:281-286.

17. Fixen KR, Zheng Y, Harris DF, Shaw S, Yang Z-Y, Dean DR, Seefeldt LC, Harwood CS. 2016. Light-driven carbon dioxide reduction to methane by nitrogenase in a photosynthetic bacterium. Proceedings of the National Academy of Sciences 113:10163-10167.

18. Lee CC, Tanifuji K, Newcomb M, Liedtke J, Hu Y, Ribbe MW. 2018. A comparative analysis of the CO-reducing activities of MoFe proteins containing Mo-and V-Nitrogenase cofactors. ChemBioChem 19:649-653.

19. Spatzal T, Perez KA, Einsle O, Howard JB, Rees DC. 2014. Ligand binding to the FeMo-cofactor: Structures of CO-bound and reactivated nitrogenase. Science 345:1620-1623.

20. Sippel D, Rohde M, Netzer J, Trncik C, Gies J, Grunau K, Djurdjevic I, Decamps L, Andrade SLA, Einsle O. 2018. A bound reaction intermediate sheds light on the mechanism of nitrogenase. Science 359:1484-1489.

21. Kirn J, Rees D. 1992. Crystallographic structure and functional implications of the nitrogenase molybdenum–iron protein from *Azotobacter vinelandii*. Nature 360:553-560.

22. Shin SH, Kim S, Kim JY, Lee S, Um Y, Oh M-K, Kim Y-R, Lee J, Yang K-S. 2012. Complete genome sequence of *Klebsiella oxytoca* KCTC 1686, used in production of 2, 3-Butanediol. Journal of Bacteriology 194:2371 - 2372.

23. Mayer SM, Lawson DM, Gormal CA, Roe SM, Smith BE. 1999. New insights into structure-function relationships in nitrogenase: a 1.6 Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe-protein. Journal of Molecular Biology 292:871-891.

24. Kim J, Woo D, Rees DC. 1993. X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0-.ANG. resolution. Biochemistry 32:7104-7115.

25. Owens CP, Katz FEH, Carter CH, Oswald VF, Tezcan FA. 2016. Tyrosine-Coordinated P-Cluster in *G. diazotrophicus* Nitrogenase: Evidence for the Importance of O-Based Ligands in Conformationally Gated Electron Transfer. Journal of the American Chemical Society 138:10124-10127.

26. Garcia AK, McShea H, Kolaczkowski B, Kaçar B. 2020. Reconstructing the evolutionary history of nitrogenases: Evidence for ancestral molybdenum-cofactor utilization. Geobiology 18:394-411.

27. Boyd ES, Peters JW. 2013. New insights into the evolutionary history of biological nitrogen fixation. Front Microbiol 4:1-12.

28. Boyd ES, Hamilton TL, Peters JW. 2011. An alternative path for the evolution of biological nitrogen fixation. Frontiers in Microbiology 2:1-11.

29. Leigh JA, Dodsworth JA. 2007. Nitrogen Regulation in Bacteria and Archaea. Annual Review of Microbiology 61:349-377.

30. Dodsworth JA, Cady NC, Leigh JA. 2005. 2-Oxoglutarate and the PII homologues $NifI_1$ and $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. Molecular Microbiology 56:1527-1538.

31. Dodsworth JA, Leigh JA. 2006. Regulation of nitrogenase by 2-oxoglutarate-reversible, direct binding of a PII-like nitrogen sensor protein to dinitrogenase. Proceedings of the National Academy of Sciences of the United States of America 103:9779-9784.

32. Kessler PS, Daniel C, Leigh JA. 2001. Ammonia Switch-Off of Nitrogen Fixation in the Methanogenic Archaeon *Methanococcus maripaludis*: Mechanistic Features and Requirement for the novel GlnB homologues, $NifI_1$ and $NifI_2$. Journal of Bacteriology 183:882-889.

33. Boyd ES, Garcia Costas AM, Hamilton TL, Mus F, Peters JW. 2015. Evolution of Molybdenum Nitrogenase during the Transition from Anaerobic to Aerobic metabolism. Journal of Bacteriology 197:1690-1699.

34. Forchhammer K, Selim KA, Huergo LF. 2022. New views on PII signaling: from nitrogen sensing to global metabolic control. Trends in Microbiology 30:722-735.

35. Huergo LF, Chandra G, Merrick M. 2013. P(II) signal transduction proteins: nitrogen regulation and beyond. FEMS Microbiology Reviews 37:251-83.

36. Müller MC, Wagner T. 2021. The Oxoglutarate Binding Site and Regulatory Mechanism Are Conserved in Ammonium Transporter Inhibitors GlnKs from *Methanococcales*. Int J Mol Sci 22:8631.

37. Contreras G, Thomsen J, Pfitzer M, Hafenbradl D, Kostner D, Holtmann D, Schmitz RA, Rother M, Molitor B. 2022. New perspectives for biotechnological applications of methanogens. Current Research in Biotechnology 4:468-474.

38. Enzmann F, Mayer F, Rother M, Holtmann D. 2018. Methanogens: biochemical background and biotechnological applications. AMB Express 8:1.

39. Mehta MP, Baross JA. 2006. Nitrogen Fixation at 92 °C by a Hydrothermal Vent Archaeon. Science 314:1783-1786.

40. Belay N, Sparling R, Daniels L. 1984. Dinitrogen fixation by a thermophilic methanogenic bacterium. Nature 312:286-288.

41. Boyd ES, Anbar AD, Miller S, Hamilton TL, Lavin M, Peters JW. 2011. A late methanogen origin for molybdenum-dependent nitrogenase. Geobiology 9:221-232.

42. Maslać N, Sidhu C, Teeling H, Wagner T. 2022. Comparative Transcriptomics Sheds Light on Remodeling of Gene Expression during Diazotrophy in the Thermophilic Methanogen *Methanothermococcus thermolithotrophicus*. mBio 13:e02443-22.

43. Einsle O, Tezcan FA, Andrade SL, Schmid B, Yoshida M, Howard JB, Rees DC. 2002. Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. Science 297:1696-1700.

44. Kang W, Lee CC, Jasniewski AJ, Ribbe MW, Hu Y. 2020. Structural evidence for a dynamic metallocofactor during N2 reduction by Mo-nitrogenase. Science 368:1381-1385.

45. Tezcan FA, Kaiser JT, Mustafi D, Walton MY, Howard JB, Rees DC. 2005. Nitrogenase complexes: multiple docking sites for a nucleotide switch protein. Science 309:1377-1380.

46. Renner KA, Howard JB. 1996. Aluminum Fluoride Inhibition of Nitrogenase: Stabilization of a Nucleotide·Fe-Protein·MoFe-Protein Complex. Biochemistry 35:5353-5358.

47. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nature Reviews Microbiology 6:579-591.

48. Rees DC, Akif Tezcan F, Haynes CA, Walton MY, Andrade S, Einsle O, Howard JB. 2005. Structural basis of biological nitrogen fixation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 363:971-984.

49. Danyal K, Dean DR, Hoffman BM, Seefeldt LC. 2011. Electron transfer within nitrogenase: evidence for a deficit-spending mechanism. Biochemistry 50:9255-9263.

50. Lanzilotta WN, Seefeldt LC. 1997. Changes in the midpoint potentials of the nitrogenase metal centers as a result of iron protein-molybdenum-iron protein complex formation. Biochemistry 36:12976-12983.

51. Howard JB, Rees DC. 1994. Nitrogenase: a nucleotide-dependent molecular switch. Annual Review of Biochemistry 63:235-264.

52. Zimmermann R, Münck E, Brill WJ, Shah VK, Henzl MT, Rawlings J, Orme-Johnson WH. 1978. Nitrogenase X: Mössbauer and EPR studies on reversibly oxidized MoFe protein from *Azotobacter vinelandii* OP. Nature of the iron centers. Biochimica et Biophysica Acta (BBA)-Protein Structure 537:185-207.

53. Keable SM, Zadvornyy OA, Johnson LE, Ginovska B, Rasmussen AJ, Danyal K, Eilers BJ, Prussia GA, LeVan AX, Raugei S. 2018. Structural characterization of the P^{1+} intermediate state of the P-cluster of nitrogenase. Journal of Biological Chemistry 293:9629-9635.

54. Pierik AJ, Wassink H, Haaker H, Hagen WR. 1993. Redox properties and EPR spectroscopy of the P clusters of *Azotobacter vinelandii* MoFe protein. European Journal of Biochemistry 212:51-61.

55. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. Acta Crystallographica Section D 67:293-302.

56. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallographica Section D: Biological Crystallography 66:486-501.

57. Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, Moriarty NW, Oeffner RD, Poon BK, Prisant MG, Read RJ, Richardson JS, Richardson DC, Sammito MD, Sobolev OV, Stockwell DH, Terwilliger TC, Urzhumtsev AG, Videau LL, Williams CJ, Adams PD. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol 75:861-877.

58. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart O, Vonrhein C. 2011. BUSTER version 2.10. 0. Global Phasing Ltd, Cambridge, UK.

4.11. Figures and legends

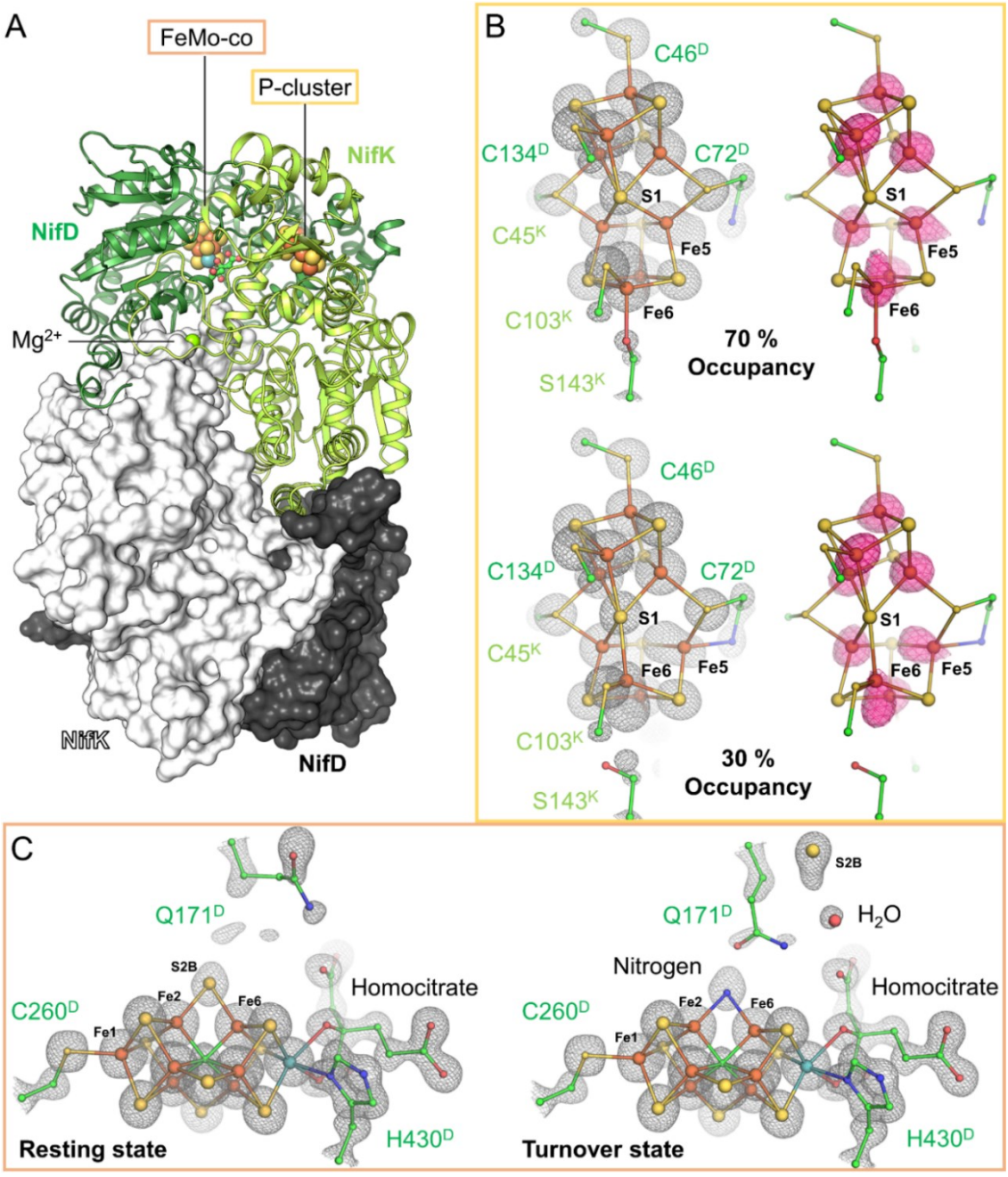


Fig. 1. Architecture of *Mi*NifDK molybdenum nitrogenase subunits and metallocofactors. A. Dimers of *Mi*NifDK presented as surface and cartoon each. The P clusters are located on the pseudo-two-fold axis between NifD and NifK, while the FeMoco is located between the Rossmann folds of NifD. **B** Architecture and coordination of the two P cluster conformations in the P^{1+} (oxidised) state. The anomalous map signal was obtained by data collection at 12.7 keV, contoured at 6-σ (pink mesh) around the P cluster. **C.** The architecture and environment of FeMoco in two alternative conformations: the resting and turnover states in which the side chain of $Gln171^{D}$ is rotated, leading to the displacement of S2B. Nitrogen was tentatively modelled between the Fe2 and Fe6 as in *Av*VnfDK and *Av*AnfDK models. Grey mesh represents the $2F_o$-F_c electron density map contoured at 2.5 σ in all panels. Metallocofactors are displayed as spheres and colored by atom: orange (Fe), yellow (S), cyan (Mo), red (O), and green (C).

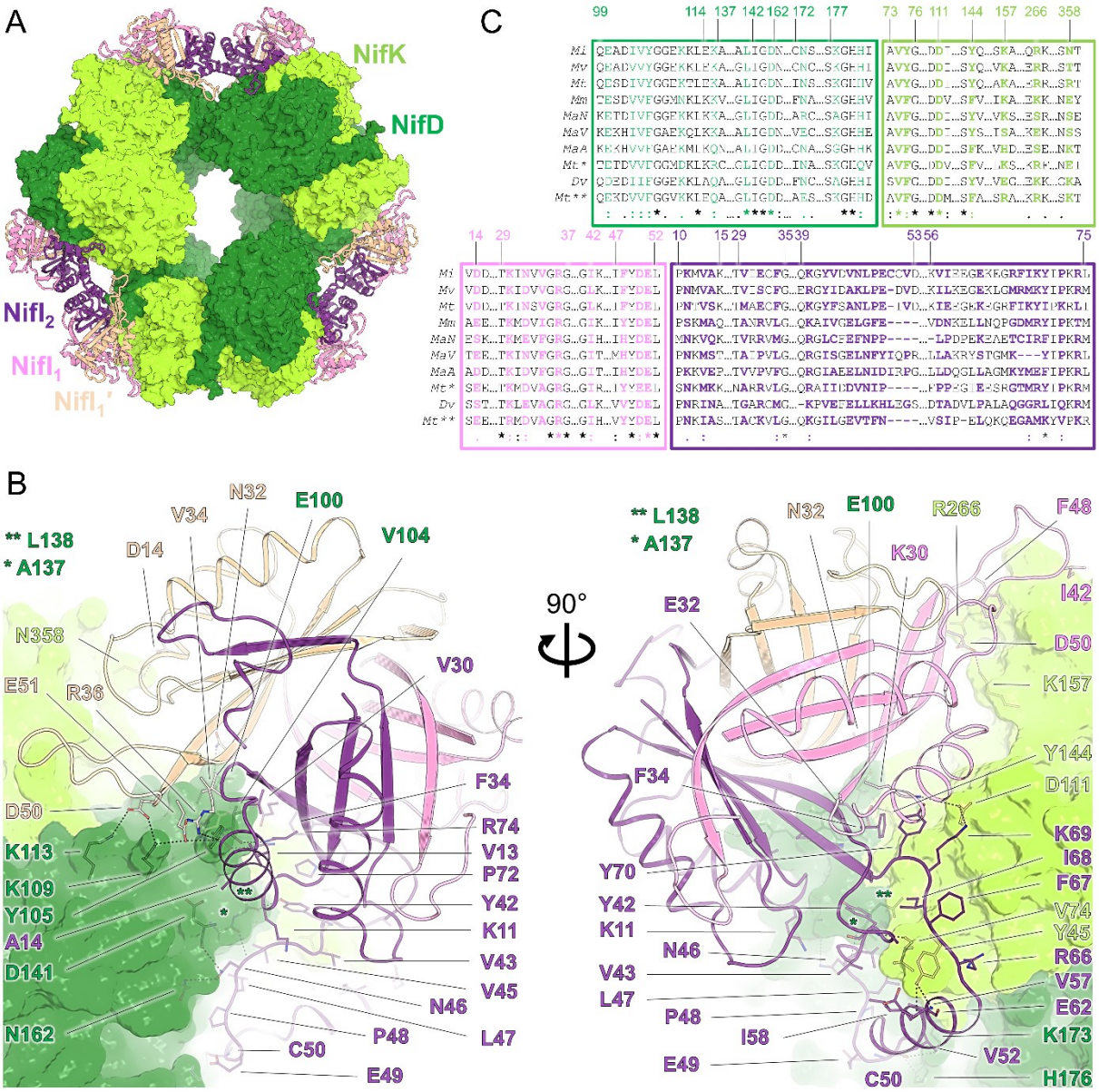


Fig. 2. The structure and binding interfaces of *Mi*NifDKI$_{1,2}$ complex. A. *Mi*NifDK are represented as surface, and *Mi*NifI$_{1,2}$ are represented as cartoons. **B.** The contact interface between *Mi*NifDK and *Mi*NifI$_{1,2}$. Residues creating the bonds are represented as sticks and color-coded as follows: blue (N), red (O), carbons are colored according to the chain with the same code as in panel A. **C.** Conservation of the residues forming the contact interface across selected microorganisms. *Mi*, *M. infernus*; *Mv*, *Methanocaldococcus vulcanius*; *Mt*, *M. thermolithotrophicus; Mm, Methanococcus maripaludis; MaN, Methanosarcina acetivorans nif; MaV, Methanosarcina acetivorans vnf; Methanosarcina acetivorans anf; Mt*, Methanothermobacter thermautotrophicus; Dv, Desulfovibrio vulgaris; Mt**, Moorella thermoacetica*

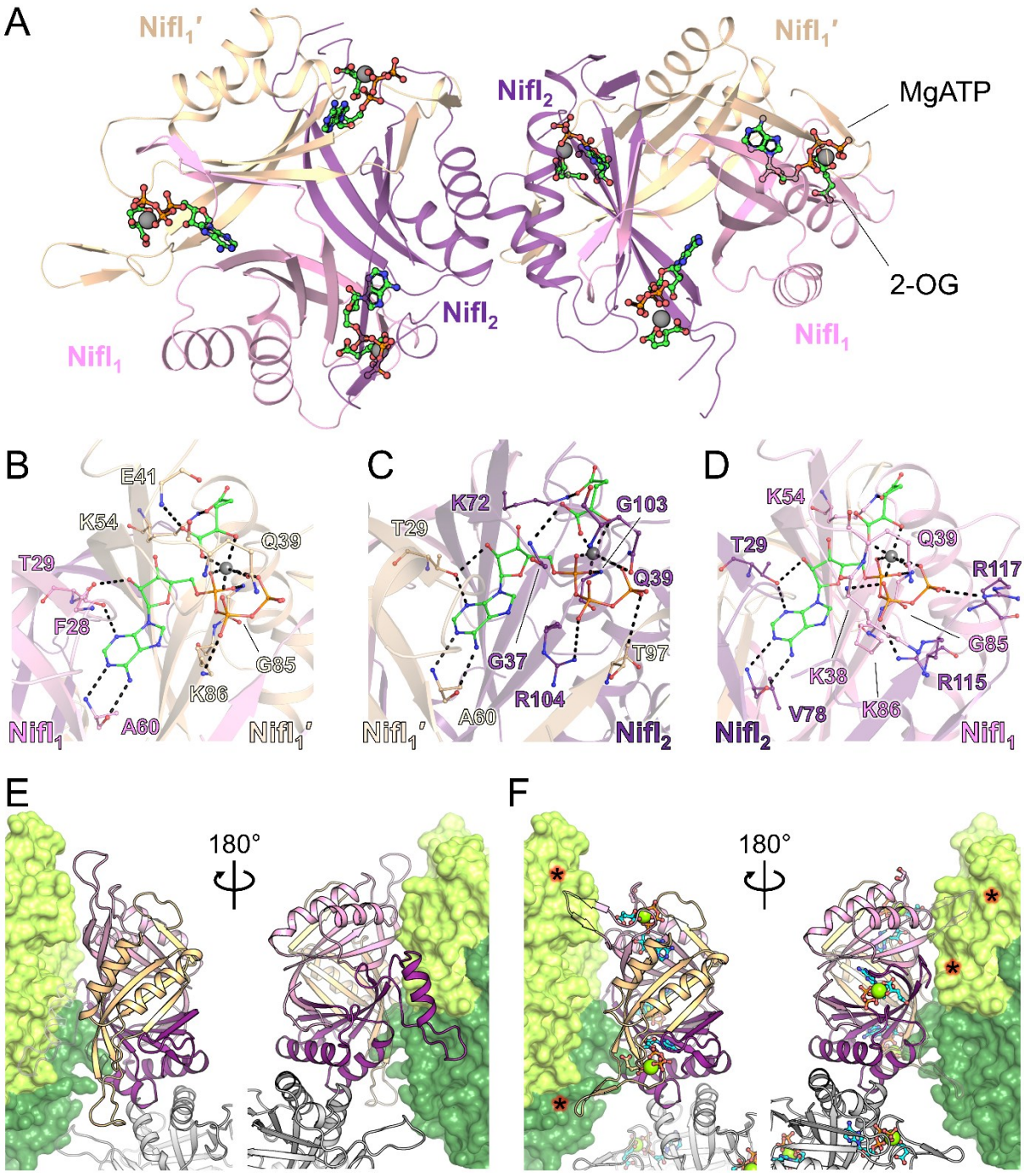


Fig. 3. The mechanism of the molecular switch in NifI$_{1,2}$ inhibitory complex. A. Cartoon representation of *Mt*NifI$_{1,2}$ heterohexamer with bound 2-OG and MgATP. **B-D.** Ligand binding sites and coordination in the NifI$_{1,2}$ heterohexamer at NifI$_1$/$_1$′ (B), NifI$_1$′/$_2$ (C), and NifI$_2$/$_1$ interfaces. **E.** Side views of *Mi*NifI$_{1,2}$ (in a cartoon) bound on *Mi*NifDK surface. **F**. Side view of superposed *Mt*NifI$_{1,2}$ holo on *Mi*NifDK surface. The superposition of *Mt*NifI$_{1,2}$ has been done only on the NifI$_2$ from *Mi*NifI$_{1,2}$ bound on the dinitrogenase. The black stars, highlighted by a red glow, indicate clashes between the T-loops and NifDK surface. In all panels, ligands and interacting residues are represented as ball and stick models and colored by atom: blue (N), green/cyan (C), red (O), orange (P), and light green (Mg).

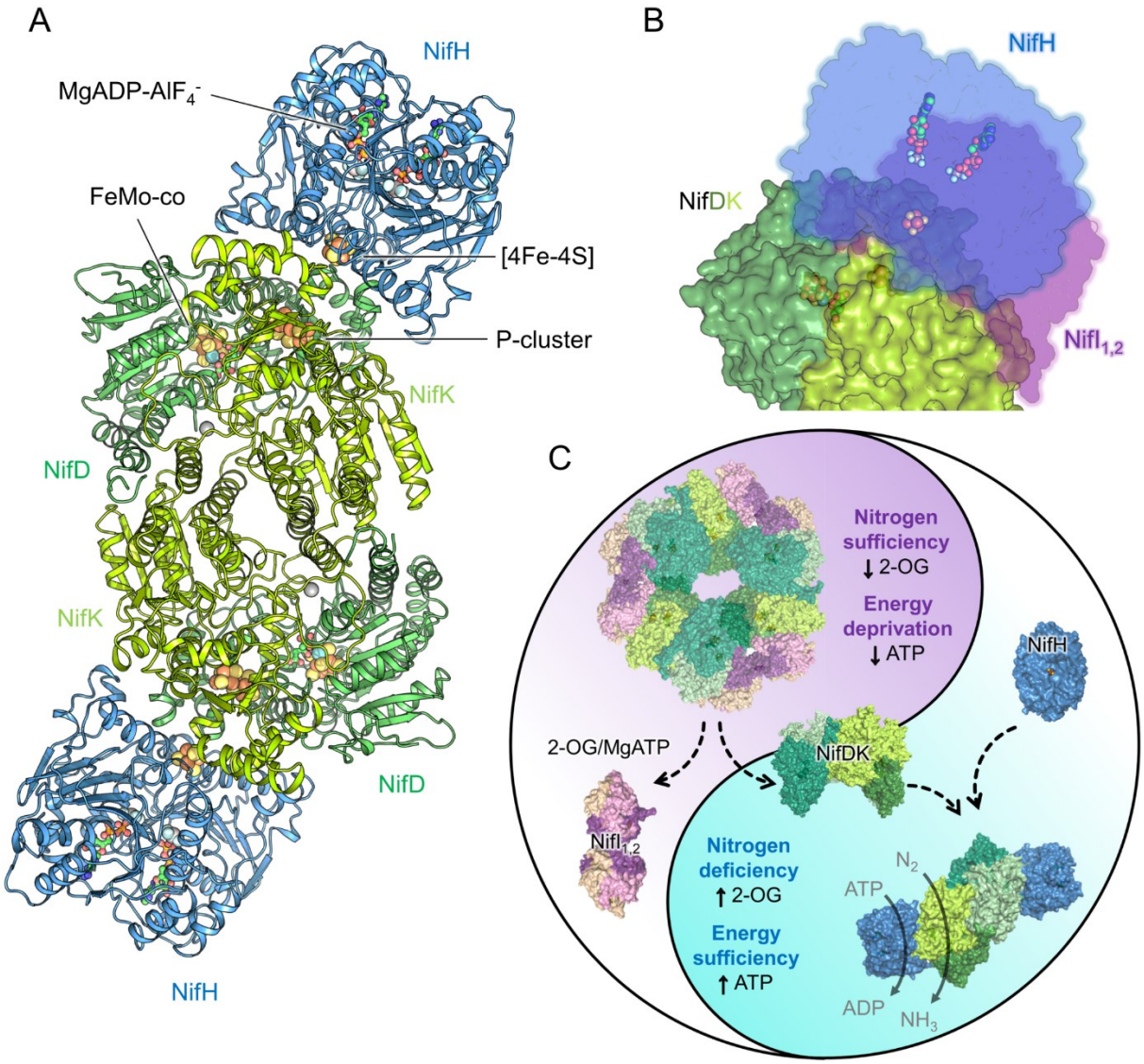


Fig. 4. The structure of MgADP-AlF$_4^-$ bound *Mi*NifDKH and NifI$_{1,2}$ operating mode. A. The structure of the MgADP-AlF$_4^-$ stabilised *Mi*NifDKH complex is represented as a cartoon. All metallocofactors are labeled once for clarity, represented as spheres and colored by atom: orange (Fe), yellow (S), blue (N), green (C), red (O), dark orange (P), grey (Al), light blue (F), cyan (Mo). **B.** Binding surfaces of NifH (blue) and NifI$_{1,2}$ mapped on the surface of *Mi*NifDK. **C.** Scheme of nitrogenase inhibition and activity in *M. infernus*. When cellular nitrogen supply is adequate, the activity of nitrogenase is inhibited by NifI$_{1,2}$ through the closed circular complex. When the cellular nitrogen levels drop, the concentration of 2-OG rises, and together with ATP, it binds to NifI$_{1,2}$, releasing the inhibition. Once NifI$_{1,2}$ dissociates, NifDK can bind NifH, forming an active nitrogenase complex that can reduce N_2 to NH_3 in an ATP-dependent way.

4.12. Supplemental Material

Supplementary Table 1. Data collection and refinement statistics for the presented structures.

Supplementary Table 2. Comparison of different NifH nucleotide-bound states.

Supplementary Table 3. Conservation of the MgADP-AlF_4^- NifH docking surface.

Supplementary Figure 1. Growth of diazotrophic *M. infenus*.

Supplementary Figure 2. SDS-PAGE of purified *M. infernus* and *M. thermolithotrophicus* proteins.

Supplementary Figure 3. Differences in crystal packing between the two *Mi*NifDK forms.

Supplementary Figure 4. Overlay between *Azotobacter vinelandii* alpha and beta nitrogenase subunits with *Mi*NifD and *Mi*NifK.

Supplementary Figure 5. Architecture and coordination of the P cluster from the *Mi*NifDK monoclinic form.

Supplementary Figure 6. The architecture and environment of FeMoco from the *Mi*NifDK in the monoclinic form.

Supplementary Figure 7. Conservation of secondary structure in the P_{II} superfamily of proteins.

Supplementary Figure 8. Electrostatic interaction between *Mi*NifDK and *Mi*$NifI_{1,2}$.

Supplementary Figure 9. Differences between the apo and holo state of $NifI_{1,2}$.

Supplementary Figure 10. Comparison of the docking geometry in MgADP-AlF_4^- stabilised nitrogenase complexes of *M. infernus* and *A. vinelandii*.

Supplementary Table 1. Data collection and refinement statistics. Refinement is in progress.

	***Mi*NifDK monoclinic form**	***Mi*NifDK mixed state**
Data collection		
Synchrotron, beamline	SLS, PXI	ESRF, FIP2
Wavelength (Å)	1.00000	0.97951
Space group	$P2_1$	$P1$
Resolution (Å)	78.85 – 1.37 (1.52 – 1.37)	101.86 – 1.21 (1.35 – 1.21)
Cell dimensions		
a, b, c (Å)	78.43, 117.18, 106.63	68.50, 80.24, 105.74
α, β, γ (°)	90, 91.48, 90	74.49, 82.25, 65.02
R_{merge}(%)[a]	9.1 (119.8)	7.4 (107.1)
R_{pim} (%)[a]	2.9 (40.4)	3.0 (42.5)
$CC_{1/2}$[a]	0.998 (0.658)	0.999 (0.669)
I/σ_I[a]	12.3 (1.9)	12.9 (1.7)
Spherical completeness[a]	68.0 (13.0)	62.1 (11.0)
Ellipsoidal completeness[a]	95.3 (68.2)	92.0 (74.5)
Redundancy[a]	10.4 (9.6)	7.1 (7.3)
Nr. unique reflections[a]	273,586 (13,675)	371,908 (18,595)
Refinement		
Resolution (Å)	22.31 – 1.37	33.62– 1.21
Number of reflections	273,461	371,066
R_{work}/R_{free}[b] (%)	14.13/16.80	12.62/15.13
Number of molecules/AU	2NifD 2NifK	2NifD 2NifK
Number of atoms		
Protein	14,965	15,112
Ligands/ions	101	155
Solvent	1392	1869
Mean B-value (Å^2)	23.69	17.49
Molprobity clash score, all atoms	5.73	4.20
Ramachandran plot		
Favoured regions (%)	96.93	97.04
Outlier regions (%)	0	0
rmsd[c] bond lengths (Å)	0.015	0.016
rmsd[c] bond angles (°)	1.387	1.426

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Supplementary Table 1. Data collection and refinement statistics. Refinement is in progress.

	***Mi*NifDK monoclinic form Fe K edge peak**	***Mi*NifDK triclinic form Mo low remote**	***Mi*NifDK triclinic form Mo K edge peak**
Data collection			
Synchrotron source	SLS, PXI	ESRF, FIP2	ESRF, FIP2
Wavelength (Å)	1.73160	0.62304	0.61684
Space group	$P2_1$	$P1$	$P1$
Resolution (Å)	78.94 – 1.84 (2.03– 1.84)	102.00 – 1.81 (1.84 – 1.81)	70.73 – 2.00 (2.03 – 2.00)
Cell dimensions			
a, b, c (Å)	78.55, 117.58, 106.56	68.60, 80.35, 105.89	68.54, 80.27, 105.80
α, β, γ (°)	90, 91.60, 90	74.49, 82.25, 65.02	74.50, 82.24, 65.00
R_{merge}(%)[a]	11.7 (76.0)	22.2 (223.6)	28.1 (233.2)
R_{pim} (%)[a]	5.4 (41.8)	13.1 (131.6)	16.5 (136.9)
$CC_{1/2}$[a]	0.996 (0.724)	0.980 (0.352)	0.979 (0.325)
I/σ_I[a]	8.9 (1.7)	3.5 (0.5)	2.2 (0.4)
Spherical completeness[a]	74.0 (14.2)	93.0 (94.1)	92.5 (92.9)
Ellipsoidal completeness[a]	93.3 (57.1)	/	/
Redundancy[a]	5.4 (4.3)	3.8 (3.8)	3.8 (3.8)
Nr. unique reflections[a]	124,152 (6,209)	168,457 (8,470)	122,976 (6,176)

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Supplementary Table 1. Data collection and refinement statistics. Refinement is in progress.

	***Mi*NifDKI$_{1,2}$**	***Mi*NifDKH**	***Mt*NifI$_{1,2}$ holo**
Data collection			
Synchrotron, beamline	ESRF, FIPII	SLS, PXI	SLS, PXI
Wavelength (Å)	0.97980	1.00000	1.00000
Space group	$P6_3$	$P2_1$	$P3_221$
Resolution (Å)	253.79 – 2.76 (2.81– 2.76)	91.03 – 2.44 (2.70 – 2.44)	46.51 – 2.73
Cell dimensions			
a, b, c (Å)	293.05, 293.05, 108.86	79.95, 100.19, 221.03	72.87, 72.87, 344.13
α, β, γ (°)	90, 90, 120	90, 99.63, 90	90, 90, 120
R_{merge}(%)[a]	19.4 (753.0)	44.7 (297.0)	/
R_{pim} (%)[a]	2.7 (104.2)	15.7 (102.9)	/
$CC_{1/2}$[a]	1.000 (0.377)	0.980 (0.358)	/
I/σ_I[a]	22.9 (0.7)	3.8 (1.6)	/
Spherical completeness[a]	100 (100)	40.1 (7.7)	/
Ellipsoidal completeness[a]	/	91.1 (67.7)	/
Redundancy[a]	50.7 (52.4)	8.6 (8.4)	/
Nr. unique reflections[a]	136,698 (6,799)	51,381 (2,570)	/
Refinement			
Resolution (Å)	30.80 – 2.76	59.33 – 2.44	46.51 – 2.73
Number of reflections	136,510	51,308	29,440
R_{work}/R_{free}[b] (%)	17.78/21.08	24.49/30.38	22.50/25.19
Number of molecules/AU	2NifI$_2$ 4NifI$_1$ 2NifD 2NifK	4NifH 2NifD 2NifK	2NifI$_2$ 4NifI$_1$
Number of atoms			
Protein	20040	23385	271
Ligands/ions	96	248	42
Solvent	/	/	39
Mean B-value (Å^2)	107.43	42.65	119.90
Molprobity clash score, all atoms	13.47	23.33	4.53
Ramachandran plot			
Favoured regions (%)	95.18	93.46	99.03
Outlier regions (%)	0.44	0.71	0
rmsd[c] bond lengths (Å)	0.006	0.005	0.010
rmsd[c] bond angles (°)	0.904	0.855	1.170

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

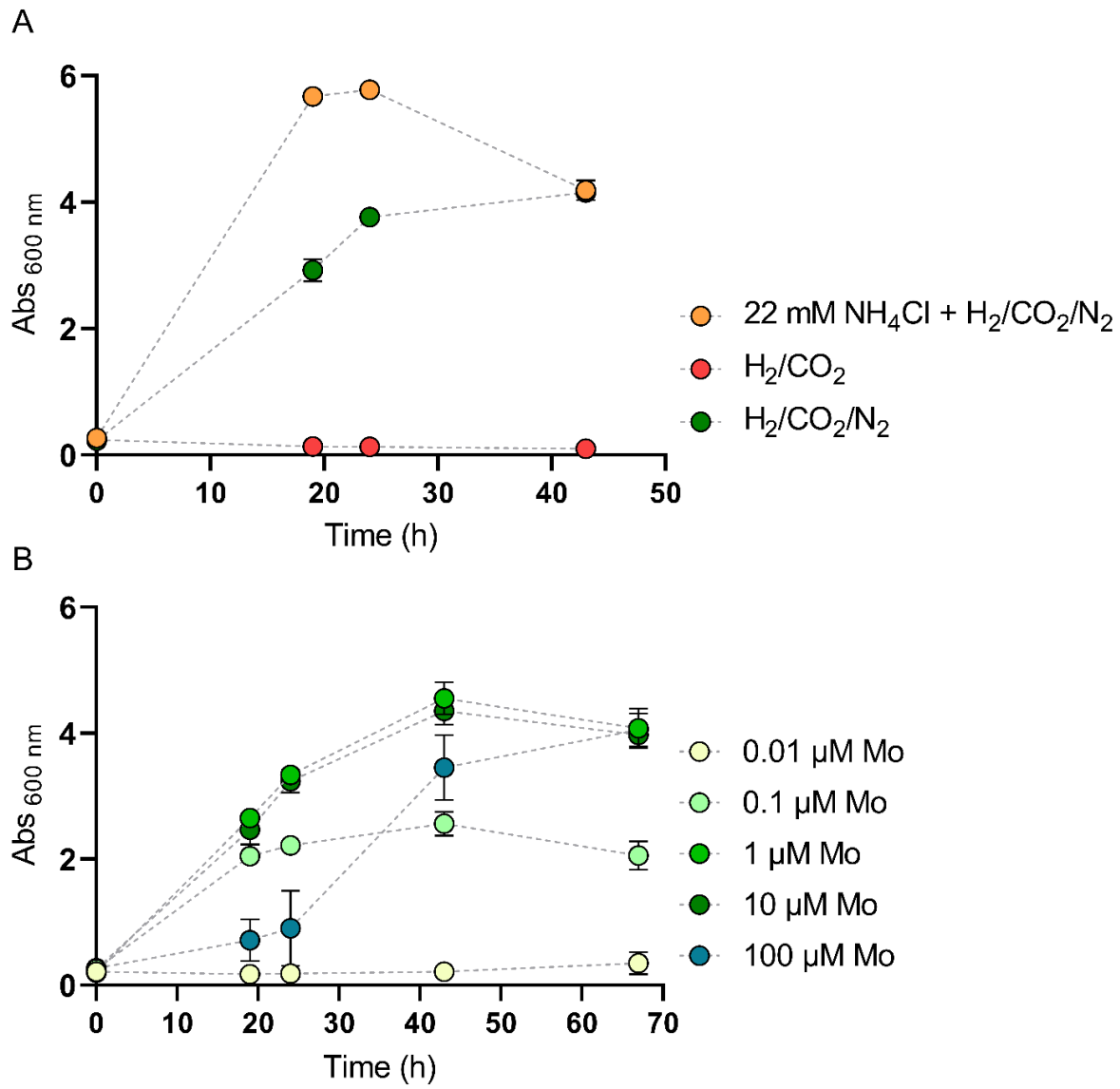
For *Mt*NifI$_{1,2}$ holo, we worked on the scaled data provided by our synchrotron colleagues. The raw data are on tape and need time to be recovered but will be reanalysed for the final submitted manuscript.

Supplementary Table 2. Comparison of different NifH nucleotide-bound states. Values in the table represent root mean square deviation (RMSD) for all atoms of aligned structures in Å. The values in bold represent the models closest to the MgADP-AlF_4^- *Mi*NifH.

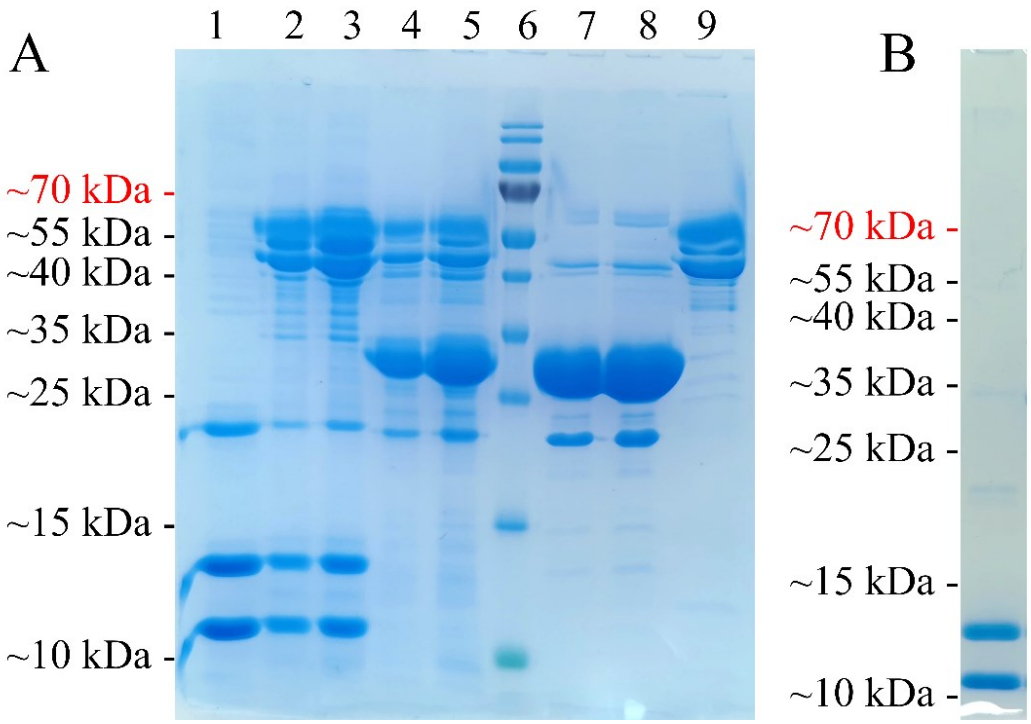
Structure	**MgADP-AlF_4^- *Mi*NifH in complex with *Mi*NifDK**	
free *Av*NifH	**Chains GH**	**Chains FI**
native *Av*NifH (1G5P)	3.890	3.987
MgADP *Av*NifH (1FP6) Chains AB	3.061	3.135
MgADP *Av*NifH (1FP6) Chains CD	3.054	3.135
***Av*NifH in complex with *Av*NifDK**		
MgADP-AlF_4 *Av*NifH (1M34) Chains GH	**0.632**	**0.640**
MgADP-AlF_4 *Av*NifH (1M34) Chains EF	**0.630**	**0.626**
MgADP-AlF_4 *Av*NifH (1M34) Chains MN	**0.625**	**0.637**
MgADP-AlF_4 *Av*NifH (1M34) Chains OP	**0.621**	**0.635**
nucleotide free *Av*NifH (2AFH) Chains EF	3.912	3.981
Mg-AMPPCP *Av*NifH (4WZB) Chains EF	2.015	2.044
Mg-AMPPCP *Av*NifH (4WZB) Chains GH	1.861	1.906
MgADP *Av*NifH (2AFI) Chains EF	2.902	2.986
MgADP *Av*NifH (2AFI) Chains GH	3.112	3.168
MgADP *Av*NifH (2AFI) Chains MN	3.792	3.870
MgADP *Av*NifH (2AFI) Chains OP	3.264	3.310

Supplementary Table 3. Conservation of the MgADP-AlF_4^- NifH docking surface. The *Av*NifH residues making the contacts with *Av*NifDK in MgADP-AlF_4^- stabilised complex (PDB 1M34) are listed in bold. Corresponding residues in *Mi*NifH are shown in green if they are making the contact with *Mi*NifDK in the MgADP-AlF_4^- stabilised complex shown in this work, and in red if they are not.

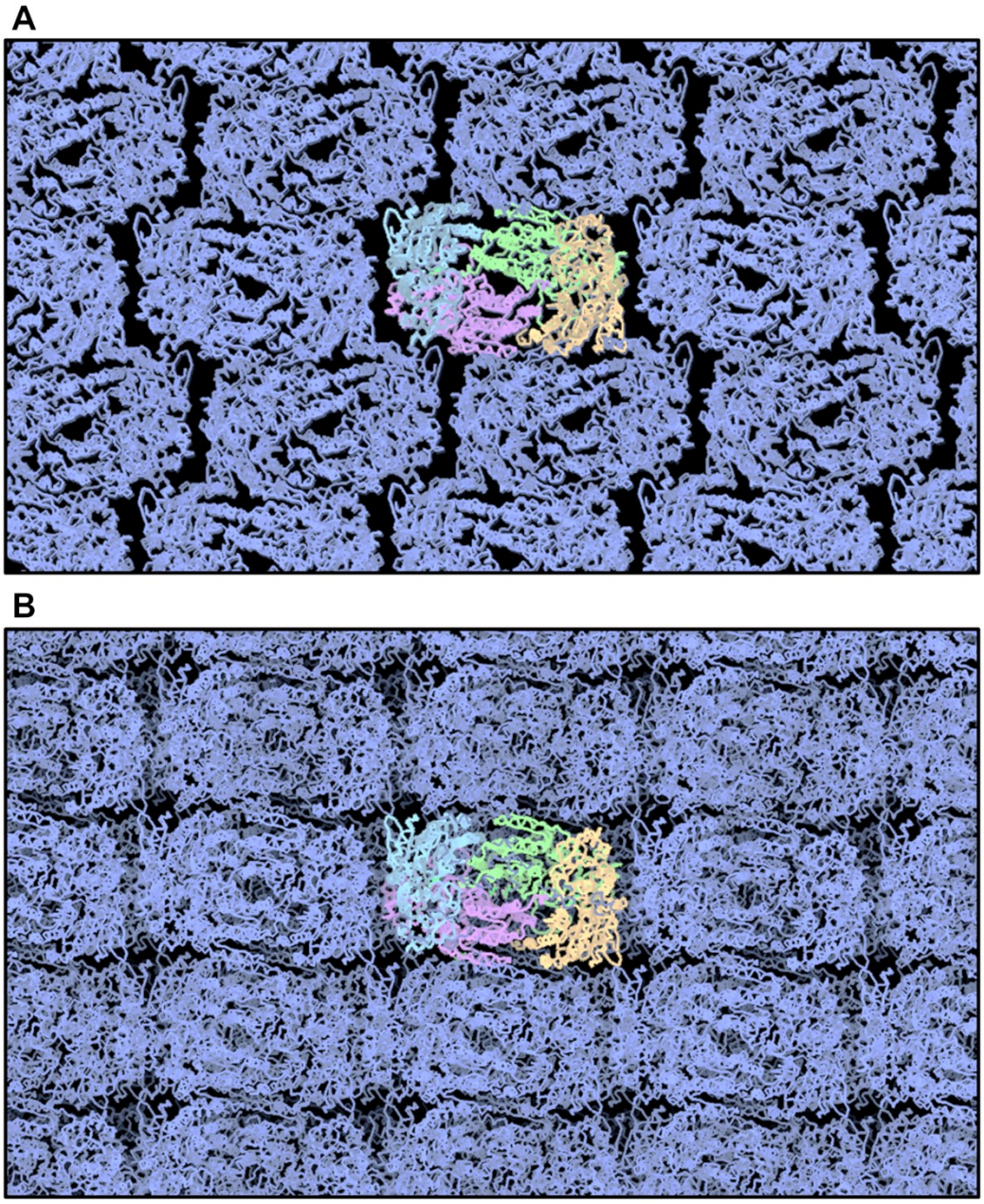
MgADP-AlF_4 *Av*NifH (1M34) Chain γ1	**MgADP-AlF_4 *Av*NifH (1M34) Chain γ2**	**MgADP-AlF_4 *Mi*NifH Chain G**	**MgADP-AlF_4 *Mi*NifH Chain H**	**MgADP-AlF_4 *Mi*NifH Chain F**	**MgADP-AlF_4 *Mi*NifH Chain I**
	Gly 66	Gly 74	Gly 74	Gly 74	Gly 74
Asp 70		Gly 77	Gly 77	Gly 77	Gly 77
Cys 98	**Cys 98**	Cys 105	Cys 105	Cys 105	Cys 105
Arg 101	**Arg 101**	Arg 108	Arg 108	Arg 108	Arg 108
Thr 105	**Thr 105**	Thr 112	Thr 112	Thr 112	Thr 112
Gly 134	**Gly 134**	Gly 141	Gly 141	Gly 141	Gly 141
Arg 141	**Arg 141**	Arg 148	Arg 148	Arg 148	Arg 148
Glu 142	**Glu 142**	Asp 149	Asp 149	Asp 149	Asp 149
Lys 171	**Lys 171**	Lys 178	Lys 178	Lys 178	Lys 178
Asn 174		Glu 181	Glu 181	Glu 181	Glu 181
Ser 175	**Ser 175**	Gln 182	Gln 182	Gln 182	Gln 182



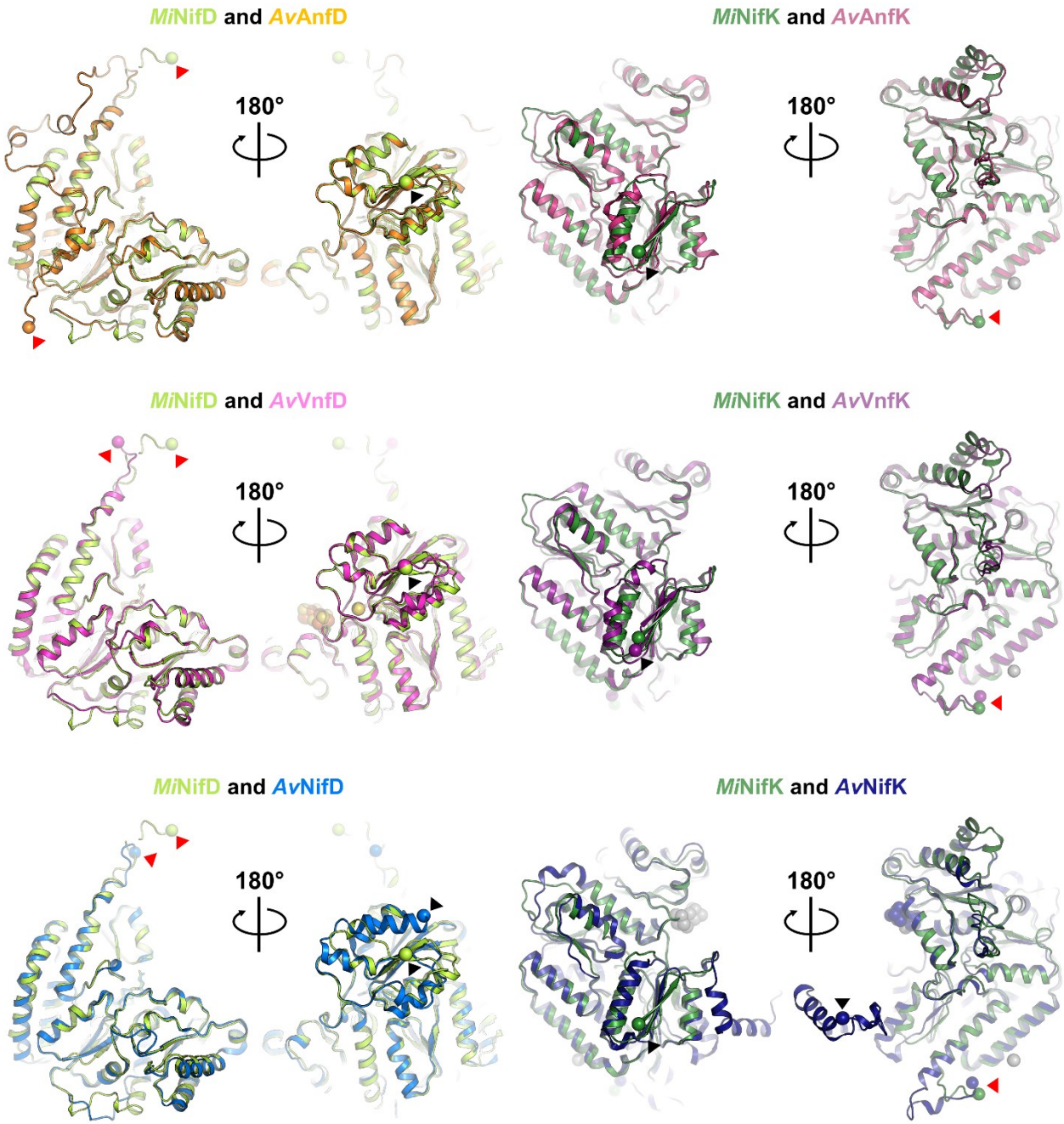
Supplementary Figure. 1. Growth of diazotrophic *M. infernus*. A. Growth curves of diazotrophic *M. infernus* grown in the medium with 10 µM Na_2MoO_4 (green circles). The culture grown with 10 µM Mo without N_2 in the gas phase was used as a negative control (red circles). The culture grown with 0.01 µM Mo with 22 mM NH_4Cl was used as a positive control (orange circles). **B.** Growth curves of diazotrophic *M. infernus* in the presence of different [Na_2MoO_4] concentrations: 0.01 µM Mo (light yellow circles), 0.1 µM Mo (light green circles), 1 µM Mo (green circles), 10 µM Mo (dark green circles) and 100 µM Mo (blue circles). All the cultures were grown at 75 ºC, in the presence of 100 µM [Na_2WO_4], 0.5 mM Na_2S, in the absence of VCl_3 with a gas phase of 0.5 bar H_2/CO_2 and 0.5 bar N_2, unless otherwise specified. All measurements have been performed in triplicates.



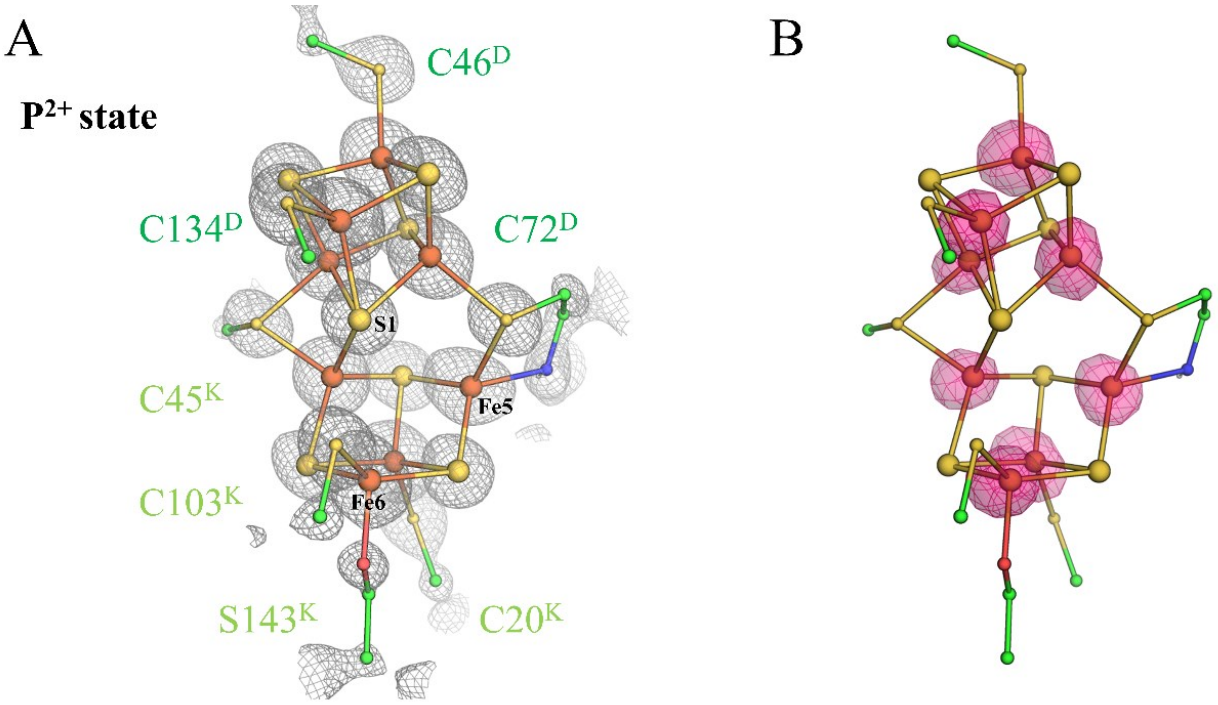
Supplementary Figure 2. SDS-PAGE of purified *M. infernus* and *M. thermolithotrophicus* proteins. A. *M. infernus* proteins: Lanes: 1-*Mi*NifI$_{1,2}$ (5 µg), 2-*Mi*NifDKI$_{1,2}$ (5 µg), 3-*Mi*NifDKI$_{1,2}$ (10 µg), 4-*Mi*NifDKH (5 µg), 5-*Mi*NifDKH (10 µg), 6-molecular weight marker, 7-*Mi*NifH (5 µg), 8-*Mi*NifH (10 µg), 9-*Mi*NifDK (4.8 µg). **B.** *M. thermolithotrophicus* native NifI$_{1,2}$ after purification. For NifI$_{1,2}$, the band below 25 kDa is a migration artifact that has already been observed for P$_{II}$-family proteins.



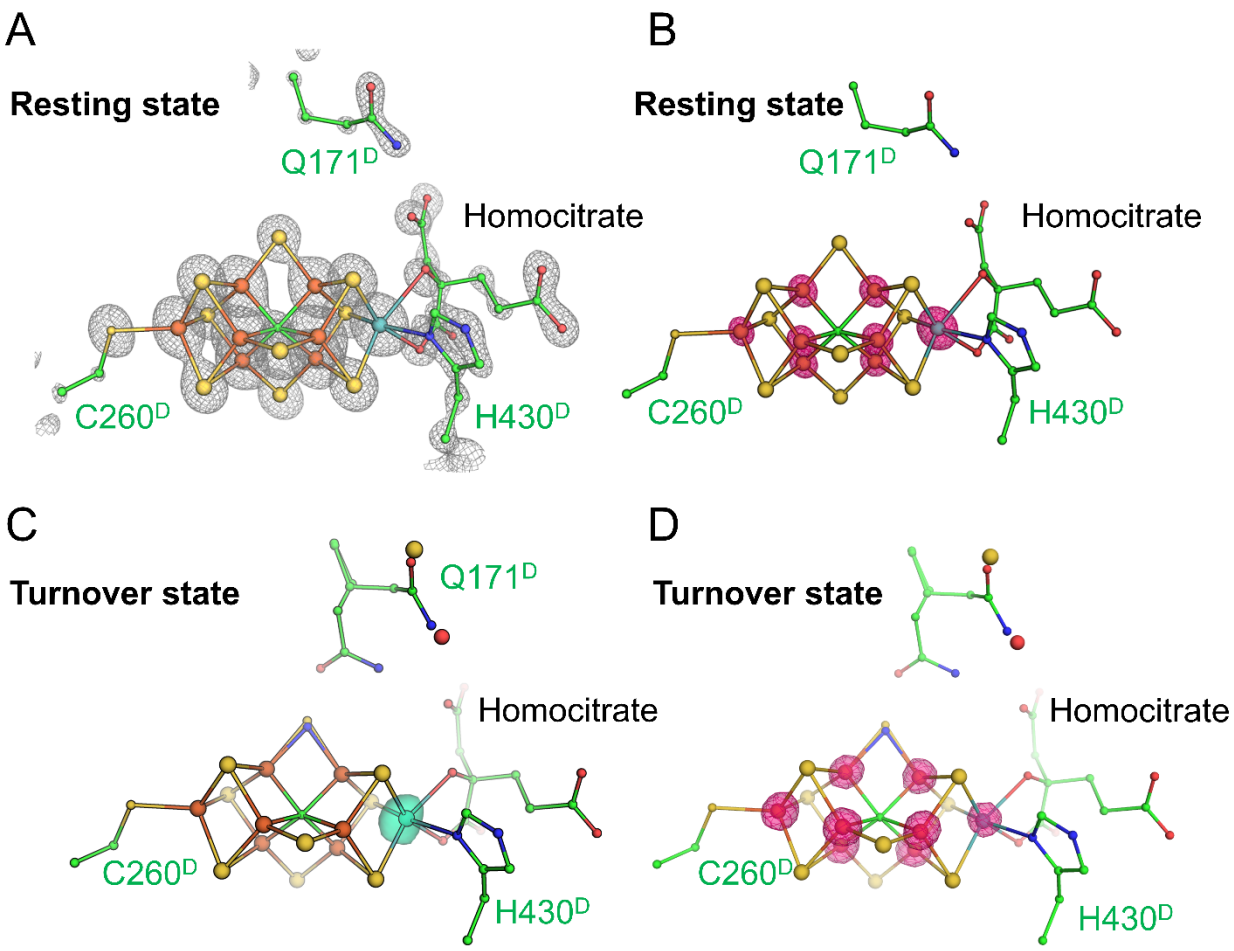
Supplementary Figure 3. Differences in crystal packing between the two *Mi*NifDK forms. **A**. Crystal packing representation for the monoclinic crystalline form. **B**. Crystal packing representation for the triclinic crystalline form. The asymmetric unit is in the center and colored by chains.



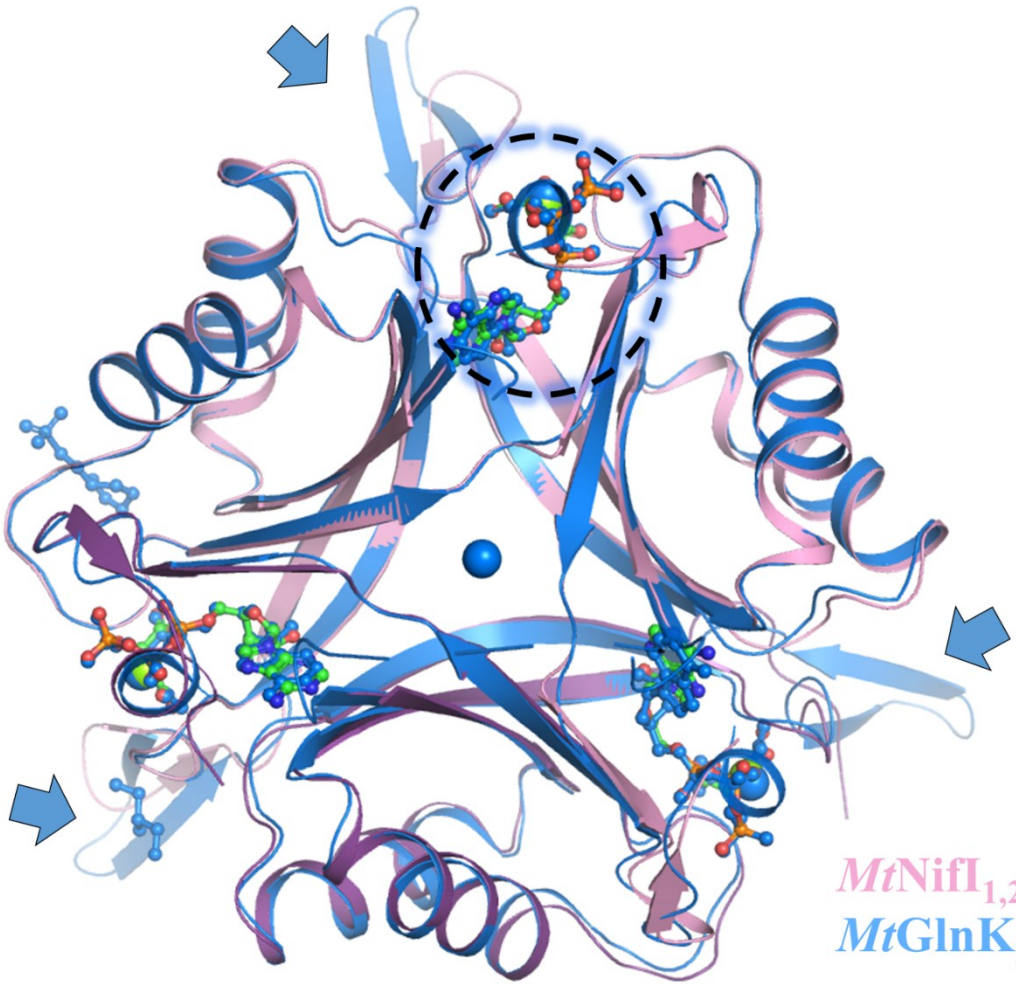
Supplementary Figure 4. Overlay between *Azotobacter vinelandii* alpha and beta nitrogenase subunits with *Mi*NifD and *Mi*NifK. All models are shown as cartoons. Chains from *A. vinelandii* isoforms were superposed on *Mi*NifD or *Mi*NifK. Black and red triangles point to the N- and C- terminal, respectively. The root mean square deviations were quantified as follows: *Mi*NifD/*Av*AnfD 0.761-Å for 362 aligned Cα, *Mi*NifD/*Av*VnfD 0.742-Å for 367 aligned Cα, *Mi*NifD/*Av*NifD 0.713-Å for 347 aligned Cα, *Mi*NifK/*Av*AnfK 0.990-Å for 338 aligned Cα, *Mi*NifK/*Av*VnfK 0.963-Å for 327 aligned Cα, *Mi*NifK/*Av*NifK 0.949-Å for 342 aligned Cα.



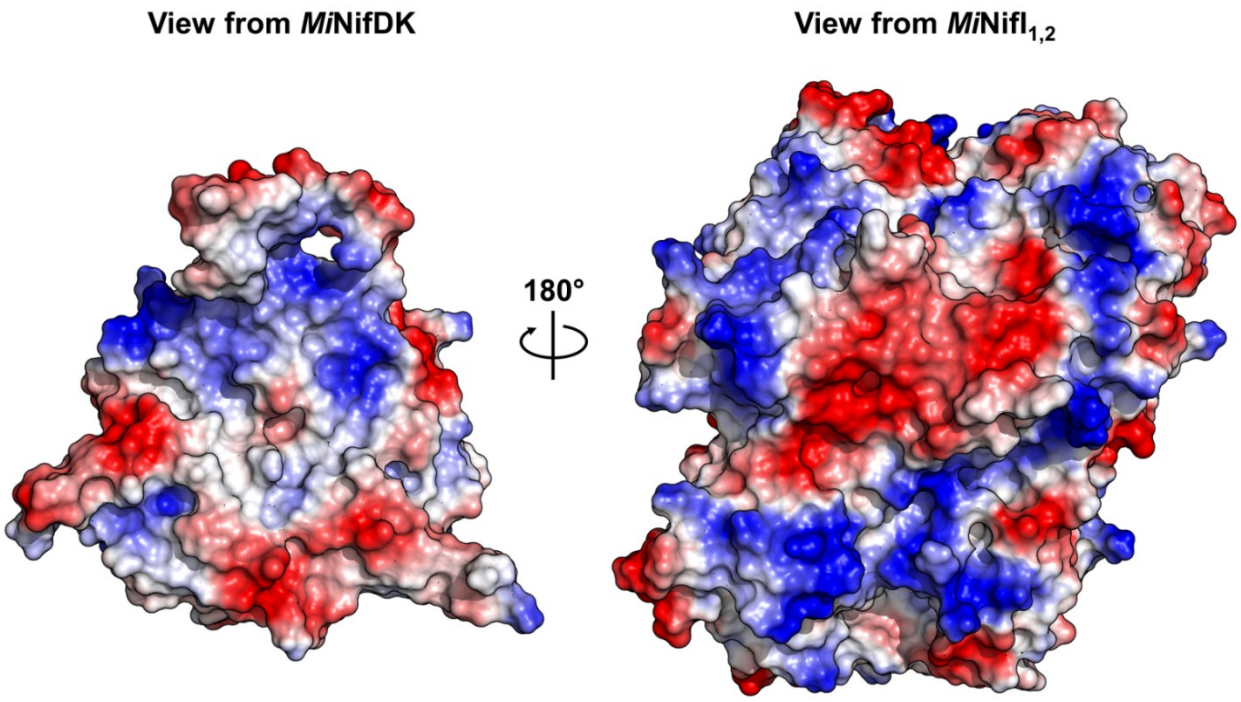
Supplementary Figure 5. Architecture and coordination of the P cluster from the *Mi*NifDK monoclinic form. **A.** Architecture and coordination of the P cluster in the P^{2+} (oxidized) state from the *Mi*NifDK. The grey mesh represents the $2F_o$-F_c electron density map contoured at 2.5 σ. **B.** Anomalous map signal obtained by data collection at 7160 eV, contoured at 8 σ (pink mesh) around the P cluster position.



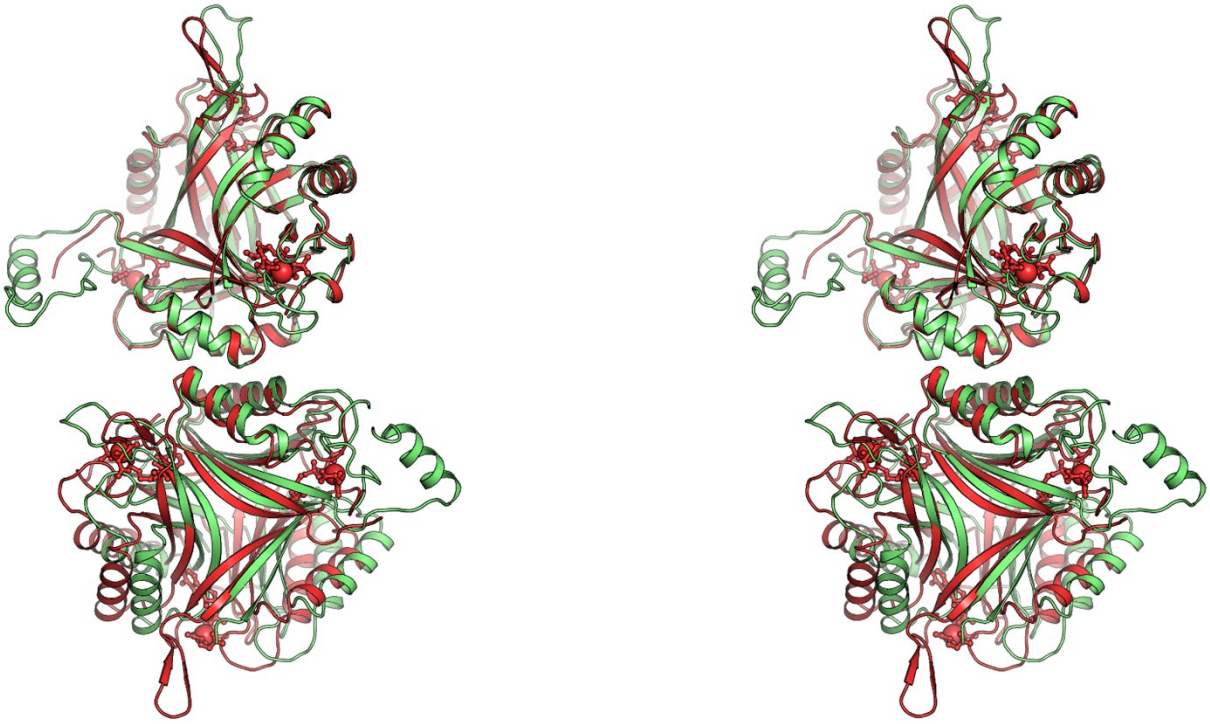
Supplementary Figure 6. The architecture and environment of FeMoco from the *Mi*NifDK in the monoclinic form. A. Sulfide S2B is bridging Fe2 and Fe6, while the Gln171^{D} occupies its holding site. The grey mesh represents the $2F_o$-F_c electron density map contoured at 2.5 σ. **B.** Anomalous map signal obtained by data collection at 7160 eV, contoured at 12 σ (pink mesh) around the FeMoco position. **C.** The difference map between anomalous data collected at 19900 eV (low remote) and 20100 eV (peak of the Mo fluorescence) contoured at 4.5 σ (cyan mesh) around FeMoco from the *Mi*NifDK in the triclinic form. **D.** Anomalous map signal obtained by data collection at 12700 eV, contoured at 10 σ (pink mesh) around the FeMoco position from the *Mi*NifDK in the triclinic form.



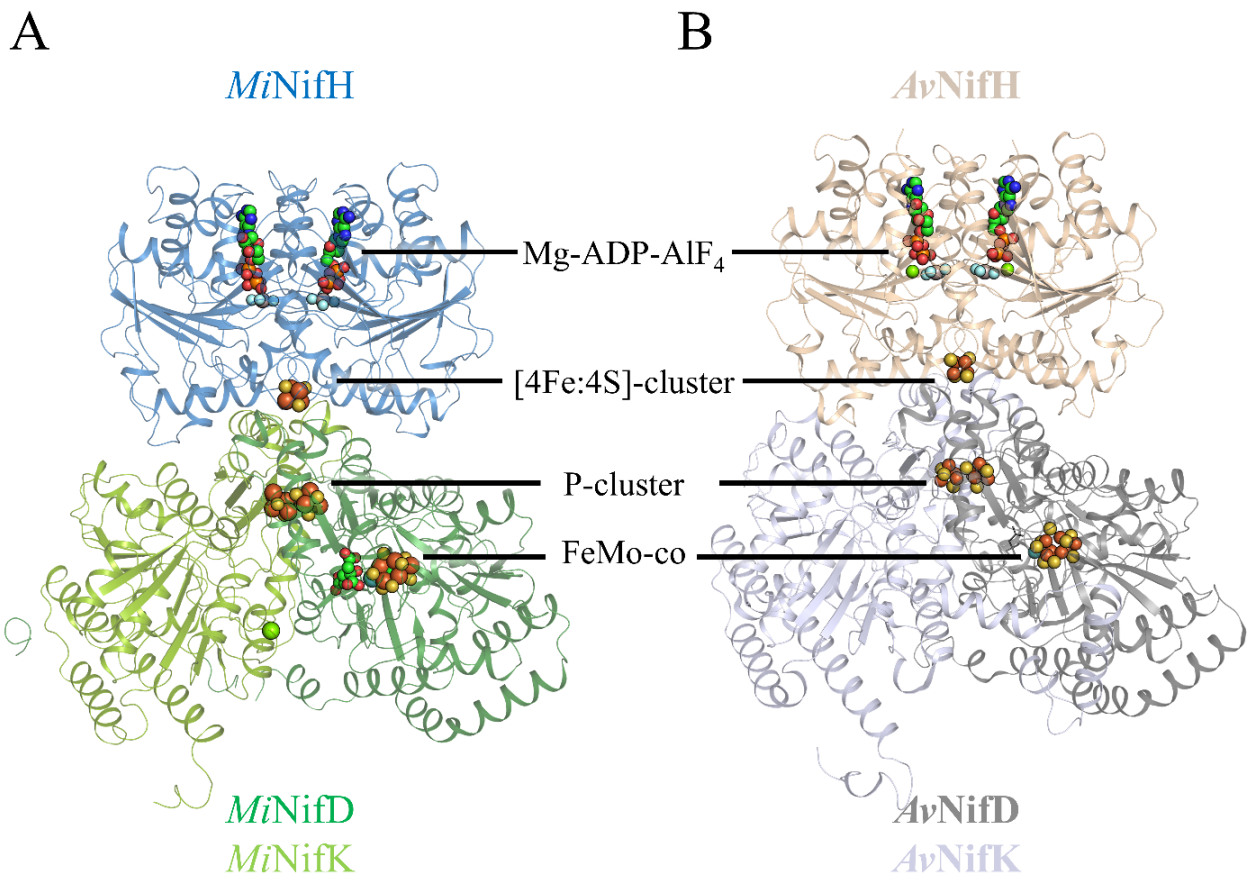
Supplementary Figure 7. Conservation of secondary structure in the PII superfamily of proteins. A. Overall superposition of *Mt*NifI$_{1,2}$ holo (pink) with *Mt*GlnK$_2$ (blue). For clarity, only one ligand binding site is highlighted with the blue-glow dashed circle in both panels. The ligands are shown as ball and stick models and color coded as follows: blue (N), green (C), red (O), orange (P), light green (Mg).



Supplementary Figure 8. Electrostatic interaction between *Mi*NifDK and *Mi*NifI$_{1,2}$. Both proteins, represented as surface, are viewed from their interaction face. The electrostatic charge profile is colored from blue (positively charged) to red (negatively charged).



Supplementary Figure 9. Differences between the apo and holo state of NifI$_{1,2}$. Wall-eyed stereo view of *Mi*NifI$_{1,2}$ apo (green) bound to *Mi*NifDK superposed to *Mt*NifI$_{1,2}$ holo (red). Both models, shown as cartoons, were superposed only on NifI$_2$ from the top heterotrimer. Ligands are shown in red as balls and sticks.



Supplementary Fig. 10. Comparison of the docking geometry in MgADP-AlF_4^- stabilised nitrogenase complexes of *M. infernus* and *A. vinelandii*. Only one NifDK and NifH dimer is shown for clarity in all panels. **A.** Protein-protein docking geometry in *M. infernus* MgADP-AlF_4^- stabilised NifDK-NifH complex. Proteins are depicted as cartoons and color coded as follows: forest green (NifD), green (NifK), and blue (NifH). **B.** Protein-protein docking geometry in *A. vinelandii* MgADP-AlF_4^- stabilised NifDK-NifH complex. Proteins are depicted as cartoons and color coded as follows: grey (NifD), light purple (NifK), and wheat (NifH). All metallocofactors are represented as spheres and colored by atom: orange (Fe), yellow (S), blue (N), green (C), red (O), dark orange (P), grey (Al), light blue (F), cyan (Mo).

CHAPTER 5: The 16th Fe of Mo-nitrogenase is not required for activity

Cécile Cadoux#[1, 2], Nevena Maslać#[3], Léa Di Luzio[1], Daniel Ratcliff[1, 2], Wenyu Gu[4], Tristan Wagner*[3] and Ross D. Milton*[1, 2]

[1] Department of Inorganic and Analytical Chemistry, Faculty of Sciences, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

[2] National Centre of Competence in Research (NCCR) Catalysis, University of Geneva, Quai Ernest Ansermet 30, 1211 Geneva 4, Switzerland

[3] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

[4] Laboratory of Microbial Physiology and Resource Biorecovery, School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

# Equally contributed to the study

Submitted to:

Angewandte Chemie

Recieved: 14th of June 2023

Correspondence to Tristan Wagner (*twagner@mpi-bremen.de*)

Correspondence to Ross D. Milton (*Ross.Milton@unige.ch*)

This chapter is displayed as the submitted manuscript.

5.1. Abstract

The biological N_2-fixation process is orchestrated by metallocofactor-containing nitrogenases. Structural and spectroscopic studies highlighted the presence of a 16th Fe in addition to the two metallocofactors required for the reaction. This ferrous Fe is located at 15-Å away from the active site, at the interface of two NifK subunits. The enigmatic function of the 16th Fe and its implication for metallocofactor installation, catalysis, electron transfer or structural stability is investigated in this work. The axial ligands coordinating the additional Fe are almost universally conserved in Mo-nitrogenases, but a detailed observation of the available structures indicates a variation in occupancy or a metal substitution. A nitrogenase variant in which the 16th Fe binding site is abrogated was generated and characterised by X-ray crystallography, biochemistry and enzymology. The crystal structure refined to 1.55-Å revealed an unambiguous loss of the metal site, also confirmed by an absence of anomalous signal for Fe. The position of the surrounding side chains and the overall architecture is superposable with the wild-type structure. Accordingly, the biochemical and enzymatic properties of the variant are similar to those of the wild-type nitrogenase, indicating that the 16th Fe does not impact nitrogenase's activity and stability *in vitro*.

5.2. Main text

Diazotrophic microbes exhibit the particularity of reducing dinitrogen (N_2) to ammonia (NH_3), a chemical reaction considered to be among the most challenging in biology due to the large activation energy (+251 kJ/mol) that must be surmounted.[1] Instead of activating the triple-bond of N_2 with high temperature and pressure, as in the Haber-Bosch process, microbes rely on energy released by the hydrolysis of adenosine triphosphate (ATP). In Mo-dependent nitrogenase, this reaction requires a minimum of 16 ATPs to reduce each N_2 to NH_3, with an approximate second-order rate constant (k_{cat}/K_m) of ~10^4 M^{-1} s^{-1}.[2] Mo-nitrogenase consists of two components, *(i)* the homodimeric Fe-protein that transfers an electron of low-potential concomitantly with ATP-hydrolysis and *(ii)* the $\alpha_2ß_2$ heterotetrameric MoFe-protein harbouring the P cluster (a [8Fe-7S] cluster) and the FeMo-cofactor (a [7Fe-Mo-9S-C] cluster bound to homocitrate) where N_2 is reduced (Figure 1A-C). The catalytic mechanism of the reaction has been studied for decades, revealing the complex ballet between the Fe protein association to the MoFe protein for electron transfer and FeMo-cofactor reactivity. Importantly, cooperativity is observed between each P cluster/FeMo-cofactor containing αß half of the MoFe protein during turnover.[3] Recently, structural investigation by cryo-electron

microscopy also highlighted asymmetry of the Fe-protein association to the MoFe-protein and its impact on local rearrangement close to the catalytic centre.[4]

While the P cluster and FeMo-co have attracted much attention, an additional Fe atom has been discovered at the ββ′ junction of the MoFe protein (the prime indicates the second protomer of the dimer, Figure 1A) close to the surface of the protein. Atomic resolution X-ray crystallography combined with anomalous information and spectroscopy analyses confirmed the identity of the metal in a ferrous (Fe^{2+}) oxidation state. The Fe has a partial occupancy, which might explain why it was mostly modelled as calcium or magnesium in previous structural work (Supporting information Table S2).[5] Since other cations might be bound to the site, the abbreviation MMB (mononuclear metal binding) site was adopted.

This hexacoordinate 16th Fe is coordinated by three carboxy groups (ß-E109, ß'-D353, ß‘-D357, using nomenclature from the MoFe-protein of *Azotobacter vinelandii*), a main chain carbonyl (ß-R108) and two aqua ligands, in ~octahedral geometry (Figure 1D). In *Bacteria*, ß-D353, ß-D357 and ß-E109 are all perfectly conserved across 255 sequences after filtering (described in Materials and Methods). In *Archaea*, these three residues are mostly conserved, and only a few sequences present substitutions (see Supporting Information). ß-R108 is not conserved in either group, possibly because it provides coordination through the peptide backbone (Figure 1D). They are also conserved in the sequences of V and Fe isoforms, the phylogenetically distant *Methanococcales* and *Methanobacteriales* sequences, and even in the BchNB protein of the homologous dark operative protochlorophyllide oxidoreductase.[6]

Despite its conservation, the possible importance of the MMB and Fe in nitrogenases has not been determined. It is hypothesized that it could be involved in optimising the loading of the metallocofactors in the enzyme, the catalytic efficiency, structural stability or electron transfer. For instance, the homologous position of the Asp357 found in BchNB facilitates proton transfer during protochlorophyllide reduction.[7] Herein, we created a MoFe protein variant deficient for the 16th Fe coordination to investigate the functional importance of the MMB through structural and biochemical characterisation.

Metal coordination at the MMB site was disrupted by substituting aspartates β-D353/D357 with glycine residues in *Azotobacter vinelandii* DJ by homologous double-reciprocal recombination.[3a] Anoxic purification of the generated β-D353G/D357G MoFe variant (via affinity, anion exchange and size-exclusion chromatography, Figure S1) yielded exploitable crystals belonging to the monoclinic *C*2 space group and containing an $\alpha_2\beta_2$ heterotetramer in the asymmetric unit (Figure 2A, Table S3). The structure, refined to 1.55-Å resolution, has

the typical MoFe-protein organisation with an excellent fit to the previously characterised atomic resolution model (root mean square deviation of 0.235-Å for 1858 Cα atoms when aligned to PDB 3U7Q, Figure 2A). The slight deviation observed at the periphery is most likely due to the difference in the crystal packing (Fig. S2).

The electron density profile presents a fully occupied P cluster, which fits with a mixture of oxidised and reduced states modelled at 90 and 10 % occupancy, respectively (Figure. 2B, Figure S3). The FeMo-cofactor also presents perfect integrity and occupies the active site at full occupancy (Figure 2C). Both metallocofactors share identical atomic positions compared to the previously described structures (Figures S3 and S4), with the Fe position corroborated by anomalous signal (Figures 2B and 2C).

As expected, the MMB site is vacant in the created β-D353G/D357G variant and is filled with water and an ethylene glycol molecule (Figure 2D). The modelled ethylene glycol presence suggests that the MMB site is not isolated from the solvent, explaining why Fe can be exchanged for other cations (*e.g.,* Ca^{2+}) or diffuse out of the active site. Accordingly, no anomalous signal could be detected in this area. The β-G353 and β-G357 atomic positions overlay with the main chain of the structures containing a partially modelled ion (Figure S5). The close environment, in particular the position of the β-E109 and β-R108, is equivalent to the wild-type.

Deficiency in the MMB site does not impact the structural integrity of the two other metallocofactors and the overall structure; however, the crystal packing may hide unsuspected features compared to the protein behaviour in solution (*i.e.*, protein dynamics during turnover).

The generated *A. vinelandii* mutant strain retained its diazotrophic phenotype (Figure S6), providing an initial indication that disruption of the MMB site did not completely abolish nitrogenase's N_2 fixation activity. *In vitro* activity assays were performed for *(i)* H_2 production under Ar and *(ii)* H_2/NH_3 production under N_2 (Figure 3). The ß-D353G/D357G MoFe protein remains fully active in comparison to the wild-type MoFe protein under these conditions, and the determined catalytic parameters confirmed that the mutant MoFe protein exhibits similar affinity towards N_2 as the wild-type MoFe protein.[3a, 8]

Electron distribution towards NH_3 was found to be ~60% under both low-flux (4 mol. eq. of Fe per MoFe) and high-flux (16 mol. eq. of Fe per MoFe) turnover conditions, suggesting that the MMB site does not impact the selectivity of nitrogenase towards N_2 under these *in vitro* conditions (Figure S7).

The intrinsic dynamics of the ß-D353G/D357G MoFe protein was also probed by size exclusion chromatography (Figure S8) and limited proteolysis (Figure S9). In both experiments, the ß-D353G/D357G variant and the wild-type showed a similar profile corroborating the structural analysis. The melting temperature of the two MoFe proteins was determined by circular dichroism to be ~57 ºC in both cases, confirming that the presence of the MMB site also does not introduce a pronounced stabilizing effect on the total secondary structure of the MoFe protein (Figure S10).[9]

Finally, protein-protein associations with this ß-D353G/D357G MoFe protein were investigated further by considering protein-mediated protection of nitrogenase against oxidative damage in the obligate aerobe *A. vinelandii*. In the presence of molecular oxygen (O_2), the Fe and MoFe proteins form a tripartite complex with a small ferredoxin-like protein ("FeSII", or "Shethna" protein) that restricts access of O_2 to its metallocofactors (*i.e.*, a switch-off protection mechanism).[10] All three component are essential for this protection mechanism. As presented in Figure S11 (Supplementary Information), 66% of nitrogenase H_2-formation activity could be preserved during exposure to 2% O_2 (for 10 minutes) when the wild-type MoFe protein was used. Importantly, only 16% of activity could be preserved in the case of the ß-D353G/D357G MoFe protein. These data suggest that the MMB site does indeed play a non-negligible role in nitrogenase catalysis *in vitro*, and possibly, *in vivo*.

In conclusion, our studies suggest that the MMB site is not involved in the P cluster and FeMo-cofactor installation, as both metallocofactors exhibit full occupancy and appropriate coordination. Nitrogenase's overall dynamic and N_2-fixation capabilities are also not impacted by the MMB site absence, even when assayed at the physiological temperature of 30 °C *in vitro*. Such observation is in accordance with the numerous structural works, which showed that an absence of Fe switches the side chain of ß-E109 without impacting the surrounding residues (Table S1). The structural similarity between the wild-type and presented ß-D353G/D357G variant might be due to a crystallographic artefact, in which the crystal packing rigidified a particular state allowing the partial release of the MMB without affecting the structure (Figure S2). Different cryo-electron microscopy snapshots recorded under turnover conditions present the additional electron density for the MMB; interestingly, the MMB site observed in the various trapped structures displays the same architecture (Figure S12), as well as its surrounding, reinforcing the concept that the site is not involved in the catalytic turnover. Nevertheless, we cannot exclude that the MMB site might influence a fine-tuning of the overall N_2-fixation process by optimising the asymmetry dynamics between the Fe-protein and the MoFe-protein or cooperativity between the subunits.[3a]

Our observation that the removal of the MMB site negatively impacts *in vitro* O_2-protection assays (Figure S11) hints at unsuspected *in vivo* implications. Since the MMB site is partially solvent exposed, a role in metal sensing could also be hypothesised in which the Fe would be exchanged for another metal that would tune nitrogenase's properties to respond to immediate physiological needs (*e.g.,* decrease ATP consumption). On the same line of thought, the MMB site might rather have a role to optimise the intracellular stability of the MoFe-protein under fluctuating conditions. Such peculiar conditions involving the combination of molecular crowding, a balance of cellular metabolites, and changes in the ionic strengths, which cannot be probed *in vitro* due to their complexity, might impact the MoFe-protein stability in the absence of a functional MMB site. It would therefore be interesting to monitor if such a Fe-occupied site exists in the phylogenetically distant archaeal homologues (*e.g.,* hyperthermophilic *Methanococcales*), with different physiology, intracellular composition, and thermotolerance.

5.3. Supporting Information

The authors have cited additional references within the Supporting Information.[3a, 5, 11] Materials and Methods, including the strategy used to generate the *A. vinelandii* DJ strain derivative producing ß-D353G/D357G MoFe protein. SDS-PAGE and Native-PAGE of the purified proteins, additional X-ray crystallography information and analyses, additional activity assays, analytical size exclusion chromatography, limited proteolysis, melting temperature determination. Raw data are freely available on the Zenodo repository: *https://doi.org/10.5281/zenodo.7973626*

5.4. Acknowledgements

We thank the Max Planck Institute for Marine Microbiology for continuous support. We acknowledge the Swiss light Source Synchrotron for the beamtime and especially the staff of beamline X06DA. We would like to thank Ramona Appel and Christina Probian for their continuous support in the Microbial Metabolism laboratory. We thank Gloria Sedoh, Plinio Maroni, Olivier Vassalli, Naomi Sakai, and Isabelle Worms for experimental and technical support. This study was funded by the Max Planck Society for T.W. and N.M. T.W. was supported by the Deutsche Forschungsgemeinschaft Schwerpunktprogram 1927 "Iron-sulfur for Life" (WA 4053/1-1). This publication was created as part of NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation.

5.5. Author contributions

C.C., W.G. and R.D.M. designed the research. C.C., L.D.L., W.G. and R.D.M. constructed and produced the ß-D353G/D357G MoFe protein mutant. C.C. performed limited proteolysis; and C.C., L.D.L. and D.R. performed activity assays. N.M. performed the crystallisation. X-ray data collection was performed by N.M. and T.W. Data processing, model building, structure refinement, validation and deposition were performed by N. M. and T.W. Structures were analysed by N. M. and T.W. N.M. performed the sequence comparison analyses. R.D.M. and T.W. acquired funding to realize the project. The paper was written by all authors.

5.6. Keywords

Nitrogenase • X-ray crystallography • Nitrogen fixation • Cooperativity • FeS clusters

5.7. References

[1] F. Mus, A. B. Alleman, N. Pence, L. C. Seefeldt, J. W. Peters, *Metallomics* **2018**, *10*, 523-538.

[2] aS. Burén, E. Jiménez-Vicente, C. Echavarri-Erasun, L. M. Rubio, *Chemical Reviews* **2020**, *120*, 4921-4968; bO. Einsle, D. C. Rees, *Chemical Reviews* **2020**, *120*, 4969-5004; cA. J. Jasniewski, C. C. Lee, M. W. Ribbe, M. W. Ribbe, Y. Hu, *Chemical Reviews* **2020**, *120*, 5107-5157; dH. L. Rutledge, F. A. Tezcan, *Chemical Reviews* **2020**, *120*, 5158-5193; eL. C. Seefeldt, Z. Y. Yang, D. A. Lukoyanov, D. F. Harris, D. R. Dean, S. Raugei, B. M. Hoffman, *Chemical Reviews* **2020**, *120*, 5082-5106; fC. Van Stappen, L. Decamps, G. E. Cutsail, R. Bjornsson, J. T. Henthorn, J. A. Birrell, S. Debeer, *Chemical Reviews* **2020**, *120*, 5005-5081.

[3] aC. Cadoux, D. Ratcliff, N. Maslać, W. Gu, I. Tsakoumagkos, S. Hoogendoorn, T. Wagner, R. D. Milton, *JACS Au* **2023**, *3*, 1521-1533; bK. Danyal, S. Shaw, T. R. Page, S. Duval, M. Horitani, A. R. Marts, D. Lukoyanov, D. R. Dean, S. Raugei, B. M. Hoffman, L. C. Seefeldt, E. Antony, *Proceedings of the National Academy of Sciences* **2016**, *113*, E5783-E5791; cS. Truscott, R. S. Lewis, G. D. Watt, *Biophysical Chemistry* **2021**, *277*, 106650-106650.

[4] H. L. Rutledge, B. D. Cook, H. P. M. Nguyen, M. A. Herzik, Jr., F. A. Tezcan, *Science* **2022**, *377*, 865-869.

[5] L. Zhang, J. T. Kaiser, G. Meloni, K.-Y. Yang, T. Spatzal, S. L. A. Andrade, O. Einsle, J. B. Howard, D. C. Rees, *Angewandte Chemie International Edition* **2013**, *52*, 10529-10532.

[6] aM. J. Bröcker, S. Schomburg, D. W. Heinz, D. Jahn, W.-D. Schubert, J. Moser, *Journal of Biological Chemistry* **2010**, *285*, 27336-27345; bE. S. Boyd, T. L. Hamilton, J. W. Peters, *Front Microbiol* **2011**, *2*, 1-11; cN. Maslać, C. Sidhu, H. Teeling, T. Wagner, *mBio* **2022**, *13*, e02443-02422.

[7] N. Muraki, J. Nomata, K. Ebata, T. Mizoguchi, T. Shiba, H. Tamiaki, G. Kurisu, Y. Fujita, *Nature* **2010**, *465*, 110-114.

[8] D. F. Harris, D. A. Lukoyanov, S. Shaw, P. Compton, M. Tokmina-Lukaszewska, B. Bothner, N. Kelleher, D. R. Dean, B. M. Hoffman, L. C. Seefeldt, *Biochemistry* **2018**, *57*, 701-710.

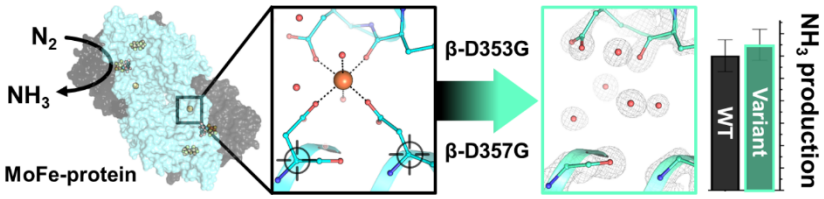
[9] N. J. Greenfield, *Nature Protocols* **2006**, *1*, 2876-2890.

[10] aY. I. Shethna, D. V. DerVartanian, H. Beinert, *Biochem Biophys Res Commun* **1968**, *31*, 862-868; bJ. Schlesier, M. Rohde, S. Gerhardt, O. Einsle, *Journal of the American Chemical Society* **2016**, *138*, 239-247.

[11] aF. Madeira, M. Pearce, A. R. N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, *Nucleic acids research* **2022**, *50*, W276-W279; bD. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, *Acta Crystallogr D Struct Biol* **2019**, *75*, 861-877; cC. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, 3rd, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, *Protein Sci* **2018**, *27*, 293-315; dE. Jimenez-Vicente, Z. Y. Yang, W. Keith Ray, C. Echavarri-Erasun, V. L. Cash, L. M. Rubio, L. C. Seefeldt, D. R. Dean, *Journal of Biological Chemistry* **2018**, *293*, 9812-9823; eR. D. Milton, R. Cai, S. Sahin, S. Abdellaoui, B. Alkotaini, D. Leech, S. D. Minteer, *Journal of the American Chemical Society* **2017**, *139*, 9044-9052; fR. D. Milton, R. Cai, S. Abdellaoui, D. Leech, A. L. De Lacey, M. Pita, S. D. Minteer, *Angewandte Chemie - International*

Edition **2017**, *56*, 2680-2683; gR. D. Milton, S. Abdellaoui, N. Khadka, D. R. Dean, D. Leech, L. C. Seefeldt, S. D. Minteer, *Energy and Environmental Science* **2016**, *9*, 2550-2554; hN. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, *Nucleic Acids Research* **2015**, *44*, D733-D745; iG. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res* **2004**, *14*, 1188-1190; jT. D. Schneider, R. M. Stephens, *Nucleic Acids Res* **1990**, *18*, 6097-6100.

5.8. Entry for the Table of Contents



The nitrogenase catalyses the biological N_2-fixation via unique metallocofactors. An additional metal binding site of unknown function coordinates a ferric iron at the dimeric interface. Here, a double mutant abrogating the metal site has been investigated. While the variant shares the same structural and enzymatic characteristics *in vitro*, the additional metal binding site might have unsuspected *in vivo* roles, such as protein-protein interaction.

5.9. Figures and legends

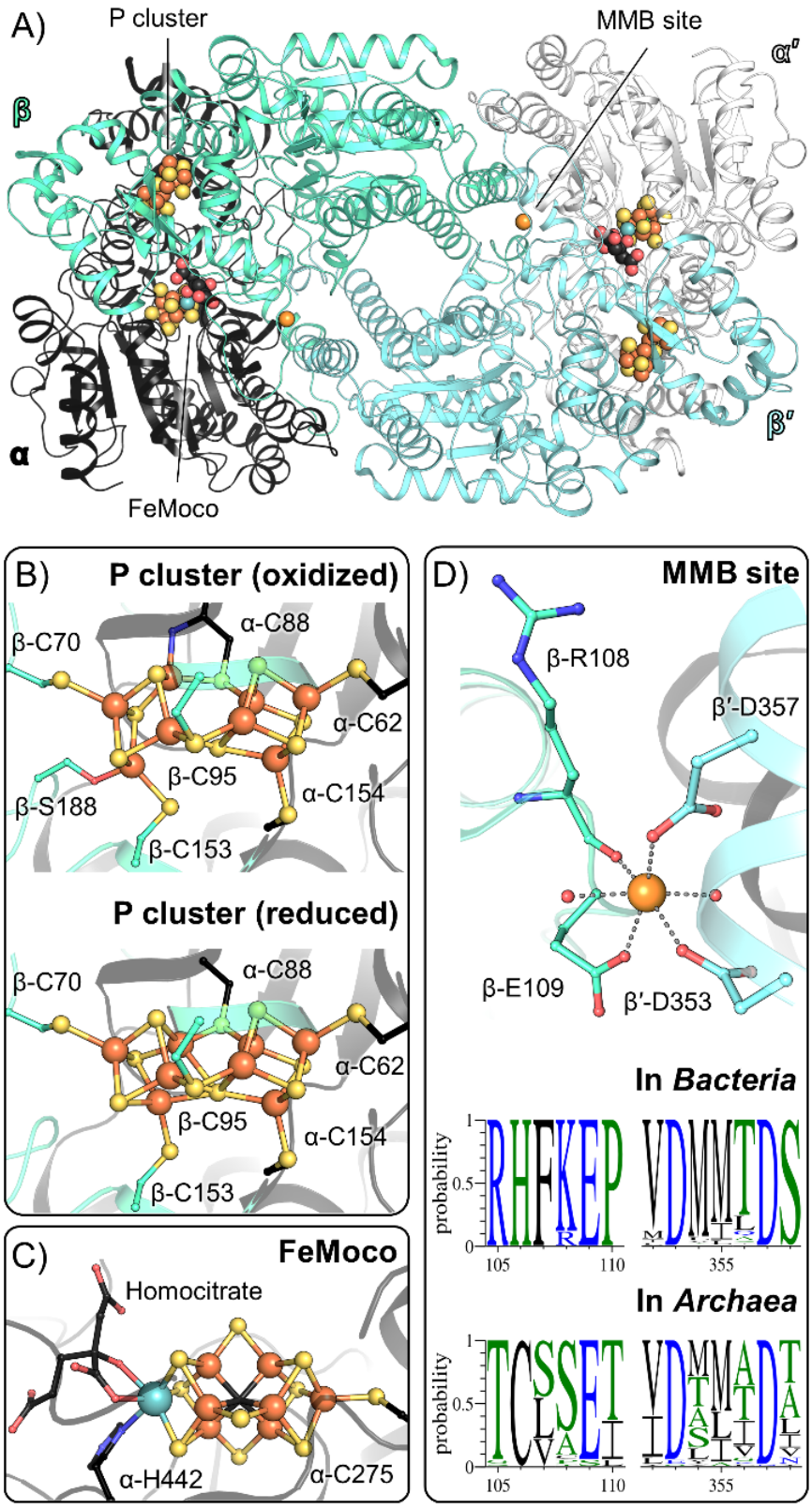


Figure 1. Overall organization of the MoFe-protein, based on the atomic resolution model (PDB 3U7Q) (A), and composition of its metallocofactors (B and C). Organization and conservation of the MMB site in nitrogenases (D). The calcium atom modelled in the structure of PDB 3U7Q was exchanged by a Fe based on the PDB 4TKU. Fe, S and Mo are shown as balls and coloured as orange, yellow and cyan, respectively.

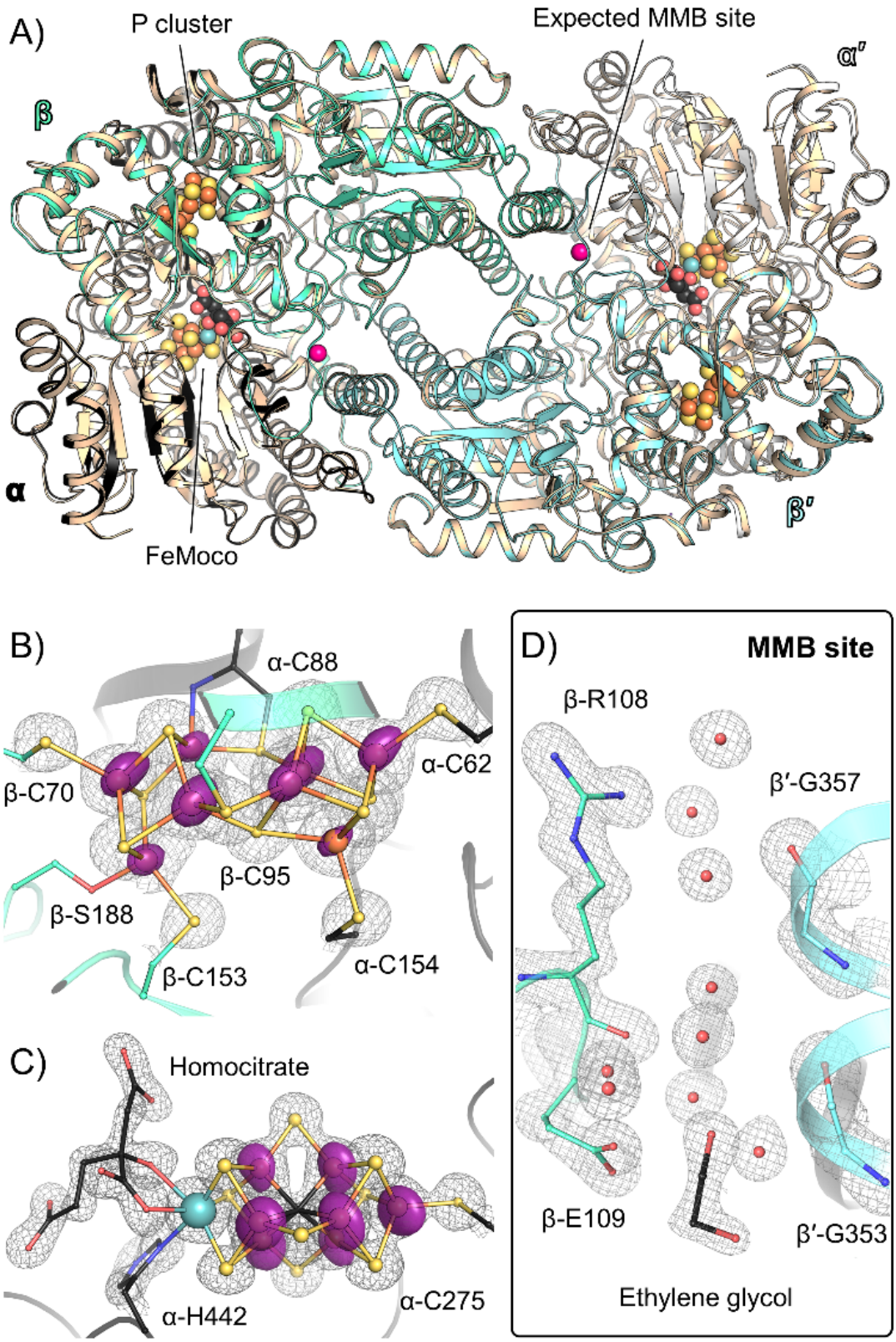


Figure 2. Overlay of the ß-D353G/D357G variant with the wild-type MoFe-protein (PDB 3U7Q, coloured in wheat) (A). Close-up of the P cluster (B), FeMo-cofactor (C), and the MMB site (D). The protein is shown as transparent cartoon and residues/metals as balls and sticks. In panels B) and C) the $2F_o$-F_c electron density (black mesh) and the anomalous (transparent purple surface) maps are contoured at 2.5 and 10 σ, respectively. In panel D, the $2F_o$-F_c electron density map is contoured at 1.5 σ, and no peaks can be detected in the anomalous map at this site.

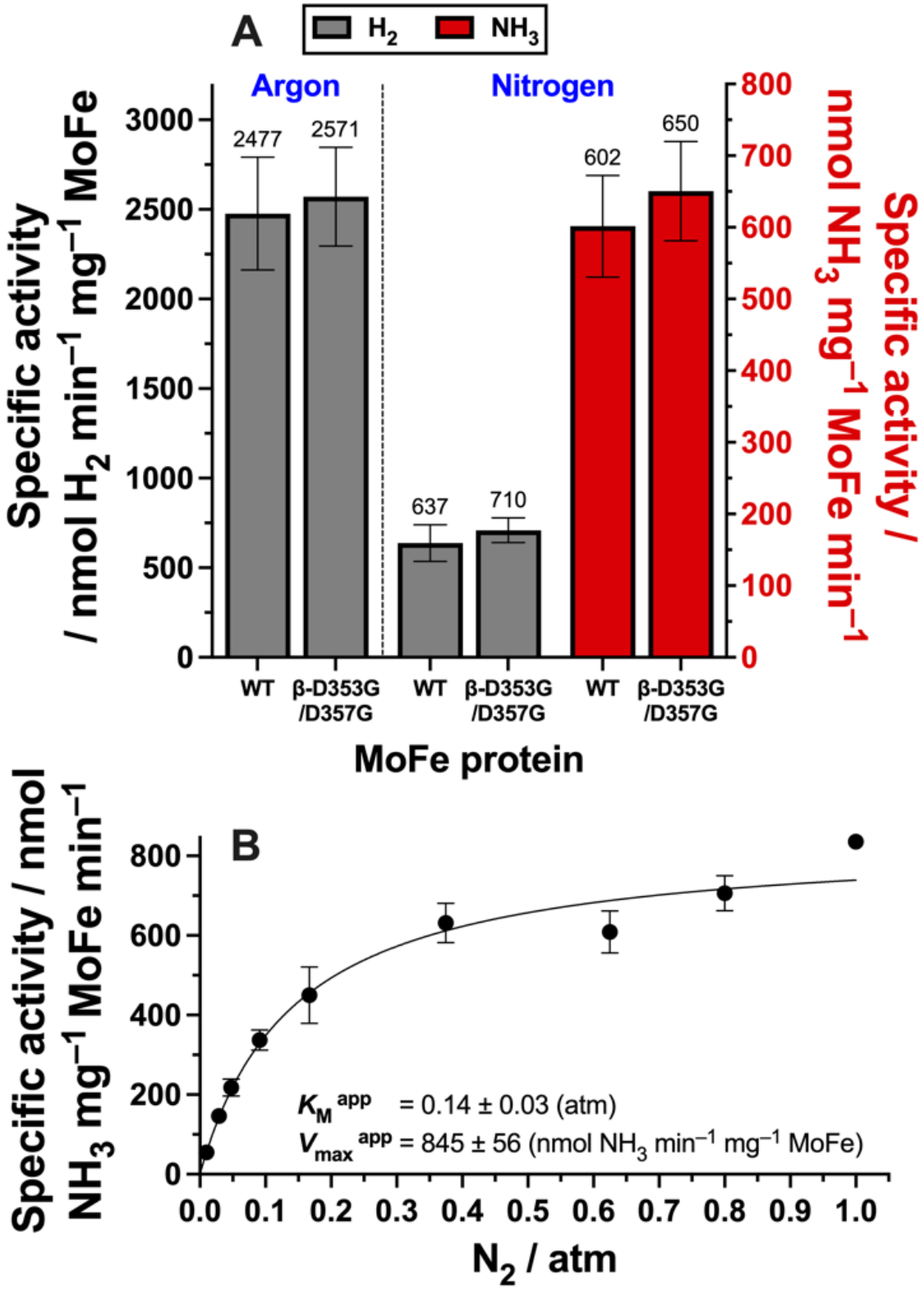


Figure 3. (A) Specific activities of wild-type (WT) and ß-D353G/D357G MoFe for H_2 and NH_3 production during turnover under Ar or N_2 (1 atm). (B) Apparent Michaelis-Menten kinetic parameters for the ß-D353G/D357G MoFe protein. All assays were performed at 30 ºC for 8 minutes and contained 0.1 mg MoFe protein with 16 molar equivalents of Fe protein. Sample quantification and additional assay information can be found in the Supporting Information. $n = 3$, $n = 2$ for the 1 atm N_2 data point in B. Assays performed on different days and using different purified protein batches.

5.10. Supporting information

Raw data are freely available on the Zenodo repository:
https://doi.org/10.5281/zenodo.7973626

Materials and methods.

Supporting Table S1. Primers used to create *Azotobacter vinelandii* DJ derivative strain producing ß-D353G/D357G MoFe protein.

Supporting Table S2. Comparison in the presence of the 16th Fe in the deposited structures.

Supporting Table S3. X-ray analysis statistics for the presented structure.

Supporting Table S4. Accession numbers for sequences used in conservation analysis.

Supporting Figure S1. SDS-PAGE and Native-PAGE of the β-D353G/D357G MoFe protein.

Supporting Figure S2. Crystal packing of the β-D353G/D357G MoFe protein.

Supporting Figure S3. Close up of the P cluster containing the two modelled redox states. Overlay of P cluster with its density in comparison to the wild type.

Supporting Figure S4. Superposition of the FeMo-co from β-D353G/D357G variant with the FeMo-co from wild-type MoFe protein.

Supporting Figure S5. Superposition of the MMB site from β-D353G/D357G variant with the MMB site from α-R96Q MoFe variant (PDB 6BBL).

Supporting Figure S6. Nitrogen-free agar plates of *Azotobacter vinelandii* strains used in this work.

Supporting Figure S7. Electron distribution of the β-D353G/D357G MoFe protein under low- and high-flux turnover.

Supporting Figure S8. Size exclusion chromatography of the variant versus the wild type.

Supporting Figure S9. Limited proteolysis profile of the variant versus the wild type.

Supporting Figure S10. Melting temperature determination of the wild-type and β-D353G/D357G MoFe proteins by circular dichroism.

Supporting Figure S11. Protection assays of wild-type and ß-D353G/D357G MoFe protein against O_2 deactivation.

Supporting Figure S12. Overlay of the MMB in the different cryo-EM snapshots (PDB 7UT6-9, 7UTA).

Materials and methods

***In silico* analysis.** The sequence logos[11] were created using the WebLogo 3 server.[12] The homologous sequences used for the logos were obtained by BlastP[13] search in RefSeq Select database[14] against *Azotobacter vinelandii* DJ NifK (NCBI accession number WP_012698833.1) with default parameters. The 250 closest homologues were filtered for redundancy and manually checked for correct annotation. The final dataset after filtering consisted of 225 bacterial and 55 archaeal sequences (accession numbers for all used sequences can be found in Supplementary Table S4). The alignment used as an input for sequence logo generation was done using MUSCLE server.[15]

***A. vinelandii* cultures and media.** The procedure outlined here was developed previously with minor modifications.[16] All buffers and solutions were prepared with MilliQ water (18.2 MΩ cm). Generally, large volumes (> 50 mL) were sterilized by autoclaving; smaller solutions were sterilized by filtration (0.45 µm syringe filters).

A modified Burke's medium was used for the cultivation of all *A. vinelandii* strains. Under sterile conditions, 6 mL of an autoclaved 100x phosphate buffer (0.46 M K_2HPO_4, 0.15 M KH_2PO_4) was added to an autoclaved molten agar solution (~18 g agar per L of water). 27.5 mL of this resulting agar buffer was added to 3 mL of an autoclaved 10x salt solution (60 mM sucrose, 8.1 mM $MgSO_4$, 6.1 mM $CaCl_2$), 250 µL of a filter-sterilized 3 M NH_4OAc solution, 30 µL of a filter-sterilized 0.1 M Fe solution ($FeCl_3.6H_2O$) and 30 µL of a filter-sterilized 10 mM Mo solution ($Na_2MoO_4.2H_2O$). The mixture was poured into a Petri dish, solidified, and the relevant *A. vinelandii* strains were plated and incubated at 30 °C for 48 h.

Under sterile conditions, 60 mL of an autoclaved 100x phosphate buffer solution (0.46 M K_2HPO_4, 0.15 M KH_2PO_4) was added to 6 L of autoclaved salt medium (60 mM sucrose, 8.1 mM $MgSO_4 \cdot 7H_2O$, 6.1 mM $CaCl_2$) after both solutions were at room temperature. Then 20 mL of a filter-sterilized 3 M NH_4OAc solution, 6 mL of a filter-sterilized 0.1 M Fe solution ($FeCl_3.6H_2O$) and 6 mL of a filter-sterilized 10 mM Mo solution ($Na_2MoO_4.2H_2O$) was added (yielding a modified Burke's medium).

For each *A. vinelandii* strain, a preculture was grown in a 250 mL baffled culture flask (glass) with a vented cap containing 60 mL of sterile liquid media. Each preculture was started by

inoculation from a fresh agar plate and grown aerobically at 200 rotations per minute (rpm) at 30 °C until $OD_{600\,nm}$ > 2.0 (usually overnight).

Under sterile conditions, 60 mL of the preculture was used to inoculate the 6 L of liquid media in 2x 5 L baffled flasks (Corning CLS431684, 3 L of culture per bottle) with vented caps. Cells were grown at 200 rpm at 30 °C overnight until $OD_{600\,nm}$ > 1.5. Cells were collected by serial centrifugation runs at 4,500 x *g*, 20 min at room temperature. The supernatant was discarded, and the cells were gently resuspended in fresh growth medium devoid of NH_4OAc or another source of fixed nitrogen (prepared fresh, but not under sterile conditions) in the incubator shaker (~400 mL media per bottle). After resuspension, the culture was divided evenly between the remaining media/bottles and incubated at 200 rpm and 30 °C for another 3 hours to facilitate derepression of the *nif* operon, and harvested by serial centrifugation collections as above. Finally, the cells were scooped into a plastic bag and stored at -80 °C until further use.

Preparation of the *A. vinelandii* strain producing β-D353G/D357G MoFe protein. The β-D353G/D357G MoFe mutant was created using the double homologous recombination method.[16] Briefly, a plasmid was constructed using the Gibson assembly (NEB, Ipswich, MA, US) method with the primers listed in Supplementary Table S1 in *E. coli* DH5α (Invitrogen, Waltham, MA, US), where flanking regions of *nifK* with a 135L - 357D in-frame deletion was cloned into a mutation plasmid pK18*mobsacB*. The constructed pK-△*nifK* was transformed into a strain of *A. vinelandii* DJ that produces the MoFe protein with a poly(histidine)$_8$ tag on the N-term of NifD [17] followed by single and double homologous recombination screenings, resulting in kanamycin-resistance, and kanamycin-sensitive, sucrose-resistance phenotypes, respectively. The final *A. vinelandii* △*nifK* is non-diazotrophic and the deletion was confirmed by Sanger sequencing.

Subsequently, a "rescue" plasmid was made where the deleted region of *nifK* was reintroduced to pK-△*nifK*, with point mutations of D353G/D357G (gac to ggc, gac to ggt, respectively) introduced on the primer (Supplementary Table S1, primer 8). Following the same transformation and screening procedures, this plasmid pK-*nifK*D353G/D357G was used to create the *A. vinelandii* β-D353G/D357G MoFe mutant, which was confirmed by Sanger sequencing. Phenotype screening showed this mutant restored diazotrophic growth (Supplementary Figure S6).

Table S1. Primers used to create the *A. vinelandii* △*nifK* and β-D353G/D357G MoFe mutants

Primers	Sequence	Notes
First step, creating pK-△*nifK* plasmid:		
1	cgttgtaaaacgacggccagtgccagtatcaaggagaagttcatcttcca	Upstream flanking region, forward
2	gcagccaggtgtgggacagaccgtccttcatgttctgct	Upstream flanking region, reverse
3	gaaggacggtctgtcccacacctggctgcacgg	Downstream flanking region, forward
4	tacgaattcgagctcggtacccgggggtatcggctcgatgaaataggtc	Downstream flanking region, reverse
Second step, creating pK-*nifK*D353G/D357G plasmid:		
5	tggtcggcatgatgaccggttcccacacctggctgcacgg*	Amplify plasmid pK-Δ*nifK*, forward
6	cttacagttctgcagcagaccgtccttcatgttctgctgg	Amplify plasmid pK-Δ*nifK*, reverse
7	agaacatgaaggacggtctgctgcagaactgtaaggctacctacaagccc	Inserting *nifK*D353G/D357G, forward
8	**ac**cggtcatcatg**c**cgaccagacggccacgctccttggtc*	Inserting *nifK*D353G/D357G, reverse

nifK in red

gac = β-D353G single-point mutation (changed to **ggc**)

gac = β-D357G single-point mutation (changed to **ggt**)

caccaccatcatcaccaccatcac = poly(histidine)$_8$ tag

Forward sequencing primer highlighted in yellow

Reverse sequencing primer highlighted in cyan

Region unconfirmed by sequencing highlighted in grey.

nifD:

atgacc**caccaccatcatcaccaccatcac**ggtatgtcgcgcgaagaggttgaatccctcatccaggaagttctggaagtttatcccg
agaaggctcgcaaggatcgtaacaagcacctggccgtcaacgacccggcggttacccagtccaagaagtgcatcatctccaacaag
aagtcccagcccggtctgatgaccatccgcggctgcgcctacgccggttccaaaggcgtggtctggggccccatcaaggacatgatc
cacatctcccacggtccggtaggctgcggccagtattcgcgcgccggccgtcgtaactactacatcggtaccaccggtgtgaacgcct
tcgtcaccatgaacttcacctcggacttccaggagaaggacatcgtgttcggtggcgacaagaagctcgccaaactgatcgacgaagt
ggaaaccctgttcccgctgaacaagggtatCtccgtccagtccgagtgcccgatcggcctgatcggcgacgacatcgaatccgtgtc
caaggtcaagggcgccgagctcagcaagaccatcgtaccggtccgttgcgaaggcttccgcggcgtttcccagtccctgggccacca
catcgccaacgacgcagtccgcgactgggtcctgggcaagcgtgacgaagacaccaccttcgccagcactccttacgatgtggccat
catcggcgactacaacatcggcggcgacgcctggtcttcccgcatcctgctggaagaaatgggcctgcgttgcgtagcccagtggtcc
ggcgacggctccatctccgaaatcgagctgaccccgaaggtcaagctgaacctggttcactgctaccgctcgatgaactacatctccc
gtcacatggaagagaagtacggtatcccatggatggagtacaacttcttcggcccgaccaagaccatcgagtcgctgcgtgccatcgc
cgccaagttcgacgagagcatccagaagaagtgcgaagaggtcatcgccaagtacaagcccgagtgggaagcggtggtcgccaag
taccgtccgcgcctggaaggcaagcgcgtcatgctctacatcggtggcctgcgtccgcgccacgtgatcggcgcctacgaagacctg
ggcatggaagtggtgggtaccggctacgagttcgcccacaacgacgactatgaccgcaccatgaaagaaatgggtgactccaccctg
ctgtacgatgacgtgaccggctacgaattcgaagaattcgtcaagcgcatcaagcccgacctgatcggctccggtatcaaggagaagt
tcatcttccagaagatgggcatccccttccgtcaaatgcactcctgggattattccggcccctaccacggcttcgatggcttcgccatcttc
gcccgtgacatggacatgaccctgaacaatccgtgctggaagaaactgcaggctccctgggaagcttccgaaggcgccgagaaagt
cgccgccagcgcctga

NifD:

MT**HHHHHHHHH**GMSREEVESLIQEVLEVYPEKARKDRNKHLAVNDPAVTQSKKCIIS
NKKSQPGLMTIRGCAYAGSKGVVWGPIKDMIHISHGPVGCGQYSRAGRRNYYIGTT
GVNAFVTMNFTSDFQEKDIVFGGDKKLAKLIDEVETLFPLNKGISVQSECPIGLIGDDI
ESVSKVKGAELSKTIVPVRCEGFRGVSQSLGHHIANDAVRDWVLGKRDEDTTFASTP
YDVAIIGDYNIGGDAWSSRILLEEMGLRCVAQWSGDGSISEIELTPKVKLNLVHCYRS
MNYISRHMEEKYGIPWMEYNFFGPTKTIESLRAIAAKFDESIQKKCEEVIAKYKPEWE
AVVAKYRPRLEGKRVMLYIGGLRPRHVIGAYEDLGMEVVGTGYEFAHNDDYDRTM
KEMGDSTLLYDDVTGYEFEEFVKRIKPDLIGSGIKEKFIFQKMGIPFRQMHSWDYSGP
YHGFDGFAIFARDMDMTLNNPCWKKLQAPWEASEGAEKVAASA*

nifK:

atgagccagcaagtcgataaaatcaaagccagctacccgctgttcctcgatcaggactacaaggacatgcttgccaagaagcgcgac
ggcttcgaggaaaagtatccgcaggacaagatcgacgaagtattccagtggaccaccaccaaggaataccaggagctgaacttccag
cgcgaagccctgaccgtcaacccggccaaggcttgccagccgctgggcgccgttctctgcgccctcggtttcgagaagaccatgccc
tacgtgcacggttcccagggttgcgtcgcctacttccgctcctacttcaaccgtcatttccgcgagccggtttcctgcgtttccgactccat
gaccgaagacgcggcagtgttcggcggccagcagaacatgaaggacggtctgcagaactgtaaggctacctacaagcccgacatg
atcgcagtgtccaccacctgcatggccgaggtcatcggtgacgacctcaacgccttcatcaacaactcgaagaaggaaggtttcattcc
tgacgagttcccggtgccgttcgcccataccccgagcttcgtgggcagccacgtgaccggctgggacaacatgttcgaaggcattgct
cgctacttcaccctgaagtccatggacgacaaggtggttggcagcaacaagaagatcaacatcgtccccggcttcgagacctacctgg
gcaacttccgcgtgatcaagcgcatgctttcggaaatgggcgtgggctacagcctgctctccgatccggaagaagtgctggacacccc
ggctgacggccagttccgcatgtacgcgggcggcaccactcaggaagagatgaaggacgctccgaacgccctcaacaccgtcctg
ctgcagccgtggcacctggagaagaccaagaagttcgtcgagggtacctggaagcacgaagtaccgaagctgaacatcccgatggg
cctggactggaccgacgagttcctgatgaaagtcagcgaaatcagcggccagccgattccggcgagcctgaccaaggagcgtggcc
gtctggtc**ggc**atgatgacc**ggt**tcccacacctggctgcacggcaagcgtttcgccctgtggggtgatccggacttcgtgatgggcct
ggtcaagttcctgctggaactgggttgcgagccggtacacattctctgccacaacggcaacaagcgttggaagaaggcggtcgacgc
catcctcgccgcttcgccctacggcaagaatgctaccgtctacatcggcaaggacctgtggcacctgcgttcgctggtcttcaccgaca
agccggacttcatgatcggcaacagctacggtaagttcatccagcgcgacaccctgcacaagggcaaggagttcgaggttccgctga
tccgtatcggcttcccgatcttcgaccgtcatcacctgcatcgctccaccaccctgggttacgagggcgccatgcagatcctgaccacc
ctggtgaactcgatcctggaacgtctggacgaggaaacccgcggtatgcaggccaccgactacaaccacgacctggtacgctaagtc
gtcggttcaagtggtatcggccggagcggcgcaagccgctccggctcttccttggcggcggccgcaggtggtcgggccttttgcccg
cgatctgcggcaaccgccaaacccgtctaaggagcaagcccatgcccagcgtcatgattcgccgcaacgacgaaggccaactgacc
ttctatatcgccaagaaagaccaggaagagatcgtggtgtccctggagcatgacagccccgaactctggggtggcgaagtcaccctc
ggcgacggttcgacctatttcatcgagccgatacc

NifK (ß-D353G/D357G):

MSQQVDKIKASYPLFLDQDYKDMLAKKRDGFEEKYPQDKIDEVFQWTTTKEYQELN
FQREALTVNPAKACQPLGAVLCALGFEKTMPYVHGSQGCVAYFRSYFNRHFREPVS
CVSDSMTEDAAVFGGQQNMKDGLQNCKATYKPDMIAVSTTCMAEVIGDDLNAFIN
NSKKEGFIPDEFPVPFAHTPSFVGSHVTGWDNMFEGIARYFTLKSMDDKVVGSNKKI
NIVPGFETYLGNFRVIKRMLSEMGVGYSLLSDPEEVLDTPADGQFRMYAGGTTQEEM
KDAPNALNTVLLQPWHLEKTKKFVEGTWKHEVPKLNIPMGLDWTDEFLMKVSEISG
QPIPASLTKERGRLV**G**MMT**G**SHTWLHGKRFALWGDPDFVMGLVKFLLELGCEPVHI
LCHNGNKRWKKAVDAILAASPYGKNATVYIGKDLWHLRSLVFTDKPDFMIGNSYGK

FIQRDTLHKGKEFEVPLIRIGFPIFDRHHLHRSTTLGYEGAMQILTTLVNSILERLDEET
RGMQATDYNHDLVR*

Purification of wild-type and ß-D353G/D357G MoFe protein. The procedures outlined here were developed previously with minor modifications.[16] Cell pellets were anaerobically thawed and resuspended for 20 minutes in lysis buffer (50 mM Tris/HCl pH 8, containing 5 mM dithionite (abbreviated as DT), and 37% v/v glycerol) with a ratio 2:1 (buffer:cells v/w), and subsequently collected by centrifugation at 12,000 x *g* (20 min at 4 °C). Once again inside the glovebox, the glycerol-containing supernatant was discarded, and cells were rapidly lysed by osmotic shock/shaking following the introduction of an equal replacement volume of glycerol-free lysis buffer (in the presence of a few µg of DNase). After incubation on ice for approximately 15 minutes, the cell lysate was clarified by centrifugation at 30,000 x *g* (4 °C, 1 hour) to yield a dark brown supernatant.

Note: From here onwards, DT was omitted from all buffers.

Within the glovebox, the supernatant was decanted and diluted with a post-lysis buffer (2 M NaCl, 234 mM Tris/HCl, pH 8.0), giving a final NaCl concentration of 0.3 M NaCl. His-tagged "wild-type" and ß-D353G/D357G MoFe protein were purified by Ni-affinity chromatography using a Ni-loaded HisTrap HP 5 mL column (Cytiva) with a flow rate of 5 mL/min using an Äkta Start system within the glovebox. The MoFe-containing supernatant was loaded onto the column, washed with an equilibration buffer for at least 3 column volumes to remove remaining DT (0.3 M NaCl, 50 mM Tris/HCl, pH 8.0), washed with a 20 mM imidazole step, and finally eluted with a 0.3 M imidazole step (0.3 M imidazole, 0.3 M NaCl, 50 mM Tris/HCl, pH 8.0). The eluted MoFe protein was then purified further with a HiPrep Q-Sepharose FF 16/10 column (Cytiva, 20 mL column volume, flow rate = 5 mL/min), which also served to remove excess imidazole and residual DT. The concentration of NaCl in the MoFe protein samples was first lowered to 0.1 M NaCl using a dilution buffer (50 mM Tris/HCl, pH 8.0) before being loaded on the pre-equilibrated Q-sepharose column (0.1 M NaCl). MoFe proteins were eluted at approximately 0.3 M NaCl using a NaCl gradient over 7 column volumes (from 0.2 to 0.65 M NaCl). ß-D353G/D357G MoFe protein prepared for crystallization experiments were subjected to an additional size-exclusion chromatography step on a HiPrep 26/60 Sephacryl S-200HR column (Cytiva) with a flow rate of 2.5 mL/min. Tris running buffer (50 mM Tris/HCl, pH 8.0, 0.5 M NaCl) was used for the elution. MoFe proteins were typically concentrated to >20 mg/mL using a Merck Millipore stirred concentrator cell (within the glovebox) equipped with a 100 kDa molecular weight cut-off

membrane (fed with ultra-high-purity N_2 5.0), flash-frozen as 20 μL pellets in liquid nitrogen and stored in liquid nitrogen until further use. Protein concentrations were initially estimated by the Biuret method using bovine serum albumin as the standard and confirmed by Mo quantification using ICP-MS.

Purification of Fe protein. This procedure was performed as recently reported.[16] In contrast to the MoFe proteins, Fe protein was purified in the presence of 2 mM DT in all buffers (5 mM DT was included in the cell lysis step, as above). Fe protein was routinely purified in tandem with the MoFe protein using an Äkta Go system. Following cell lysis, the supernatant was diluted with a post-lysis buffer (2 M NaCl, 234 mM Tris/HCl, 2 mM DT, pH 8.0) in order to obtain NaCl concentration of 0.1 M NaCl. WT Fe protein was purified from His-column flow-through fractions (above), where the flow-through fraction was first diluted to 0.1 M NaCl using the dilution buffer above (containing DT). The Fe protein was first purified using a HiPrep Q-Sepharose HP 16/10 column with a flow rate of 5 mL/min. Fe protein elution was performed using a NaCl gradient (7 column volumes, from 0.2 to 0.65 M NaCl) and concentrated to <10 mL using a Merck-Millipore stirred concentrator cell (30 kDa molecular weight cut-off membrane). The Fe protein was then purified further by size-exclusion chromatography (HiPrep 26/60 Sephacryl S-200 HR, Cytiva) with a flow rate of 2.5 mL/min, using a Tris running buffer (50 mM Tris/HCl, pH 8.0, 0.5 M NaCl). Eluted Fe protein sample was concentrated to >20 mg/mL, flash frozen in liquid nitrogen as 20 μL pellets and stored in liquid nitrogen until further use. Protein concentrations were determined by the Biuret method using bovine serum albumin as the standard.

Activity assays. Activity assays were performed as recently reported.[16] Substrate reduction activity assays were conducted in triplicate 1 mL reactions in 13 mL septum-sealed glass vials (Wheaton) containing deoxygenated buffer and an ATP-regenerating system (5 mM ATP, 30 mM phosphocreatine, 1.3 mg BSA, 0.2 mg creatine phosphokinase (from Rabbit muscle, Merck Switzerland), 10 mM sodium dithionite and 100 mM MOPS/NaOH at pH 7.0). Activity assays contained 0.1 mg MoFe protein and either 16.6 molar equivalents of Fe protein (0.48 mg, “high-flux”), or 4 molar equivalents of Fe protein (0.12 mg, “low-flux”). All reaction vials were sealed within an Ar-filled glovebox (Jacomex, France) and vented to atmospheric pressure. Where necessary, vials were flushed with ultra-high-purity N_2 (or desired quantities were introduced using gas-tight syringes); all vials were vented to atmospheric pressure before reactions were started. Reactions were performed within a shaking water bath (30 °C) and initiated by the addition of $MgCl_2$ (from a 1 M stock, 10 mM final concentration) using a gas-tight syringe. Reactions were quenched after 8 minutes by the

addition of 300 µL of 400 mM EDTA (pH 8.0). H_2 quantification was performed on all reactions using a calibrated GC-TCD equipped with a molecular sieve 5 Å column (Ar carrier, SRI Instruments model 8610C). Ammonia was quantified by the *ortho*-phthalaldehyde method (corrected to controls and assays performed under 1 atm Ar) using NH_4Cl as the standard, as reported previously.[17-18]

O_2 protection assays. Activity assays were performed using the "FeSII" (or "Shethna") protein of *A. vinelandii*, heterologously produced in *E. coli*. The protein was expressed and purified as previously published, with a single modification: in this work, the gene corresponding to the FeSII protein was introduced to a pET21a vector (Merck Millipore).[19] The ability of the FeSII protein to protect nitrogenase from oxidation by O_2 was determined based on the quantity of H_2 produced in a nitrogenase assay following exposure to O_2 for 10 minutes. Assays employed the same buffer components as listed above, except that DT was omitted and $MgCl_2$ (10 mM) was included in the activity assay buffer. All protein components (when used) were present at a concentration of 0.4 µM for these assays. All reaction vials were first assembled within an Ar-filled glovebox (Jacomex, France) and vented to atmospheric pressure. When required, O_2 (final concentration of 2% v/v) was added using gas-tight syringes, and the vials were incubated for 10 min at 30 °C. Vials were then flushed with ultra-high purity Ar gas (5.0, consisting of 3x vacuum/flushing cycles). All reaction vials were once again vented to atmospheric pressure. Reactions were performed in a shaking water bath (30 °C) and initiated by the addition of sodium dithionite (from a 1 M stock, 20 mM final concentration) using a gas-tight syringe. Reactions were quenched after 8 minutes by the addition of 300 µL of 400 mM EDTA (pH 8.0).

Analytical Gel Filtration. Analytical gel filtration was performed using an Äkta Go within an anoxic glovebox.[16] The proteins were diluted with the running buffer (50 mM Tris/HCl, 500 mM NaCl, pH 8, 0.5 mM DT) to a concentration of 20 µmol/mL for MoFe protein. 200 µL aliquots were prepared and incubated for one hour before being loaded (100 µL) onto the pre-equilibrated column.

Limited Proteolysis. A 50 mM Tris/HCl (pH 7.6) buffer containing 10 mM $CaCl_2$ buffer was prepared and deoxygenated in an Ar-filled glovebox (Jacomex, France) overnight. The buffer was split into 2 bottles, and 0.2 mg/mL fresh trypsin (Sigma) was added to the first. This trypsin-containing buffer was then activated at 30 °C for 20 minutes using a heating block. Proteins were brought into the glovebox and diluted to a concentration of 1 mg/mL using the buffer that did not contain trypsin. The proteins were placed in LC/MS vials with stirring in a

water bath at 37 °C, and the activated trypsin was then added to the protein samples to a final concentration of 2 μg/mL (trypsin, 500:1 MoFe:trypsin) as the starting point for the reaction.

The protein reactions were sampled at time 0 (before trypsin addition), 2, 5, 30, 60 and 90 minutes after trypsin addition. 4 μL of samples of the reactions were transferred to Eppendorf tubes that already contained 4 μL TruPAGE™ LDS (4X) sample buffer (Merck, Switzerland), then immediately placed at 95 °C on a heating block for 5 minutes to deactivate the trypsin and terminate proteolysis. 1.6 μL 1 M dithiothreitol (DTT) and 6.4 μL MilliQ water were added afterwards. 4-20% mPAGE® Bis-tris gels (Merck, Switzerland) were used for analysis. 7 μL aliquots of the treated protein samples were loaded in each well (1.75 μg MoFe protein). The gel was run at 200 V and stained with One-Step Blue® (Biotium, Brunschwig Switzerland).

Circular dichroism. All circular dichroism experiments were performed using a sealed cuvette with a 1 mm pathlength. The anoxic protein samples were loaded in the cuvette within an anoxic glovebox (Ar atmosphere, <1 ppm O_2). Protein samples were diluted to 0.2 mg/mL in 50 mM Tris/HCl, 0.5 M NaCl pH 8.0 buffer. Circular dichroism spectra were recorded firstly to determine lambda max (224 nm for both MoFe proteins). For the melting point experiments, the temperature was ramped 1°C per minute starting from 20 to 80 °C. Reactions were conducted as triplicates for both MoFe proteins.

Crystallization of the β-D353G/D357G MoFe-protein. The β-D353G/D357G MoFe-protein was crystallized anaerobically at 17.5 mg/mL^{-1} under 100% N_2 atmosphere using the OryxNano Crystallization Robot (Douglas Instruments Ltd, Berkshire, United Kingdom). Prior to crystallization, the sample was centrifuged at 13,000 x *g* for 3 min to remove macro-aggregates and dust. The protein was spotted as a sitting drop to 96-Well MRC 2-Drop polystyrene Crystallization Plates (SWISSCI) containing 90 μL of crystallization solution in the reservoir. The following crystallization screens were used to harvest crystals JCSG++, XP Screen, PEP, Wizard 1/2, Wizard 3/4, and PACT++ (Jena Bioscience, Germany). Each drop contained 0.5 μL of protein sample and 0.5 μL of crystallization solution. Crystals were obtained in the crystallization solution containing 10.00 % w/v Polyethylene glycol 10,000, 2.00 % v/v 1,4-Dioxane, 100 mM tri-Sodium citrate; pH 5.6, and 1 mM polyoxotungstate $[TeW_6O_{24}]^{6-}$ (TEW). Sealed plates were stored inside a Coy anaerobic chamber (Coy Laboratory Products Inc, Grass Lake, USA) with an atmosphere of N_2:H_2 97:3 at 20 °C. Crystals were soaked in the crystallization solution supplemented with 30% v/v ethylene glycol for a few seconds before freezing in liquid nitrogen.

Data processing and refinement. Crystals were tested and collected at 100 K at the Swiss Light Source X06DA (Villigen, Switzerland). Datasets were collected at 1.73648-Å (7140 eV) for the single-wavelength anomalous dispersion experiment. It must be specified here that the 16th Fe has an absorption edge 5 eV higher in comparison to the Fe atoms composing the two other metallocofactors, and the data presented here were collected at an energy of 7140 eV for this purpose (the theoretical absorption of Fe as its K-edge is 7112 eV).[20] Native datasets were collected at a wavelength of 1.00002-Å on the same crystal. Data were processed with autoPROC run with staraniso due to the high anisotropy.[21] For the Fe-K edge dataset, the diffraction limits along the principal axes are the following: a = 2.43 Å, b= 2.27 Å, and c= 1.67 Å. For the dataset collected close to the Se K edge, the diffraction limits along the principal axes are the following: a = 2.33 Å, b= 2.04 Å , and c= 1.53 Å. The structure of the variant was solved by using PHENIX with the following template: 3U7Q. The model was manually rebuilt with COOT and further refined with PHENIX.[22] During the refinement, a translational-libration screw was applied with the model containing hydrogens added in the riding position during the last refinement cycles. Hydrogens were removed in the final deposited model. The model was validated through the MolProbity server.[23] Data collection and refinement statistics for the deposited model and structure factors are listed in Table S3. The model is accessible via the PDB ID 8P8G. Figures were generated with PyMOL (Schrödinger, LLC).

Table S2. Comparison in the presence of the 16th Fe in the deposited structures.

Entity ID	Resolution [Å]	PDB title	Organism	MMB site	Metal occupancy	Year of release
1MIO	3.00	X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0 angstroms resolution	*Clostridium pasteurianum*	Ca^{2+}	1.00	1993
1N2C	3.00	Nitrogenase complex from *Azotobacter vinelandii* stabilized by ADP-tetrafluoroaluminate	*Azotobacter vinelandii*	Ca^{2+}	1.00	1997
2MIN	2.03	nitrogenase MoFe protein from *Azotobacter vinelandii*, oxidized state	*Azotobacter vinelandii*	Ca^{2+}	1.00	1997
3MIN	2.03	nitrogenase MoFe protein from *Azotobacter vinelandii*, oxidized state	*Azotobacter vinelandii*	Ca^{2+}	1.00	1997
1QGU	1.60	nitrogenase MoFe protein from *Klebsiella pneumoniae*, dithionite-reduced state	*Klebsiella pneumoniae*	Mg^{2+}	1.00	1999
1QH1	1.60	nitrogenase MoFe protein from *Klebsiella pneumoniae*, phenosafranin oxidized state	*Klebsiella pneumoniae*	Mg^{2+}	1.00	1999
1QH8	1.60	nitrogenase MoFe protein from *Klebsiella pneumoniae*, as crystallized (mixed oxidation) state	*Klebsiella pneumoniae*	Mg^{2+}	1.00	1999
1G20	2.20	MgATP-bound and nucleotide-free structures of a nitrogenase protein complex between the Leu127Del-Fe-protein and the MoFe-protein	*Azotobacter vinelandii*	Ca^{2+}	1.00	2001
1G21	3.00	MgATP-bound and nucleotide-free structures of a nitrogenase protein complex between the Leu127Del-Fe-protein and the MoFe-protein	*Azotobacter vinelandii*	Ca^{2+}	1.00	2001
1FP4	2.50	crystal structure of the alpha-H195Q mutant of nitrogenase	*Azotobacter vinelandii*	Ca^{2+}	1.00	2002
1L5H	2.30	FeMo-cofactor deficient nitrogenase MoFe Protein	*Azotobacter vinelandii*	Ca^{2+}	1.00	2002
1H1L	1.90	nitrogenase MoFe protein from *Klebsiella pneumoniae*, NifV mutant	*Klebsiella pneumoniae*	Mg^{2+}	1.00	2002
1MIN	1.16	nitrogenase MoFe protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Ca^{2+}	1.00	2002
1M1Y	3.20	chemical crosslink of nitrogenase MoFe Protein and Fe Protein	*Azotobacter vinelandii*	Ca^{2+}	1.00	2003
1M34	2.30	nitrogenase complex from *Azotobacter Vinelandii* stabilized By ADP-tetrafluoroaluminate	*Azotobacter vinelandii*	Ca^{2+}	1.00	2003
2AFH	2.10	crystal structure of nucleotide-free Av2-Av1 complex	*Azotobacter vinelandii*	Ca^{2+}	1.00	2005
2AFI	3.10	crystal structure of MgADP bound Av2-Av1 complex	*Azotobacter vinelandii*	Ca^{2+}	1.00	2005
3K1A	2.23	insights into substrate binding at FeMo-cofactor in nitrogenase from the structure of an alpha-70Ile MoFe protein variant	*Azotobacter vinelandii*	Ca^{2+}	1.00	2010
3U7Q	1.00	*Azotobacter vinelandii* nitrogenase MoFe protein at atomic resolution	*Azotobacter vinelandii*	Ca^{2+}	1.00/0.60	2011
4ND8	2.00	Av nitrogenase MoFe protein high pH form	*Azotobacter vinelandii*	Fe^{3+}	1.00	2014
4TKU	1.43	reactivated nitrogenase MoFe-protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{2+}	0.50	2014
4TKV	1.50	CO-bound nitrogenase MoFe-protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{2+}	0.50	2014
4WZA	1.90	asymmetric nucleotide binding in the nitrogenase complex	*Azotobacter vinelandii*	Fe^{3+}	0.61/0.59	2014
4WES	1.08	nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 1.08 A resolution	*Clostridium pasteurianum*	Fe^{2+}	0.47/0.52	2015
4WZB	2.30	crystal structure of MgAMPPCP-bound Av2-Av1 complex	*Azotobacter vinelandii*	Fe^{2+}	0.4/0.57	2015
4WN9	1.90	structure of the nitrogenase MoFe protein from Clostridium pasteurianum pressurized with xenon	*Clostridium pasteurianum*	Fe^{3+}	0.59	2015
4WNA	2.00	structure of the nitrogenase MoFe protein from *Azotobacter vinelandii* pressurized with xenon	*Azotobacter vinelandii*	Fe^{3+}	0.50	2015
5CX1	1.75	nitrogenase molybdenum-iron protein beta-K400E mutant	*Azotobacter vinelandii*	Ca^{2+}	1.00	2015
4XPI	1.97	Fe protein independent substrate reduction by nitrogenase variants altered in intramolecular electron transfer	*Azotobacter vinelandii*	Ca^{2+}	1.00	2015
5BVG	1.60	selenium incorporated nitrogenase MoFe-protein (Av1-Se2B) from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{2+}	0.50	2015
5BVH	1.53	CO-bound form of selenium incorporated nitrogenase MoFe-protein (Av1-Se-CO) from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{2+}	0.50	2015
5KOJ	2.59	nitrogenase MoFeP protein in the IDS oxidized state	*Gluconacetobacter diazotrophicus* PA1 5	Fe^{3+}	1.00	2016
5KOH	1.83	nitrogenase MoFeP protein from *Gluconacetobacter diazotrophicus* in dithionite reduced state	*Gluconacetobacter diazotrophicus* PA1 5	Fe^{3+}	1.00	2016
5N6Y	1.35	*Azotobacter vinelandii* vanadium nitrogenase	*Azotobacter vinelandii*	Mg^{2+}	1.00	2017
5VPW	1.85	Nitrogenase Cp1 at pH 5	*Clostridium pasteurianum*	Fe^{2+}	0.50	2017
5VQ3	1.72	Nitrogenase Cp1 at pH 6.5	*Clostridium pasteurianum*	Fe^{2+}	0.50	2017
5VQ4	2.30	nitrogenase Av1 at pH 5	*Azotobacter vinelandii*	Fe^{3+}	0.50	2017
6BBL	1.68	crystal structure of the a-96Gln MoFe protein variant in the presence of the substrate acetylene	*Azotobacter vinelandii*	Fe^{3+}	1.00/0.94	2018
6FEA	1.20	*A. vinelandii* vanadium nitrogenase, turnover state	*Azotobacter vinelandii*	Mg^{2+}	1.00	2018
6CDK	2.10	characterization of the P1+ intermediate state of nitrogenase P cluster	*Azotobacter vinelandii*	Fe^{3+}	1.00	2018
6O7L	2.26	nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	*Azotobacter vinelandii*	Fe^{3+}	0.59/0.48	2019
6O7M	1.40	nitrogenase MoFeP mutant F99Y from *Azotobacter vinelandii* in the indigo carmine oxidized state	*Azotobacter vinelandii*	Fe^{3+}	0.67	2019
6O7N	1.75	nitrogenase MoFeP mutant F99Y/S188A from *Azotobacter vinelandii* in the indigo carmine oxidized state	*Azotobacter vinelandii*	Fe^{3+}	0.77	2019
6O7O	1.89	nitrogenase MoFeP mutant F99Y/S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	*Azotobacter vinelandii*	Fe^{3+}	0.61	2019
6O7P	1.70	nitrogenase MoFeP mutant F99Y from *Azotobacter vinelandii* in the dithionite reduced state	*Azotobacter vinelandii*	Fe^{3+}	0.70/0.69	2019
6O7Q	2.00	nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	*Azotobacter vinelandii*	Fe^{3+}	0.81/0.77	2019
6O7R	2.27	nitrogenase MoFeP mutant F99Y, S188A from *Azotobacter vinelandii* in the dithionite reduced state	*Azotobacter vinelandii*	Fe^{3+}	0.65/0.58	2019
6O7S	2.27	nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the indigo carmine oxidized state	*Azotobacter vinelandii*	Fe^{3+}	1.00/0.86	2019
6OP1	1.70	selenium incorporated, carbon monoxide inhibited FeMo-cofactor of	*Azotobacter vinelandii*	Ca^{2+}	1.00	2019

		Azotobacter vinelandii				
6OP2	1.90	selenium incorporated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii* at high concentration of selenium	*Azotobacter vinelandii*	Ca^{2+}	1.00	2019
6OP3	1.60	selenium incorporated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii* with low concentration of selenium	*Azotobacter vinelandii*	Ca^{2+}	1.00	2019
6OP4	2.30	selenium-incorporated, carbon monoxide-inhibited, reactivated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Ca^{2+}	1.00	2019
6UG0	1.83	N_2-bound nitrogenase MoFe-protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{3+}	0.19/0.14	2020
6VXT	1.74	activated nitrogenase MoFe-protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Fe^{3+}	0.45/0.41	2020
7ADR	1.00	CO bound as bridging ligand at the active site of vanadium nitrogenase VFe protein	*Azotobacter vinelandii*	Mg^{2+}	1.00	2020
7ADY	1.05	CO-removed state of the active site of vanadium nitrogenase VFe protein	*Azotobacter vinelandii*	Mg^{2+}	1.00	2020
7JRF	1.33	CO-CO-bound nitrogenase MoFe-protein from *Azotobacter vinelandii*	*Azotobacter vinelandii*	Ca^{2+}	0.70/0.60	2021
7AIZ	1.05	vanadium nitrogenase Vfe protein, high CO state	*Azotobacter vinelandii*	Mg^{2+}	1.00	2021
7MCI	1.65	MoFe protein from *Azotobacter vinelandii* with a sulfur-replenished cofactor	*Azotobacter vinelandii*	Ca^{2+}	0.56/0.55	2022
7UT6	1.91	C1 symmetric cryoEM structure of *Azotobacter vinelandii* MoFeP under non-turnover conditions	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
7UT7	1.91	C2 symmetric cryoEM structure of *Azotobacter vinelandii* MoFeP under non-turnover conditions	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
7UT8	2.43	cryoEM structure of *Azotobacter vinelandii* nitrogenase complex (1:1 FeP:MoFeP, ADP/ATP-bound) during catalytic N_2 reduction	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
7UT9	2.44	cryoEM structure of Azotobacter vinelandii nitrogenase complex (1:1 FeP:MoFeP, ATP-bound) during catalytic N_2 reduction	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
7UTA	2.40	cryoEM structure of Azotobacter vinelandii nitrogenase complex (2:1 FeP:MoFeP) inhibited by BeFx during catalytic N_2 reduction	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
8PDN	2.49	cryoEM structure of Azotobacter vinelandii nitrogenase MoFeP during catalytic N_2 reduction	*Azotobacter vinelandii*	Fe^{3+}	1.00	2022
8E3T	2.20	gallium-reconstituted nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* after IDS oxidation	*Azotobacter vinelandii*	Fe^{3+}	0.78/0.76	2022
8E3U	1.99	nickel-reconstituted nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* after IDS oxidation	*Azotobacter vinelandii*	Fe^{3+}	0.84/0.80	2022
8E3V	2.00	cobalt-reconstituted nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* after IDS oxidation	*Azotobacter vinelandii*	Fe^{3+}	0.72	2022
8CRS	2.04	cryoEM structure of partially oxidized MoFe-protein on ultrathin carbon	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023
8DBX	1.92	cryoEM structure of nitrogenase MoFe-protein in detergent	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023
8ENL	2.37	cryoEM structure of the high pH turnover-inactivated nitrogenase MoFe-protein	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023
8ENM	2.14	cryoEM structure of the high pH nitrogenase MoFe-protein under non-turnover conditions	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023
8ENN	2.58	homocitrate-deficient nitrogenase MoFe-protein from *Azotobacter vinelandii nifV* knockout	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023
8ENO	2.71	homocitrate-deficient nitrogenase MoFe-protein from *Azotobacter vinelandii nifV* knockout in complex with NafT	*Azotobacter vinelandii*	Fe^{3+}	1.00	2023

Table S3. X-ray analysis statistics for the presented structure.

	Fe K edge	Native
Data collection		
Wavelength (Å)	1.73648	1.00002
Space group	*C*2	*C*2
Resolution (Å)	101.48 – 1.70 (1.97 – 1.70)	101.46 – 1.55 (1.76– 1.55)
Cell dimensions		
a, b, c (Å)	167.39, 74.43, 208.45	167.49, 74.47, 208.47
α, β, γ (°)	90, 103.17, 90	90, 103.24, 90
R_{merge}(%)[a]	19.7 (97.0)	17.7 (189.0)
R_{pim} (%)[a]	6.8 (42.9)	6.6 (65.2)
$CC_{1/2}$[a]	0.973 (0.522)	0.995 (0.566)
I/σ_I[a]	8.6 (2.0)	8.1 (1.8)
Spherical completeness[a]	47.0 (6.6)	48.3 (7.2)
Ellipsoidal completeness[a]	87.8 (62.8)	94.0 (80.5)
Redundancy[a]	9.4 (7.2)	8.1 (9.3)
Nr. unique reflections[a]	129551 (6479)	176043 (8391)
Refinement		
Resolution (Å)		48.32 – 1.55
Number of reflections		175938
R_{work}/R_{free}[b] (%)		15.53/18.73
Number of atoms		
Protein		16000
Ligands/ions		285
Solvent		2176
Mean B-value (Å^2)		18.62
Molprobity clash score, all atoms		2.32
Ramachandran plot		
Favoured regions (%)		
Outlier regions (%)		0.20
rmsd[c] bond lengths (Å)		0.012
rmsd[c] bond angles (°)		1.035
PDB ID code		8P8G

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Table S4. Accession numbers for sequences used in conservation analysis.

ARCHAEA

Accession numbers for amino acid sequences

Taxon	NCBI ID
Methanosarcina spelaei	WP_095642672.1
Methanosarcina vacuolata	WP_048118825.1
Methanolobus halotolerans	WP_135389265.1
Methanosarcina soligelidi	WP_048050481.1
Methanosarcina horonobensis	WP_048137468.1
Methanosarcina acetivorans	WP_011023795.1
Methanolobus psychrotolerans	WP_094228262.1
Methanosarcina mazei	WP_011032674.1
Methanosarcina siciliae	WP_048173728.1
Methanolobus chelungpuianus	WP_256621605.1
Methanolobus profundi	WP_091937239.1
Methanolobus tindarius	WP_023845584.1
Methanolobus vulcani	WP_091709561.1
Methanolobus zinderi	WP_176964182.1
Methanolobus bombayensis	WP_209619861.1
Methanocella conradii	WP_014404756.1
Methanosphaerula palustris	WP_012617249.1
Methanomassiliicoccus luminyensis	WP_026069119.1
Methanospirillum lacunae	WP_245926443.1
Methanocella arvoryzae	WP_012035709.1
Methanogenium organophilum	WP_268187572.1
Methanofollis liminatans	WP_004037776.1
Methanolacinia paynteri	WP_048152030.1
Methanolacinia petrolearia	WP_013328216.1
Methanoregula boonei	WP_012106692.1
Methanoculleus taiwanensis	WP_128693523.1
Methanospirillum stamsii	WP_109942296.1
Methanolinea mesophila	WP_209673767.1
Methanocaldococcus vulcanius	WP_015733163.1
Methanocaldococcus infernus	WP_013099455.1
Methanobacterium lacus	WP_013645978.1
Methanotorris igneus	WP_013799861.1
Methanobacterium formicicum	WP_023991282.1
Methanobrevibacter curvatus	WP_067088896.1
Methanothermococcus thermolithotrophicus	WP_018154781.1
Methanobacterium paludis	WP_013826518.1
Methanococcus vannielii	WP_011971879.1
Methanotorris formicicus	WP_007044852.1
Methanobacterium petrolearium	WP_245247905.1
Methanobacterium alkalithermotolerans	WP_211534065.1
Methanothermococcus okinawensis	WP_013867269.1
Methanothermobacter wolfeii	WP_074358289.1
Methanothermobacter marburgensis	WP_013294987.1
Methanothermobacter thermautotrophicus	WP_010877173.1
Methanococcus maripaludis	WP_104837233.1
Methanococcus aeolicus	WP_011974138.1
Methanobrevibacter cuticularis	WP_067259997.1
Methanobacterium congolense	WP_231916352.1
Methanobacterium alcaliphilum	WP_248611961.1
Methanobacterium aggregans	WP_245315979.1
Methanobrevibacter arboriphilus	WP_221061946.1
Methanobrevibacter filiformis	WP_066971018.1
Methanobacterium alcaliphilum	WP_248612368.1
Methanobrevibacter arboriphilus	WP_244987881.1
Methanobrevibacter filiformis	WP_066973458.1

Asp353[β], Asp357[β] and Glu109[β] not conserved
Asp353[β] and Glu109[β] not conserved
Glu109[β] not conserved

BACTERIA

Accession numbers for amino acid sequences

Taxon	NCBI ID
Azotobacter vinelandii	WP_012698833.1
Azotobacter salinestris	WP_152387826.1
Azotobacter chroococcum	WP_089166738.1
Azotobacter beijerinckii	WP_090619180.1
Pseudomonas oryzae	WP_090349184.1
Stutzerimonas azotifigens	WP_028240786.1
Pseudomonas linyingensis	WP_090313067.1
Pseudomonas aromaticivorans	WP_217681573.1
Pseudomonas lopnurensis	WP_193682425.1
Pseudomonas benzenivorans	WP_090444290.1
Pseudomonas sagittaria	WP_092432159.1
Azomonas agilis	WP_144570100.1
Azomonas macrocytogenes	WP_183166844.1
Pseudomonas tumuqii	WP_172148992.1
Stutzerimonas kirkiae	WP_131183836.1
Halomonas endophytica	WP_102654334.1
Teredinibacter turnerae	WP_045825325.1
Methylogaea oryzae	WP_221047314.1
Sessilibacter corallicola	WP_233089742.1
Teredinibacter haidensis	WP_075185652.1
Oceanobacter mangrovi	WP_221797276.1
Gynuella sunshinyii	WP_044616594.1
Candidatus Methylospira mobilis	WP_153248009.1
Amphritea atlantica	WP_091356309.1
Agaribacterium haliotis	WP_096085048.1
Amphritea pacifica	WP_205213473.1
Marinobacterium lutimaris	WP_104002707.1
Thiorhodospira sibirica	WP_006786684.1
Marinobacterium ramblicola	WP_217414969.1
Marinobacterium mangrovicola	WP_132286901.1
Marinobacterium zhoushanense	WP_188748033.1
Acidihalobacter aeolianus	WP_070071969.1
Methylomonas methanica	WP_013818144.1
Parasulfuritortus cantonensis	WP_206199457.1
Thioalbus denitrificans	WP_114278828.1
Imhoffiella purpurea	WP_043753449.1
Thiorhodococcus mannitoliphagus	WP_164652993.1
Methylotuvimicrobium buryatense	WP_017842413.1
Thiorhodococcus drewsii	WP_007041617.1
Parazoarcus communis	WP_110523395.1
Thioflexithrix psekupsensis	WP_086487136.1
Methylomonas rhizoryzae	WP_150047214.1
Marinobacterium litorale	WP_027854784.1
Thiocapsa bogorovii	WP_232409645.1
Methylicorpusculum oleiharenae	WP_159653136.1
Acidihalobacter prosperus	WP_038093040.1
Thiocapsa roseopersicina	WP_093032656.1
Thiocapsa rosea	WP_120795823.1
Thiobaca trueperi	WP_132977828.1
Thiocystis violascens	WP_014779898.1
Thiocystis minor	WP_201097629.1
Beggiatoa leptomitoformis	WP_201800126.1
Thiothrix caldifontis	WP_093070871.1
Pseudothauera hydrothermalis	WP_102041592.1
Plasticicumulans acidivorans	WP_110019778.1
Thiothrix unzii	WP_210217628.1
Allochromatium humboldtianum	WP_176975874.1
Allochromatium warmingii	WP_091333974.1
Chromatium weissei	WP_200206373.1
Ferrigenium kumadai	WP_212786957.1
Chromatium okenii	WP_200153176.1
Lamprocystis purpurea	WP_020505236.1
Methylomonas lenta	WP_066980561.1
Methylomonas koyamae	WP_054761834.1
Sideroxydans lithotrophicus	WP_013029019.1
Sideroxyarcus emersonii	WP_237248192.1
Thiocapsa marina	WP_007192137.1
Thiocystis violacea	WP_200374461.1
Allochromatium vinosum	WP_012971076.1
Pseudothauera lacus	WP_107492235.1
Thiothrix subterranea	WP_202715972.1
Thiothrix litoralis	WP_210223912.1
Ferriphaselus amnicola	WP_062627001.1
Allochromatium palmeri	WP_155450425.1
Thiorhodococcus minor	WP_164453656.1
Candidatus Methylobacter oryzae	WP_127029523.1
Candidatus Methylobacter favarea	WP_174625051.1
Candidatus Accumulibacter contiguus	WP_169071201.1
Rhabdochromatium marinum	WP_200284055.1
Marichromatium purpuratum	WP_005224877.1
Methylobacter tundripaludum	WP_006892493.1
Magnetovibrio blakemorei	WP_069958445.1
Candidatus Thiodictyon syntrophicum	WP_100918828.1
Lamprobacter modestohalophilus	WP_200245958.1
Pseudothauera rhizosphaerae	WP_136386576.1
Methylomonas paludis	WP_215582523.1
Thiococcus pfennigii	WP_200383095.1
Ectothiorhodospira shaposhnikovii	WP_200177000.1
Marichromatium bheemlicum	WP_168670444.1
Thiospirillum jenense	WP_182583739.1
Ectothiorhodospira magna	WP_090203320.1
Halochromatium glycolicum	WP_200345181.1
Thiothrix nivea	WP_002710381.1
Thioflavicoccus mobilis	WP_015282049.1
Methylobacter marinus	WP_020159951.1
Thiorhodovibrio winogradskyi	WP_201065493.1
Sedimenticola thiotaurini	WP_046860623.1
Thermochromatium tepidum	WP_153975377.1
Pseudothauera nasutitermitis	WP_136348403.1
Pseudomonas kuykendallii	WP_090231403.1
Methylomonas albis	WP_192376221.1
Methylobacter luteus	WP_027158572.1
Rhodobium orientis	WP_111435550.1
Halochromatium roseum	WP_201220328.1
Methyloprofundus sedimenti	WP_080523747.1
Methylovulum miyakonense	WP_019865783.1
Thiohalocapsa halophila	WP_200232875.1
Azoarcus indigens	WP_133593340.1

BACTERIA (continued)

Accession numbers for amino acid sequences

Taxon	NCBI ID
Sulfurivermis fontis	WP_127477633.1
Mariprofundus ferrooxydans	WP_018294619.1
Methylovulum psychrotolerans	WP_088617647.1
Azospira restricta	WP_203388505.1
Rivicola pingtungensis	WP_110389901.1
Ectothiorhodospira mobilis	WP_090485544.1
Halochromatium salexigens	WP_201246178.1
Candidatus Thiosymbion oneisti	WP_133511476.1
Cohaesibacter haloalkalitolerans	WP_119308534.1
Solemya velesiana gill symbiont	WP_078486469.1
Caldichromatium japonicum	WP_246237554.1
Azonexus fungiphilus	WP_121459394.1
Candidatus Thiodiazotropha endolucinida	WP_069124652.1
Neiella holothuriorum	WP_220104180.1
Marinospirillum perlucidum	WP_114418418.1
Cohaesibacter marisflavi	WP_090068489.1
Marinospirillum alkaliphilum	WP_084662180.1
Candidatus Thiodiazotropha endoloripes	WP_069005826.1
Martelella endophytica	WP_045685338.1
Varunaivibrio sulfuroxidans	WP_132939447.1
Ferribacterium limneticum	WP_226498258.1
Cohaesibacter intestini	WP_114011043.1
Nitrogeniibacter aestuarii	WP_230970094.1
Nitrincola alkalisediminis	WP_193494681.1
Agarivorans albus	WP_016403296.1
Agarivorans litoreus	WP_220721311.1
Agarivorans gilvus	WP_055731568.1
Candidatus Accumulibacter aalborgensis	WP_186407934.1
Nitrincola iocasae	WP_151053119.1
Nitrogeniibacter mangrovi	WP_173768238.1
Dechloromonas hortensis	WP_153130690.1
Quatrionicoccus australiensis	WP_226442692.1
Neiella marina	WP_087506926.1
Tolumonas osonensis	WP_188025070.1
Tolumonas lignilytica	WP_024872643.1
Rhodocyclus gracilis	WP_153472810.1
Candidatus Competibacter phosphatis	WP_169249285.1
Rhodocyclus tenuis	WP_153117591.1
Candidatus Propionivibrio aalborgensis	WP_186410055.1
Candidatus Contendobacter odensis	WP_034434632.1
Cognatazoarcus halotolerans	WP_159682678.1
Dechloromonas denitrificans	WP_066879933.1
Agarivorans aestuarii	WP_221073737.1
Halorhodospira halophila	WP_011813089.1
Echinimonas agarilytica	WP_251261869.1
Tolumonas auensis	WP_012728880.1
Azonexus hydrophilus	WP_028996170.1
Reinekea marinisedimentorum	WP_132702266.1
Motiliproteus sediminis	WP_210394706.1
Reinekea thalattae	WP_147711941.1
Shewanella mangrovi	WP_037441804.1
Azovibrio restrictus	WP_026687659.1
Aliagarivorans taiwanensis	WP_026958339.1
Aliagarivorans marinus	WP_026971406.1
Shewanella dokdonensis	WP_213680929.1
Shewanella yunxiaonensis	WP_212594381.1
Uliginosibacterium gangwonense	WP_018607904.1
Shewanella avicenniae	WP_207355365.1
Halorhodospira abdelmalekii	WP_200194248.1
Vibrio diazotrophicus	WP_102962304.1
Vibrio ziniensis	WP_165313368.1
Aestuariirhabdus litorea	WP_125017132.1
Psychromonas antarctica	WP_238742049.1
Vibrio xiamenensis	WP_093275485.1
Halorhodospira neutriphila	WP_200257700.1
Celerinatantimonas diazotrophica	WP_131911474.1
Sodalis ligni	WP_132921729.1
Propionivibrio dicarboxylicus	WP_218122589.1
Vibrio albus	WP_109319112.1
Vibrio mangrovi	WP_087479241.1
Niveibacterium microcysteis	WP_206252505.1
Niveibacterium umoris	WP_183632637.1
Halorhodospira halochloris	WP_096409263.1
Vibrio cincinnatiensis	WP_078924457.1
Magnetococcus marinus	WP_011712866.1
Candidatus Dactylopiibacterium carminicum	WP_095523325.1
Dickeya dianthicola	WP_024107449.1
Propionivibrio limicola	WP_153108755.1
Vibrio astriarenae	WP_164648054.1
Brenneria alni	WP_121575299.1
Vibrio tritonius	WP_068716988.1
Vibrio aerogenes	WP_073601921.1
Dickeya fangzhongdai	WP_100850215.1
Mangrovibacter phragmitis	WP_064601338.1
Uliginosibacterium aquaticum	WP_170021172.1
Musicola keenii	WP_192457812.1
Vibrio quintilis	WP_073583165.1
Dickeya undicola	WP_033570862.1
Brenneria tiliae	WP_249243873.1
Dickeya solani	WP_022635032.1
Mangrovibacter yixingensis	WP_226569862.1
Musicola paradisiaca	WP_012764123.1
Vibrio viridaestus	WP_124938092.1
Mangrovibacter plantisponsor	WP_110025353.1
Formivibrio citricus	WP_091193706.1
Vibrio hippocampi	WP_237485982.1
Rahnella perminowiae	WP_217233633.1
Rahnella aquatilis	WP_014333919.1
Vibrio salinus	WP_232768100.1
Brenneria roseae	WP_109054135.1
Rouxiella aceris	WP_169402327.1
Vibrio hangzhouensis	WP_103878865.1
Affinibrenneria salicis	WP_150437453.1
Acerihabitans arboris	WP_162365675.1
Klebsiella indica	WP_138361699.1
Pectobacterium peruviense	WP_048259546.1
Brenneria salicis	WP_113865218.1
Pectobacterium atrosepticum	WP_039292652.1
Magnetofaba australis	WP_085442126.1
Brenneria izadpanahii	WP_208231071.1
Brenneria rubrifaciens	WP_137713596.1
Kosakonia quasisacchari	WP_131411155.1
Raoultella terrigena	WP_076947536.1
Kosakonia oryzae	WP_064566547.1
Kosakonia arachidis	WP_090120215.1
Klebsiella quasipneumoniae	WP_004203556.1
Kosakonia sacchari	WP_065368546.1

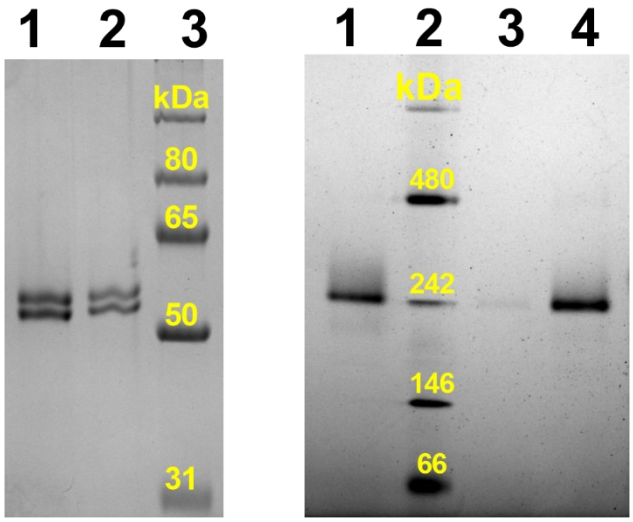


Figure S1. (Left) SDS-PAGE (12%, 200 V) analysis of wild-type and ß-D353G/D357G MoFe proteins, purified as detailed above. Coomassie blue stained and scanned in grayscale. Lanes: 1 = β-D353G/D357G MoFe, 2 = WT MoFe, 3 = molecular weight markers. 1.2 µg of protein was loaded in each well. (Right) Native-PAGE (3-12% acrylamide gradient, 150 V) analysis of the wild-type and ß-D353G/D357G MoFe proteins, performed under anoxic conditions (using deoxygenated buffers and performed within an anoxic glovebox filled with >95% N_2 and <5% H_2 atmosphere). Lanes: 1 = β-D353G/D357G MoFe, 2 = molecular weight markers, 3 = WT MoFe, 4= WT MoFe. Lanes 1/4 contained 1 µg protein, and lane 3 contained 0.4 µg protein.

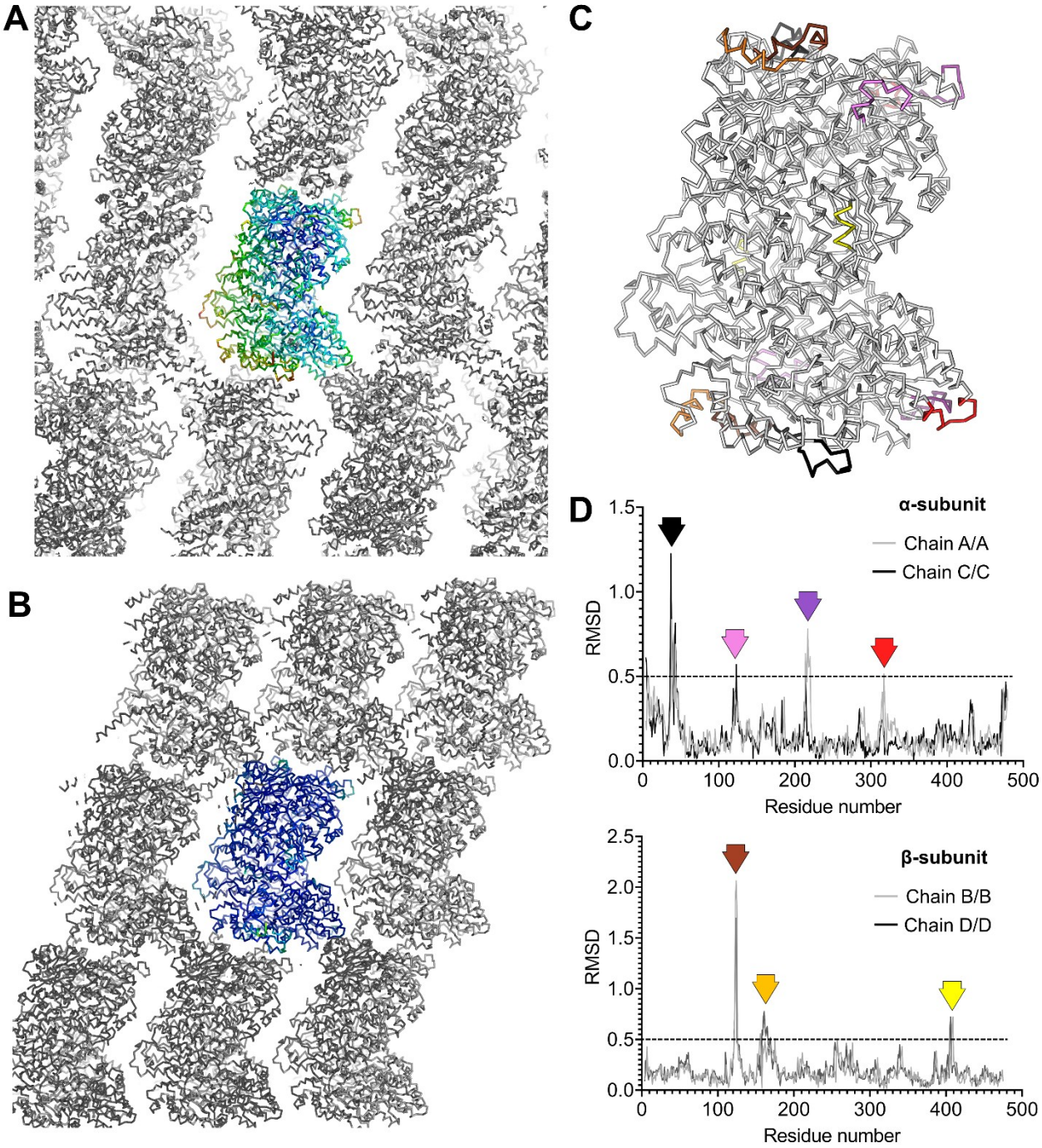


Figure S2. Crystal packing of the β-D353G/D357G MoFe variant (A), displaying prominent anisotropy, compared to the atomic resolution of the wild-type (B, PDB 3U7Q). The Cα backbone of the protein is represented as a ribbon and colored from low (blue) to high (red) B-factor value for the central unit and in grey for the symmetry mates. (C) Representation of one MoFe-protein variant in the same pose as (A). The protein is shown as a ribbon in which the highest root mean square deviation (in Å) with the wild type is color-coded based on the panel's (D) arrows. The highest deviations are at the periphery and mostly at the crystalline contacts.

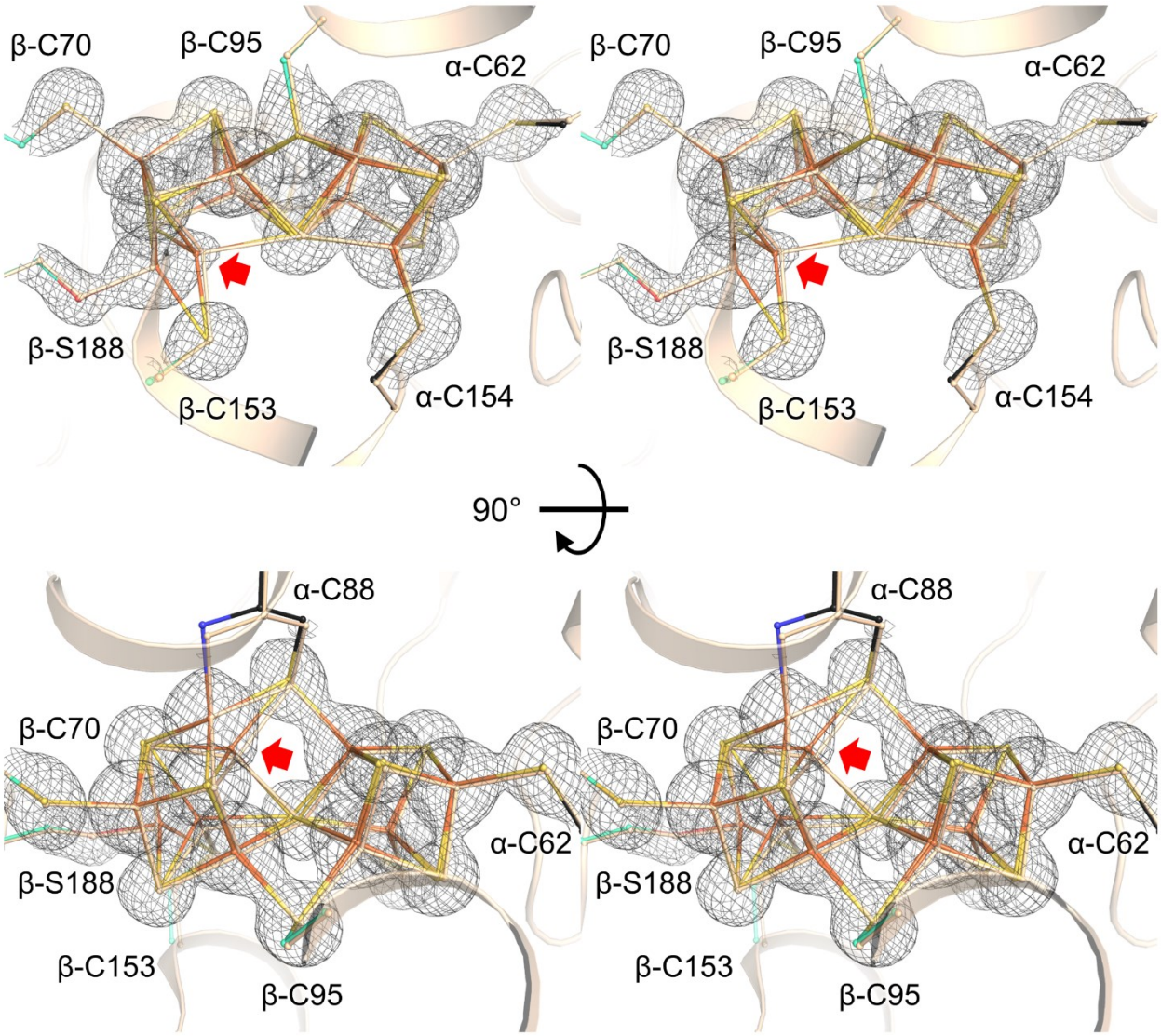


Figure S3. Stereo view of the P cluster from β-D353G/D357G variant (colored by atom) with the P cluster from wild-type MoFe protein (in wheat, PDB 3U7Q). The protein backbones are shown as transparent cartoons. Cluster-coordinating residues and P clusters are in balls and sticks. The $2F_o$-F_c electron density (black mesh) is contoured at 2.5 σ. Both redox states are represented, and red arrows highlight the modelled reduced state at an occupancy of 10%.

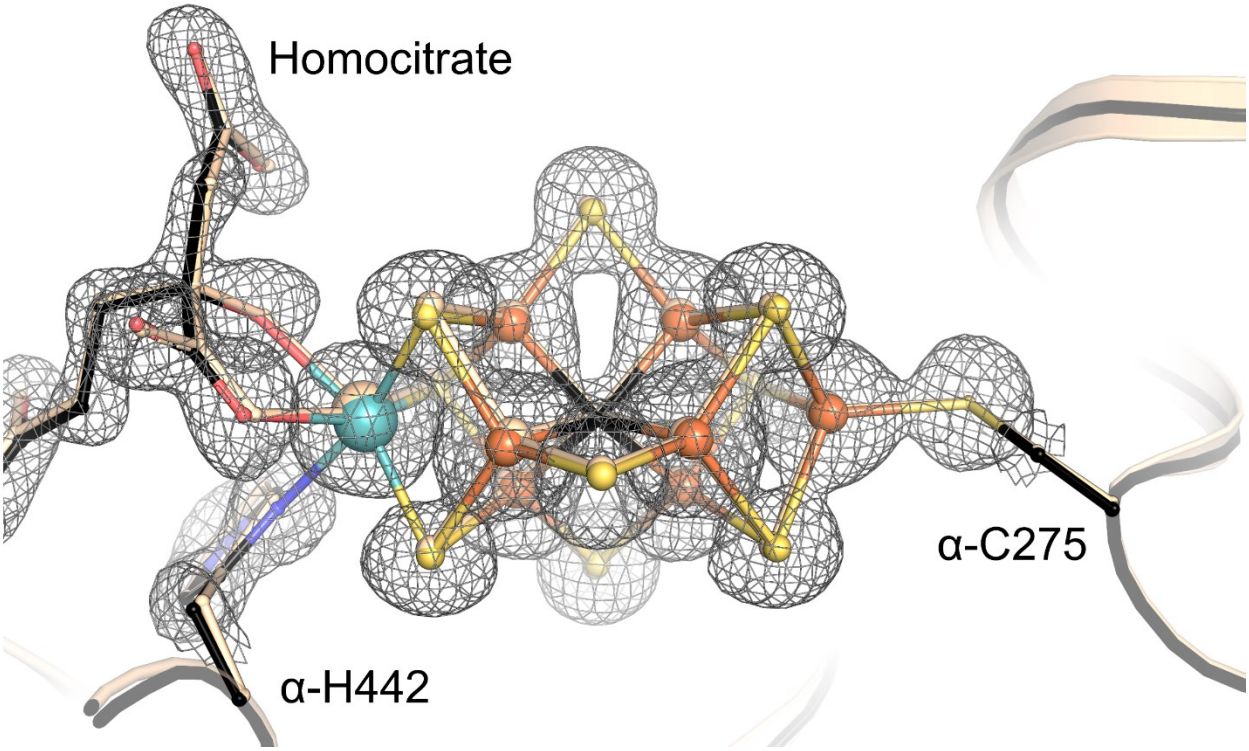


Figure S4. Superposition of the FeMo-co from β-D353G/D357G variant (colored by atom) with the FeMo-co from wild-type MoFe protein (in wheat, PDB 3U7Q). The protein backbones are shown as transparent cartoons. Cluster-coordinating residues and FeMo-co as balls and sticks. The $2F_o$-F_c electron density (black mesh) is contoured at 2.5 σ.

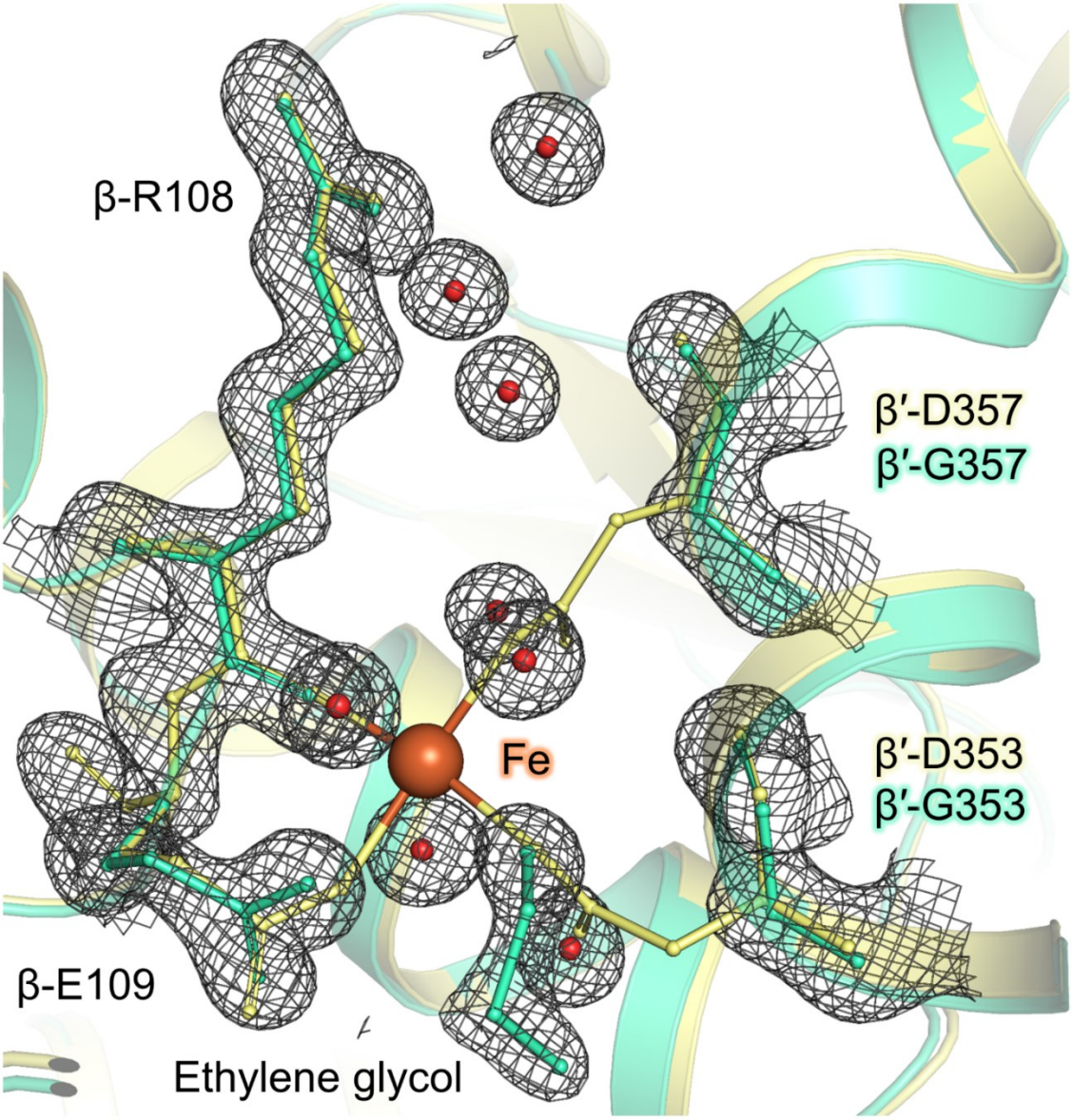


Figure S5. Superposition of the MMB site from β-D353G/D357G variant (green cyan) with the MMB site from α-R96Q MoFe variant (pale yellow, PDB 6BBL) in which the 16th Fe atom was modelled at 100 % occupancy. The proteins are shown as transparent cartoons and coordinating residues/metals/water (the latter colored in red) as balls and sticks. The $2F_o$-F_c electron density (grey mesh) is contoured at 1.5 σ. The modelled ethylene glycol could come from the cryoprotection step (30 % v/v final) or as a degradation product from the PEG 10,000 present in the crystallization solution.

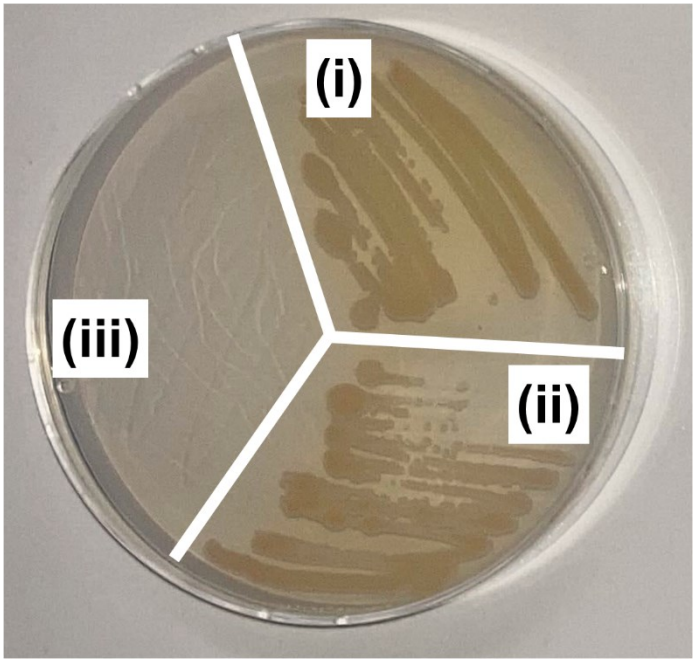


Figure S6. Nitrogen-free agar plates of *Azotobacter vinelandii* strains used in this work. (i) = "wild-type" MoFe-producing *A. vinelandii* DJ derivative (produces MoFe with an 8x His-tag on the N-term of NifD,[17] (ii) the *A. vinelandii* DJ derivative developed here for the production of ß-D353G/D357G (derived from (i)), and (iii) *A. vinelandii* DJ2107 (Δ*nifB*, kindly shared by Dennis Dean, Virginia Tech).[24]

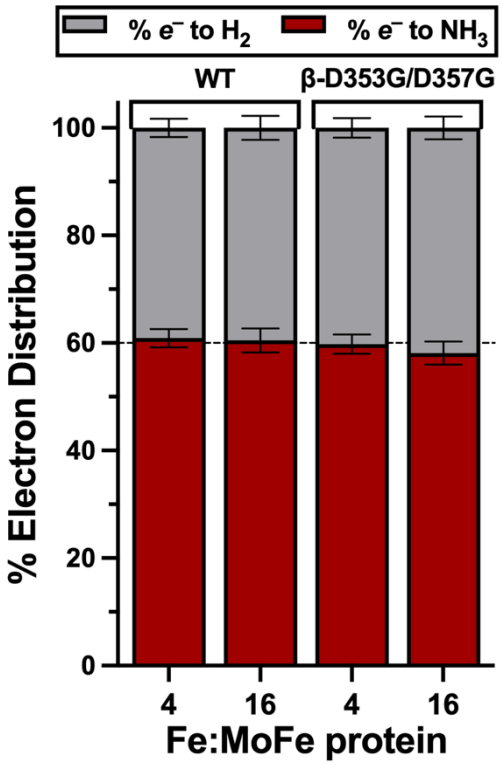


Figure S7. Percentage electron distribution between H_2 formation (gray) and NH_3 formation (red) by wild-type and ß-D353G/D357G MoFe proteins (stated above the bars) under 1 atm of N_2. Electron distribution was calculated by the detection of H_2 and NH_3, considering that $2e^-$ are required for the formation of each H_2 and N_2 reduction to NH_3 requires $6e^-$ ($3e^-$ for each NH_3 quantified). n = 3 (technical repeats) and error bars represent standard deviations (SD, propagated). Assays contained 0.1 mg of the respective MoFe protein (~240,000 kDa), alongside the stated molecular equivalents of Fe protein (~66,000 kDa). All assays were performed for 8 minutes at 30 ºC and terminated with the injection of 300 µL of 400 mM EDTA (neutral pH).

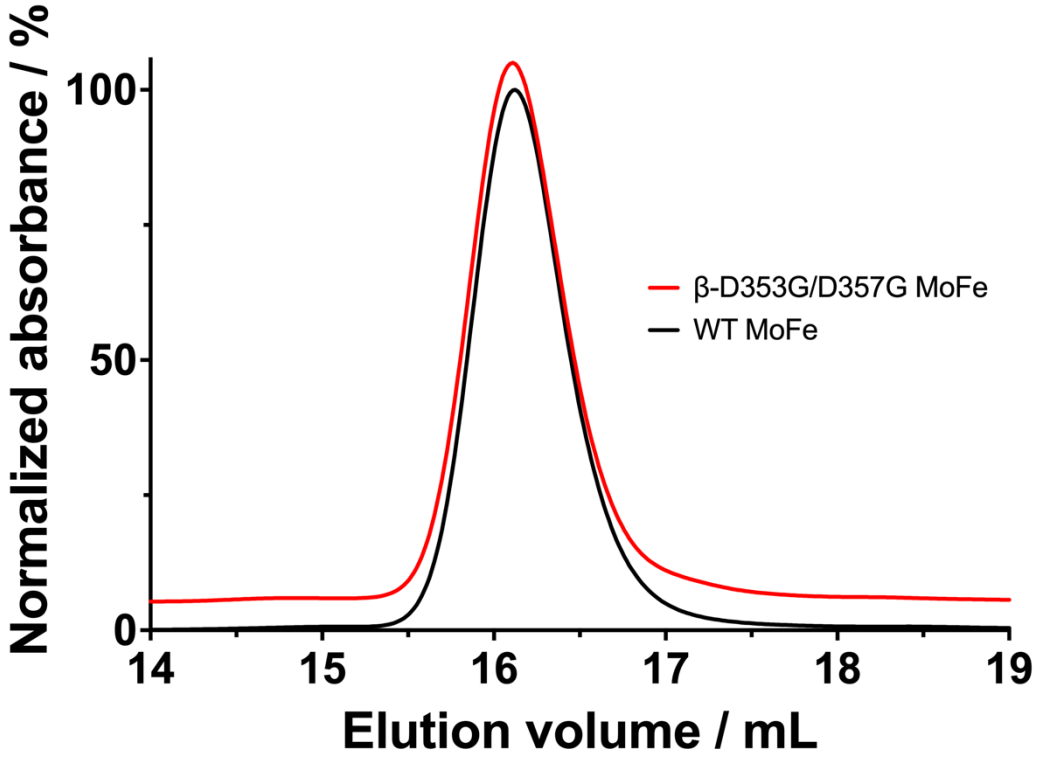


Figure S8. Analytical gel filtration of wild-type and ß-D353G/D357G MoFe proteins, performed using a Superose 6 Increase 10/300 GL column (Cytiva), as described above. 200 µL aliquots of protein (diluted to a concentration of 20 µmol/mL) were injected onto the column. The profile of the ß-D353G/D357G MoFe protein was offset by +5 y-axis units for clarity.

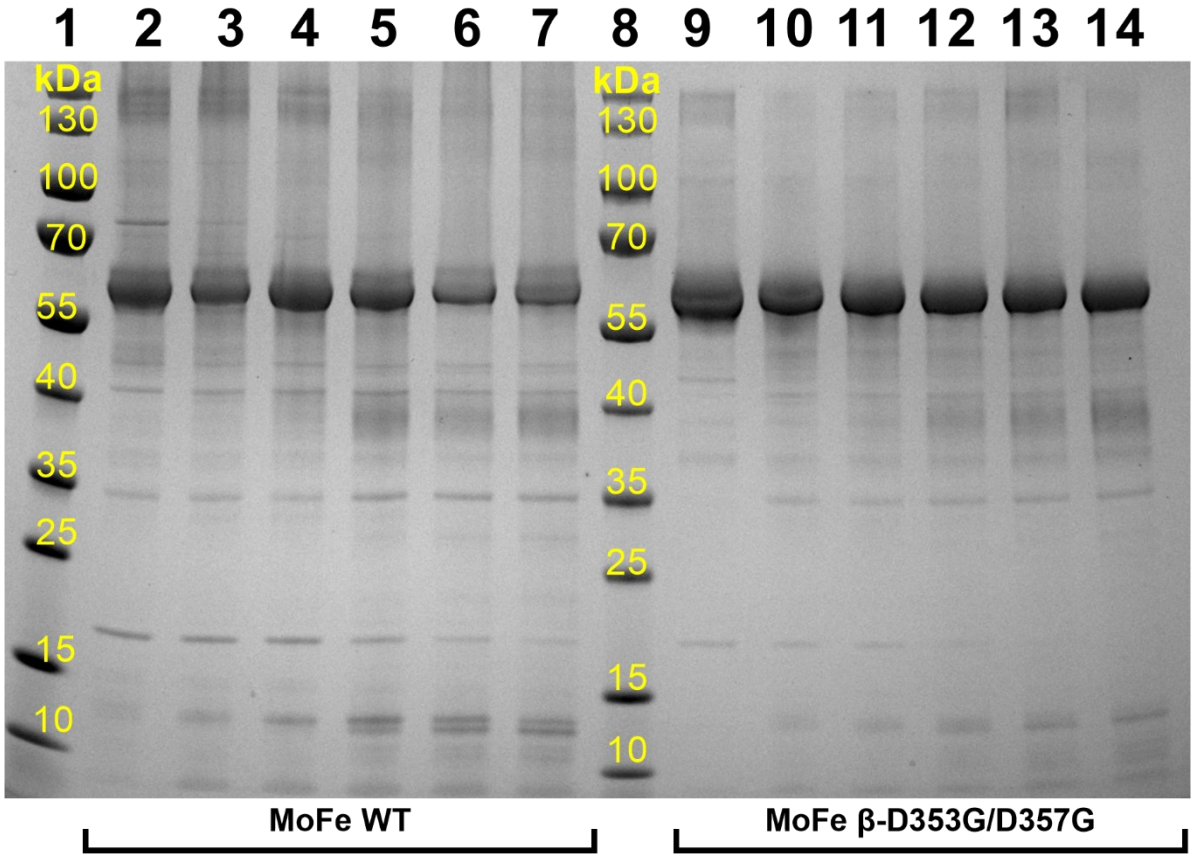


Figure S9. SDS-PAGE (4-12% precast gel, 200 V) analysis for temperature-controlled limited proteolysis (Coomassie blue staining). Lanes: 1 and 8 = molecular weight markers, 2 and 9 = MoFe proteins at t = 0 min after trypsin addition, 3 and 10 = MoFe proteins at t = 2 min, 4 and 11 = MoFe proteins at t = 5 min, 5 and 12 = MoFe proteins at t = 30 min, 6 and 13 = MoFe proteins at t = 60 min, 7 and 14 = MoFe proteins at t = 90 min. Each well contained a total of 1.75 µg protein.

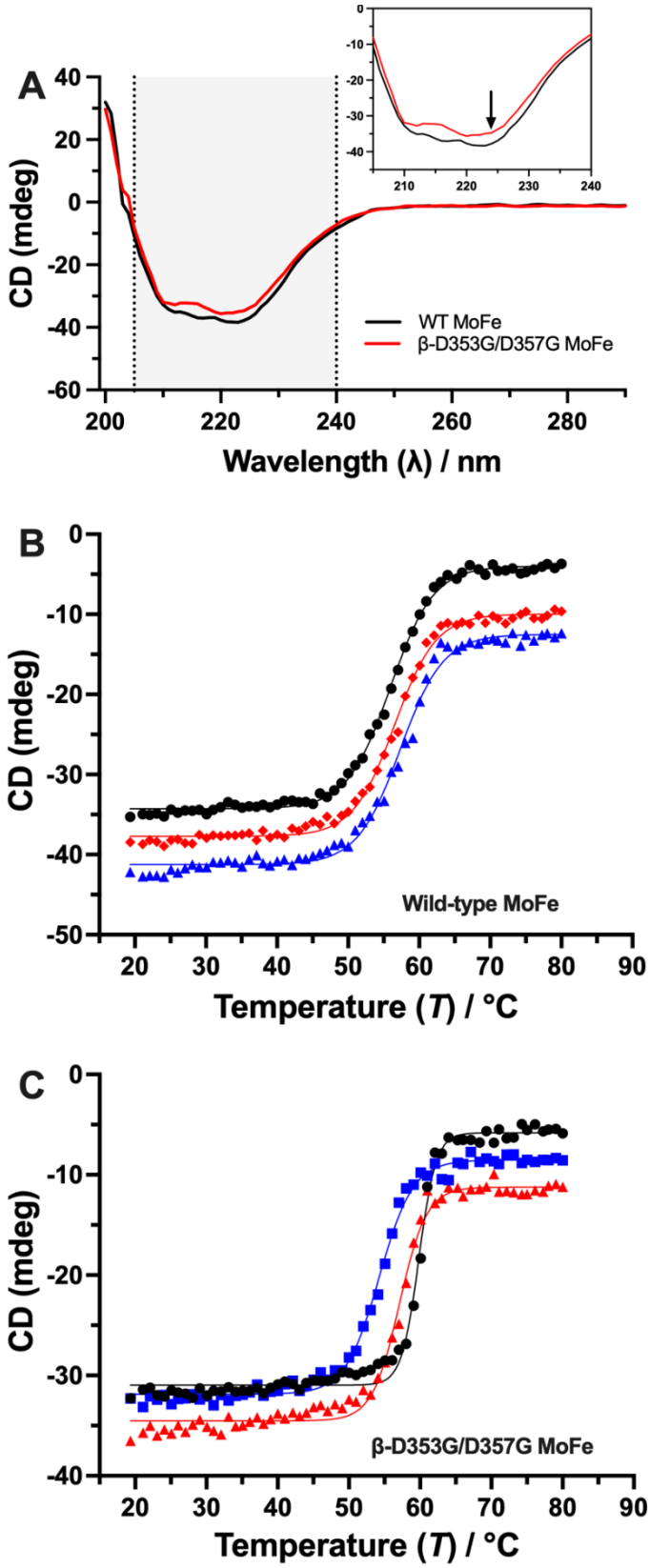


Figure S10. Melting temperature determination using circular dichroism. B and C report the CD response recorded at 224 nm. Red/blue/black present the data from three repeats.

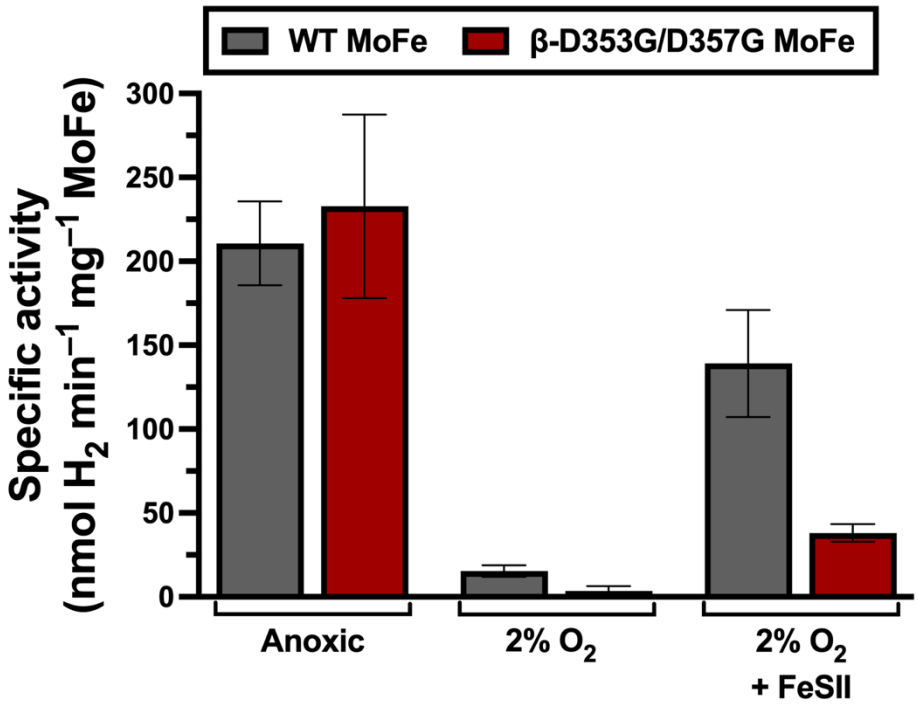


Figure S11. Nitrogenase FeSII protection assays. Specific activities for H_2 production by wild-type and ß-D353G/D357G MoFe proteins under 1 atm Ar atmosphere after exposure to 2% for 10 minutes. *n* = 3 (technical repeats), and error bars represent SD (error propagated). Assays contained 0.1 mg (0.4 µM) of the respective MoFe protein (~240 kDa), with one molecular equivalent of Fe protein (~66 kDa). Where necessary, 1 molar equivalent of the FeSII protein (Shethna protein) was included prior to exposure to O_2. All assays were performed at 30 ºC for 8 minutes and terminated with the injection of 300 µL of 400 mM EDTA (neutral pH). These specific activities are lower than those reported in Figure 3 (main article) as the Fe:MoFe protein ratio has been lowered to 1:1 (from 16.6:1) for O_2 protection assays.

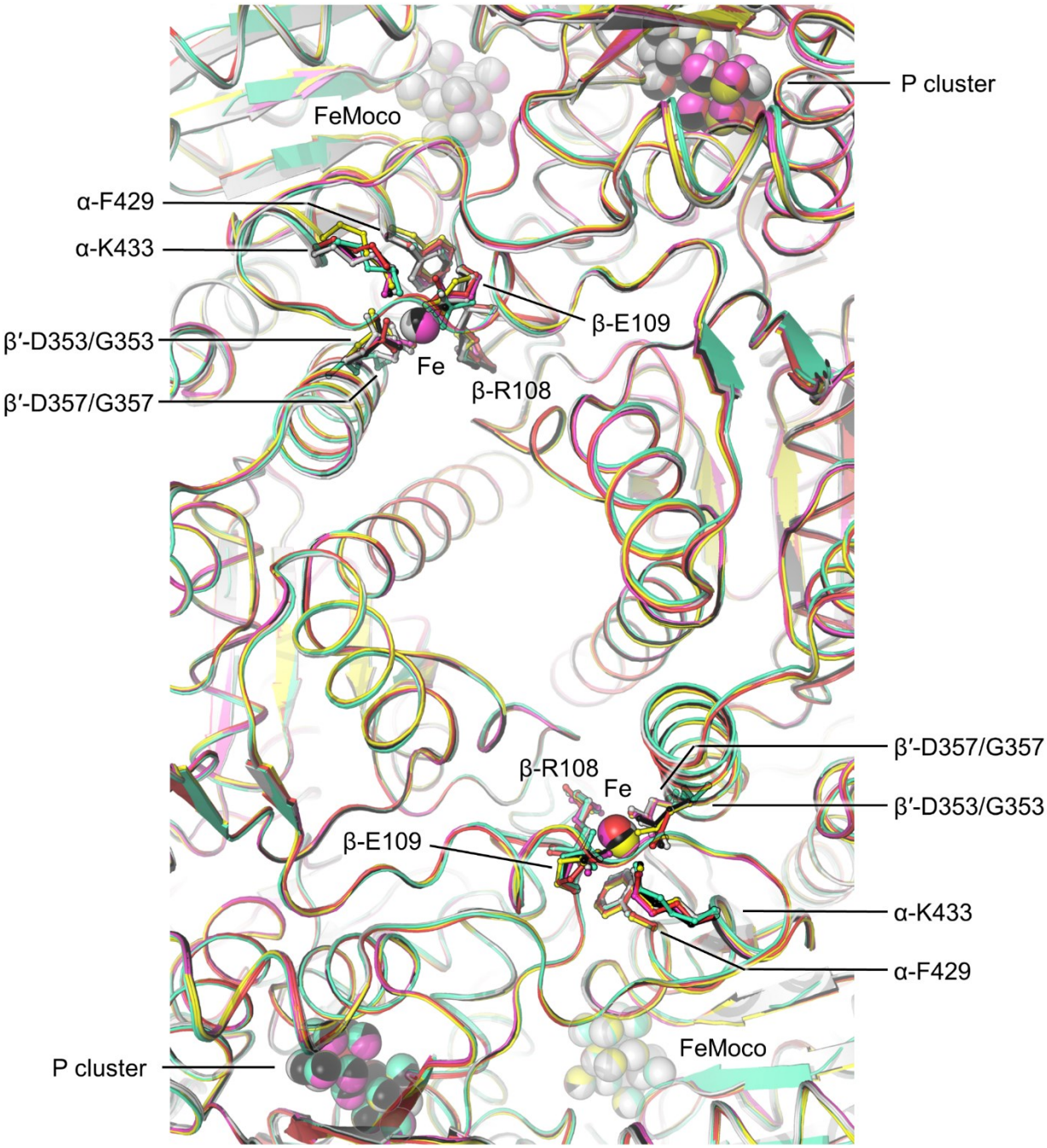


Figure S12. Overall view of the MoFe-proteins from the cryo-electron microscopy snapshots under turnover conditions. The MoFe-proteins are represented as cartoons with FeMo-cofactors, P clusters, and residues surrounding the MMB sites, and the 16th Fe are represented as balls and sticks. The calcium atom modelled in the structure of PDB 3U7Q was exchanged by a Fe based on the PDB 4TKU. The atomic resolution structure of the wild-type (PDB 3U7Q) and the variant presented in this work are colored white and green cyan, respectively. The cryo-electron microscopy structures of the MoFe-proteins alone (7UT6), in complex with one Fe-protein loaded with ATP (7UT8), in complex with one Fe-protein loaded with ADP/ATP (7UT9) and in complex with two Fe-proteins loaded with BeFx are colored in black, red, yellow, and pink, respectively.

Supplementary references.

[11] T. D. Schneider, R. M. Stephens, *Nucleic Acids Res* **1990**, *18*, 6097-6100.

[12] G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, *Genome Res* **2004**, *14*, 1188-1190.

[13] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, *J Mol Biol* **1990**, *215*, 403-410.

[14] N. A. O'Leary, M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C. M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V. S. Joardar, V. K. Kodali, W. Li, D. Maglott, P. Masterson, K. M. McGarvey, M. R. Murphy, K. O'Neill, S. Pujar, S. H. Rangwala, D. Rausch, L. D. Riddick, C. Schoch, A. Shkeda, S. S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R. E. Tully, A. R. Vatsan, C. Wallin, D. Webb, W. Wu, M. J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T. D. Murphy, K. D. Pruitt, *Nucleic Acids Research* **2015**, *44*, D733-D745.

[15] F. Madeira, M. Pearce, A. R. N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, *Nucleic acids research* **2022**, *50*, W276-W279.

[16] C. Cadoux, D. Ratcliff, N. Maslać, W. Gu, I. Tsakoumagkos, S. Hoogendoorn, T. Wagner, R. D. Milton, *JACS Au* **2023**, *3*, 1521-1533.

[17] R. D. Milton, R. Cai, S. Abdellaoui, D. Leech, A. L. De Lacey, M. Pita, S. D. Minteer, *Angewandte Chemie - International Edition* **2017**, *56*, 2680-2683.

[18] R. D. Milton, S. Abdellaoui, N. Khadka, D. R. Dean, D. Leech, L. C. Seefeldt, S. D. Minteer, *Energy and Environmental Science* **2016**, *9*, 2550-2554.

[19] R. D. Milton, R. Cai, S. Sahin, S. Abdellaoui, B. Alkotaini, D. Leech, S. D. Minteer, *Journal of the American Chemical Society* **2017**, *139*, 9044-9052.

[20] L. Zhang, J. T. Kaiser, G. Meloni, K. Y. Yang, T. Spatzal, S. L. A. Andrade, O. Einsle, J. B. Howard, D. C. Rees, *Angewandte Chemie - International Edition* **2013**, *52*, 10529-10532.

[21] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallographica Section D* **2011**, *67*, 293-302.

[22] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, *Acta Crystallogr D Struct Biol* **2019**, *75*, 861-877.

[23] C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, 3rd, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, *Protein Sci* **2018**, *27*, 293-315.

[24] E. Jimenez-Vicente, Z. Y. Yang, W. Keith Ray, C. Echavarri-Erasun, V. L. Cash, L. M. Rubio, L. C. Seefeldt, D. R. Dean, *Journal of Biological Chemistry* **2018**, *293*, 9812-9823.

CHAPTER 6: Nitrogen fixation and hydrogen evolution by sterically encumbered Mo-nitrogenase

Cécile Cadoux[1,2], Daniel Ratcliff[1,2], Nevena Maslać[3], Wenyu Gu[4], Ioannis Tsakoumagkos[5], Sascha Hoogendoorn[5], Tristan Wagner[3] and Ross D. Milton[1,2]

[1] Department of Inorganic and Analytical Chemistry, Faculty of Sciences, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

[2] National Centre of Competence in Research (NCCR) Catalysis, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

[3] Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany

[4] Department of Chemical Engineering, Stanford University, Stanford, California 94305, United States

[5] Department of Organic Chemistry, National Center of Competence in Research (NCCR) Chemical Biology, University of Geneva, Quai Ernest-Ansermet 30, 1211 Geneva 4, Switzerland

Published in:

JACS Au Volume 3 • Number 5 • 09th May 2023

DOI: https://doi.org/10.1021/jacsau.3c00165

Recieved: 4th of April 2023

Accepted: 20th of April 2023

Published online: 9th of May 2023

Correspondence to Ross D. Milton (*Ross.Milton@unige.ch*)

The publication can be accesed at:

https://pubs.acs.org/doi/10.1021/jacsau.3c00165

DOI: *https://doi.org/10.1021/jacsau.3c00166.1.*

6.1. Keywords

Nitrogenase; cooperativity; nitrogen fixation; ammonia; hydrogen; metalloenzyme.

6.2. Abstract

The substrate-reducing proteins of all nitrogenases (MoFe, VFe, and FeFe) are organized as α2ß2(γ2) multimers with two functional halves. While their dimeric organization could afford improved structural stability of nitrogenases in vivo, previous research has proposed both negative and positive cooperativity contributions with respect to enzymatic activity. Here, a 1.4 kDa peptide was covalently introduced in the proximity of the P cluster, corresponding to the Fe protein docking position. The Strep-tag carried by the added peptide simultaneously sterically inhibits electron delivery to the MoFe protein and allows the isolation of partially inhibited MoFe proteins (where half-inhibited MoFe protein was targeted). We confirm that partially functional MoFe protein retains its ability to reduce N_2 to NH_3, with no significant difference in selectivity over obligatory/parasitic H_2 formation. Our experiment concludes that wild-type nitrogenase exhibits negative cooper-ativity during steady state regarding H_2 and NH_3 formation (under Ar or N_2). These results confirm that turnover in one half of the MoFe protein inhibits turnover in the second half, emphasizing the importance and presence of long-range (>95 Å) protein-protein communication in biological N_2-fixation in *Azotobacter vinelandii*.

6.3. Introduction

The fixation of kinetically inert atmospheric dinitrogen (N_2) to ammonia (NH_3) is catalyzed in some specific microbes by a single family of enzymes known as nitrogenases, with turnover frequencies of around one N_2 fixed per second and a second-order rate constant of ~10^4 M^{-1} s^{-1} (k_{cat}/K_M).[1–6] The Mo-dependent nitrogenase consists of an N_2-reducing molybdenum-iron (MoFe) protein and a corresponding reductase called the iron (Fe) protein (Figure 1). The MoFe protein is a ~230 kDa $(\alpha\beta)_2$ heterotetramer (NifDK), where each αß half contains an electron-transferring [8Fe-7S] P cluster and a [7Fe-9S-C-Mo]:homocitrate FeMo cofactor (FeMoco).[2] The Fe protein is a homodimeric $NifH_2$ protein of ~66 kDa, containing a single [4Fe-4S] cluster and two MgATP binding sites. During turnover, each αß half of the MoFe protein accepts electrons from the ATP-hydrolyzing Fe protein, which are transferred *via* the P cluster to the FeMoco for N_2 fixation to NH_3 (**Equation 1**):[4]

Equation 1

$$16MgATP + 8e^- + 8H^+ + N_2 \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$$

It is important to note that the above equation represents the optimized stoichiometry of one H_2 produced per N_2 fixed by the Lowe-Thorneley mechanism, with additional unproductive H_2 production taking place under non-optimal turnover.[1,7–10] Each transient association of the Fe protein (Fe protein cycle) ultimately results in the transfer of $1e^-$ to the FeMoco, where each Fe protein cycle consists of at least (i) MgATP-bound Fe:MoFe association, (ii) electron transfer (ET) from the P cluster to the FeMoco, (iii) ET from the Fe protein's [4Fe-4S] cluster to the P cluster, (iv) 2ATP hydrolysis, (v) the release of two inorganic phosphate (P_i) and (vi) MgADP-bound Fe:MoFe dissociation. The rate-limiting step of nitrogenase catalysis is thought to be the release of P_i by the Fe protein, taking place with a rate constant of 25-27 s^{-1}.[11] This, in turn, implies that each Fe protein electron delivery cycle takes place with an overall rate constant of ~13 s^{-1}.

Importantly, it has been shown that the two αß halves of the MoFe protein do not function independently during their Fe protein electron delivery cycles, suggesting that communication between the Fe proteins ~100 Å apart takes place during turnover.[12–14] In 2016, Danyal *et al.* reported that the Fe protein electron delivery cycles experience negative cooperativity in the pre-steady-state.[12] In other words, the Fe protein electron delivery cycle on one αß half suppresses ET in the other half. This was observed by quantitative measurements of Fe protein oxidation, ATP hydrolysis and P_i release.[12] In 2021, Truscott *et al*. reported positive cooperativity in the steady-state for the reduction of acetylene, a non-physiological yet historically prominent substrate of nitrogenase.[13] In this approach, cooperativity was investigated by forming inactive Fe:MoFe complexes on one αß half of the MoFe protein (using AlF_4^- or nonstandard, tightly associating Fe proteins). Notably, cooperativity was not observed for $2H^+$ reduction to H_2, a physiologically relevant reaction that is catalyzed by nitrogenase both in the absence or presence of N_2. In 2011, Eady and coworkers also observed that MoFe protein containing only one FeMoco (half-populated) undergoes additional, non-electron-transferring interactions with the half-reactive MoFe protein, although the absence of the FeMoco was previously found to introduce a large change in the conformation of the MoFe protein.[15,16] More recently, cryogenic electron microscopic investigation into Fe-MoFe interactions during turnover identified a potential preference for the MoFe protein to associate to one Fe protein at a time, refocusing the spotlight on the MoFe protein's arrangement as a heterotetramer with two functional αß halves.[17]

An important open question is therefore: how does cooperativity (negative, positive, or indeed none) impact N_2 fixation by nitrogenase during continued turnover? Harris *et al*. recently proposed that decreased selectivity towards N_2 fixation in the alternative vanadium-dependent

and iron-only nitrogenases is due to decreased rate constants for the reductive elimination of H_2 (an activational step for N_2 fixation).[18] This reductive elimination requires the delivery of at least $4e^-$ to the FeMoco for sufficient activation, which in turn requires ATP-hydrolysis-coupled electron transfer from the Fe protein in one αß half of the MoFe protein. Stalled electron transfer to the FeMoco (by, for example, negative cooperativity induced by the second αß half of the MoFe protein) then provides time for the non-productive evolution of H_2 (and the loss of reducing equivalents) by the protonolysis of metal-hydrides on FeMoco.[10] This competition between reductive elimination and metal-hydride protonolysis explains the "optimal" stoichiometry of $1H_2$ evolved per N_2 fixed.[7] We hypothesized that the liberation of inhibited electron delivery to one αß half of the MoFe protein could therefore yield the stoichiometrically optimized production of one H_2 per N_2 fixed during continuous nitrogenase turnover, provided that cooperativity is not strictly necessary for N_2 fixation. However, in order to observe this, we deemed it of utmost importance to study half-reactivity on a MoFe protein that (i) contained FeMoco in both αß halves (retaining its native conformation[16]), and (ii) was not half-inhibited by a tightly associating Fe protein on one αß half, given that conformational changes transmitted between the Fe proteins bound to both αß halves is considered essential to cooperativity.[13,14]

Here we report on the reactivity of MoFe protein from *Azotobacter vinelandii*) where we sought to selectively inhibit Fe protein association on only one αß half by steric inhibition. To achieve this, we employed a MoFe protein mutant possessing a single solvent-exposed Cys residue in proximity to the P cluster (α-C45A/L158C, NifD = α); this mutant was previously employed to conjugate a Ru-based photosensitizer in the place of the Fe protein, enabling photo-excited electron transfer to the P cluster.[19] X-ray crystallography confirms the solvent accessibility (albeit somewhat geometrically hidden) of this cysteine residue. We subsequently employed the reactivity of this cysteine in a thiol-maleimide Michael addition to introduce a large synthetic Strep-tag-containing peptide (1'380 g mol^{-1}) to facilitate both (i) Fe protein steric inhibition and (ii) the separation of inhibited MoFe proteins from unmodified MoFe protein.

This population of partially inhibited MoFe protein (lacking uninhibited MoFe protein) was confirmed to be active for N_2 fixation, where a maximum velocity (V_{max}) of 66% was determined, consistent with negative cooperativity during N_2 fixation. Importantly, we observed that the selectivity (product/electron distribution) of this partially inhibited MoFe was practically unchanged between 0-1 atm N_2, suggesting that cooperativity may not contribute towards nitrogenase's remarkable selectivity for N_2. We conclude that negative

cooperativity, globally, is employed by this nitrogenase for both H_2 production and N_2 fixation.

6.4. Results and discussion

6.4.1. Structure of the α-C45A/L158C MoFe protein. As shown in Figure 1, residue α-L158 is located at the Fe:MoFe protein interface, and we therefore hypothesized that the functionalization of a Cys residue in this location with a steric inhibitor could prevent access of the Fe protein (and therefore, nitrogenase catalysis). Indeed, a nitrogenase α-C45A/L158C MoFe mutant has been previously reported, yielding a single solvent-exposed Cys in proximity to the P cluster.[19] We first prepared this α-C45A/L158C MoFe with an N-terminal 8xHis tag on NifD (α subunit) for affinity purification using a sacB-based markerless mutagenesis approach (Figure S1, Supporting Information).[20–22]

After verifying the introductions of the mutations by DNA sequencing (Supporting Information), we next crystallized the purified mutant MoFe protein and elucidated its structure by X-ray crystallography (Figure 2). The X-ray crystal structure belonging to the $P2_1$ space group was solved by molecular replacement using the PDB 3MIN as a template. The structure was refined to a resolution of 3.03 Å and contained 2 MoFe proteins in the asymmetric unit (Table S2). The cell dimensions are fitting none of the previously solved structures (Table S3) and might be due to the introduced mutations that impacted the crystal packing (Figure S2, Supporting Information). The α-C45A/L158C MoFe protein was found to overlay well with high-resolution structures previously reported for wild-type MoFe protein from *A. vinelandii*, suggesting a minimal impact of the introduced mutations on the overall conformation of the MoFe protein (Table S4 and Figure 2A).[23,24] Despite the rather low resolution, some distinct changes in the rotamers of amino acids around the introduced mutations were also observed (Figure 2BC), once again being consistent with the mutations having been successfully introduced. Importantly, the P cluster and FeMoco of the α-C45A/L158C MoFe protein were both found to be present and intact (Supporting Information, Figures S3).

6.4.2. Desthiobiotin-maleimide steric inhibitor. As depicted in Figure 1, the α-L158C is positioned at the Fe protein-binding interface on the MoFe protein. To interrogate the half-reactivity of nitrogenase's MoFe protein, we sought to modify only one-half of the α-C45A/L158C MoFe protein at this position with a steric inhibitor that would enable (i) inhibition of the Fe protein on this half of the MoFe protein, and (ii) affinity purification of the hybrid α-C45A/L158C MoFe protein.

Previously, iodoacetamide-Cys reactivity was employed to attach a Ru-based photosensitizer to this α-C45A/L158C MoFe protein.[19] In this work, we elected to utilize maleimide-Cys thiol-Michael addition chemistry to modify the α-158C residue due to its improved chemoselectivity over iodoacetamides.[25] Initially, a steric inhibitor was synthetized by coupling a poly(ethylene glycol)$_3$-modified desthiobiotin moiety with N-aminoethylmaleimide *via* a peptide/amide bond formation (Figure 3A; addition details are provided in the Supporting Information, Figures S4-S6). The maleimide moiety was incorporated for the site-selective modification of the α-C158 residue of the α-C45A/L158C MoFe protein, whereas the desthiobiotin moiety was included as a binding motif for avidin-based affinity purification post conjugation of the MoFe protein. It was hypothesized that the poly(ethyleneglycol) repeating units would both increase the size-in-space of the steric inhibitor (and therefore its potency) as well as increase its solubility during protein conjugation. Initially, this maleimide-modified desthiobiotin-containing inhibitor (referred to subsequently as "DTB") was incubated with both wild-type and α-C45A/L158C MoFe proteins, and western blotting with a streptavidin-horseradish peroxidase (HRP) conjugate to confirm successful modification of the MoFe proteins (Figure 3B, Figure S7-8).

Sodium dithionite (DT) is commonly used as a reducing agent during the purification and handling of nitrogenases, due to their deactivation by molecular oxygen (O2). We hypothesized that, much like thiol-based reducing agents, DT could reduce the maleimide functional group of DTB and thus quash subsequent Cys functionalization. Therefore, DT-free MoFe protein samples were prepared (Supporting Information) and Cys-maleimide labelling was evaluated in the presence and absence of DT.26 As qualitatively shown in Figure 3B, the presence of DT in the MoFe protein samples (additional DT was not included during the reaction) was observed to lower the overall labelling of the MoFe proteins with DTB (also Figure S8, Supporting Information). Importantly, this issue could not be completely resolved with the use of tris(2-carboxylethyl)phosphine (TCEP) as a replacement reducing agent, commonly used in thiol-maleimide Michael addition reactions.

DT was therefore removed from all subsequent MoFe protein preparations prior to maleimide functionalization reactions, only being reintroduced to terminate the maleimide-Cys reaction and maintain reducing conditions after the incubation period. Importantly, the omission of DT from the purification procedure (removed during the first purification column) did not result in a statistically lowered specific activity of the α-C45A/L58C MoFe protein (t test, P = 0.93).

We next evaluated the residual activities of α-C45A/L158C MoFe protein (DT-free) following incubation with varying molar equivalents of the DTB inhibitor, using the Fe protein as the electron donor for H^+ reduction under Ar and N_2 reduction (Figure 3C; specific activities are provided in Figure S9). Although a marked decreased in activity was observed globally, the residual activities of the α-C45A/L158C MoFe proteins treated with only 0.5 – 1.5 molar equivalents of DTB were found to range from approximately 55-75%, indicating that DTB effectively inhibits the Fe protein and subsequent substrate reduction by the MoFe protein. Increasing molar equivalents of DTB were found to further decrease the specific activity of the MoFe protein (<20% with 5 molar equivalents). As previously shown, 10 mM DT significantly impeded the Cys labelling of the α-C45A/L158C MoFe protein with the DTB inhibitor, therefore highlighting the necessity to remove DT from the α-C45A/L158C MoFe protein prior to the reaction. DTB steric inhibition was also performed on the WT MoFe protein (exploiting the α-C45 residue), which exhibited a less pronounced decreased in specific activity (Figure S9, Supporting Information).

Having observed a decrease in both H_2 formation (under Ar and N_2) and NH_3 formation (under N_2), we next determined whether the steric inhibition on one αß half of the α-C45A/L158C MoFe protein impacts the distribution of electrons between H^+ and N_2 reduction (*i.e.*, does a cooperativity mechanism contribute to nitrogenase's selectivity towards N_2 fixation?). According to the modified Lowe-Thorneley model of nitrogenase's enzymatic mechanism, the reductive elimination of one H_2 enables the binding and subsequent reduction of each N_2 at the FeMoco, leading to the reaction stoichiometry in Equation 1 (Figure 4).[1,10] This has been experimentally observed with H_2 remaining persistent under a high pressure of N_2 (50 atm).[7] The $6e^-$ reduction of N_2 therefore requires a total of $8e^-$ (the reductive elimination of H_2 requires $2e^-$), with 75% of the electrons delivered to the MoFe protein ultimately being detected as NH_3. However, in practice under typical laboratory conditions (*i.e.*, 1 atm of N_2), the distribution of electrons towards N_2 fixation typically reaches an upper limit of 60%.[20,27–30] This can be explained by the non-productive release of H_2 from reduced FeMoco states by the protonolysis of metal-hydrides, which also explains nitrogenase's H_2-production activity in the absence of N_2 (Figure 4).[8] Here, we observed an electron distribution between 50-60% for N_2, implying that at least 2 or 3 H_2 molecules are released for each N_2 molecule reduced (presumably including at least one H_2 reductive elimination step) (Figure 3D). Interestingly, the observed distribution of electrons towards N_2 fixation remained around 50-60% upon the titration of increasing equivalents of the DTB inhibitor even though the overall specific activities were observed to decrease (Figure 3D). This result

provided an initial indication that both halves of the αß MoFe protein are not strictly required to function in order to fix N_2 in any given αß domain. However, it is necessary to control the homogeneity of DTB-bound FeMo population to study the cooperativity of FeMo (*i.e.,* to purify half-inhibited MoFe proteins from uninhibited proteins). Therefore, a purification protocol for the half-inhibited MoFe was established.

Although the functionalization of the α-C45A/L158C MoFe protein with the DTB inhibitor was detected by western blotting with a streptavidin-HRP conjugate (Figure 3B), our attempts to purify this functionalized protein with commercial StrepTactin-based affinity columns (*i.e.*, "StrepTrap" by Cytiva) were unsuccessful. This was hypothesized to be due to poor affinity of this engineered StrepTactin protein for desthiobiotin over the conventional "StrepTag", and we therefore reoriented our strategy.

6.4.3. A Strep-containing, cysteine-reactive peptide permits the purification of partially reactive MoFe protein conjugates. We next sought to replace the DTB inhibitor with an alternative steric inhibitor that would additionally enable affinity-based purification. Given the widespread success of "Strep-tag" peptides for affinity purification, we elected to employ an N-terminal-modified Strep-tag peptide for the purification of α-C45A/L158C MoFe proteins that had been successfully modified (sequence: maleimide-GGGWSHPQFEK, referred to herein as the "Strep" inhibitor) (Figure 5AB).

The Strep inhibitor was reacted with α-C45A/L158C MoFe with a 0.5:1 molar equivalent of Strep:α-C45A/L158C MoFe protein (four accessible Cys158 residues per Strep-maleimide), to favor the formation of half-functionalized MoFe protein over doubly inhibited MoFe protein (additional discussion can be found in the Supporting Information, Figure S10). After 4h of reaction, the maleimide was quenched by the addition of 1 mM DT and the mixture was loaded onto a commercial pre-packed StrepTactin column (Figure S11, Supporting Information). Unmodified (uninhibited) MoFe protein was collected in the flowthrough fraction, and a dark band was observed to bind to the top of column, consistent with the successful functionalization of the α-C45A/L158C MoFe protein. This functionalization reaction was performed in triplicate on the same sample of α-C45A/L158C MoFe protein. Figure 5C highlights the purity of the uninhibited (flow-through) and Strep-inhibited MoFe proteins (StrepTactin-bound), where western blotting with a StrepTactin-HRP conjugate highlighted the presence of the Strep moiety on only the inhibited (modified) MoFe protein samples. A lower molecular weight impurity with high affinity to the StrepTactin-conjugate was identified, although this was not expected to impact subsequent studies of the Strep-

inhibited MoFe proteins due to its comparatively low abundance on the SDS-PAGE gel (this was also not identified during subsequent proteomics analysis discussed below). Quantification of the total Strep-inhibited MoFe protein fraction revealed a functionalization efficiency of 14% ± 1; this reflects the inefficiency of the maleimide-Cys labelling reaction that may be in part due to the inward-facing geometry of the solvent-exposed α-L158C residue, which further reduces the probability of obtaining di-functionalized MoFe proteins (Figure 2C). LC-ESI-MS/MS was performed on a sample of Strep-inhibited α-C45A/L158C MoFe protein to confirm the presence of the maleimide-Strep modification on the intended α-C158 residue. LC-ESI-MS/MS analysis of the SDS-PAGE-excised sample (81% α-subunit coverage, 91% ß-subunit coverage) confirmed that the Strep-modification was principally performed on the peptide fragment (trypsin-digested) containing the α-C158 residue, although an additional P cluster-coordinating Cys residue is also present on this fragment (α-C154). A potential functionalization was also identified on the ß-subunit/NifK (a single partially solvent-exposed peptide fragment), although this low-score hit could not be fully assigned (further discussion in the Supporting Information). Thus, WT MoFe protein incubated with an excess of the Strep-maleimide inhibitor per MoFe (5 molar equivalents, subsequently purified over a StrepTactin column) was also analyzed by LC-ESI-MS/MS (88% α-subunit coverage, 92% ß-subunit coverage). The Strep-inhibitor modification was not detected on the α-C154-containing fragment and is consistent with good selectivity of the Strep inhibitor to the surface-exposed α-C158 residue. Further, the potential modification on the ß-subunit was not identified, further supporting its identification on the α-C45A/L158C MoFe protein to be a false positive due to its low-quality spectra (further discussion in the Supporting Information). While our data are indicative of partial MoFe protein labelling, it is not possible to quantify the fraction of half-inhibited MoFe (the target) *vs*. doubly inhibited MoFe with this approach.

Anoxic Native-PAGE analysis was performed to confirm that the Strep-inhibited α-C45A/L158C MoFe protein remained its heterotetrameric organization (Figure S12, Supporting Information). Interestingly, the unmodified MoFe protein appears as a major product at ~240 kDa although a faint product with a slightly lower molecular weight was also identified systematically during repeated Native-PAGE analysis of both the WT and α-C45A/L158C proteins. Due to the high purity of the MoFe proteins (by SDS-PAGE analysis), we hypothesized that this second product was due to a different conformation of the MoFe proteins that was realized predominantly during electrophoretic analysis. Different conformations of MoFe proteins have previously been observed during native-PAGE analysis.[31] Interestingly, Strep-inhibited MoFe protein samples (WT and α-C45A/L158C)

qualitatively exhibited more of this second conformation than the proteins that did not bind to the StrepTactin solid phase during purification. Subsequent limited proteolytic analysis (Figure S13, Supporting Information) indicated that uninhibited and Strep-inhibited α-C45A/L158C MoFe did not exhibit a significant difference in their conformational flexibilities and we concluded that this additional conformational state is therefore induced during electrophoretic analysis, which is further pronounced following the anchoring of this 1.4 kDa Strep-inhibitor peptide to the α-subunit of the MoFe protein (in either the WT α-C45 or α-C158 position).

These results were confirmed by gel filtration in which uninhibited and Strep-inhibited α-C45A/L158C proteins share very similar elution volumes (Figure 5D). We also used this method to investigate whether a Fe_2:MoFe transition state complex could be formed between the Strep-inhibited α-C45A/L158C MoFe protein and the Fe protein, as evidence for inhibition of the Fe protein association in proximity to the P cluster site on the MoFe protein (Figure 5D). A non-ATP-hydrolyzing L127Δ Fe protein mutant was employed, which was previously found to form a tight complex with the MoFe protein.[32] To confirm the possibility of the L127Δ Fe protein to associate to the α-C45A/L158C Strep-inhibited MoFe protein, both proteins were incubated and the complex was subsequently purified using the N-terminal His-tag on the α-subunit of the MoFe protein. SDS-PAGE analysis of the obtained sample confirmed the presence of a protein bands corresponding to the MoFe and Fe proteins (Figure S14, Supporting Information). After incubation of the uninhibited α-C45A/L158C MoFe protein with 5 molar equivalents of the L127Δ Fe protein, virtually all of the MoFe protein was found to shift to a complex of increased molecular weight, consistent with the formation of a tight Fe_2:MoFe complex (Figure 5D). On the other hand, incubation of the Strep-inhibited α-C45A/L158C MoFe protein with the L127Δ Fe protein resulted in a broadened MoFe protein peak; further, a comparatively increased quantity of non-complexed L127Δ Fe protein was observed. We hypothesize that this broadening corresponds to a mixed population of (i) non-complexed Strep-inhibited MoFe protein, (ii) Fe_1:MoFe protein complex and (iii) doubly inhibited MoFe protein, consistent with inhibition of Fe protein association to the MoFe protein in the presence of the Strep-inhibitor at the α-C158 position. Multi-Gaussian peak analysis of this gel-filtration profile was performed in an attempt to provide indicative quantification of these different fractions (Figure S15). We calculated that this broad feature may be comprised of ~68% half-inhibited MoFe:Fe_1 complex and ~32% of either non-complexed half-inhibited MoFe or doubly inhibited MoFe (we cannot discriminate between the latter two). Recent cryogenic electron microscopic investigation into Fe:MoFe complex

formation during turnover suggests that the MoFe protein preferentially docks with only one Fe protein at a time.[17]

The formation of a Fe_2:MoFe protein complex (uninhibited and strep-inhibited) was also evaluated by Native-PAGE analysis, also suggesting that strep-inhibited α-C45A/L158C MoFe protein does not form a tight Fe_2:MoFe complex (Figure S16, Supporting Information). Uninhibited WT and α-C45A/L158C MoFe proteins were observed to form two different complexes of apparent larger size. Keeping in-mind the proposed alternative conformation of the Strep-inhibited α-C45A/L158C MoFe protein during Native-PAGE analysis, incubation with 5 molar equivalents of L127Δ Fe protein yielded one major (and one significantly weaker) complex of apparent increased size, consistent with inhibition of Fe protein access to primarily one half of the MoFe protein.

6.4.4. Reactivity of strep-inhibited MoFe protein. After having successfully purified the Strep-inhibited α-C45A/L158C MoFe protein, we next evaluated its remaining residual specific activity. Importantly, the mean specific activities for both H_2 (1 atm Ar) and NH_3 (1 atm N_2) production of the MoFe proteins that did not react with the strep-inhibitor (the flow-through fractions) were found not to differ from the unreacted control protein (one-way ANOVA, P = 0.0833 and 0.3323), indicating that neither the reaction conditions nor the handling of the samples drastically impacted enzymatic activity.

We then compared the specific activities of the uninhibited and strep-inhibited α-C45A/L158C MoFe proteins for H^+ reduction under 1 atm Ar. As shown in Figure 6A, the specific activities of the strep-inhibited MoFe proteins were found to be 79 ± 4% (mean of the three reactions) of their uninhibited counterparts. In addition, the specific activities of the three strep-inhibited α-C45A/L158C MoFe protein samples were not found to significantly differ from one another (H_2 production under 1 atm Ar, one-way ANOVA, P = 0.8755). Our observation of >50% residual H^+ reduction activity per MoFe protein following strep-inhibition suggests that uninhibited α-C45A/L158C MoFe protein (more globally, nitrogenase) employs a negative cooperativity mechanism for H^+ reduction in the steady-state/continued turnover. This is consistent with the previous observation that the electron delivery cycle of the Fe protein also experiences negative cooperativity (cooperativity in terms of product formation in the steady-state was not evaluated).[12]

Another recent study observed that the MoFe protein does not exhibit cooperativity when producing H_2 under Ar.[13,33] As mentioned in the introduction, this approach inhibited Fe protein access to one αß half by introducing the L127Δ Fe protein to competitively form a

non-dissociating complex (or by using a mismatched Fe protein from a different organism). Our gel filtration data (Figure 5D) indicates that MoFe proteins half-inhibited by L127Δ Fe protein may in fact not form non-dissociating complexes on only one αß half of the MoFe protein, leading to a population of Fe_2:MoFe and free MoFe proteins (*i.e.*, in the case of a 1:1 $Fe_{L127\Delta}$:MoFe protein ratio). The Strep-inhibition approach reported here is not anticipated to introduce a long-range conformational change at the second Fe protein-binding site. Further, we observed that the affinity of the Fe protein to the MoFe protein is relatively unchanged (H_2 production under 1 atm Ar) in the presence of the Strep inhibitor on the MoFe protein (Supporting Information, Figure S17).

The specific activity for NH_3 formation under 1 atm N_2 was investigated after strep-inhibition of the MoFe protein, where a residual specific activity of 53 ± 4% (mean of the three reactions) was observed in comparison to the uninhibited MoFe protein (Figure 6B, and Figure S18, Supporting Information). The specific activities of these three functionalization reactions were found to differ for NH_3 production only weakly under 1 atm N_2 (one-way ANOVA, P = 0.0467). As shown in Figure 6C the % electron distribution towards N_2 fixation remains between 50-60% for all three strep-inhibited MoFe proteins under 1 atm N_2 (one-way ANOVA, P = 0.1896), confirming under these conditions that nitrogenase's selectivity towards N_2 on both αß halves is not a result of cooperativity. The determination of whether a cooperativity mechanism is at play during N_2 fixation is more delicate and is treated in the following section.

A strep-inhibition reaction was performed on a larger batch of α-C45A/L158C MoFe protein (~18 mg of Strep-inhibited protein obtained) to evaluate Michaelis-Menten kinetic parameters and the distribution of electrons towards N_2 fixation under a range of N_2 partial pressures (Figure 7A). Neither the Michaelis constant (K_M) nor maximum velocity (V_{max}) for the α-C45A/L158C MoFe protein were found to differ statistically from the wild-type MoFe protein (P = 0.2065 and P = 0.3873, Figure S19/Table S5, Supporting Information). Importantly, the quantity of H_2 produced under increasing pressures of N_2 is consistent with increased N_2 fixation by nitrogenase (electron allocation towards N_2 fixation), with uninhibited and strep-inhibited MoFe proteins exhibiting similar trends. Figure 7B reports the percentage e^- distribution between N_2 fixation and H_2 formation by uninhibited and strep-inhibited α-C45A/L158C MoFe proteins with increasing N_2 pressures. A maximum electron distribution of approximately 65% towards N_2 fixation was observed under these conditions, with no clear difference in electron distribution between uninhibited or inhibited α-C45A/L158C MoFe observed over this range of N_2 pressures. In contrast to the k_{cat} and k_{cat}/K_M parameters, this

suggests that a cooperativity mechanism, or the arrangement of the MoFe protein as an $\alpha_2ß_2$ heterotetramer, does not contribute towards the selectivity of nitrogenase towards N_2 under these conditions.

In agreement with previous reports, the total electron flux (total electrons consumed determined by production quantification, where $H_2 = 2e^-$ and $NH_3 = 3e^-$) of uninhibited and strep-inhibited α-C45A/L158C MoFe proteins was observed to decrease following the introduction of increasing N_2 partial pressure in the reaction vials (Figure 8A, 100% electron flux represents activity under 1 atm Ar).[27–29] However, no clear difference in the decrease of electron flux with or without strep-inhibition over 0-1 atm of N_2 was observed, where Figure 8B compares the percentage remaining total electron flux over 0-1 atm N_2 after having strep-inhibited the α-C45A/L158C MoFe protein (in comparison to uninhibited α-C45A/L158C MoFe protein). Importantly, this remains <100% regardless of the N_2 partial pressure (mean = 75%), suggesting that the MoFe protein indeed follows a negative cooperativity mechanism with respect to electron delivery. This is consistent with the earlier finding that the Fe protein electron delivery cycle exhibits negative cooperativity in the pre-steady-state.[12] We therefore conclude that negative cooperativity for electron delivery does in fact propagate itself in the formation of total products (including the production of NH_3), although this does not significantly impact the selectivity of N_2 fixation by this enzyme complex. Considering the proposal that negative cooperativity may arise from the MoFe protein only interacting with one Fe protein at a time during turnover, and our observations of >50% activity strep-inhibited MoFe proteins, we hypothesize that negative cooperativity is introduced after the association of a second Fe protein to a Fe_1:MoFe protein turnover complex.[17]

Badalyan *et al*. employed voltammetry to determine the rate constant for electron consumption by nitrogenase to be 14 s^{-1}; importantly, this value was observed to remain constant in the presence of N_2, in fact suggesting that electron flux to nitrogenase remains constant.[30] In this study 23% of electrons were unaccounted for (not detected as H_2 or NH_3). More recently, Lee *et al*. observed that nitrogenase proteins purified in the strict absence of DT were unable to undergo continued turnover, where it was proposed that sulfite (resulting from DT decomposition/oxidation) plays an additional sulfur-recharging role upon reduction at the FeMoco.[26] Such a reaction may explain the apparent decrease in total electron flux upon N_2 fixation by nitrogenase.

6.5. Conclusions

Long-range (>95 Å) communication between the Fe proteins interacting with the two αß halves of nitrogenase's MoFe protein is thought to be of mechanistic importance.[14] Previous studies have observed (i) negative cooperativity for electron delivery in the pre-steady-state, and (ii) no cooperativity for H^+ reduction by nitrogenase upon "locking" one αß MoFe half in an inactive Fe protein-bound state.[12,13]

We sought to inhibit one αß half of the MoFe protein by an approach that was not expected to impact Fe protein transient association behavior on the remaining uninhibited active site. More importantly, we sought to determine whether cooperativity is strictly necessary for N_2 fixation by nitrogenase. We conclude that negative cooperativity plays a role in both H^+ reduction to H_2 as well as overall electron delivery to the MoFe protein (and thus, also for N_2 fixation), and that cooperativity across the MoFe protein is not strictly necessary for the fixation of N_2 to NH_3. The k_{cat} per FeMoco site for NH_3 production was observed to increase by 31% upon strep-inhibition (from 1.3 to >1.7 s^{-1}), consistent with negative cooperativity in N_2 fixation by nitrogenase in terms of electron delivery (selectivity towards N_2 is not impacted). It remains difficult to precisely determine the magnitude of negative cooperativity induced suppression of nitrogenase activity in a single αß MoFe protein half due to the possible presence of doubly inhibited MoFe protein in our strep-inhibited sample. A potential method to surmount this issue could be the co-expression of two copies of *nifD* (*i.e., nifD* and *nifD**) followed by tandem affinity purification.

While *in vitro* investigation into nitrogenase's cooperativity mechanism provides insight into its catalytic mechanism, it remains important to determine the importance and magnitude of negative cooperativity of nitrogenase in the context of *in vivo* turnover (in *A. vinelandii*), where (i) Fe:MoFe protein ratios could be more dynamic or (ii) additional partners could play a role (such as maturases, activity modulating proteins such as CowN,[35] O_2-protection proteins such as the Shethna FeSII protein[36,37] or indeed an as-of-yet unidentified allosteric effector).

6.6. Materials and methods

6.6.1. Culturing of *Azotobacter vinelandii*. Extensive procedures are reported in the Supporting Information. Briefly, MoFe and Fe nitrogenase proteins were isolated from various *A. vinelandii* strains were cultivated on a modified Burke's medium. The *nif* operon was derepressed after overnight growth by centrifugation and resuspension into fresh Burke's

medium lacking fixed N (NH_4^+). Cells were harvested by centrifugation after ~3.5h and stored at −80 ºC until use.

6.6.2. Fe protein purification. All Fe protein purifications were performed in a COY anaerobic chamber (<5% H_2/>95% N_2, Michigan, USA). Wild-type Fe protein was purified from a strain derived from *A. vinelandii* DJ that was modified to introduce an 8x His-tag to the N-terminus of NifD (*A. vinelandii* RS1).[20] L127Δ Fe protein was purified from *A. vinelandii* strain DJ1065 (provided by Dennis Dean, Virginia Tech). Cells from a 12 L culture were thawed and resuspended in anoxic lysis buffer (50 mM Tris/HCl, pH 8.0) containing 5 mM dithionite (DT) and 37% v/v glycerol. All Fe protein purification buffers contained 2 mM DT, except for the lysis buffer mentioned above. After incubation for 20 mins, the cells were collected by centrifugation and resuspended in glycerol-free lysis buffer to induce cell lysis by osmotic shock. The lysate was incubated on ice for a further 15 minutes and clarified by centrifugation at 30'000 x *g* (4 ºC, 1 h). A post-lysis buffer (2 M NaCl, 234 mM Tris/HCl, 2 mM DT, pH 8.0) was added to the cell-free supernatant to achieve a final NaCl concentration of 0.3 M. The cell-free supernatant was next passed over a HisTrap HP column (5 mL column volume, Cytiva) to remove His-tagged MoFe protein (below). The column flow-through (containing Fe protein) was next diluted with NaCl-free buffer (50 mM Tris/HCl, pH 8.0, 2 mM DT) to obtain a final NaCl concentration of 0.1 M. The Fe protein was first purified over a HiPrep Q-Sepharose HP 16/10 column (20 mL column volume, Cytiva) and eluted over a linear NaCl gradient of 0.2 – 0.65 M NaCl. Following concentration to <10 mL using a Merck-Millipore stirred concentration cell (30 kDa molecular weight cut-off membrane), the Fe protein was next purified by size-exclusion chromatography over a HiPrep 26/60 Sephacryl S-200 HR column (320 mL column volume, Cytiva) using a Tris running buffer (50 mM Tris/HCl, pH 8.0, 0.5 M NaCl). The eluted Fe protein was concentrated further to >20 mg/mL and flash-frozen in liquid nitrogen until use.

6.6.3. MoFe protein purification. All MoFe protein purifications were performed in a COY anaerobic chamber (<5% H_2/>95% N_2, Michigan, USA). Wild-type MoFe protein carrying an 8x His-tag on the N-terminus of NifD was purified from *A. vinelandii* strain RS1. α-C45A/L158C MoFe protein was purified from a derivative of *A. vinelandii* RS1 (*A. vinelandii* strain "M1", Supporting Information), which carried the same 8x His-tag on the N-terminus of NifD. The cell-free supernatant (prepared as above) was first passed over a HisTrap HP column (5 mL column volume, Cytiva). The running buffers for this step (50 mM Tris/HCl, pH 8.0, ± 0.3 M imidazole) were incubated overnight in the anaerobic chamber and did not contain DT (all subsequent buffers did not contain DT). MoFe proteins bound to the His-resin

were first washed with >3 column volumes of DT-free buffer (0 mM imidazole) to remove DT introduced during cell lysis. After a 20 mM imidazole washing step, MoFe proteins were eluted using 0.3 M imidazole. The eluted MoFe proteins were next diluted with Tris buffer (50 mM Tris/HCl, pH 8.0) to reach a final NaCl concentration of 0.1 M prior to being loaded onto a pre-equilibrated Q-Sepharose FF 16/10 column (20 mL column volume, flow-rate = 20 mL/min). DT-free MoFe proteins were eluted over a linear gradient of 0.2 – 0.65 M NaCl and concentrated to >20 mg/mL using a 100 kDa stirred concentrator cell (fed with ultra-high-purity N_2 5.0) prior to being flash-frozen in liquid nitrogen until use.

6.6.4. MoFe protein functionalization with DTB inhibitor. DT-free MoFe protein was treated with DTB-inhibitor (Supporting Information) freshly prepared in 0.1 M MOPS/NaOH buffer (pH 7.0) for 4 h at room temperature within the COY anoxic chamber. After 4 h, unreacted maleimide was quenched and MoFe proteins were reduced by the addition of DT to a final concentration of 2 mM.

6.6.5. MoFe protein functionalization with Strep-tag inhibitor and conjugate purification. A lyophilized synthetic Strep-tag peptide containing an N-terminal maleimide functionality (sequence = GGGWSHPQFEK) was obtained from GenScript (USA) and used as received. This 1.4 kDa peptide as as both (i) a steric inhibitor of Fe protein association to the MoFe protein and (ii) a Strep-tag for purification of the MoFe protein conjugates. Strep-inhibited α-C45A/L158C MoFe protein was obtained following the incubation of DT-free α-C45A/L158C MoFe protein with 0.5 molar equivalents of the Strep-tag inhibitor (0.5 inhibitor per MoFe protein, resulting in a 1:4 Strep-tag to surface Cys ratio, freshly dissolved in 0.1 M MOPS/NaOH pH 7.0 to a final concentration of 1-3 mg/mL) for 4 h within the anoxic glovebox. The reaction was quenched (and the MoFe protein reduced) by the addition of DT to a final concentration of 2 mM. Functionalized/inhibited MoFe protein was next purified over a StrepTrap XT column (5 mL column volume, Cytiva) pre-equilibrated with MOPS buffer (0.1 M MOPS/NaOH, 0.2 M NaCl, 2 mM DT, pH 7.0). Unreacted MoFe proteins were collected from the column flow-through and strep-inhibited Strep-MoFe proteins were collected following the application of 50 mM biotin to the column. MoFe proteins were concentrated as above and flash-frozen in liquid nitrogen until use.

6.6.6. Nitrogenase activity assays. Briefly, nitrogenase activity assays were performed in 13 mL septum-sealed glass vials containing 1 mL of an ATP-regenerating MOPS buffer (100 mM MOPS/NaOH, pH 7.0, 5 mM ATP, 30 mM phosphocreatine, 1.3 mg bovine serum albumin, 0.2 mg creatine phosphokinase from rabbit muscle, and 10 mM DT). All activity

assays employed DT as the electron donor. All reactions were assembled within an Ar-filled anoxic glovebox (Jacomex, France). MoFe proteins (0.1 mg) and Fe proteins (0.48 mg) were included with a 16.6:1 Fe to MoFe protein ratio, sealed, and vented to atmospheric pressure. The reaction vials were then heated to 30 ºC in a shaking water bath and initiated by the injection of $MgCl_2$ to a final concentration of 10 mM using a gas-tight syringe. After 8 minutes, reactions were terminated by the addition of 300 µL of 400 mM EDTA. H_2 was quantified using a GC-TCD (molecular sieve 5 Å column, Ar carrier gas, SRI Instruments model 8610C). NH_3 was quantified by the ortho-phthalaldehyde method (corrected to controls and assays performed under 1 atm Ar) using NH_4Cl as the standard.[20,38]

6.6.7. Crystallization of α-C45A/L158C MoFe protein. The purified enzyme in Tris/HCl buffer (50 mM Tris/HCl, 300 mM NaCl, 2 mM DT, pH 8.0) was crystallized anaerobically at 8 mg/mL^{-1} (under 100% N_2 atmosphere) with an OryxNano (Douglas Instrument, UK). The initial screening was performed at 20 °C using the sitting drop method on 96-Well MRC 2-drop crystallization plates in polystyrene (SWISSCI) containing 90 µL of crystallization solution in the reservoir. The protein sample (0.5 µL) was mixed with 0.5 µL reservoir solution. Crystals were transferred and stored anaerobically in a Coy tent (N_2/H_2, 97:3). Thin brown plate crystals appeared after a few weeks. The reservoir solution contained 25 % w/v Polyethylene glycol 3,350, 100 mM BIS-TRIS at pH 5.5, and 200 mM Lithium sulfate.

6.6.8. X-ray crystallography data collection and refinement. Crystal handling was done inside a Coy tent under an anaerobic atmosphere (N_2/H_2, 97:3). Crystals were soaked in the crystallization solution supplemented with 15 % v/v glycerol as a cryo-protectant before being frozen in liquid nitrogen. Crystals were tested and collected at 100 K at the Swiss Light Source, X06DA – PXIII. Due to the high anisotropy, data were processed and scaled with *autoPROC*.[39] The relative resolution limits along the unit cell axis are: a = 4.34 Å, b = 4.07 Å, and c = 3.03 Å. The molecular replacement was done with *PHASER* from the *PHENIX* package.[40] The model was then manually built with *COOT* and further refined with *PHENIX* without hydrogens.[41] The last cycles of refinement were performed with Buster, and the final one with *PHENIX*. The model was validated by the Molprobity server (used on the 15th of August 2022).[42,43]

6.7. Associated content and Supporting Information

Supporting information: protein production and purification methodology; activity assays; analytical gel filtration methodology; limited proteolysis methodology; additional X-ray crystallography data; synthesis of DTB inhibitor; *sacB*-based markerless mutagenesis;

additional SDS-PAGE and western blotting; activity assays for DTB-inhibited wild-type MoFe protein; binding and elution of strep-inhibited MoFe on StrepTrap XT (streptactin) columns; native-PAGE; Fe protein titration of MoFe protein; specific activities for NH_3 and H_2 formation by strep-inhibited MoFe protein; Michaelis-Menten kinetic parameters for wild-type MoFe protein. Raw research data and sample reports have been deposited on Zenodo (https://doi.org/10.5281/zenodo.6865680) and contains: specific activities, proteomics reports, raw proteomics data, unprocessed X-ray crystallography, sequencing data for *A. vinelandii* strain producing α-C45A/L158C MoFe protein, and sequences of plasmids used to generate this mutant strain (pK18mobsacB).

6.8. Author information

Corresponding Author * ross.milton@unige.ch

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This publication was created as part of NCCR Catalysis (grant number 180544), a National Centre of Competence in Research funded by the Swiss National Science Foundation. RDM thanks the Ernst and Lucie Schmidheiny foundation for support. This research was funded by the Swiss National Science Foundation (Project Grant SNSF 310030_189246, to S.H.) and the Swiss National Centre of Competence in Research (NCCR) Chemical Biology (S.H.). The research was funded by the Max-Planck Gesellschaft and was supported by the Deutsche Forschungsgemeinschaft (DFG) Schwerpunktprogram 1927 “Iron-sulfur for Life“ (WA 4053/1-1).

6.9. Notes

One-way ANOVA tests (Brown-Forsythe) and *t* tests were performed using GraphPad Prism. The following notation was employed for statistical significance: ns (not significant) for $P > 0.05$, * for $P \leq 0.05$, ** for $P \leq 0.01$, and *** for $P \leq 0.001$. Graphs were plotted using GraphPad Prism. Figures were created using a combination of Affinity Designer, ChemDraw, ChimeraX and PyMol. RDM is an Early Career Advisory Board member of JACS Au for 2022. The structural model of the α-C45A/L158C MoFe protein has been deposited on the Protein Data Bank (www.pdb.org), available under the accession code PDB 8BTS.

6.10. Acknowledgements

We thank Plinio Maroni, Olivier Vassalli, Darren Martin, Yashar Sadian, Léa Di Luzio, CHIAM technical staff, Isabelle Worms and Yibo Wu for discussions, supporting experiments and technical support. We thank Aleksandar Salim for assistance with mass spectrometry measurements. We thank Dennis Dean and Valerie Cash for sharing *A. vinelandii* strains DJ2102 (Strep-tagged MoFe protein (N-term of NifD)), DJ1065 (NifH L127Δ) and DJ2192 (α-C45S MoFe protein, N-terminal His-tag on NifD). We thank the Proteomics Core Facility within the Faculty of Medicine at the University of Geneva.

6.11. Abbreviations

DTB, desthiobiotin-maleimide inhibitor; SDS-PAGE, sodium dodecyl sulfate poly(acrylamide) gel electrophoresis; ATP, adenosine triphosphate; ADP, adenosine diphosphate; FeMoco, FeMo cofactor; DTB, desthiobiotin; DT, dithionite; LC-ESI-MS/MS; liquid chromatography – electrospray ionization – tandem mass spectrometry.

6.12. References

(1) 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. https://doi.org/10.1021/acs.chemrev.9b00556.

(2) Einsle, O.; Rees, D. C. Structural Enzymology of Nitrogenase Enzymes. *Chem. Rev.* **2020**, *120* (12), 4969–5004. https://doi.org/10.1021/acs.chemrev.0c00067.

(3) Jasniewski, A. J.; Lee, C. C.; Ribbe, M. W.; Ribbe, M. W.; Hu, Y. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. *Chem. Rev.* **2020**, *120* (12), 5107–5157. https://doi.org/10.1021/acs.chemrev.9b00704.

(4) Rutledge, H. L.; Tezcan, F. A. Electron Transfer in Nitrogenase. *Chem. Rev.* **2020**, *120* (12), 5158–5193. https://doi.org/10.1021/acs.chemrev.9b00663.

(5) Burén, S.; Jiménez-Vicente, E.; Echavarri-Erasun, C.; Rubio, L. M. Biosynthesis of Nitrogenase Cofactors. *Chem. Rev.* **2020**, *120* (12), 4921–4968. https://doi.org/10.1021/acs.chemrev.9b00489.

(6) 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. https://doi.org/10.1021/acs.chemrev.9b00650.

(7) Simpson, F. B.; Burris, R. H. A Nitrogen Pressure of 50 Atmospheres Does Not Prevent Evolution of Hydrogen by Nitrogenase. *Science (80-.).* **1984**, *224* (4653), 1095–1097. https://doi.org/10.1126/science.6585956.

(8) Harris, D. F.; Lukoyanov, D. A.; Kallas, H.; Trncik, C.; Yang, Z. Y.; Compton, P.; Kelleher, N.; Einsle, O.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism to Achieve N2 Reduction. *Biochemistry* **2019**, *58* (30), 3293–3301. https://doi.org/10.1021/acs.biochem.9b00468.

(9) Thorneley, R.; Lowe, D. Kinetics and Mechanism of the Nitrogenase Enzyme System. In *Molybdenum Enzymes*; Spiro, T., Ed.; Wiley-Interscience: New York, 1985; pp 221–284.

(10) Rohde, M.; Sippel, D.; Trncik, C.; Andrade, S. L. A.; Einsle, O. The Critical E4 State of Nitrogenase Catalysis. *Biochemistry* **2018**, *57* (38), 5497–5504. https://doi.org/10.1021/acs.biochem.8b00509.

(11) Yang, Z. Y.; Ledbetter, R.; Shaw, S.; Pence, N.; Tokmina-Lukaszewska, M.; Eilers, B.; Guo, Q.; Pokhrel, N.; Cash, V. L.; Dean, D. R.; Antony, E.; Bothner, B.; Peters, J. W.; Seefeldt, L. C. Evidence That the Pi Release Event Is the Rate-Limiting Step in the Nitrogenase Catalytic Cycle. *Biochemistry* **2016**, *55* (26), 3625–3635. https://doi.org/10.1021/acs.biochem.6b00421.

(12) Danyal, K.; Shaw, S.; Page, T. R.; Duval, S.; Horitani, M.; Marts, A. R.; Lukoyanov, D.; Dean, D. R.; Raugei, S.; Hoffman, B. M.; Seefeldt, L. C.; Antony, E. Negative Cooperativity in the Nitrogenase Fe Protein Electron Delivery Cycle. *Proc. Natl. Acad. Sci.* **2016**, *113* (40), E5783–E5791. https://doi.org/10.1073/pnas.1613089113.

(13) Truscott, S.; Lewis, R. S.; Watt, G. D. Positive Cooperativity during Azotobacter Vinelandii Nitrogenase-Catalyzed Acetylene Reduction. *Biophys. Chem.* **2021**, *277*, 106650. https://doi.org/10.1016/j.bpc.2021.106650.

(14) Huang, Q.; Tokmina-Lukaszewska, M.; Johnson, L. E.; Kallas, H.; Ginovska, B.; Peters, J. W.; Seefeldt, L. C.; Bothner, B.; Raugei, S. Mechanical Coupling in the Nitrogenase Complex. *PLoS Comput. Biol.* **2021**, *17* (3), e1008719. https://doi.org/10.1371/journal.pcbi.1008719.

(15) Clarke, T. A.; Fairhurst, S.; Lowe, D. J.; Watmough, N. J.; Eady, R. R. Electron

Transfer and Half-Reactivity in Nitrogenase. In *Biochemical Society Transactions*; Portland Press, 2011; Vol. 39, pp 201–206. https://doi.org/10.1042/BST0390201.

(16) Grossmann, J. G.; Hasnain, S. S.; Yousafzai, F. K.; Eady, R. R. Evidence for the Selective Population of FeMo Cofactor Sites in MoFe Protein and Its Molecular Recognition by the Fe Protein in Transition State Complex Analogues of Nitrogenase. *J. Biol. Chem.* **2001**, *276* (9), 6582–6590. https://doi.org/10.1074/jbc.M005350200.

(17) Rutledge, H. L.; Cook, B. D.; Nguyen, H. P. M.; Herzik, M. A.; Akif Tezcan, F. Structures of the Nitrogenase Complex Prepared under Catalytic Turnover Conditions. *Science (80-.).* **2022**, *377* (6608), 865–869. https://doi.org/10.1126/science.abq7641.

(18) Harris, D. F.; Yang, Z. Y.; Dean, D. R.; Seefeldt, L. C.; Hoffman, B. M. Kinetic Understanding of N2 Reduction versus H2 Evolution at the E4(4H) Janus State in the Three Nitrogenases. *Biochemistry* **2018**, *57* (39), 5706–5714. https://doi.org/10.1021/acs.biochem.8b00784.

(19) Roth, L. E.; Tezcan, F. A. ATP-Uncoupled, Six-Electron Photoreduction of Hydrogen Cyanide to Methane by the Molybdenum–Iron Protein. *J. Am. Chem. Soc.* **2012**, *134* (20), 8416–8419. https://doi.org/10.1021/ja303265m.

(20) Milton, R. D.; Cai, R.; Abdellaoui, S.; Leech, D.; De Lacey, A. L.; Pita, M.; Minteer, S. D. Bioelectrochemical Haber–Bosch Process: An Ammonia-Producing H2/N2 Fuel Cell. *Angew. Chemie - Int. Ed.* **2017**, *56* (10), 2680–2683. https://doi.org/10.1002/anie.201612500.

(21) Gu, W.; Ul Haque, M. F.; Baral, B. S.; Turpin, E. A.; Bandow, N. L.; Kremmer, E.; Flatley, A.; Zischka, H.; DiSpirito, A. A.; Semrau, J. D. A TonB-Dependent Transporter Is Responsible for Methanobactin Uptake by Methylosinus Trichosporium OB3b. *Appl. Environ. Microbiol.* **2016**, *82* (6), 1917–1923. https://doi.org/10.1128/AEM.03884-15.

(22) Puri, A. W.; Owen, S.; Chu, F.; Chavkin, T.; Beck, D. A. C.; Kalyuzhnaya, M. G.; Lidstrom, M. E. Genetic Tools for the Industrially Promising Methanotroph Methylomicrobium Buryatense. *Appl. Environ. Microbiol.* **2015**, *81* (5), 1775–1781. https://doi.org/10.1128/AEM.03795-14.

(23) 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), 1696–1700. https://doi.org/10.1126/science.1073877.

(24) 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*. American Association for the Advancement of Science November 18, 2011, p 940. https://doi.org/10.1126/science.1214025.

(25) Sornay, C.; Vaur, V.; Wagner, A.; Chaubet, G. An Overview of Chemo- and Site-Selectivity Aspects in the Chemical Conjugation of Proteins. *R. Soc. Open Sci.* **2022**, *9*, 211563. https://doi.org/10.1098/rsos.211563.

(26) Lee, C. C.; Kang, W.; Jasniewski, A. J.; Stiebritz, M. T.; Tanifuji, K.; Ribbe, M. W.; Hu, Y. Evidence of Substrate Binding and Product Release via Belt-Sulfur Mobilization of the Nitrogenase Cofactor. *Nat. Catal.* **2022**, *5*, 443–454. https://doi.org/10.1038/s41929-022-00782-7.

(27) Harris, D. F.; Lukoyanov, D. A.; Shaw, S.; Compton, P.; Tokmina-Lukaszewska, M.; Bothner, B.; Kelleher, N.; Dean, D. R.; Hoffman, B. M.; Seefeldt, L. C. Mechanism of N2 Reduction Catalyzed by Fe-Nitrogenase Involves Reductive Elimination of H2. *Biochemistry* **2018**, *57* (5), 701–710. https://doi.org/10.1021/acs.biochem.7b01142.

(28) Hageman, R. V.; Burris, R. H. Electron Allocation to Alternative Substrates of Azotobacter Nitrogenase Is Controlled by the Electron Flux through Dinitrogenase. *BBA - Bioenerg.* **1980**, *591* (1), 63–75. https://doi.org/10.1016/0005-2728(80)90220-0.

(29) Wherland, S.; Burgess, B. K.; Stiefel, E. I.; Newton, W. E. Nitrogenase Reactivity: Effects of Component Ratio on Electron Flow and Distribution during Nitrogen Fixation. *Biochemistry* **1981**, *20* (18), 5132–5140. https://doi.org/10.1021/bi00521a006.

(30) Badalyan, A.; Yang, Z. Y.; Seefeldt, L. C. A Voltammetric Study of Nitrogenase Catalysis Using Electron Transfer Mediators. *ACS Catal.* **2019**, *9* (2), 1366–1372. https://doi.org/10.1021/acscatal.8b04290.

(31) Tal, S.; Chun, T. W.; Gavini, N.; Burgess, B. K. The ΔnifB (or ΔnifE) FeMo Cofactor-Deficient MoFe Protein Is Different from the ΔnifH Protein. *J. Biol. Chem.* **1991**, *266* (16), 10654–10657. https://doi.org/10.1016/S0021-9258(18)99273-2.

(32) Chiu, H.; Peters, J. W.; Lanzilotta, W. N.; Ryle, M. J.; Seefeldt, L. C.; Howard, J. B.;

Rees, D. C. MgATP-Bound and Nucleotide-Free Structures of a Nitrogenase Protein Complex between the Leu 127 Delta-Fe-Protein and the MoFe-Protein. *Biochemistry* **2001**, *40*. https://doi.org/10.1021/bi001645e.

(33) Ryle, M. J.; Seefeldt, L. C. Elucidation of a MgATP Signal Transduction Pathway in the Nitrogenase Iron Protein: Formation of a Conformation Resembling the MgATP-Bound State by Protein Engineering. *Biochemistry* **1996**, *35* (15), 4766–4775. https://doi.org/10.1021/bi960026w.

(34) Malviya, A.; Vrabec, J. Henry's Law Constant of Nitrogen, Oxygen, and Argon in Ternary Aqueous Alcoholic Solvent Mixtures. *J. Chem. Eng. Data* **2020**, *65* (3), 1189–1196. https://doi.org/10.1021/acs.jced.9b00571.

(35) Medina, M. S.; Bretzing, K. O.; Aviles, R. A.; Chong, K. M.; Espinoza, A.; Garcia, C. N. G.; Katz, B. B.; Kharwa, R. N.; Hernandez, A.; Lee, J. L.; Lee, T. M.; Lo Verde, C.; Strul, M. W.; Wong, E. Y.; Owens, C. P. CowN Sustains Nitrogenase Turnover in the Presence of the Inhibitor Carbon Monoxide. *J. Biol. Chem.* **2021**, *296*, 100501. https://doi.org/10.1016/j.jbc.2021.100501.

(36) Shethna, Y. I. Non-Heme Iron (Iron-Sulfur) Proteins of Azotobacter Vinelandii. *Biochim. Biophys. Acta, Bioenerg.* **1970**, *205* (1), 58–62. https://doi.org/http://dx.doi.org/10.1016/0005-2728(70)90061-7.

(37) Schlesier, J.; Rohde, M.; Gerhardt, S.; Einsle, O. A Conformational Switch Triggers Nitrogenase Protection from Oxygen Damage by Shethna Protein II (FeSII). *J. Am. Chem. Soc.* **2016**, *138* (1), 239–247. https://doi.org/10.1021/jacs.5b10341.

(38) Corbin, J. L. Liquid Chromatographic-Fluorescence Determination of Ammonia from Nitrogenase Reactions: A 2-Min Assay. *Appl. Environ. Microbiol.* **1984**, *47* (5), 1027–1030.

(39) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data Processing and Analysis with the AutoPROC Toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 293–302. https://doi.org/10.1107/S0907444911007773.

(40) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, 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. https://doi.org/10.1107/S2059798319011471.

(41) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.

(42) Bricogne, G.; Blanc, E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.; Sharff, A.; Smart, O. S.; Vonrhein, C.; Womack, T. O. Buster. *Buster version 2.10.4* **2017**, No. Cambridge, United Kingdom:, Global Phasing Ltd.

(43) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

6.13. Figures

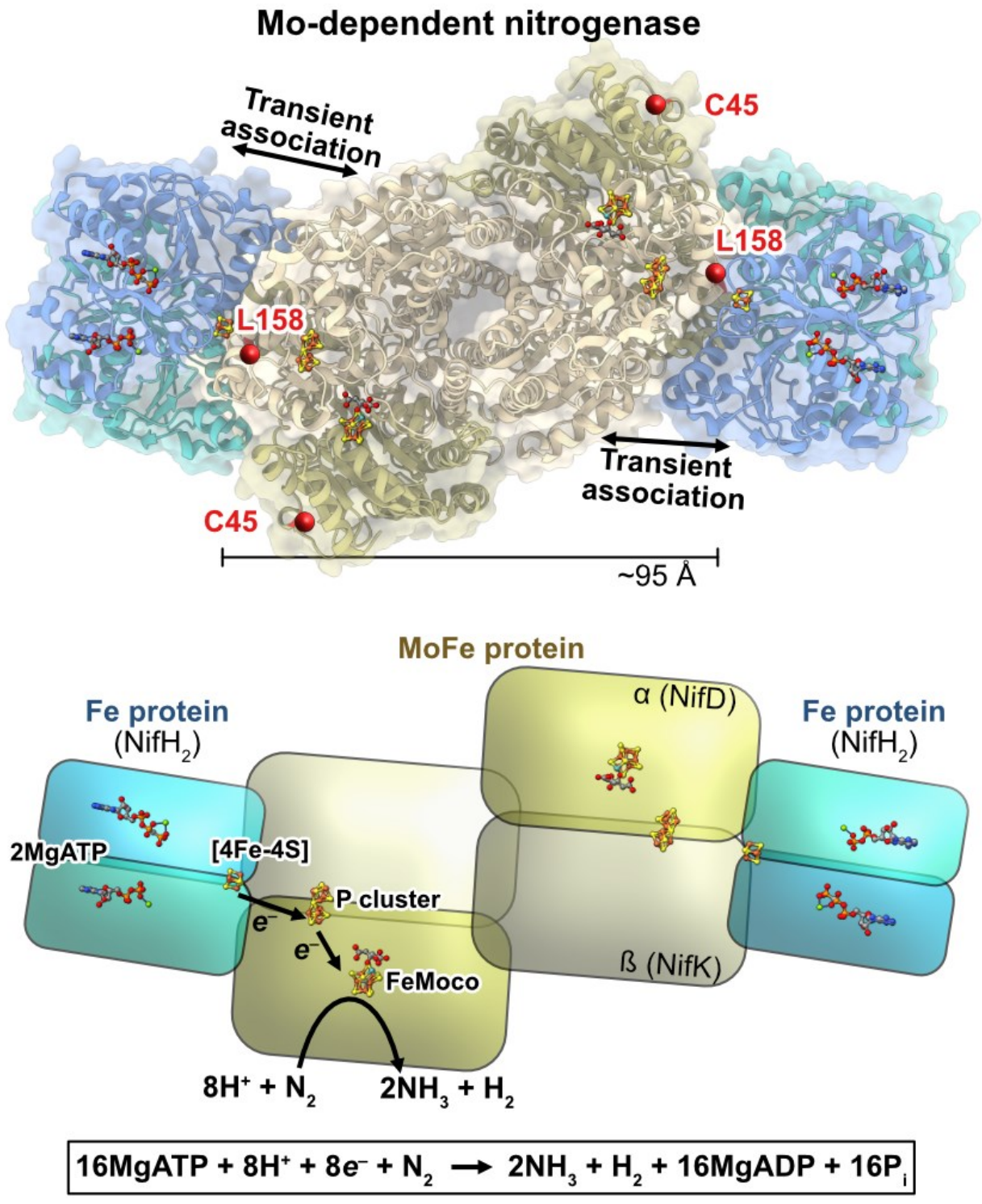


Figure 1: Structure of the Mo-dependent nitrogenase transition state complex between the Fe protein and MoFe protein (PDB: 4WZA). In this complex, 2MgATP was replaced with MgADP and non-hydrolyzable MgAMPPCP to form enable stable complex formation. The α-C45 and α-L158 residues are highlighted with red spheres. Fe = rust, S = yellow, Mo = cyan, C = gray, O = red, N = blue, Mg = green, P = orange.

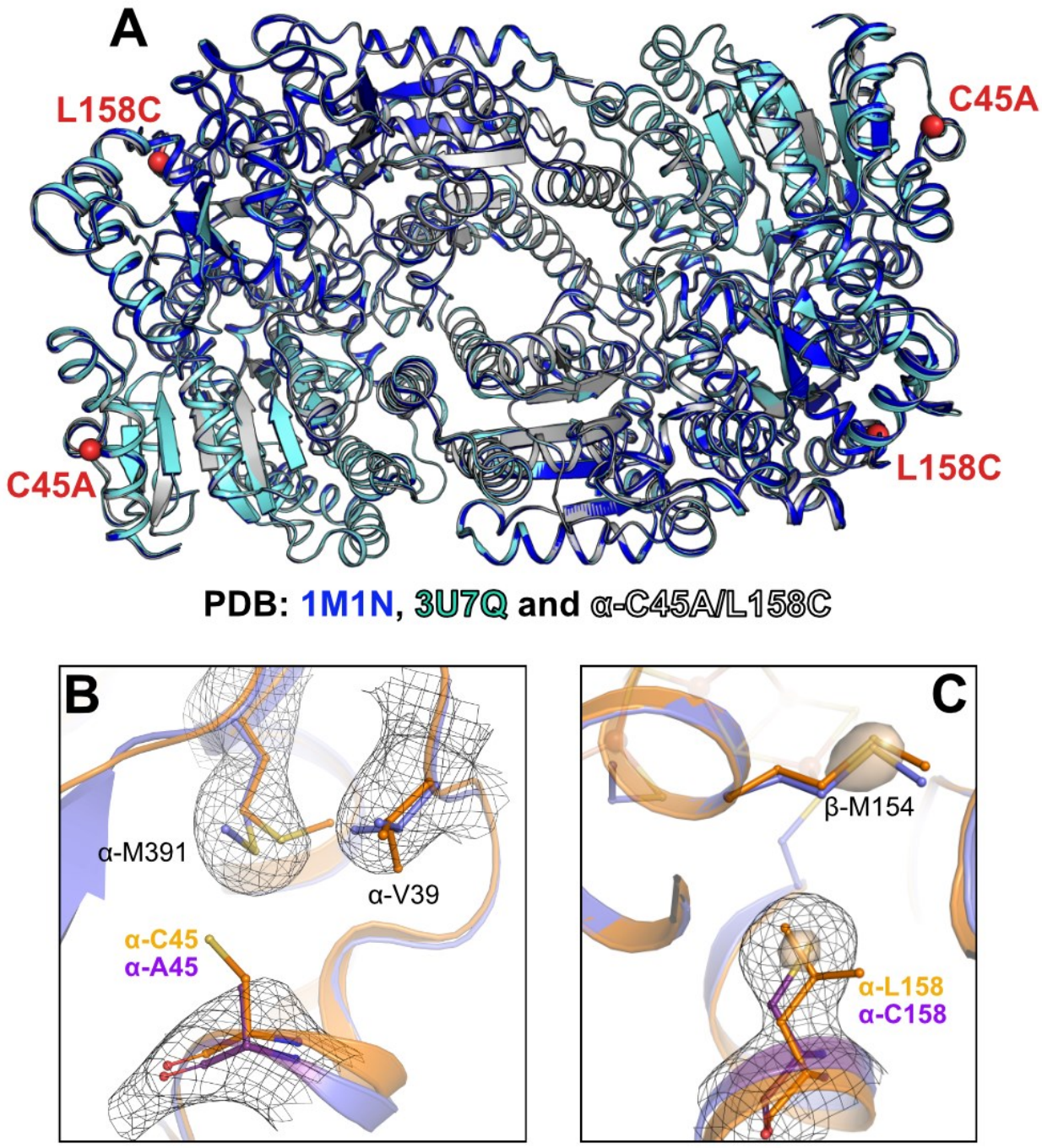


Figure 2: (A) X-ray crystal structure overlay of *A. vinelandii* MoFe proteins PDB 1M1N (1.16 Å resolution, blue), PDB 3U7Q (1 Å resolution, cyan) and the α-C45A/L158C MoFe protein from this work (3.03 Å resolution, white, PDB 8BTS). Electron densities of the (B) α-A45 (violet) and (C) α-C158 regions, overlayed with the X-ray crystal structure of wild-type MoFe protein (PDB 3U7Q). The P cluster and FeMoco were omitted for clarity (Supporting Information). The $2F_o$-F_c electron density maps, contoured at 1 σ, are shown by a black mesh. In panel C, an additional $2F_o$-F_c electron density contoured at 4.5 σ was superposed and represented as a transparent surface.

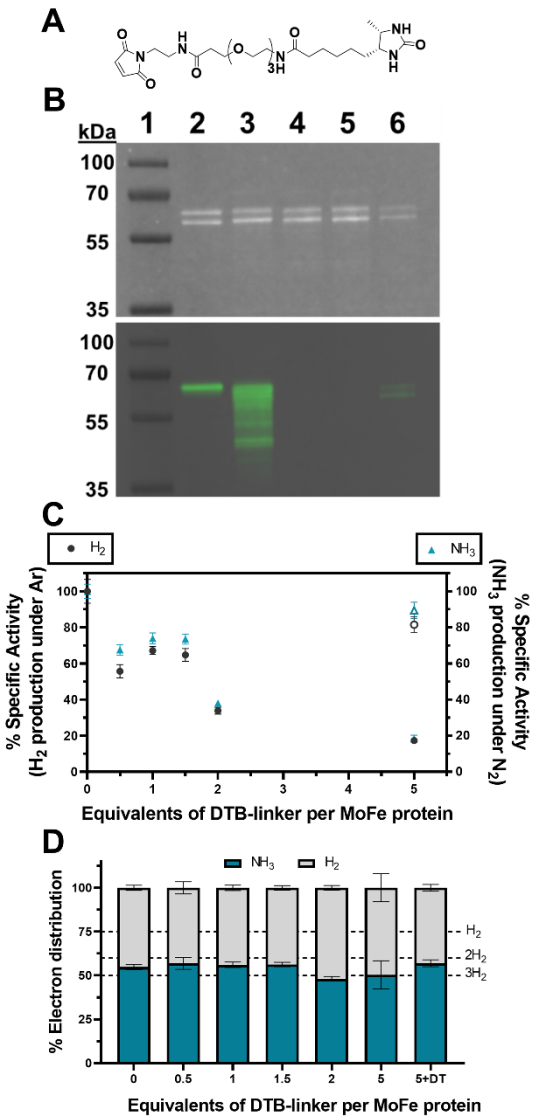
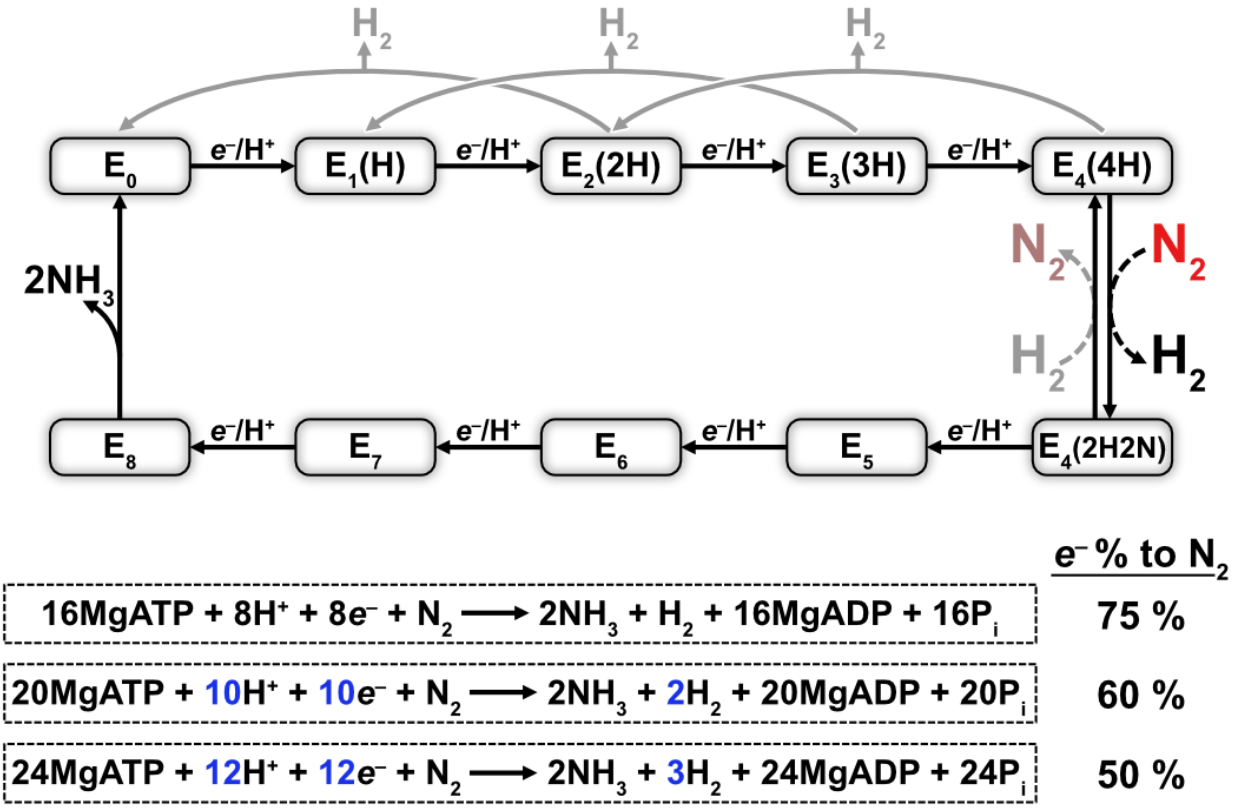


Figure 3: Analysis of maleimide based DTB steric inhibitor reaction with MoFe protein (A) Structure of maleimide-DTB inhibitor. (B) SDS-PAGE analysis (top) and western blotting with streptavidin-HRP (bottom) of α-C45A/L158C MoFe proteins after treatment with the DTB inhibitor. Lanes: 1 = molecular weight marker; 2 = Strep-tagged WT MoFe protein (N_{term} Strep-II tag on α-subunit); 3 = α-C45A/L158C MoFe protein treated with DTB inhibitor (DT-free); 4 = "3" although the reaction was performed in the presence of 1 mM DT, 5 = "3" although the reaction was performed in the presence of 10 mM DT, 6 = "3" although the reaction was performed in the presence of 1 mM TCEP. 1 µg protein per well. In-gel stain-free imaging (BioRad Laboratories) was employed. (C) Residual specific activities with increasing equivalents of DTB inhibitor per one α-C45A/L158C MoFe (left y axis labeled as black dots: for H_2 evolution under 1 atm Ar; right y axis labeled as blue triangles: NH_3 evolution under 1 atm N_2). Empty symbols for α-C45A/L158C treated with 10 mM DT before DTB inhibitor addition. All activity assays were performed for 8 minutes at 30 ºC with 0.1 mg mL^{-1} MoFe protein and 16.6 molar equivalents of Fe protein. (D) Percentage electron distribution between H_2 formation (gray, $2e^-$ per H_2) *vs*. N_2 fixation (turquoise, $3e^-$ per NH_3) under 1 atm N_2 with increasing equivalents of DTB inhibitor. (C and D) n = 3 and error bars represent SD (propagated where necessary).



	e⁻ % to N_2
$16MgATP + 8H^+ + 8e^- + N_2 \longrightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$	75 %
$20MgATP + 10H^+ + 10e^- + N_2 \longrightarrow 2NH_3 + 2H_2 + 20MgADP + 20P_i$	60 %
$24MgATP + 12H^+ + 12e^- + N_2 \longrightarrow 2NH_3 + 3H_2 + 24MgADP + 24P_i$	50 %

Figure 4: Simplified Lowe-Thorneley scheme for nitrogenase. N_2 associates to the FeMoco at the E4 state along with the reductive elimination of H_2 by a productive pathway. Non-productive H_2 release by the FeMoco is colored in gray. The release of $2NH_3$ is placed after the E8 state in this representation. Percentage electron distribution is calculated assuming that $2e^-$ are required for $2H^+$ reduction to H_2, and that $6e^-$ are required for N_2 reduction to $2NH_3$.

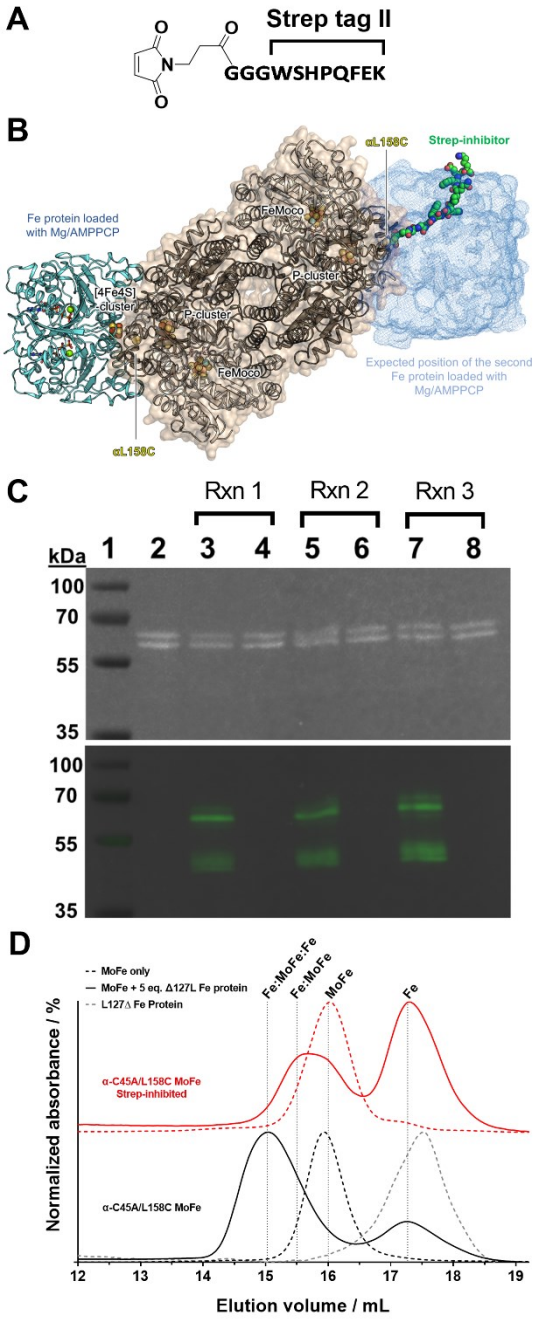


Figure 5: (A) Structure of the maleimide-containing Strep inhibitor. (B) Illustrative model of the Strep inhibitor-modified α-C45A/L158C MoFe protein. (C) SDS-PAGE (upper) and western blot (lower) analysis of α-C45A/L158C modification reactions with the Strep inhibitor. Lanes: 1 = molecular weight marker; 2 = untreated α-C45A/L158C MoFe protein; 3/5/7 = Strep-inhibited α-C45A/L158 MoFe protein (eluted from the StrepTactin solid phase); 4/6/8 = unmodified α-C45A/L158C MoFe protein that did not tightly bind to the column solid phase (flow-through). 3/4, 5/6 and 7/8 represent three independent functionalization reactions. Western blotting was performed using a StrepTactin-HRP conjugate (artificially colored image). 0.5 µg protein per well. In-gel stain-free imaging (BioRad Laboratories) was employed. (D) Gel-filtration of untreated (black dashed line) and Strep-inhibited (red dashed line) α-C45A/L158C MoFe protein. These MoFe proteins were also analyzed after pre-treatment with 5 molar equivalents of L127Δ Fe protein (solid lines) in order to form α-C45A/L158C MoFe:Fe protein complexes.

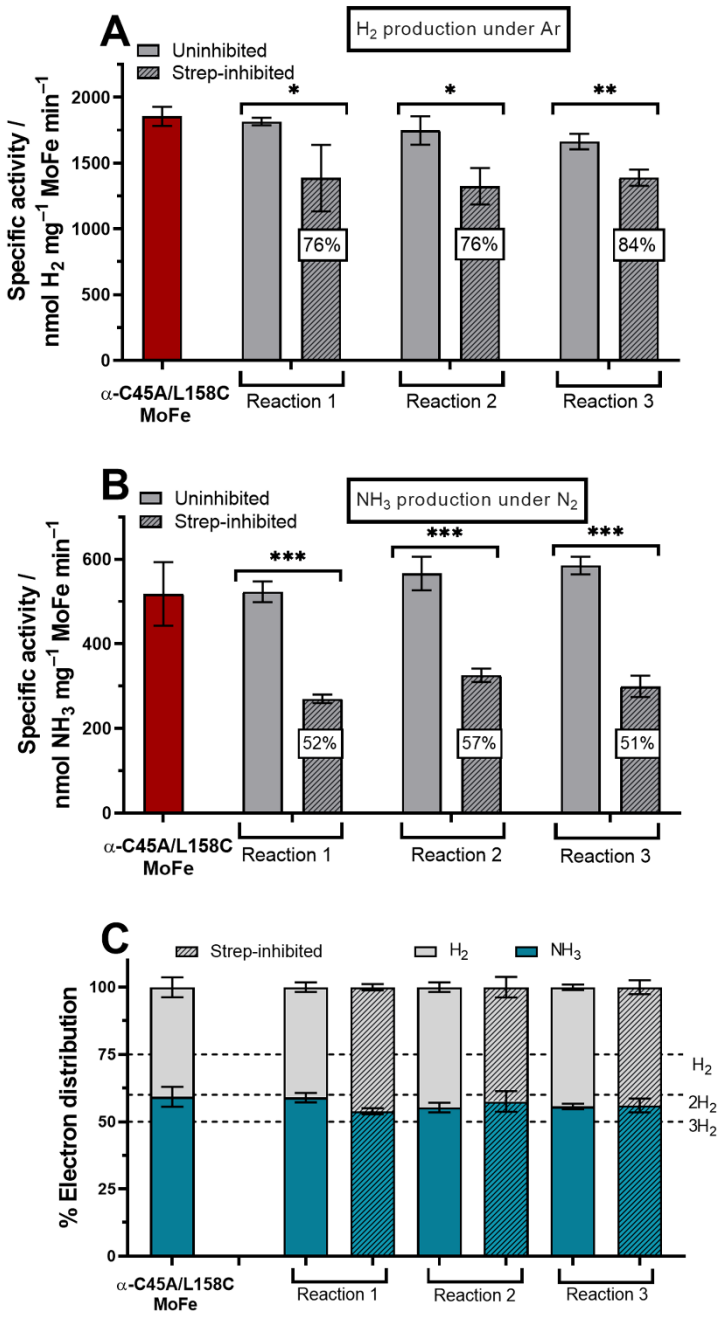


Figure 6: (A) Specific activity for H_2 evolution under 100% Ar headspace. Red is the standard α-C45A/L158C MoFe purified in absence of DT. Light grey bars represent the activities obtained for α-C45A/L158C MoFe protein from the strep-inhibitor functionalization reactions that did not bind to the StrepTrap XT solid phase. The hatched bars represents strep-inhibited α-C45A/L158C MoFe (eluted from the StrepTrap XT solid phase). Percentage value in white square corresponds to specific activity of strep-inhibited MoFe divided by specific activity to related unreacted MoFe or FT fraction. (B) Specific activity for NH_3 evolution under 100% N_2 atmosphere. (C) Electron distribution in percentage. Blue is for electrons consumed to make NH_3 ($3e^-$ per NH_3), grey is for electrons consumed to make H_2 ($2e^-$ per H_2) under N_2 atmosphere. All activity assays were performed for 8 minutes at 30 ºC with 0.1 mg mL^{-1} MoFe protein and 16.6 molar equivalents of Fe protein. (A-C) $n = 3$ and error bars represent SD (propagated where necessary). Significance: ns (not significant) for $P > 0.05$, * for $P \leq 0.05$, ** for $P \leq 0.01$, and *** for $P \leq 0.001$.

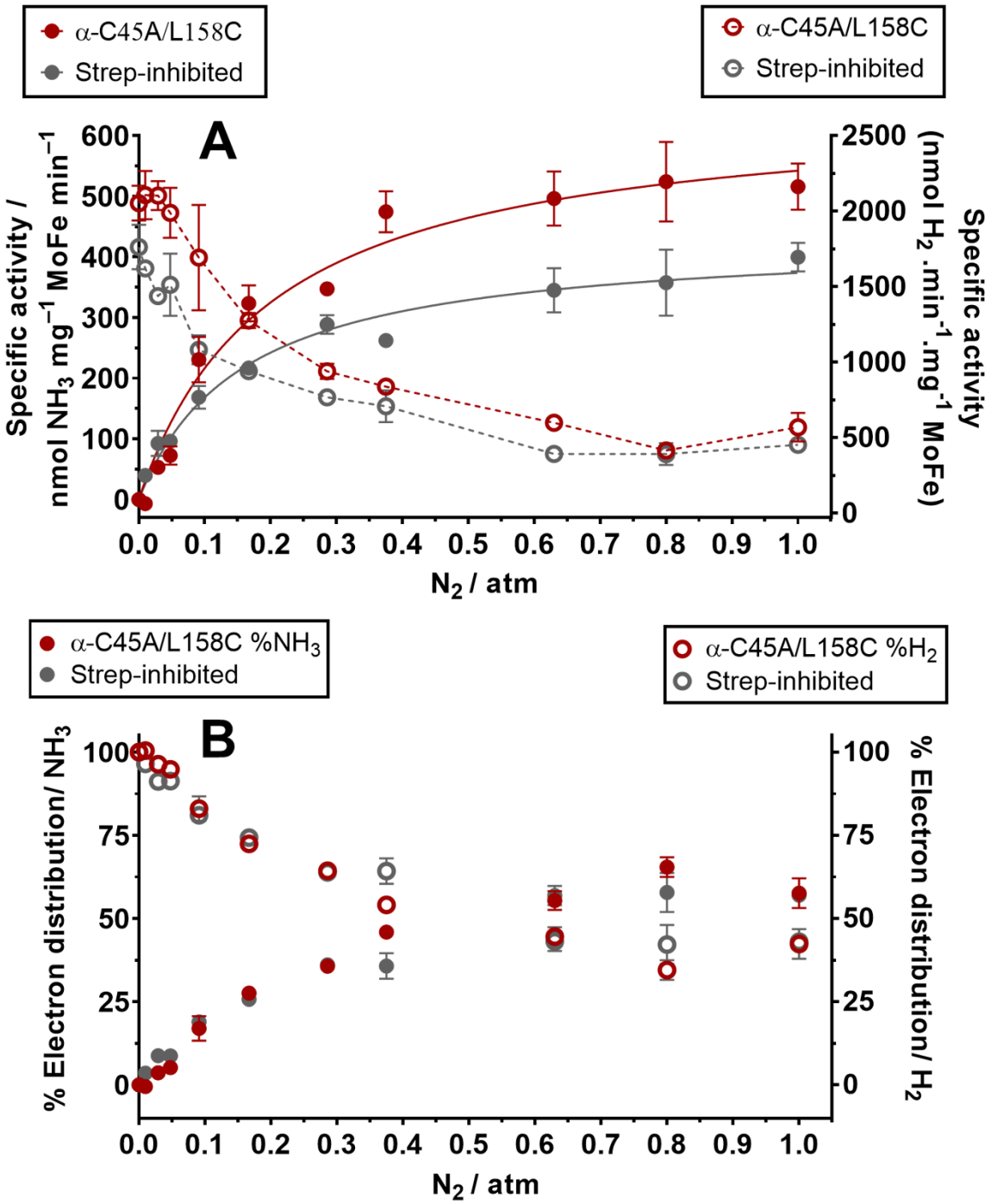


Figure 7: (A) Specific activity for H_2 and NH_3 versus N_2 partial pressure. The left axis represents specific activity in nmol NH_3 min^{-1} mg^{-1} MoFe (solid lines, solid points). The right axis represents specific activity in nmol H_2 min^{-1} mg^{-1} MoFe (dashed lines, hollow points). Red data: uninhibited α-C45A/L158C MoFe; grey data: strep-inhibited α-C45A/L158C MoFe. (B) Percentage of electron distribution towards N_2 fixation (solid points) or H_2 production (hollow points). Red data: uninhibited α-C45A/L158C MoFe; grey data: strep-inhibited α-C45A/L158C MoFe. Percentage electron distributions are determined by the assumption that H_2 formation requires $2e^-$ and N_2 fixation requires $6e^-$. All activity assays were performed for 8 minutes at 30 ºC with 0.1 mg mL^{-1} MoFe protein and 16.6 molar equivalents of Fe protein. (A-B) n = 3 (technical repeats) and error bars represent SD (propagated where necessary); n = 2 (technical repeats) for 0.8 atm N_2 data points.

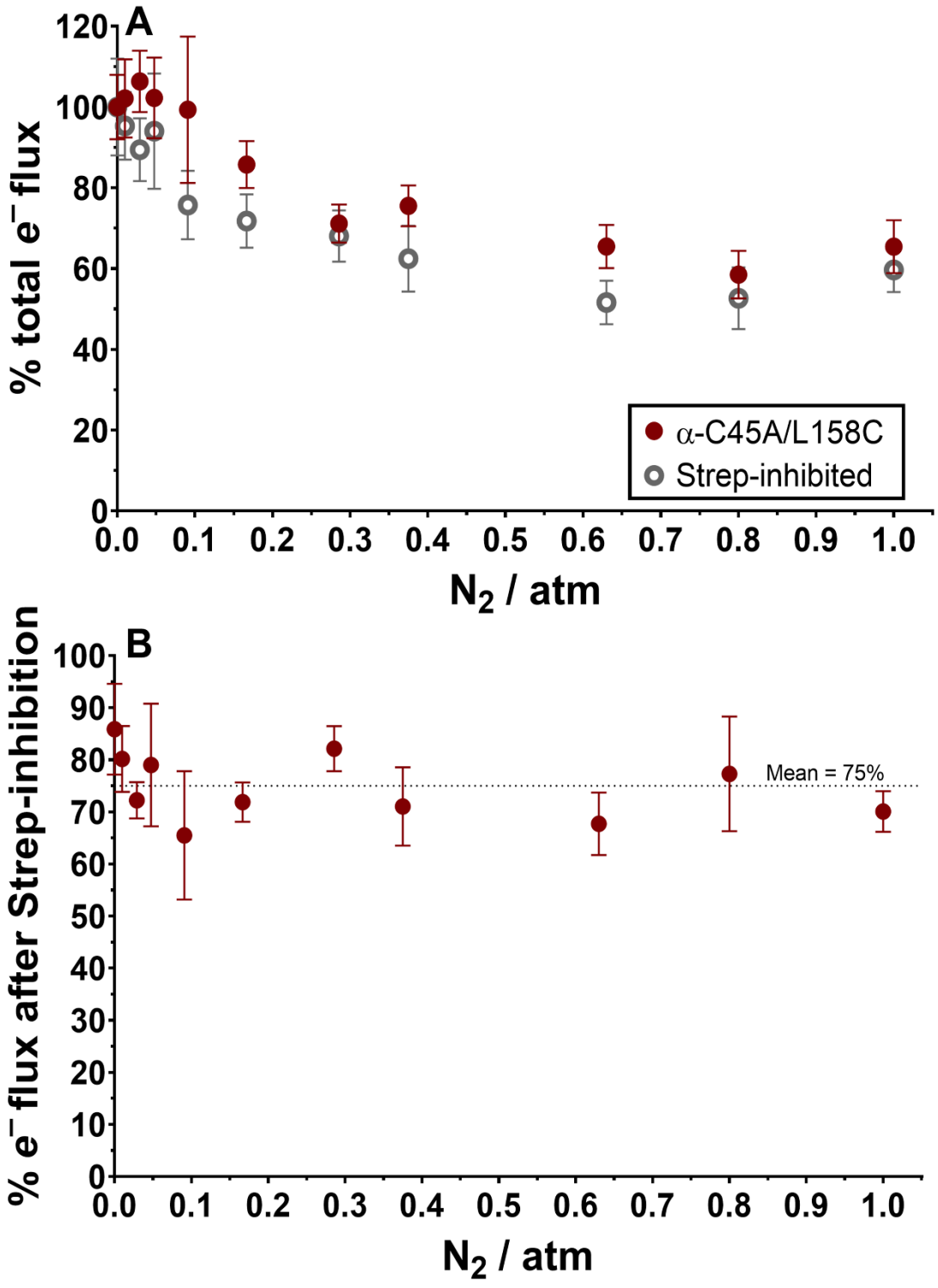


Figure 8: (A) Percentage total electron flux of uninhibited and strep-inhibited α-C45A/L158C MoFe proteins upon the introduction of increasing N2 partial pressures, with 1 atm Ar resulting in the largest total electron flux for both proteins. (B) Remaining total electron flux after strep-inhibition of the α-C45A/L158C MoFe protein between 0-1 atm of N2. Percentage electron distributions are determined by the assumption that H2 formation requires 2e– and N2 fixation requires 6e−. All activity assays were performed for 8 minutes at 30 ºC with 0.1 mg mL−1 MoFe protein and 16.6 molar equivalents of Fe protein. (A-B) n = 3 and error bars represent SD (propagated where necessary).

Table 1: Michaelis-Menten kinetic parameters for α-C45A/L158C and strep-inhibited α-C45A/L158C MoFe proteins.

Michaelis-Menten parameter (N_2 fixation)	α-C45A/L158C	α-C45A/L158C strep-inhibited
K_M^{app} (atm)	0.20 ± 0.05	0.16 ± 0.03
V_{max}^{app} (nmol NH_3 min^{-1} mg^{-1})	651 ± 55	432 ± 29
k_{cat} (s^{-1})	1.3 ± 0.1	>1.7 ± 0.1**
k_{cat}/K_M (s^{-1} atm^{-1})	6.5 ± 1.7	11.0 ± 2.5
k_{cat}/K_M (x10^4 s^{-1} M^{-1})	1.0 ± 0.3	1.7 ± 0.4

kcat is defined here as turnover frequency per active αß half (nmol NH_3 $FeMoco^{-1}$ s^{-1}).

**The possible presence of doubly inhibited MoFe protein in this sample could result in a larger kcat value per FeMoco (αß half). Partial pressures of N_2 were from atm to M^{-1} using a Henry's law conversion factor of 6.4 x10^{-4}.[34]

6.14. Supporting information

Azotobacter vinelandii cultures

All buffers and solutions were prepared with MilliQ water (18.2 MΩ cm). Generally, large volumes (> 50 mL) were sterilized by autoclaving; smaller solutions were sterilized by filtration (0.45 µm syringe filters).

Agar culture medium

A modified Burke's medium was used for the cultivation of all *Azotobacter vinelandii* strains. Under sterile conditions, 6 mL of an autoclaved 100x phosphate buffer (0.46 M K_2HPO_4, 0.15 M KH_2PO_4) was added to an autoclaved molten agar solution (~18 g agar per L of water). 27.5 mL of this resulting agar buffer was added to 3 mL of an autoclaved 10x salt solution (60 mM sucrose, 8.1 mM $MgSO_4$, 6.1 mM $CaCl_2$), 250 µL of a filter-sterilized 3 M NH_4OAc solution, 30 µL of a filter-sterilized 0.1 M Fe solution ($FeCl_3.6H_2O$) and 30 µL of a filter-sterilized 10 mM Mo solution ($Na_2MoO_4.2H_2O$). The mixture was poured in a Petri dish, solidified, and the relevant *A. vinelandii* strains were plated and incubated at 30 °C for 48 h.

Liquid culture medium

Under sterile conditions, 60 mL of an autoclaved 100x phosphate buffer solution (0.46 M K_2HPO_4, 0.15 M KH_2PO_4) was added to 6 L of autoclaved salt medium (60 mM sucrose, 8.1 mM $MgSO_4 \cdot 7H_2O$, 6.1 mM $CaCl_2$) after both solutions were at room temperature. Then 20 mL of a filter-sterilized 3 M NH_4OAc solution, 6 mL of a filter-sterilized 0.1 M Fe solution ($FeCl_3.6H_2O$) and 6 mL of a filter-sterilized 10 mM Mo solution ($Na_2MoO_4.2H_2O$) was added (yielding a modified Burke's medium).

Large scale cultures

For each *A. vinelandii* strain, a preculture was grown in a 250 mL baffled culture flask (glass) with a vented cap containing 60 mL of sterile liquid media. Each preculture was started by inoculation from a fresh agar plate and grown aerobically at 200 rpm at 30 °C until $OD_{600\ nm}$ > 2.0 (usually overnight). Under sterile conditions, 60 mL of the preculture was used to inoculate the 6 L of liquid media in 2x 5 L baffled flasks (Corning CLS431684, 3 L of culture per bottle) with vented caps. Cells were grown at 200 rpm at 30 °C overnight until $OD_{600\ nm}$ > 1.5 (usually overnight). Cells were collected by serial centrifugation runs at 4,500 x *g*, 20 min at room temperature. The supernatant was discarded, and the cells were gently resuspended in

fresh growth medium devoid of NH_4OAc or another source of fixed nitrogen (prepared fresh, but not under sterile conditions) in the incubator shaker (~400 mL media per bottle). After resuspension, the culture was divided evenly between the remaining media/bottles and incubated at 200 rpm and 30 °C for another 3 hours to facilitate derepression of the *nif* operon, and harvested by serial centrifugation collections as above. Finally, the cells were scooped into a plastic bag and stored at -80 °C until further use.

Azotobacter vinelandii strains

"Wild-type" MoFe protein contains an N-terminal poly(histidine)$_8$ tag between codons 2 and 3 of *nifD*. The strain containing this construct was prepared previously by markerless mutagenesis (yielding *A. vinelandii* RS1, derived from strain DJ).[1,2] α-C45A/L158C MoFe protein was prepared from this strain (detailed below), and therefore also carries a poly(histidine)$_8$ tag between codons 2 and 3 of *nifD*.[2] L127Δ Fe protein was purified from *A. vinelandii* strain DJ1065, which was generously shared by Dennis Dean (Virginia Tech). A strep tag-containing MoFe protein (N-terminal of NifD) was purified from *A. vinelandii* strain DJ2102, which was generously shared by Dennis Dean (Virginia Tech); this protein was used as a control only in western blotting experiments.

Preparation of A. vinelandii strain producing α-C45A/L158C MoFe protein

All *A. vinelandii* strains were maintained in a modified Burke's medium. When required, kanamycin was added at a final concentration of 3 μg/mL (*A. vinelandii*). *Escherichia coli* DH5α was cultivated in Luria-Bertani medium at 37 °C. When required, kanamycin was added to 25 μg/mL (*E. coli*). Plasmid constructions were performed using *E. coli* DH5α (Invitrogen, USA). A *sacB*-based system was used to create the α-C45A/L158C double-mutation in *A. vinelandii via* double homologous recombination.[3,4] During this process, *A. vinelandii* cells were maintained in Burke's medium with glucose as the substrate. Briefly, pK18*mobsacB* based vectors were transformed into *A. vinelandii* as described before.[1,2] We thank Professor Jeremy D. Semrau at University of Michigan (US) for providing plasmid pK18*mobsacB*. Successful transformants were kanamycin-resistant, suggesting successful single homologous recombination and incorporation of the plasmid, and transferred at least one more time on kanamycin-plates before counter selection. The transformants were then grown in liquid medium without antibiotics for a few hours to overnight for outgrowth. Various amounts of cells were then spread on agar plates with glucose as the substrate and with addition of 3% sucrose for counter selection of the second homologous recombination. Sucrose-resistant colonies were transferred at least one more time on 3% sucrose selection

plates before further phenotype screening. Desired clones were sucrose-resistant and kanamycin-sensitive suggesting loss of pK18*mobsacB* backbone and were first unable of diazotrophic growth (for *A. vinelandii* Δ*nifD*) and had restored diazotrophic growth ability for the *A. vinelandii nifD*:C45A/L158C mutant. Mutations were also confirmed by PCR and Sanger sequencing.

To create the *A. vinelandii nifD*:C45A/L158C mutant, an *A. vinelandii* Δ*nifD* mutant was first constructed with the in-frame deletion from 45C to 158L (Figure S1) in strain *A. vinelandii* RS1. Briefly, two flanking regions were amplified using primer sets del*nifD*_F1/R1 and del*nifD*_F2/R2 (Table S1) and inserted into the suicidal vector pK18*mobsacB* using NEBuilder® HiFi DNA Assembly (New England Biolabs, US) to create vector pK18del*nifD* (Figure S1). The pK18*mobsacB* plasmid was linearized *via* digestion by HindIII and BamHI enzymes (New England Biolabs, US). This vector was used to create a Δ*nifD* mutant in *A. vinelandii* RS1 *via* double homologous recombination.

Subsequently, a second mutation vector was constructed by re-inserting the deleted *nifD* region carrying the desired point mutations (Figure S1). Briefly, vector pK18del*nifD* was linearized by HindIII restriction digestion previously included in del*nifD*_R1 and del*nifD*_F2 primers (Table S1). Partial *nifD* was amplified using primers pm*nifD*_InsF/R with point mutations introduced on both primers (Table S1). This fragment was inserted into pK18del*nifD* *via* Gibson assembly to create vector pK18del*nifD*:C45A/L158C (Figure S1). This vector was transformed into *A. vinelandii* Δ*nifD* for double homologous recombination, restoration of diazotrophic growth, and introduction of the desired point mutations. *Note:* The wild-type (non-His-tag-containing) numbering is used here. Including the poly(histidine)$_8$ tag, α-C45A/L158C becomes α-C53A/L166C.

Table S1: Primers, vectors, and strains used in this study.

Primers	
(underlined regions are overhangs for Gibson assembly, ***bolded capital*** *letters are point mutations, sequences in parentheses are HindIII restriction sites inserted for linearizing plasmid)*	
del*nifD*_F1	cgttgtaaaacgacggccagtgccatcgtattctacgacgtgctgggcg
del*nifD*_R1	cacggattcgatgtcgtcgccgat(aagctt)cttcttggactgggtaaccgccg
del*nifD*_F2	cccggcggttacccagtccaagaag(aagctt)atcggcgacgacatcgaatccg
del*nifD*_R2	cgaattcgagctcggtacccgggcggagccgatcaggtcgggctt
pm*nifD*_InsF	cggttacccagtccaagaag**GC**catcatctccaacaagaagtcccagcc
pm*nifD*_InsR	ggattcgatgtcgtcgccgat**GCA**gccgatcgggcactcggactg
Av-genome-sequencing_1	aacaacatctccaagggcatcg
Av-genome-sequencing_2	gtatcaaggagaagttcatcttcca
Av-genome-sequencing_3	gcagtttcttccagcacggatt
Av-genome-sequencing_4	tttccgactccatgaccgaa
Av-genome-sequencing_5	ggtatcggctcgatgaaataggtc
Vectors	
pK18*mobsacB*	Km^r RP4-mob, mobilizable cloning vector containing *sacB* from *Bacillus subtilis*.[5]
pK18del*nifD*	pK18*mobsacB* carrying flanking regions to knock out *nifD*.
pK18del*nifD*: C45A/L158C	pK18*mobsacB* carrying flanking regions to create point mutations in *nifD*: C45A/L158C.
Strains	
E. coli DH5α	F^- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1 (Invitrogen, USA)
A. vinelandii RS1	8xHIS-tag between codons 2/3 of *nifD*.[1,2] Comparable to strain DJ1141.[6] Produced from *A. vinelandii* DJ (BTCC).
A. vinelandii Δ*nifD*	In-frame deletion of *nifD* from 45C to 158L from *A. vinelandii* RS1.
A. vinelandii nifD: C45A/L158C	Point mutations of *nifD*:C45A/L158C. Analogous to MoFe protein in Ref[2]
A. vinelandii DJ1065	Strain of *A. vinelandii* producing NifH with residue 127L deleted (Dennis Dean).[7,8]
A. vinelandii DJ2102	Strain of *A. vinelandii* producing MoFe protein with an Nterm strep-tag on NifD (Dennis Dean).[9]

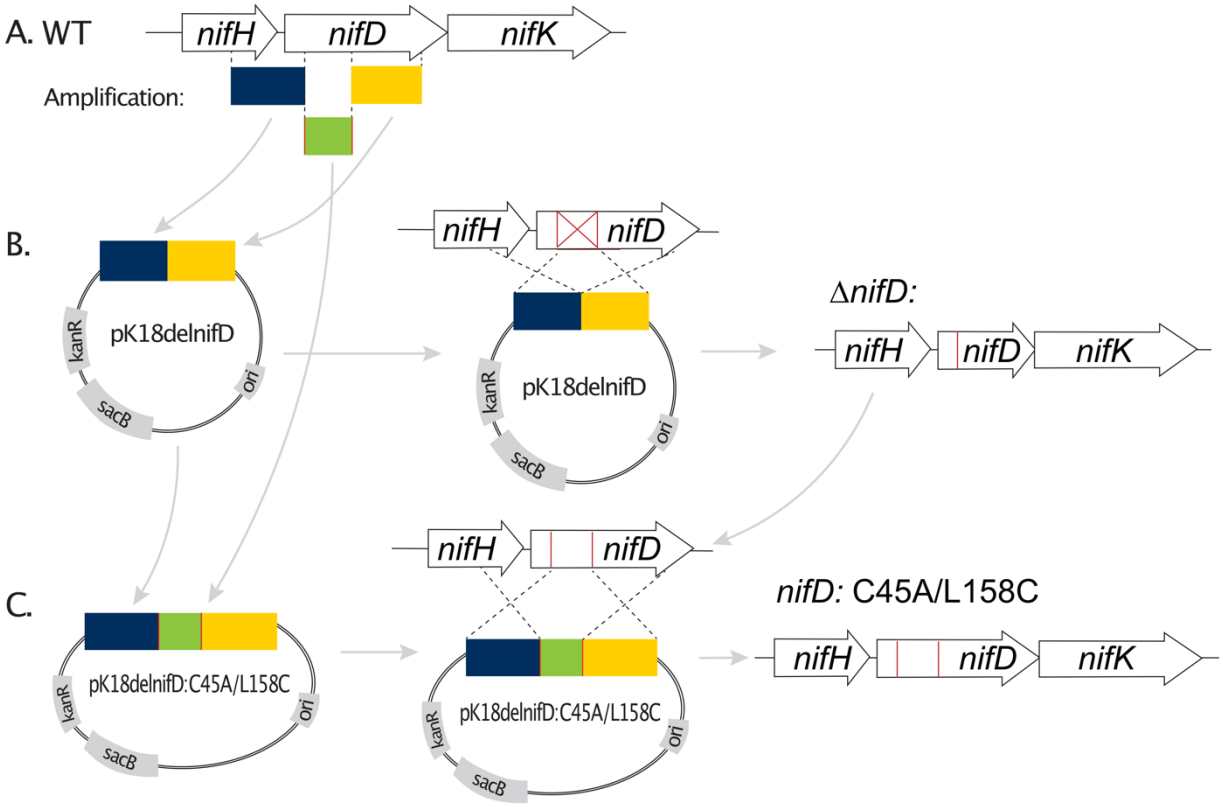


Figure S1: Graphical scheme of deletion of *nifD* in *A. vinelandii* (steps A and B), and the point mutations of C45A/L158C in *nifD* (steps A and C). Red lines across *nifD* represents either deletion or point mutation.

***Sequence of nifD encoding α-C45A/L158C* NifD**

nifD in red

nifK in orange

gcc = α-C45A single-point mutation

tgc = α-L158C single-point mutation

caccaccatcatcaccaccatcac = poly(histidine)$_8$ tag

Forward sequencing primers are highlighted in yellow

Reverse sequencing primers are highlighted in cyan

Regions unconfirmed by sequencing are highlighted in grey.

aacaacatctccaagggcatcgtgaagtatgccaactccggcagcgtgcgtctgggcggcctgatctgcaacagccgtaacaccgac
cgcgaagacgagctgatcatcgctctggccaacaagctgggcacccagatgatccacttcgtgccgcgtgacaacgtcgtgcagcgc
gccgaaatccgccgcatgaccgtgatcgaatacgatccgaaagccaagcaagccgacgaataccgcgctctggcccgcaaggtcgt
cgacaacaaactgctggtcatcccgaacccgatcaccatggacgagctcgaagagctgctgatggaattcggcatcatggaagtcga
agacgaatccatcgtcggcaaaaccgccgaagaagtctgatagccgctccggtttcagaaggacgggacagggcagattggctctgt
cggggtggcgccccccgcatggggcgggcgccccacccgttacccgcatatgaacgctaaggcaagaggagtcatacccatgacc
caccaccatcatcaccaccatcacggtatgtcgcgcgaagaggttgaatccctcatccaggaagttctggaagtttatcccgagaaggc
tcgcaaggatcgtaacaagcacctggccgtcaacgacccggcggttacccagtccaagaag**gcc**atcatctccaacaagaagtccca
gcccggtctgatgaccatccgcggctgcgcctacgccggttccaaaggcgtggtctggggccccatcaaggacatgatccacatctc
ccacggtccggtaggctgcggccagtattcgcgcgccggccgtcgtaactactacatcggtaccaccggtgtgaacgccttcgtcacc
atgaacttcacctcggacttccaggagaaggacatcgtgttcggtggcgacaagaagctcgccaaactgatcgacgaagtggaaacc
ctgttcccgctgaacaagggtatctccgtccagtccgagtgcccgatcggc**tgc**atcggcgacgacatcgaatccgtgtccaaggtca
agggcgccgagctcagcaagaccatcgtaccggtccgttgcgaaggcttccgcggcgtttcccagtccctgggccaccacatcgcca
acgacgcagtccgcgactgggtcctgggcaagcgtgacgaagacaccaccttcgccagcactccttacgatgtggccatcatcggcg
actacaacatcggcggcgacgcctggtcttcccgcatcctgctggaagaaatgggcctgcgttgcgtagcccagtggtccggcgacg
gctccatctccgaaatcgagctgaccccgaaggtcaagctgaacctggttcactgctaccgctcgatgaactacatctcccgtcacatg
gaagagaagtacggtatcccatggatggagtacaacttcttcggcccgaccaagaccatcgagtcgctgcgtgccatcgccgccaagt
tcgacgagagcatccagaagaagtgcgaagaggtcatcgccaagtacaagcccgagtgggaagcggtggtcgccaagtaccgtcc
gcgcctggaaggcaagcgcgtcatgctctacatcggtggcctgcgtccgcgccacgtgatcggcgcctacgaagacctgggcatgg
aagtggtgggtaccggctacgagttcgcccacaacgacgactatgaccgcaccatgaaagaaatgggtgactccaccctgctgtacg
atgacgtgaccggctacgaattcgaagaattcgtcaagcgcatcaagcccgacctgatcggctccggtatcaaggagaagttcatcttc
cagaagatgggcatccccttccgtcaaatgcactcctgggattattccggcccctaccacggcttcgatggcttcgccatcttcgcccgt
gacatggacatgaccctgaacaatccgtgctggaagaaactgcaggctccctgggaagcttccgaaggcgccgagaaagtcgccgc
cagcgcctgatagcagagcaatcgtacgcaacgtccgctgcgggcggtttccgcccgcggccgacatccgctaacgccgttcacag
atgagtgaggcgtaggagagagtcatgagccagcaagtcgataaaatcaaagccagctacccgctgttcctcgatcaggactacaag
gacatgcttgccaagaagcgcgacggcttcgaggaaaagtatccgcaggacaagatcgacgaagtattccagtggaccaccaccaa
ggaataccaggagctgaacttccagcgcgaagccctgaccgtcaacccggccaaggcttgccagccgctgggcgccgttctctgcg
ccctcggtttcgagaagaccatgccctacgtgcacggttcccagggttgcgtcgcctacttccgctcctacttcaaccgtcatttccgcga
gccggtttcctgcgtttccgactccatgaccgaagacgcggcagtgttcggcggccagcagaacatgaaggacggtctgcagaactg
taaggctacctacaagcccgacatgatcgcagtgtccaccacctgcatggccgaggtcatcggtgacgacctcaacgccttcatcaac
aactcgaagaaggaaggtttcattcctgacgagttcccggtgccgttcgcccatacccgagcttcgtgggcagccacgtgaccggct
gggacaacatgttcgaaggcattgctcgctacttcaccctgaagtccatggacgacaaggtggttggcagcaacaagaagatcaacat
cgtccccggcttcgagacctacctgggcaacttccgcgtgatcaagcgcatgctttcggaaatgggcgtgggctacagcctgctctcc
gatccggaagaagtgctggacaccccggctgacggccagttccgcatgtacgcgggcggcaccactcaggaagagatgaaggacg

ctccgaacgccctcaacaccgtcctgctgcagccgtggcacctggagaagaccaagaagttcgtcgagggtacctggaagcacgaa
gtaccgaagctgaacatcccgatgggcctggactggaccgacgagttcctgatgaaagtcagcgaaatcagcggccagccgattccg
gcgagcctgaccaaggagcgtggccgtctggtcgacatgatgaccgactcccacacctggctgcacggcaagcgtttcgccctgtgg
ggtgatccggacttcgtgatgggcctggtcaagttcctgctggaactgggttgcgagccggtacacattctctgccacaacggcaacaa
gcgttggaagaaggcggtcgacgccatcctcgccgcttcgccctacggcaagaatgctaccgtctacatcggcaaggacctgtggca
cctgcgttcgctggtcttcaccgacaagccggacttcatgatcggcaacagctacggtaagttcatccagcgcgacaccctgcacaag
ggcaaggagttcgaggttccgctgatccgtatcggcttcccgatcttcgaccgtcatcacctgcatcgctccaccaccctgggttacgag
ggcgccatgcagatcctgaccaccctggtgaactcgatcctggaacgtctggacgaggaaacccgcggtatgcaggccaccgacta
caaccacgacctggtacgctaagtcgtcggttcaagtggtatcggccggagcggcgcaagccgctccggctcttccttggcggcggc
cgcaggtggtcgggccttttgcccgcgatctgcggcaaccgccaaacccgtctaaggagcaagcccatgcccagcgtcatgattcgc
cgcaacgacgaaggccaactgaccttctatatcgccaagaaagaccaggaagagatcgtggtgtccctggagcatgacagccccga
actctggggtggcgaagtcaccctcggcgacggttcgacctatttcatcgagccgatacc

Purification of MoFe and Fe proteins:

All buffers were prepared with MilliQ water and deoxygenated inside of a COY glovebox (>95% N_2/<5% H_2, Coy Laboratory Products, USA) before use (at least overnight). All subsequent procedures were conducted within the glovebox, with the exception of centrifugation steps (performed in deoxygenated and anoxically sealed poly(propylene) bottles).

Resuspension and lysis

Cell pellets were anaerobically thawed and resuspended for 20 minutes in lysis buffer (50 mM Tris/HCl pH 8, containing 5 mM dithionite (abbreviated as DT), and 37% v/v glycerol) with a ratio 2:1 (buffer:cells v/w), and subsequently collected by centrifugation at 12,000 x *g* (20 min at 4 °C). Once again inside of the glovebox, the glycerol-containing supernatant was discarded, and cells were rapidly lysed by osmotic shock/shaking following the introduction of an equal replacement volume of glycerol-free lysis buffer (in presence of few µg of DNase). After incubation on ice for approximately 15 minutes, the cell lysate was clarified by centrifugation at 30,000 x *g* (4 °C, 1 hour) to yield a dark brown supernatant.

Purification of MoFe (WT and α-C45A/L158C) proteins

Note: From here onwards, DT was omitted from all buffers.

Within the glovebox, the supernatant was decanted and diluted with a post-lysis buffer (2 M NaCl, 234 mM Tris/HCl, pH 8.0) in order to obtain a final NaCl concentration of 0.3 M NaCl. His-tagged "wild-type" and α-C45A/L158C MoFe protein were purified by Ni-affinity chromatography using a Ni-loaded HisTrap HP 5 mL column (Cytiva) with a flow rate of 5 mL/min using an Äkta Start system within the glovebox. The MoFe-containing supernatant was loaded onto the column, washed with an equilibration buffer for at least 3 column volumes to remove remaining DT (0.3 M NaCl, 50 mM Tris/HCl, pH 8.0), washed with a 20 mM imidazole step, and finally eluted with a 0.3 M imidazole step (0.3 M imidazole, 0.3 M NaCl, 50 mM Tris/HCl, pH 8.0). The eluted MoFe protein was then purified further with a HiPrep Q-Sepharose FF 16/10 column (Cytiva, 20 mL column volume, flow rate = 5 mL/min), which also served to remove excess imidazole and residual DT. The concentration of NaCl in the MoFe protein samples was first lowered to 0.1 M NaCl using a dilution buffer (50 mM Tris/HCl, pH 8.0) before being loaded on the pre-equilibrated Q-sepharose column (0.1 M NaCl). MoFe proteins were eluted at approximately 0.3 M NaCl using a NaCl gradient over 7 column volumes (from 0.2 to 0.65 M NaCl). MoFe proteins were typically

concentrated to >20 mg/mL using a Merck Millipore stirred concentrator cell (within the glovebox) equipped with a 100 kDa molecular weight cut-off membrane (fed with ultra-high-purity N_2 5.0). MoFe proteins were subsequently flash-frozen as 20 µL pellets in liquid nitrogen and stored in liquid nitrogen until further use. Protein concentrations were determined by the Biuret method using bovine serum albumin as the standard.

Purification of Fe (WT and L127Δ) proteins

In contrast to the MoFe proteins, Fe protein samples were purified in the presence of 2 mM DT in all buffers (5 mM DT was included in the cell lysis step, as above). Fe proteins were routinely purified in tandem to the MoFe proteins using an Äkta Go system. Following cell lysis (L127Δ Fe protein) the supernatant was diluted with a post lysis buffer (2 M NaCl, 234 mM Tris/HCl, 2 mM DT, pH 8.0) in order to obtain NaCl concentration of 0.3 M NaCl. WT Fe proteins were purified from His-column flow-through fractions (above), where the flow-through fraction was first diluted to 0.1 M NaCl using the dilution buffer above (containing 2 mM DT). The Fe protein was first purified using a HiPrep Q-Sepharose HP 16/10 column with a flow rate of 5 mL/min. Fe protein elution was performed using a NaCl gradient (7 column volumes, from 0.2 to 0.65 M NaCl) and concentrated to <10 mL using a Merck-Millipore stirred concentrator cell (30 kDa molecular weight cut-off membrane). The Fe protein was then purified further by size-exclusion chromatography (HiPrep 26/60 Sephacryl S-200 HR, Cytiva) with a flow rate of 2.5 mL/min, using a Tris running buffer (50 mM Tris/HCl, pH 8.0, 0.5 M NaCl). Eluted Fe protein samples were concentrated to >20 mg/mL, flash frozen in liquid nitrogen as 20 µL pellets and stored in liquid nitrogen until further use. Protein concentrations were determined by the Biuret method using bovine serum albumin as the standard.

Functionalization with maleimide-desthiobiotin ("DTB") inhibitor

DT-free MoFe protein was reacted in the presence of increasing molar equivalents of DTB-inhibitor (synthesis below, fresh stocks prepared at 1 – 3 mg/mL within the glovebox using 0.1 M MOPS/NaOH buffer, pH 7) for 4 hours at room temperature within the glovebox. Time reactions were quenched by the addition of 2 mM DT (final concentration) to the reaction to rapidly reduce the maleimide functional group. Activity assays were performed immediately (see below).

A control reaction was performed in which 26 µM DTB-maleimide inhibitor (confirmed by mass spectrometry) was incubated in 2 mM DT-containing MOPS/NaOH buffer (0.1 M, pH

7) for 30 minutes at room temperature. Upon analysis of the reaction mixture was analyzed by LC-MS after this time the starting material was no longer visible, consistent with the reduction of the maleimide by DT (consistent with the results presented in Figure S8).

As discussed in the main article, we attempted to purify the DTB-conjugated MoFe proteins using commercial Streptactin-based solid phases (specifically, "StrepTrap" by Cytiva). After failing to observe the association of DTB-modified MoFe proteins to this solid phase, we deemed it prohibitively costly to employ an alternative streptavidin-containing stationary phase; streptavidin cannot be efficiently regenerated following incubation with biotin (to elute the DTB-functionalized protein).

Functionalization with Strep-tag inhibitor and conjugate purification

A synthetic peptide with the sequence GGGWSHPQFEK and containing an N-terminal maleimide functional group was synthesized by Genscript (USA). DT-free MoFe protein was reacted with 0.5 molar equivalents (per MoFe protein) of Strep-maleimide inhibitor to minimize the formation of di-functionalized MoFe proteins for 4 hours at room temperature within the glovebox (additional information with respect to these molar ratios is presented below, Figure S10). Stock solutions of the Strep-maleimide peptide were prepared fresh within the glovebox using 0.1 M MOPS/NaOH buffer (pH 7), with concentrations between 1 – 3 mg/mL. Functionalization reactions were quenched by the addition of DT (2 mM final concentration). The strep-inhibited (strep-functionalized) α-C45A/L158C MoFe protein was subsequently purified from unreacted protein using a StrepTrap-XT 1 mL column (Cytiva) equilibrated with MOPS buffer (0.1 M MOPS/NaOH, 0.2 M NaCl, 2 mM DT, pH 7.0). Unreacted MoFe was collected as the flow-through (FT) fraction. Strep-functionalized protein was subsequently eluted in a single step with biotin-containing buffer (50 mM biotin, 0.1 M MOPS/NaOH, 0.2 M NaCl, 2 mM DT, pH 7.0). Samples were concentrated separately with a Merck-Millipore stirred concentrator cell equipped with a 100 kDa molecular-weight cut-off membrane. Protein samples were flash frozen in liquid nitrogen as 20 μL pellets and stored in liquid nitrogen until further use.

Activity assays

Reduction activity assays were conducted in triplicate 1 mL reactions in 13 mL septum-sealed glass vials (Wheaton) containing deoxygenated buffer and an ATP-regenerating system (5 mM ATP, 30 mM phosphocreatine, 1.3 mg BSA, 0.2 mg creatine phosphokinase (from Rabbit muscle), 10 mM sodium dithionite and 100 mM MOPS/NaOH at pH 7.0). All

reactions contained 0.1 mg MoFe protein and 16.6 molar equivalents of Fe protein (0.48 mg). All reaction vials were sealed within an Ar-filled glovebox (Jacomex, France) and vented to atmospheric pressure. Where necessary, vials were flushed with ultra-high-purity N_2 (or desired quantities were introduced using gas-tight syringes); all vials were vented to atmospheric pressure before reactions were started. Reactions were performed within a shaking water bath (30 °C) and initiated by the addition of $MgCl_2$ (from a 1 M stock, 10 mM final concentration) using a gas-tight syringe. Reactions were quenched after 8 minutes by the addition of 300 µL of 400 mM EDTA (pH 8.0). H_2 quantification was performed on all reactions using a calibrated GC-TCD equipped with a molecular sieve 5 Å column (Ar carrier, SRI Instruments model 8610C). Ammonia was quantified by the *ortho*-phthalaldehyde method (corrected to controls and assays performed under 1 atm Ar) using NH_4Cl as the standard, as reported previously.[1,10]

Analytical Gel filtration

Analytical gel filtration was performed using a Superose 6 Increase 10/300 GL column (Cytiva), using an Äkta Start within an anoxic glovebox. The proteins were diluted with the running buffer (50 mM Tris/HCl, 500 mM NaCl, pH 8, 0.5 mM DT) to a concentration of 20 µmol/mL for MoFe protein and 100 µmol/mL for L127Δ Fe protein. 200 µL aliquots were prepared and incubated for one hour before being loaded (100 µL) onto the pre-equilibrated column.

Limited Proteolysis

A 50 mM Tris/HCl (pH 7.6) buffer containing 10 mM $CaCl_2$ buffer was prepared and deoxygenated in a Ar-filled glovebox (Jacomex, France) overnight. The buffer was split in 2 bottles and 0.2 mg/mL fresh trypsin (Sigma) was added to the first. This trypsin-containing buffer was then activated at 30 °C for 20 minutes using a heating block. Proteins were brought in the glovebox and diluted to a concentration of 1 mg/mL using the buffer that did not contain trypsin. The proteins were placed in LC/MS vials with stirring in a water bath at 37 °C and the activated trypsin was then added to the protein samples to a final concentration of 2 µg/mL (trypsin, 500:1 MoFe:trypsin) as the starting point for the reaction. The protein reactions were sampled at time 0 (before trypsin addition), 2, 5, 30, 60 and 90 minutes after trypsin addition. 4 µL of samples of the reactions was transferred to Eppendorf tubes containing that already contained 4 µL TruPAGE™ LDS (4X) sample buffer (Merck, Switzerland), 1.6 µL 1 M dithiothreitol (DTT) and 6.4 µL MilliQ water. The tube was then immediately placed at 95 °C on a heating block for 5 minutes to deactivate the trypsin and

terminate proteolysis. 4-12% mPAGE® Bis-tris gels (Merck, Switzerland) were used for analysis. 7 µL aliquots of the treated protein samples were loaded in each well. The gel was run at 200 V and stained with One-Step Blue® (Biotium, Brunschwig Switzerland).

Native PAGE

All the reagents and equipment were purchased from Invitrogen™/ThermoFisher scientific (Switzerland). The buffers were prepared using NativePAGE™ Running Buffer (20X) and NativePAGE™ Cathode Buffer Additive (20X). Buffers were diluted in an N_2-filled glovebox with anoxic MilliQ water. 0.5 mM DT was added to the cathode buffer. The gel (NativePAGE™ 3 to 12%, Bis-Tris) was placed in the cassette which was then filled with the two separate buffers. The gel was then run for 5 minutes prior to sample loading at 150 V. The protein samples were then diluted to a concentration of 0.2 mg/mL with 50 mM Tris (containing 0.5 mM DT) in the glovebox. 2.5 µL of NativePAGE™ Sample Buffer (4X) was then added to 7.5 µL of protein, and then 10 µL of the resulting samples (and 5 µL of protein ladder, NativeMark™ unstained protein standard) were loaded on the gel which was subsequently run in the glovebox at 150 V. The gel was stained with Coomassie following the protocol from Invitrogen™.

Western blotting

Protein samples (0.5 - 1 µg per well) were loaded on a 4-15% TGX Stainfree gel (BioRad, Switzerland) and run for 40 min at 220 V. The gel was irradiated with UV light for 1 min to activate the Stainfree component, before transferring the protein to PVDF membranes using the Trans-Blot Turbo Transfer System (Bio-Rad). The membranes were scanned for total protein (Stainfree), before being blocked with 10% Sea Block blocking buffer (ThermoFisher Scientific, Switzerland) in PBS for 1h at room temperature. Membranes were incubated overnight with 1:10000 HRP-Streptavidin (BioLegend, Lucerna-Chem Switzerland) or 1:10000 HRP-Strep-tactin (IBA lifesciences, Lucerna-Chem Switzerland) in 10% Sea Block/TBST (TRIS/HCl-buffered (20 mM pH 7.4-7.6) saline (150 mM) with 0.1% tween-20) at 4 °C, washed with TBST (4x 5 min), and the antibodies detected by chemiluminescence using Supersignal West Pico Plus chemiluminescent substrate and a Fusion imager (Vilber).

Proteomics analysis and specificity of the maleimide-Cys reaction

Raw proteomics data and detailed proteomics reports provided by the Proteomics Core Facility within the Faculty of Medicine at the University of Geneva have been uploaded to Zenodo (link on the first page). Briefly, strep-inhibited proteins were purified over a

StrepTrap XT (Cytiva Switzerland, Streptactin solid phase) to separate unmodified proteins. Next, these proteins were run into an SDS-PAGE gel, the band excised and stored in 10% acetic acid, and sent to the core proteomics facility for analysis. This was performed for strep-inhibited α-C45A/L158C MoFe protein and for wild-type MoFe protein; the wild-type MoFe protein was prepared as above for the α-C45A/L158C MoFe protein, although 5 molar equivalents of strep-maleimide were used per MoFe protein. After in-gel digestion with trypsin, samples were analyzed by LC-ESI-MS/MS using a Q-Exactive HF Hybrid Quadrupole-Orbitrap Mass Spectrometer (ThermoFisher Scientific) equipped with an Easy nLC 1000 Liquid Chromatography system (ThermoFisher Scientific) for modified cysteine residues. A peaklist was generated using ProteoWizard and searched against the Uniprot database for *A. vinelandii* DJ, with the addition of a modified version of *nifD* due to the expected α-C45A/L158C mutation. A subsequent search was also performed in order to identify potential strep-inhibitor modifications on any amino acids. A Mascot search of the results was analyzed using Scaffold 5.0.0 (Proteome software); identified peptides were accepted only in the case of having a Percolator posterior error probability of >95%. Under these conditions, the Strep-maleimide modification was detected on the expected trypsin-digested peptide of NifD, GISVQSE<u>C</u>PIG**C**IGDDIESVSK, where the bold Cys represents the target α-C158 residue and the underlined Cys represents a buried P cluster-coordinating α-C154 residue. Further, both unmodified and modified version of this peptide fragment were identified, consistent with the partial functionalization of $\alpha_2\beta_2$ α-C45A/L158C MoFe protein. The subsequent broader search for the potential modification of the strep-inhibitor on any amino acid identified three low-quality hits for a single partially solvent-exposed peptide fragment on the ß-subunit (NifK).

In order to clarify the selectivity of the strep-inhibitor maleimide for our targeted cysteine residue, analyses were repeated using wild-type MoFe protein functionalized with an excess (5 molar equivalents) of the strep-maleimide inhibitor (StrepTactin affinity purified as above). A strep-modified GISVQSE<u>C</u>PIG**L**IGDDIESVSK peptide was not detected (the unmodified peptide was detected), consistent with the strep-maleimide inhibitor's selectivity towards the target α-C158 residue. Neither an unmodified nor a strep-modified peptide fragment containing the original surface-exposed α-C45 residue was detected in this sample (although the sample was purified over the StrepTrap XT column, as above); this was hypothesized to be due to the small size of the target peptide (outside of the MS detection window). Importantly, the potential ß-subunit (NifK) modification identified above was not identified in

the presence of excess strep-inhibitor, consistent with the above identification indeed being an experimental artefact.

Two additional database searches were performed on the raw data by Dr Yibo Wu within the Department of Organic Chemistry at the University of Geneva. Interestingly, the search engine SpectroMine did not return a strep-functionalized peptide fragment for the α-C45A/L158C MoFe protein. An additional search on the Proteome Discover database (v2.4) corroborated the desired functionalization of the α-C158 residue, although an additional modification was identified on the FeMoco-coordinating α-C275 residue. Since (i) the α-C45A/L158C MoFe protein has high specific activity values (this work and in Ref[2]), (ii) X-ray crystallography confirms the double occupancy of the MoFe protein with FeMoco and (iii) a Strep-functionality at this buried residue is unlikely (and would not permit purification over the StrepTrap XT column), this is anticipated to be a minor artefact due to reversible maleimide-Cys degradation during peptide digestion/sample preparation.[11]

An additional control experiment was performed using an additional α-C45S MoFe protein mutant (purified from *A. vinelandii* strain DJ2192, producing a MoFe protein with a His-tag on the N-term). The strep-inhibition strep-modification protocol outlined above was repeated and the reaction product was purified as above. Importantly, the reaction efficiency decreased from 14% (for the α-C45A/L158C MoFe mutant) to 7% (for the α-C45S MoFe mutant), indicating that the maleimide-Cys may not be entirely selective. We hypothesize that this promiscuous reactivity could have been further enhanced in the absence of a solvent exposed Cys residue on this mutant. Additional analysis of this α-C45S MoFe mutant (Figures S18-20) also indicated that this non-specific strep-inhibitor modification may also inhibit Fe protein access to the MoFe protein interaction site. This is easily explained by the relatively large size and potential flexibility of the strep inhibitor.

Table S2: X-ray analysis statistics.

Data collection	
Wavelength (Å)	1.00003
Space group	$P2_1$
Resolution (Å)	143.49 – 3.03 (3.49 – 3.03)
Cell dimensions	
a, b, c (Å)	148.39, 73.44, 211.25
α, β, γ (°)	90.00, 104.76, 90.00
R_{merge}(%)[a]	29.7 (51.4)
R_{pim} (%)[a]	17.5 (33.4)
$CC_{1/2}$ [a]	0.957 (0.806)
I/σ_I [a]	3.4 (1.6)
Spherical completeness[a]	46.2 (6.8)
Ellipsoidal completeness[a]	86.6 (69.4)
Redundancy[a]	3.7 (3.3)
Nr. unique reflections[a]	39,936 (1,998)
Refinement	
Resolution (Å)	95.50 – 3.03
Number of reflections	39,924
R_{work}/R_{free} [b] (%)	20.37/22.27
Number of atoms	
Protein	31,844
Ligands/ions	314
Solvent	0
Mean B-value (Å^2)	35.11
Molprobity clash score, all atoms	2.14
Ramachandran plot	
Favored regions (%)	96.58
Outlier regions (%)	0.15
rmsd[c] bond lengths (Å)	0.004
rmsd[c] bond angles (°)	0.664
PDB ID code	**8BTS**

[a] Values relative to the highest resolution shell are within parentheses. [b] R_{free} was calculated as the R_{work} for 5 % of the reflections that were not included in the refinement. [c] rmsd, root mean square deviation.

Table S3: Comparison of the unit cell dimensions between the crystal structures of MoFe proteins.

Entity ID	Length a (Å)	Length b (Å)	Length c (Å)	Angle α (°)	Angle β (°)	Angle γ (°)	Molecules/AU	Space Group
5BVG	77.582	130.825	107.128	90	108.85	90	4	*P* 1 21 1
5BVH	81.071	130.833	107.318	90	110.64	90	4	*P* 1 21 1
4ND8	76.527	127.915	107.082	90	108.9	90	4	*P* 1 21 1
4TKU	80.941	130.785	107.005	90	110.58	90	4	*P* 1 21 1
4TKV	81.172	130.623	106.996	90	110.65	90	4	*P* 1 21 1
1G20	264.24	111.46	121.59	90	97.4	90	8	*C* 1 2 1
1G21	110.5	121.5	264.9	90	90	90	8	*P* 21 21 21
6CDK	107.88	130.78	80.789	90	110.85	90	4	*P* 1 21 1
6UG0	84.362	156.967	202.343	90	90	90	4	*P* 21 21 2
5VQ4	81.315	128.859	108.359	90	110.89	90	4	*P* 1 21 1
4WNA	77.123	129.78	107.543	90	108.94	90	4	*P* 1 21 1
3U7Q	81.19	130.696	107.224	90	110.67	90	4	*P* 1 21 1
6O7R	76.61	128.81	107.64	90	109.07	90	4	*P* 1 21 1
6O7S	76.775	128.05	107.539	90	108.97	90	4	*P* 1 21 1
6O7P	76.847	128.439	107.707	90	109.06	90	4	*P* 1 21 1
6O7Q	77.437	130.413	107.751	90	109.09	90	4	*P* 1 21 1
6O7N	76.337	127.96	107.509	90	109.01	90	4	*P* 1 21 1
6O7O	76.746	128.678	107.536	90	108.94	90	4	*P* 1 21 1
6O7L	80.562	131.029	107.616	90	110.85	90	4	*P* 1 21 1
6O7M	76.457	127.788	107.539	90	109	90	4	*P* 1 21 1
7JRF	77.052	129.974	107.223	90	109.11	90	4	*P* 1 21 1
6VXT	163.535	203.435	83.857	90	103.945	90	4	*C* 1 2 1
1M34	326.1	75.8	312.2	90	102.6	90	16	*C* 1 2 1
3MIN	108	131.3	81	90	110.7	90	4	*P* 1 21 1
6BBL	76.781	128.251	107.272	90	109.11	90	4	*P* 1 21 1
1M1N	108.31	131.63	159.159	90	108.37	90	8	*P* 1 21 1
1L5H	150.452	191.922	102.339	90	90	90	2	*C* 2 2 21
2AFH	170.911	75.893	223.665	90	90	90	6	*P* 21 21 21
1M1Y	113.267	214.937	320.466	90	90	90	16	*P* 21 21 21
2MIN	107.7	130.2	81.3	90	110.8	90	4	*P* 1 21 1
2AFI	72.915	141.432	165.549	73.69	79.37	76.58	16	*P* 1
1N2C	79	299.7	334.5	90	90	90	8	*C* 2 2 21
3K1A	77.019	129.458	107.088	90	109.01	90	4	*P* 1 21 1
1FP4	107.2	130.2	80.4	90	111.2	90	4	*P* 1 21 1
4XPI	80.804	130.83	108.11	90	111.14	90	4	*P* 1 21 1
5CX1	175.502	144.595	177.747	90	114.27	90	16	*P* 1 21 1
4WZA	110.201	120.412	264.318	90	90	90	8	*P* 21 21 21
4WZB	110.533	120.894	264.834	90	90	90	8	*P* 21 21 21
6OP4	76.981	129.326	106.755	90	108.84	90	4	*P* 1 21 1
6OP3	80.872	131.023	106.957	90	110.62	90	4	*P* 1 21 1
6OP2	80.761	130.38	106.896	90	110.54	90	4	*P* 1 21 1
6OP1	77.191	128.003	107.633	90	109.15	90	4	*P* 1 21 1
αC45A αL158C	**148.392**	**73.444**	**211.253**	**90**	**104.76**	**90**	**8**	***P* 1 21 1**

MoFe protein crystalline form 1

MoFe protein crystalline form 2

MoFe protein other crystalline forms

Fe:MoFe complexes

Table S4: Structural alignment between the α-C45A L158C and MoFe-protein crystal structures. Models with high similarity (root mean square deviation, RMSD < 0.205 Å) to the α-C45A L158C model are bolded.

PDB code	Structure name	RMSD Å and (Superposed Cα)	PDB code	Structure name	RMSD Å and (Superposed Cα)
1FP4	Crystal structure of the alpha-H195Q mutant of nitrogenase	0.258 (1896)	5BVH	CO-bound form of Selenium incorporated nitrogenase MoFe-protein (Av1-Se-CO) from *A. vinelandii*	0.268 (1900)
1G20	Nucleotide-free structure of a nitrogenase protein complex between Leu127del-Fe protein and the MoFe protein	0.228 (1851)	**5CX1**	Nitrogenase molybdenum-iron protein beta-K400E mutant	**0.185 (1855)**
1G21	MG-ATP-bound structure of a nitrogenase protein complex between leu127del-Fe protein and the MoFe protein	0.439 (1908)	5VQ4	Nitrogenase Av1 at pH 5	0.253 (1876)
1L5H	FeMo-cofactor Deficient Nitrogenase MoFe Protein	0.400 (792) one NifDK in AU	6BBL	Crystal structure of the a-96Gln MoFe protein variant in the presence of the substrate acetylene	0.221 (1862)
1M1N	Nitrogenase MoFe protein from *Azotobacter vinelandii*	0.288 (1916)	6CDK	Characterization of the P1+ intermediate state of nitrogenase P-cluster	0.243 (1858)
1M1Y	Chemical Crosslink of Nitrogenase MoFe Protein and Fe Protein	0.304 (1908)	6O7L	Nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	0.223 (1849)
1M34	Nitrogenase Complex From *Azotobacter vinelandii* Stabilized By ADP-Tetrafluoroaluminate	0.283 (1912)	6O7M	Nitrogenase MoFeP mutant F99Y from *Azotobacter vinelandii* in the indigo carmine oxidized state	0.235 (1832)
1N2C	Nitrogenase complex from *Azotobacter vinelandii* stabilized by ADP-tetrafluoroaluminate	0.369 (1912)	6O7N	Nitrogenase MoFeP mutant F99Y/S188A from *Azotobacter vinelandii* in the indigo carmine oxidized state	0.220 (1811)
2AFH	Crystal Structure of Nucleotide-Free Av2-Av1 Complex	**0.204 (1867)**	6O7O	Nitrogenase MoFeP mutant F99Y/S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	0.216 (1840)
2AFI	Crystal Structure of MgADP bound Av2-Av1 Complex	0.280 (1889)	6O7P	Nitrogenase MoFeP mutant F99Y from *Azotobacter vinelandii* in the dithionite reduced state	0.224 (1839)
2MIN	Nitrogenase mofe protein from *Azotobacter vinelandii*, oxidized state	0.222 (1885)	6O7Q	Nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the dithionite reduced state after redox cycling	0.215 (1846)
3K1A	Insights into substrate binding at FeMo-cofactor in nitrogenase from the structure of an alpha-70Ile MoFe protein variant	0.235 (1859)	6O7R	Nitrogenase MoFeP mutant F99Y, S188A from *Azotobacter vinelandii* in the dithionite reduced state	0.237 (1871)
3MIN	Nitrogenase MoFe protein from *Azotobacter vinelandii*, oxidized state	0.238 (1865)	**6O7S**	Nitrogenase MoFeP mutant S188A from *Azotobacter vinelandii* in the indigo carmine oxidized state	**0.206 (1842)**
3U7Q	*A. vinelandii* nitrogenase MoFe protein at atomic resolution	0.270 (1887)	6OP1	Selenium incorporated, carbon monoxide inhibited FeMo-cofactor of *Azotobacter vinelandii*	0.226 (1839)
4ND8	Av Nitrogenase MoFe Protein High pH Form	0.239 (1817)	6OP2	Selenium incorporated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii* at high concentration of selenium	0.269 (1876)
4TKU	Reactivated Nitrogenase MoFe-protein from *A. vinelandii*	0.267 (1867)	6OP3	Selenium incorporated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii* with low concentration of selenium	0.271 (1869)
4TKV	CO-bound Nitrogenase MoFe-protein from *A. vinelandii*	0.251 (1874)	6OP4	Selenium-incorporated, carbon monoxide-inhibited, reactivated FeMo-cofactor of nitrogenase from *Azotobacter vinelandii*	0.301 (1915)
4WNA	Structure of the Nitrogenase MoFe Protein from *Azotobacter vinelandii* Pressurized with Xenon	0.243 (1860)	6UG0	N2-bound Nitrogenase MoFe-protein from *Azotobacter vinelandii*	0.232 (1920)
4WZA	Asymmetric Nucleotide Binding in the Nitrogenase Complex	**0.203 (1850)**	6VXT	Activated Nitrogenase MoFe-protein from *Azotobacter vinelandii*	0.205 (1910)
4WZB	Crystal Structure of MgAMPPCP-bound Av2-Av1 complex	0.234 (1887)	7JRF	CO-CO-bound nitrogenase MoFe-protein from *A. vinelandii*	0.250 (1816)
4XPI	Fe protein independent substrate reduction by nitrogenase variants altered in intramolecular electron transfer	0.248 (1919)	5BVG	Selenium incorporated nitrogenase MoFe-protein (Av1-Se2B) from *A. vinelandii*	0.255 (1874)

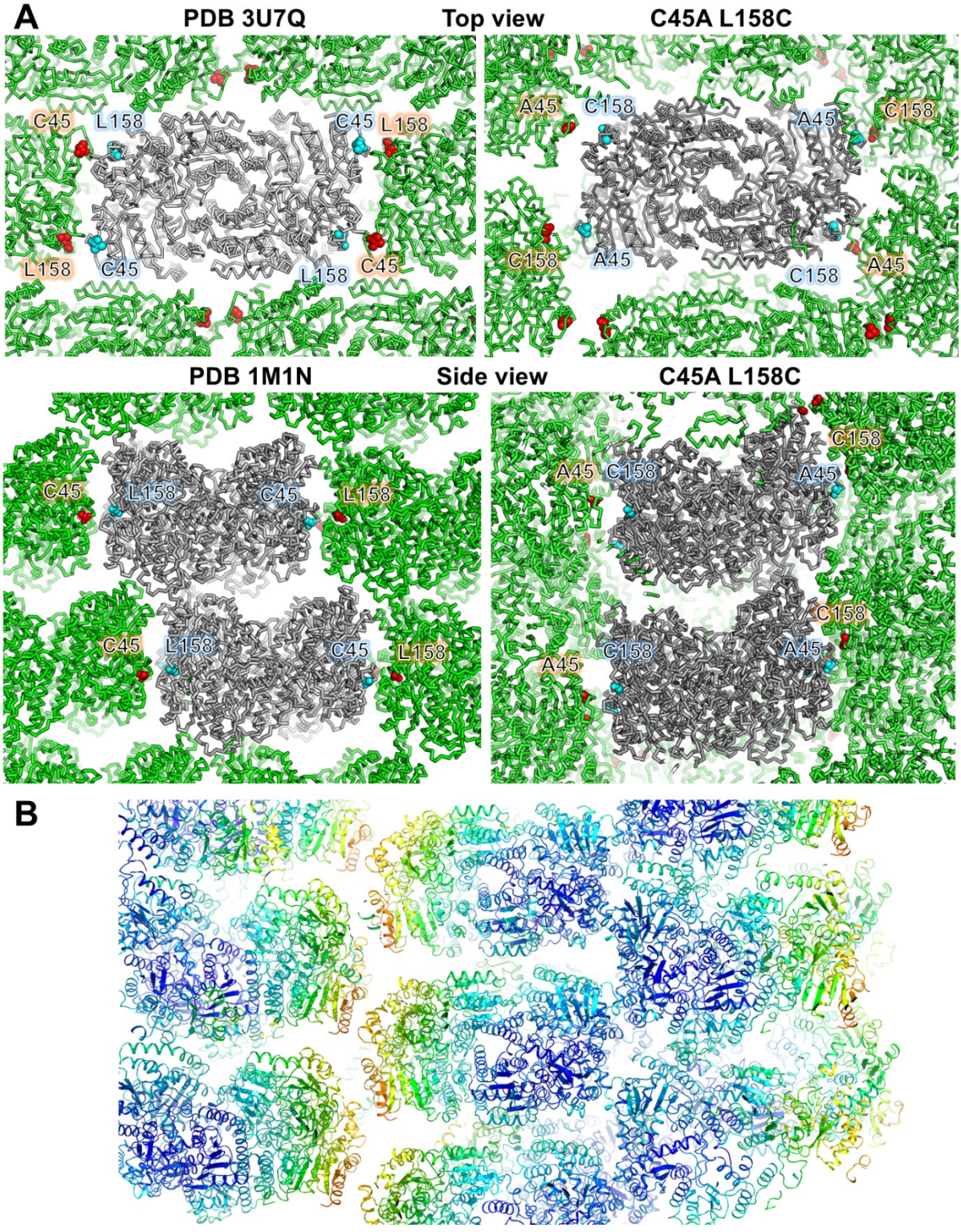


Figure S2: Influence of the mutations on the crystal packing. (**A)** The Cα backbone of the protein is represented as a ribbon. Residues at positions 45 and 158 on the α-subunit are shown as balls. The two nitrogenases of the asymmetric unit are colored in grey, while the nitrogenases constituting the crystal packing are colored in green. (**B)** Crystal packing of C45A L158C double mutant. The anisotropy is illustrated by the B-factor color code from blue (low values) to red (high values).

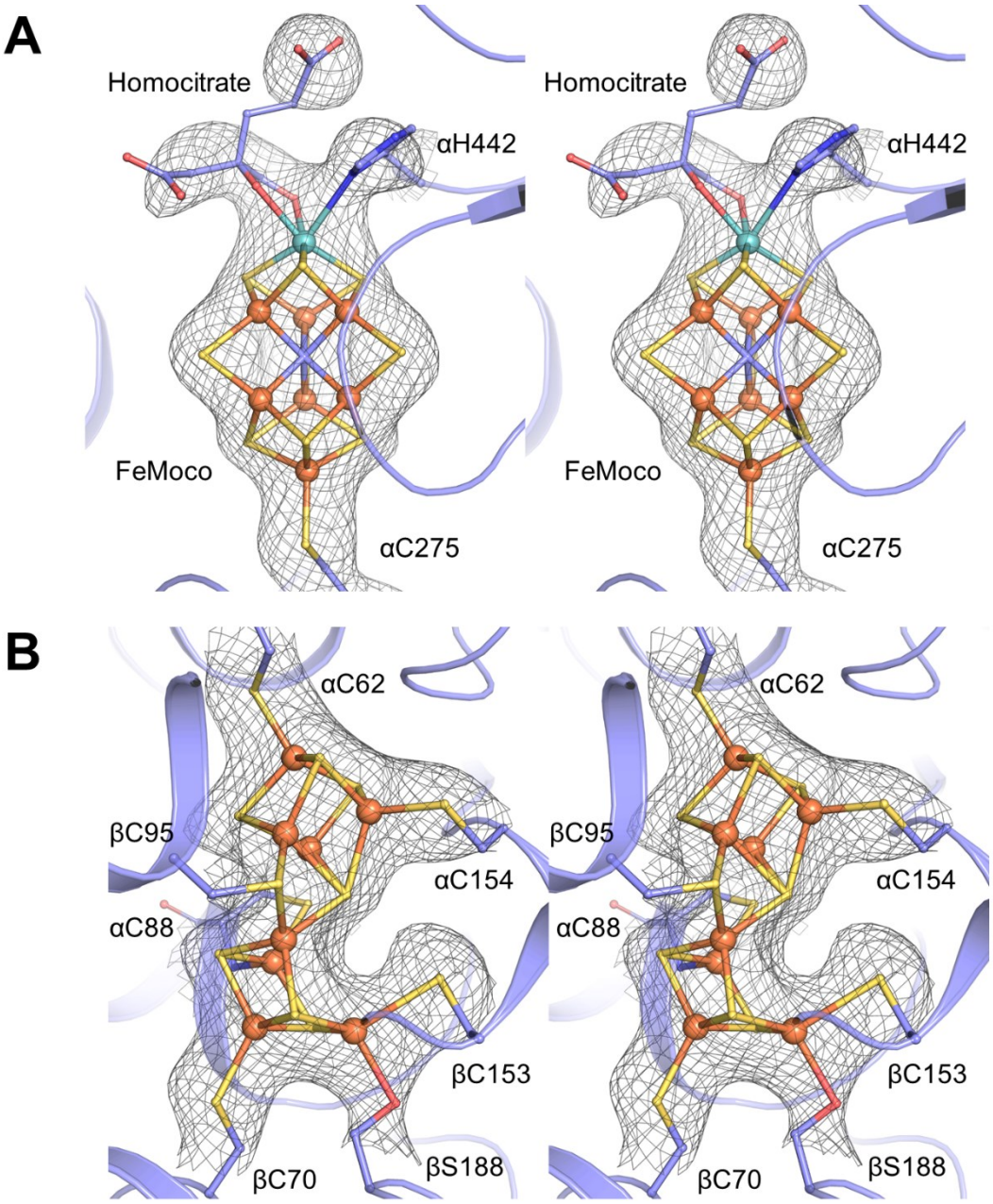


Figure S3: Integrity of the FeMo cofactors and P clusters in the crystallized α-C45A/L158C MoFe proteins. Stereo view of the FeMoco (**A**) and P cluster (**B**). The $2F_o$-F_c map, contoured at 1.5 is shown in black mesh and the residues coordinating the clusters are in sticks.

Synthesis of the maleimide-desthiobiotin inhibitor ("DTB inhibitor")

General Information: Anhydrous conditions were set up under an inert atmosphere (Ar) utilizing glassware that were oven dried and cooled under Ar purging or under N_2 using standard Schlenk line apparatus. Starting materials were purchased directly from commercial suppliers (Sigma Aldrich, Acros, Alfa Aesar, Fluorochem) and used without further purification unless otherwise stated. All solvents were dried according to standard procedures or bought from commercial suppliers. Reactions were monitored by LC-MS (DIONEX Ultimate 3000 UHPLC with a Thermo LCQ Fleet Mass Spectrometer System using PINNACLE DB C18 column (1.9 µm, 50 x 2.1 mm) with Thermo Xcalibur 2.2.SP1.48 software and analysis with Thermo Xcalibur Qual Browser 2.2.Sp1.48). Reverse phase chromatography on SNAP Ultra 4.5g C18 column using **Biotage®** Isolera^{TM} Four, **15 mL/min**.

NMR characterisation data (^{1}H NMR, ^{13}C NMR and 2D spectra) were collected at 300 K on a Bruker DRX400 (400 MHz) using $CDCl_3$ as solvent. Data for ^{1}H NMR are reported as follows: chemical shift (δ ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad), coupling constant (Hz), integration with the solvent resonance as internal standard ($CDCl_3$: δ = 7.26 ppm for ^{1}H, δ = 77.16 ppm for ^{13}C).

For the characterization of new compounds HR ESI-MS was performed on a Xevo G2-S Tof (Waters) and are reported as mass per charge ratio (m/z) calculated and observed.

N-(15-(2,5-dioxo-2,5-dihydro-1H-pyrrol-1-yl)-12-oxo-3,6,9-trioxa-13-azapentadecyl)-6-((4S,5S)-5-methyl-2-oxoimidazolidin-4-yl)hexanamide (maleimide-desthiobiotin ("DTB") inhibitor)

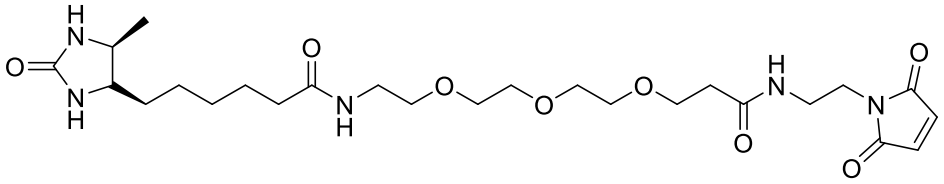


Figure S4: Chemical structure of the DTB inhibitor.

To a solution of 1-(2-aminoethyl) maleimide hydrochloride (13 mg, 0.07 mmol, 1.2 eq.) in dry DMF (1.5 mL), 2,5-dioxopyrrolidin-1-yl 19-((4S,5S)-5-methyl-2-oxoimidazolidin-4-yl)-14-oxo-4,7,10-trioxa-13-azanonadecanoate (31 mg, 0.06 mmol, 1 eq.) and DIPEA (8 mg, 11 µL, 0.06 mmol, 1 eq.) were added at 0 °C. The reaction mixture was stirred at room temperature, under argon, for 15 minutes before the solvent was removed under reduced pressure. The crude reaction mixture was purified by RP flash chromatography (linear gradient 0 – 60% CH_3CN in H_2O with 0.1% TFA) to yield the title compound as a yellow oil (14 mg, 0.025 mmol, 42%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.92 (s, 1H), 6.71 (s, 2H), 6.40 (s, 1H), 5.59 (s, 1H), 4.98 (s, 1H), 3.95 – 3.83 (m, 1H), 3.77 – 3.72 (m, 1H), 3.71 – 3.68 (m, 2H), 3.68 – 3.66 (m, 2H), 3.65 – 3.59 (m, 8H), 3.56 (t, *J* = 5.1 Hz, 2H), 3.49 – 3.39 (m, 4H), 2.43 (t, *J* = 5.8 Hz, 2H), 2.20 (t, *J* = 7.4 Hz, 2H), 1.72 – 1.60 (m, 2H), 1.54 – 1.44 (m, 2H), 1.43 – 1.23 (m, 4H), 1.15 (d, *J* = 6.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.35, 172.21, 171.08, 163.95, 134.34, 70.48, 70.29, 70.21, 70.16, 70.12, 67.15, 56.30, 51.77, 39.32, 38.46, 37.80, 36.78, 36.18, 29.52, 28.88, 26.02, 25.41, 15.83. LC/MS analysis (linear gradient 30→90% ACN) t_R: 1.54 min, ESI-MS (m/z): $[M+H]^+$: 540.25. HRMS (ESI): calc. for $C_{25}H_{41}N_5O_8$ $[M+Na]^+$ 562.2853, found 562.2867.

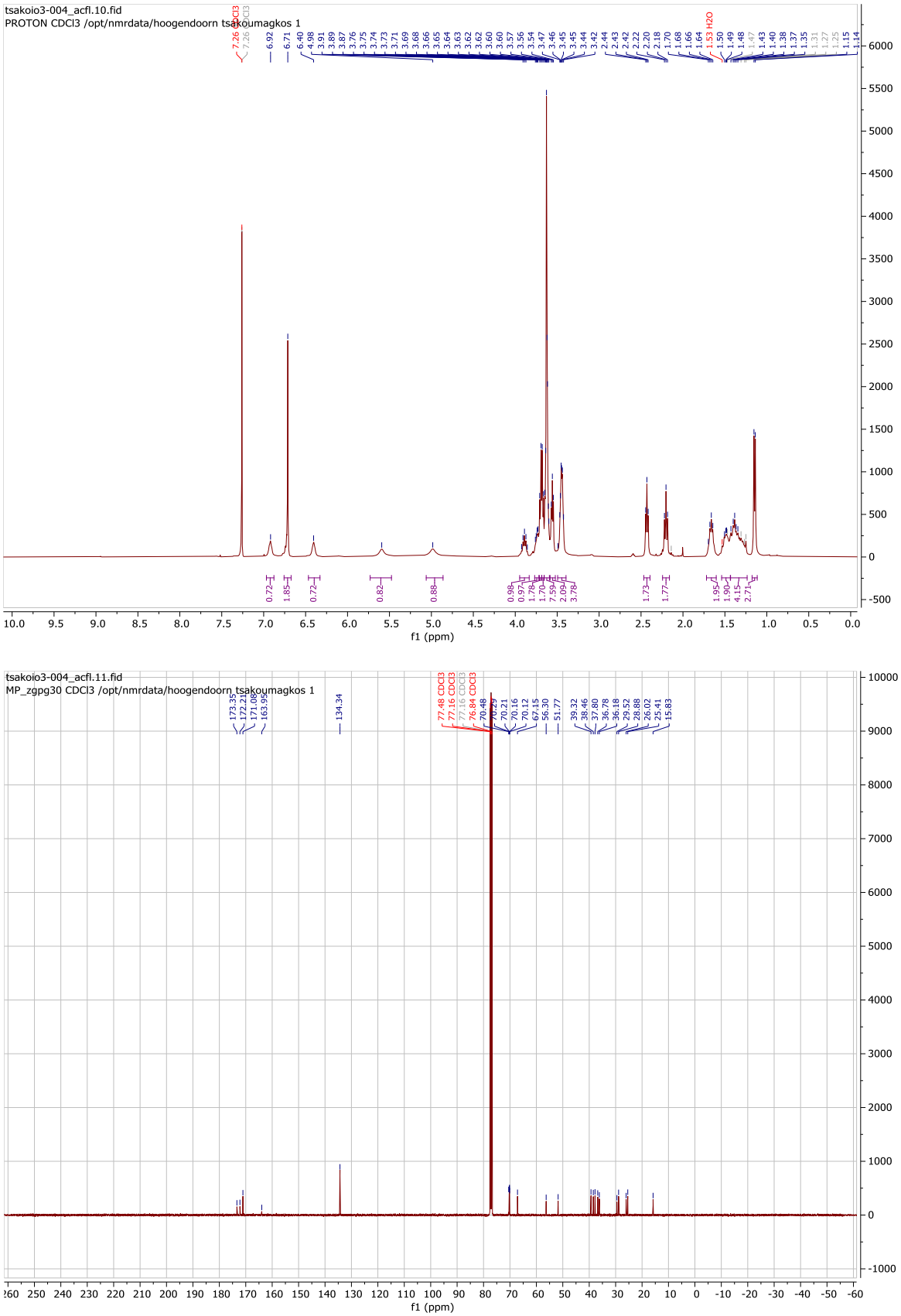


Figure S5: ^{1}H NMR (400 MHz) ^{13}C NMR (101 MHz) spectra of the DTB inhibitor ($CDCl_3$).

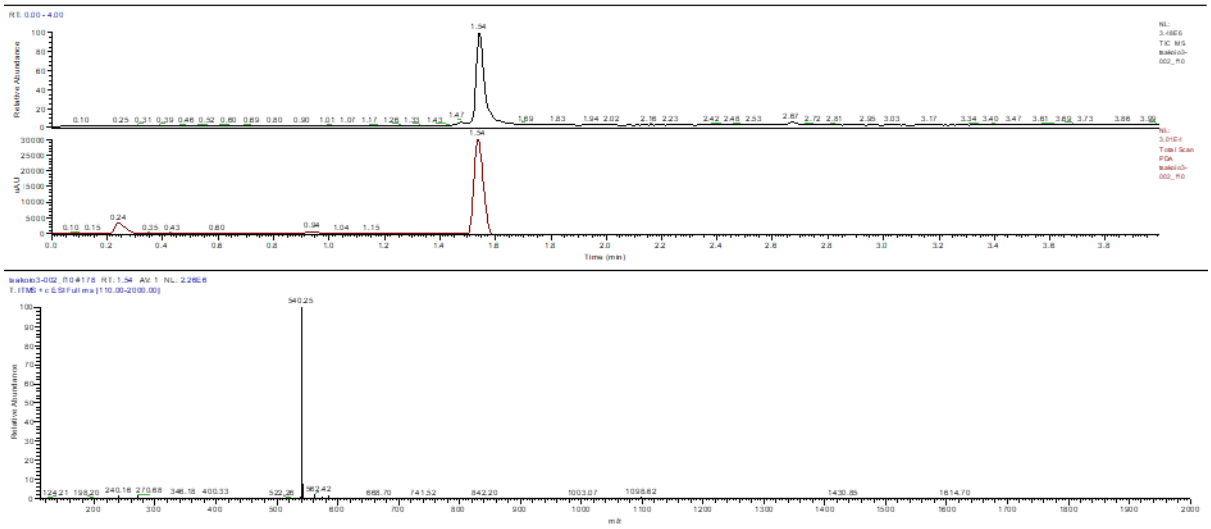


Figure S6: LC-MS of the title compound.

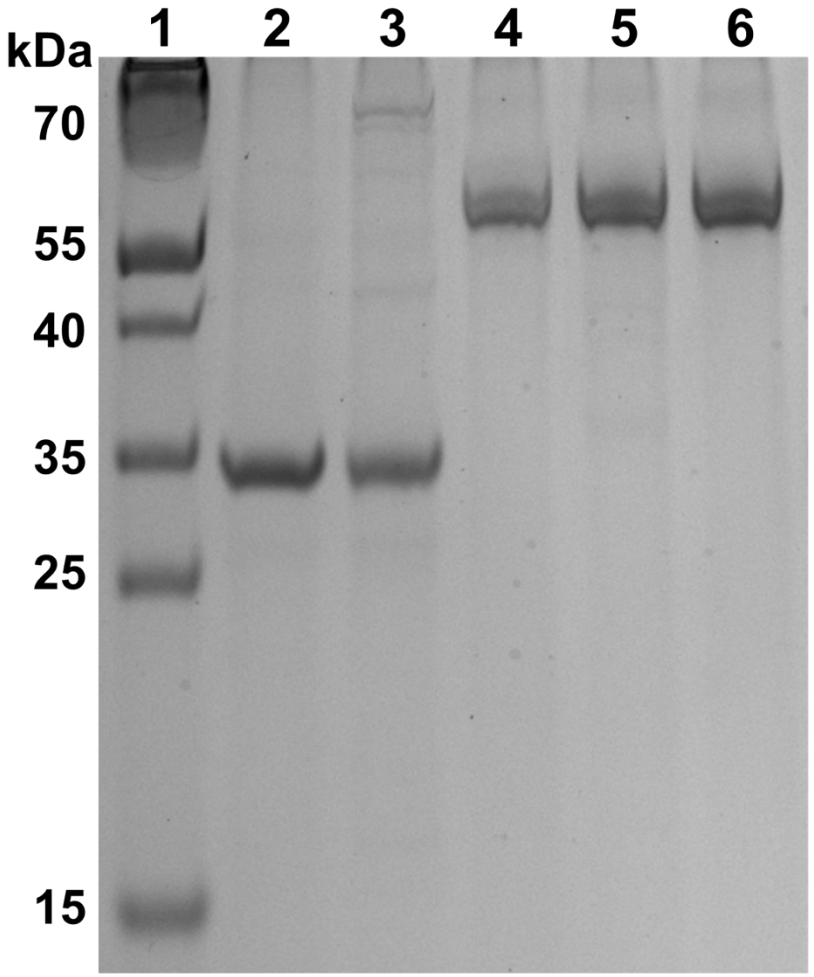


Figure S7: SDS-PAGE (12%), 200 V, Coomassie blue stained. Lanes: 1 = molecular weight markers, 2 = L127Δ Fe protein, 3 = WT Fe protein, 4 = α-C45A/L158C MoFe protein, 5 = WT MoFe protein, 6 = Strep-tagged MoFe protein (N_{term} Strep-II tag on NifD). 400 ng of protein per well.

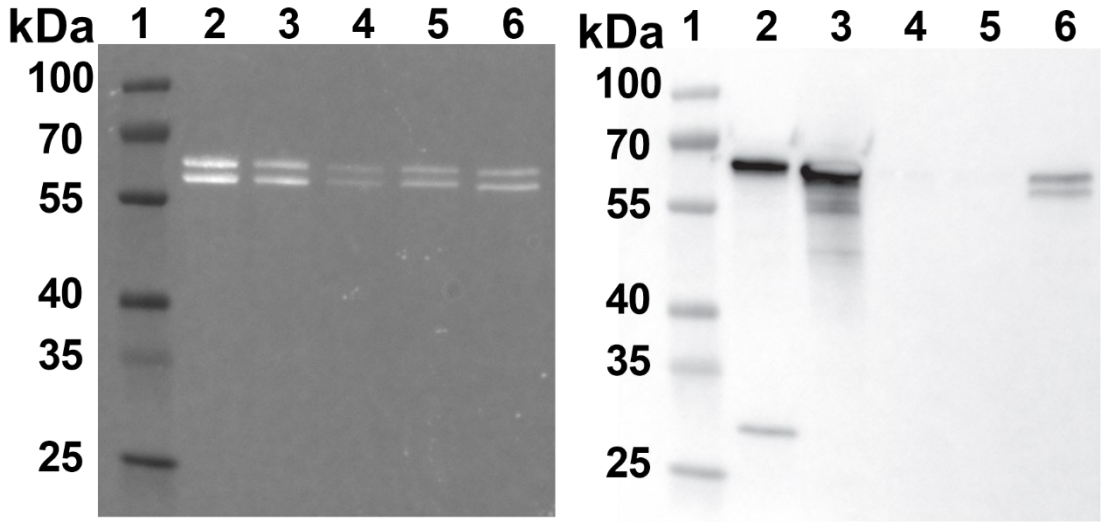


Figure S8: SDS-PAGE analysis (left, in-gel staining, BioRad) and western blotting with streptavidin-HRP on WT MoFe protein (right): Lanes: 1 = molecular weight marker; 2 = Strep-tagged WT MoFe protein (N_{term} Strep-II tag on NifD, from *A. vinelandii* strain DJ2102, shared by Dennis Dean – Virginia Tech); 3 = WT MoFe protein treated with DTB inhibitor (DT-free); 4 = "3" although the reaction was performed in the presence of 1 mM DT, 5 = "3" although the reaction was performed in the presence of 10 mM DT, 6 = "3" although the reaction was performed in the presence of 1 mM TCEP. 1 μg of protein per well.

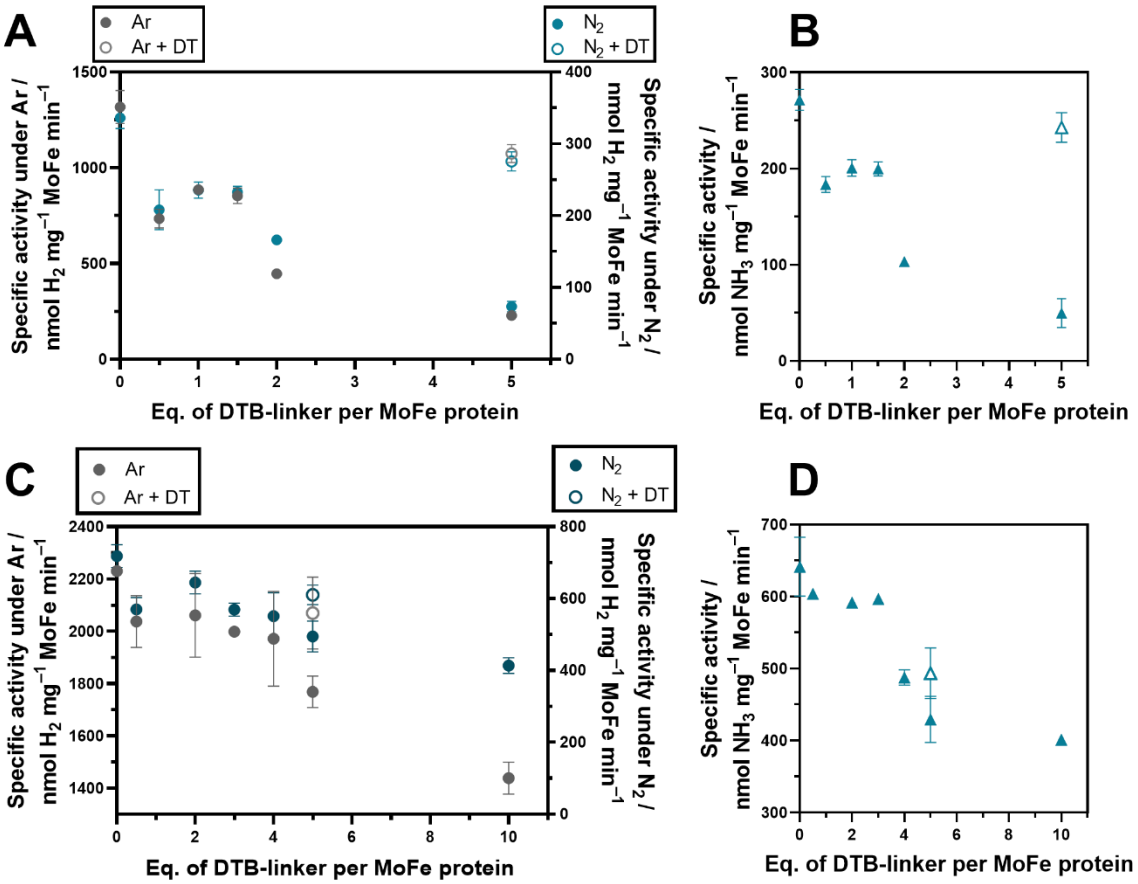


Figure S9: Specific activities of the α-C45A/L158C (top) and WT (bottom) MoFe proteins upon treatment with increasing equivalents of DTB inhibitor (per MoFe protein). **(A)** Specific activity for H_2 evolution, labelled as black dots on the left y axis (under 1 atm Ar) and as blue dots on the right y axis (under 1 atm N_2). Empty dots represent the specific activities obtained when treating the α-C45A/L158C MoFe protein with 10 mM DT before the addition of the DTB inhibitor (to reduce the maleimide). **(B)** Specific activity for NH_3 evolution, labeled as blue triangles under 1 atm N_2 and an empty triangle for α-C45A/L158C MoFe treated with 10 mM DT prior to DTB inhibitor addition. **(C)** Specific activity for H_2 evolution, labelled as black dots on the left y axis (under 1 atm Ar) and as blue dots on the right y axis (under 1 atm N_2). Empty dots represent the specific activities obtained when treating the WT MoFe protein with 10 mM DT before the addition of the DTB inhibitor (to reduce the maleimide and inhibit the Cys conjugation reaction). **(D)** Specific activity for NH_3 evolution, labeled as blue triangles under 1 atm N_2 and an empty triangle for WT MoFe treated with 10 mM DT prior to DTB inhibitor addition. In all cases $n = 3$.

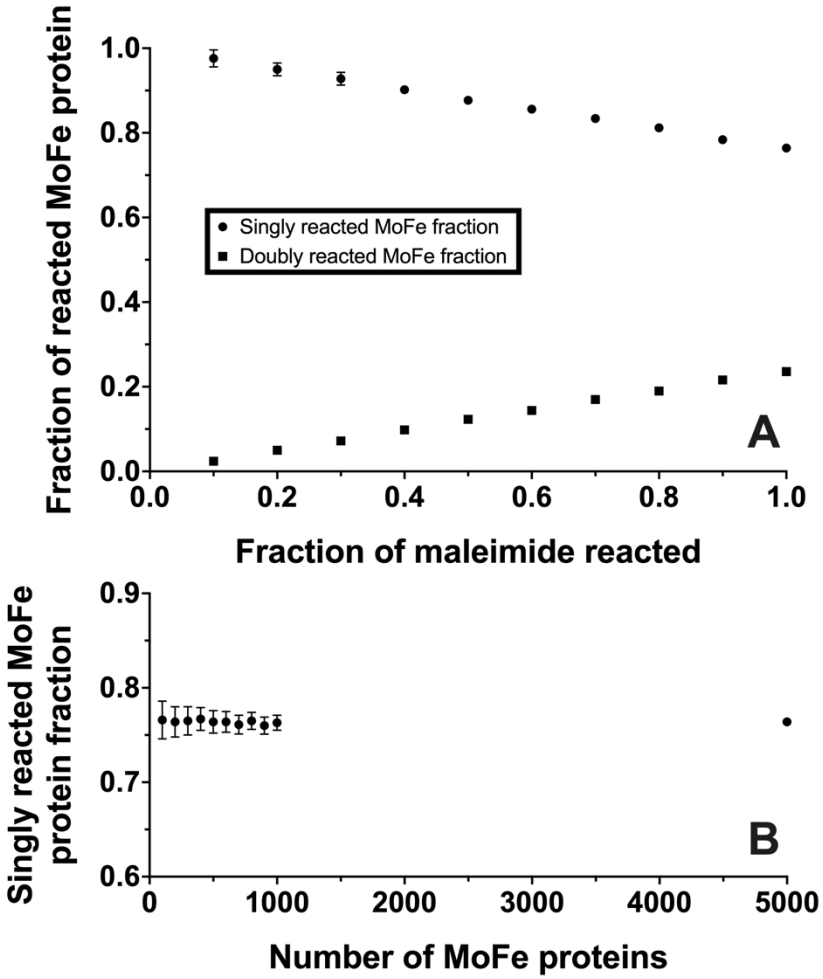


Figure S10: Simulation of singly and doubly reacted α-C45A/L158C MoFe protein fraction upon treatment with 0.5 molar equivalents of maleimide-streptag inhibitor. (**A**) Simulation of 1000 proteins reacting with 500 maleimide-Streptags (100 simulations for each point, mean and SD reported), represented the fractions of "reacted" MoFe protein that contain singly and doubly reacted MoFe (where a singly reacted MoFe = half-inhibited MoFe). Assumptions: (i) a reaction takes place each time a maleimide-streptag inhibitor reaches the α-158C position, (ii) no reaction occurs if the doubly reacted α-C45A/L158C MoFe is encountered, (iii) the maleimide-streptag reaction is completely selective for the α-C158 residue, and (iv) the maleimide is not hydrolyzed or reduced prior to reaching the α-C158 residue. (**B**) Fraction of singly reacted α-C45A/L158C MoFe protein upon scaling up of the simulation at the same ratio of molar equivalents.

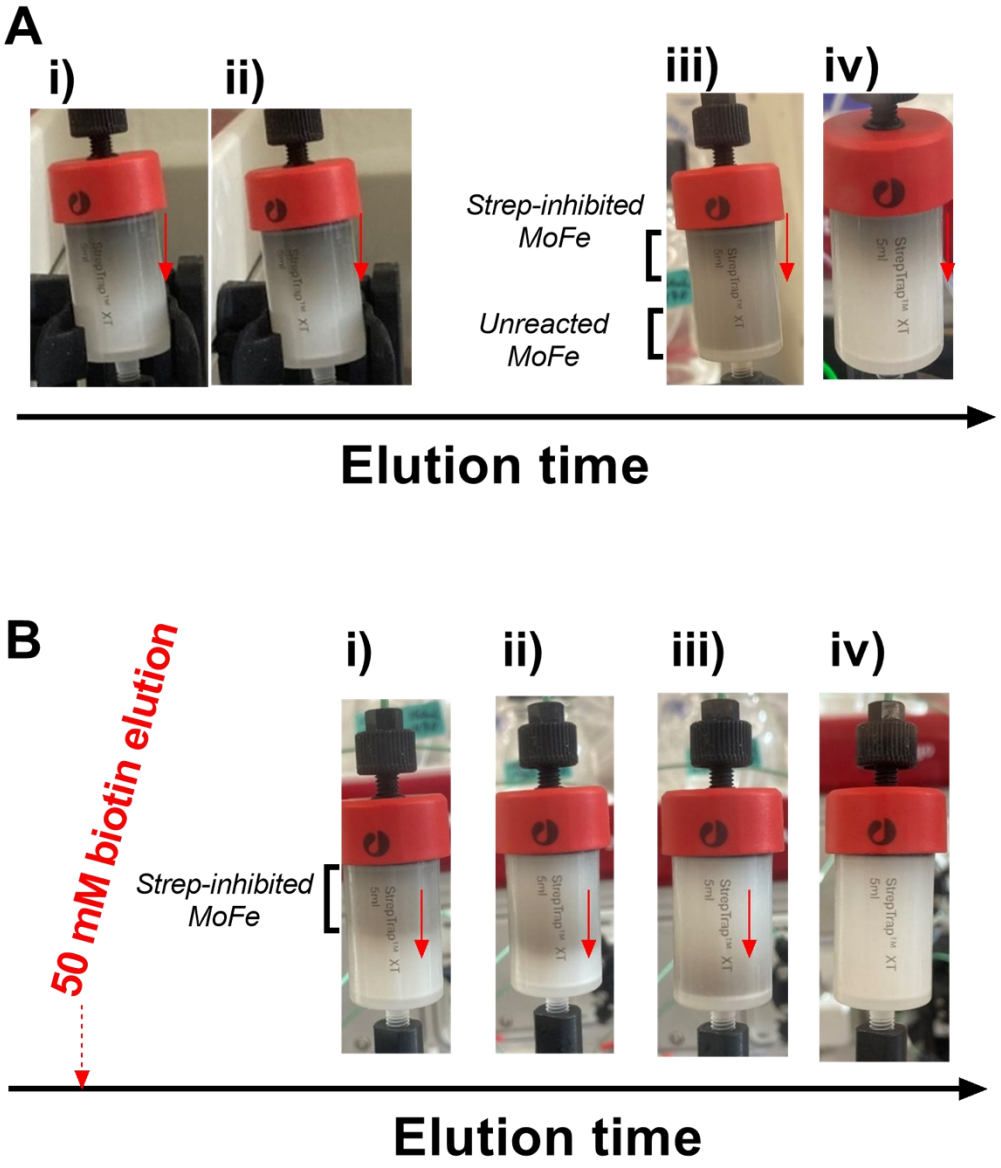


Figure S11: Purification of strep-inhibited MoFe protein after reaction with the Strep-peptide inhibitor. (**A**) Pictures of the StrepTrap XT column after sample loading (sample contains untreated and Strep-inhibited α-C45A/L158C MoFe protein). Red arrow indicates the direction of elution, and (iii) represents the differentiation between Strep-inhibited MoFe protein binding to the column and unreacted MoFe protein passing as the column flow-through. (**B**) Pictures of the column after elution of the Strep-inhibited MoFe protein with buffer containing 50 mM biotin.

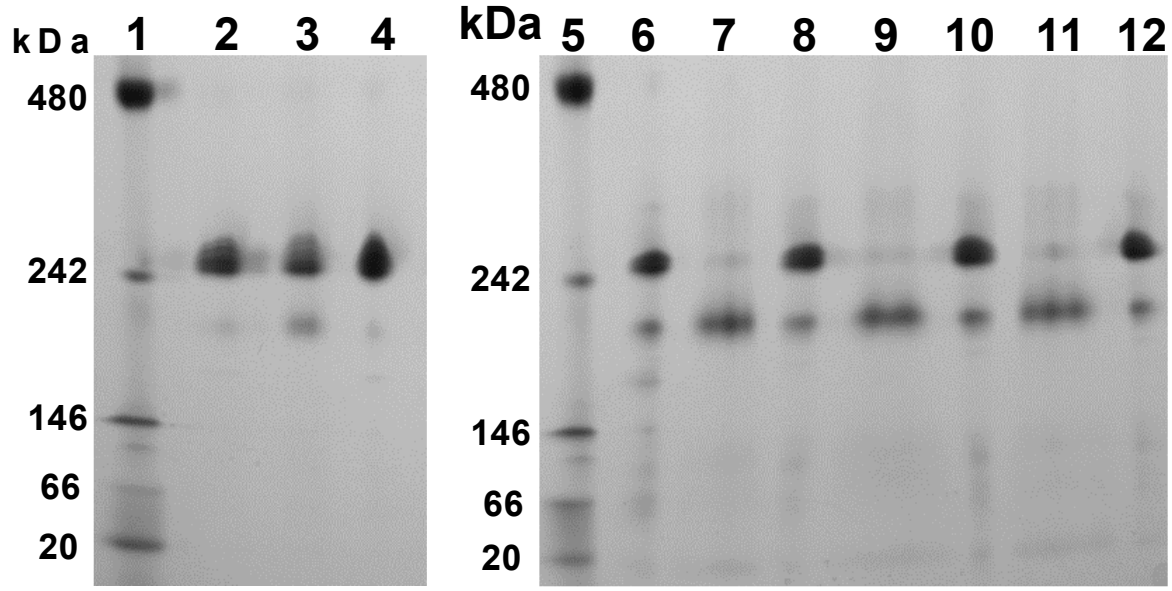


Figure S12: Native PAGE 3-12% acrylamide gradient, anoxic atmosphere (>95% N_2 and <5% H_2), 150 V. Lanes: 1 and 5 = molecular weight markers, 2 = untreated WT MoFe protein, 3 = Strep-inhibited WT MoFe protein (eluted from the StrepTactin XT solid phase), 4= unmodified WT MoFe protein that did not tightly bind to the column solid phase (unreacted flow-through), 6 = unreacted α-C45A/L158C MoFe protein, 7/9/11 = Strep-inhibited α-C45A/L158C MoFe protein (eluted from the StrepTactin XT solid phase); 8/10/12 = unmodified α-C45A/L158C MoFe protein that did not tightly bind to the column solid phase (unreacted flow-through). 7/8, 9/10 and 11/12 represent three independent α-C45A/L158C MoFe protein functionalization reactions. Strep-inhibited wild-type MoFe protein was prepared as detailed for the α-C45A/L158C MoFe protein. 1.5 µg protein per well.

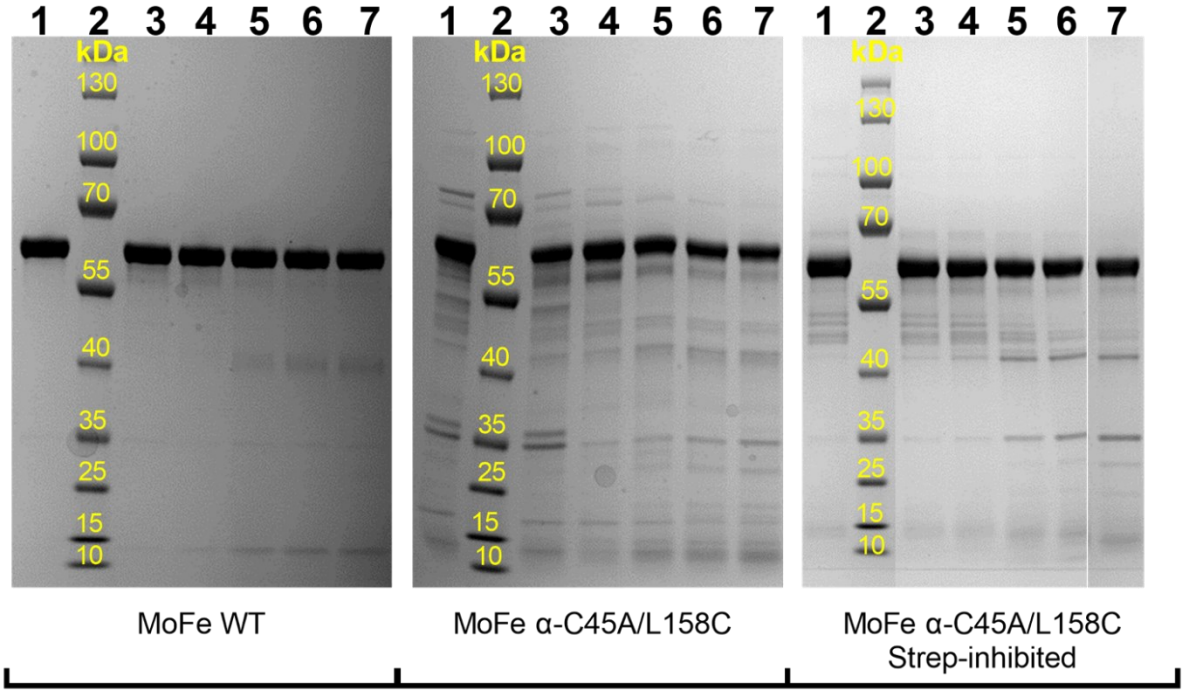


Figure S13: SDS-PAGE analysis for temperature-controlled limited proteolysis (Coomassie blue staining). Lanes: 1 = MoFe protein at t = 0 min after trypsin addition, 2 = molecular weight markers, 3 = MoFe protein at t = 2 min, 4 = MoFe protein at t = 5 min, 5 = MoFe protein at t = 30 min, 6 = MoFe protein at t = 60 min, 7 = MoFe protein at t = 90 min. *Note*: Lane 7 in the image of the SDS-PAGE gel for the strep-inhibited α-C45A/L158C MoFe protein was cut and reshuffled after analysis, in order to clearly present the time-course of the experiment (due to the sample loading sequence on the gel). 1.5-1.75 µg of protein per well.

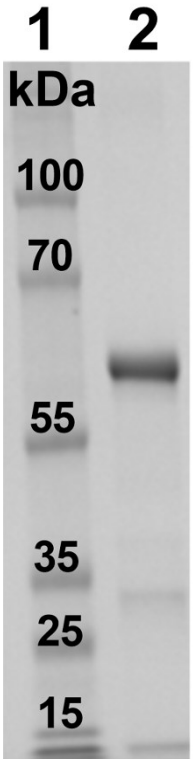
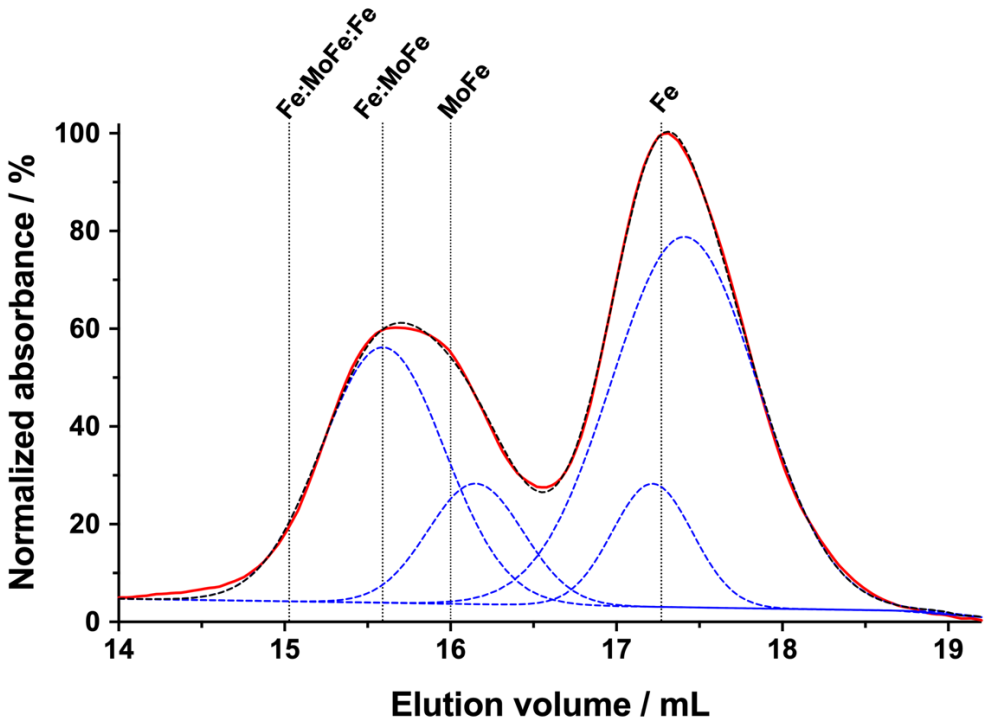


Figure S14: SDS-PAGE analysis of Strep-inhibited α-C45A/L158C MoFe protein was incubated with 5 molar equivalents of L127Δ Fe protein further purified by His Trap column. Lanes: 1 = molecular weight markers, 2 = Strep-inhibited α-C45A/L158C MoFe protein was incubated with 5 molar equivalents of L127Δ Fe protein further purified by His Trap column. 2 μg of protein per well.



Fitted Peak / mL	Max. Peak Height (%)
15.59	52.3
16.15	24.7
17.22	25.2
17.41	75.8

Figure S15: Multi-Gaussian peak fitting (performed using Origin Pro) of the gel filtration profile obtained for strep-reacted α-C45A/L158C MoFe protein (half-inhibited MoFe protein targeted, as detailed in the experimental section) after incubation with 5 molar equivalents of L127Δ Fe protein, extracted from Figure 5D in the main article. The experimental profile is shown in red, whereas the individual fitted peaks are shown in blue dashed lines and the cumulative predicted profile is shown by the black dashed line. The vertical lines indicating the elution volumes of the four protein components were extracted from Figure 5D in the main article (the apparent peaks in the red trace).

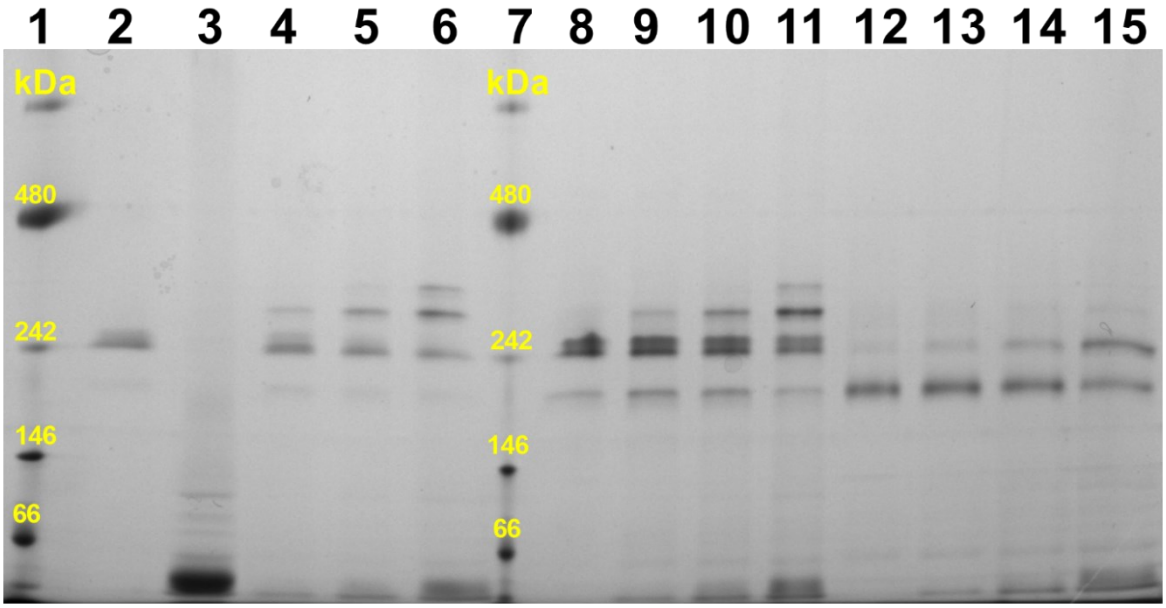


Figure S16: Anoxic Native PAGE 3-12%, 150 V. Lanes: 1,7 = molecular weight markers, 2 = WT MoFe protein, 3 = L127Δ Fe protein, 4 = WT MoFe + L127Δ Fe (1:1), 5 = WT MoFe + L127Δ Fe (1:2), 6 = WT MoFe + L127Δ Fe (1:5), 8 = α-C45A/L158C MoFe, 9 = α-C45A/L158C MoFe + L127Δ Fe (1:1), 10 = α-C45A/L158C MoFe + L127Δ Fe (1:2), 11 = α-C45A/L158C MoFe + L127Δ Fe (1:5), 12 = Strep-inhibited α-C45A/L158C MoFe, 13 = Strep-inhibited α-C45A/L158C MoFe + L127Δ Fe (1:1), 14 = Strep-inhibited α-C45A/L158C MoFe + L127Δ Fe (1:2), 15 = Strep-inhibited α-C45A/L158C MoFe + L127Δ Fe (1:5).

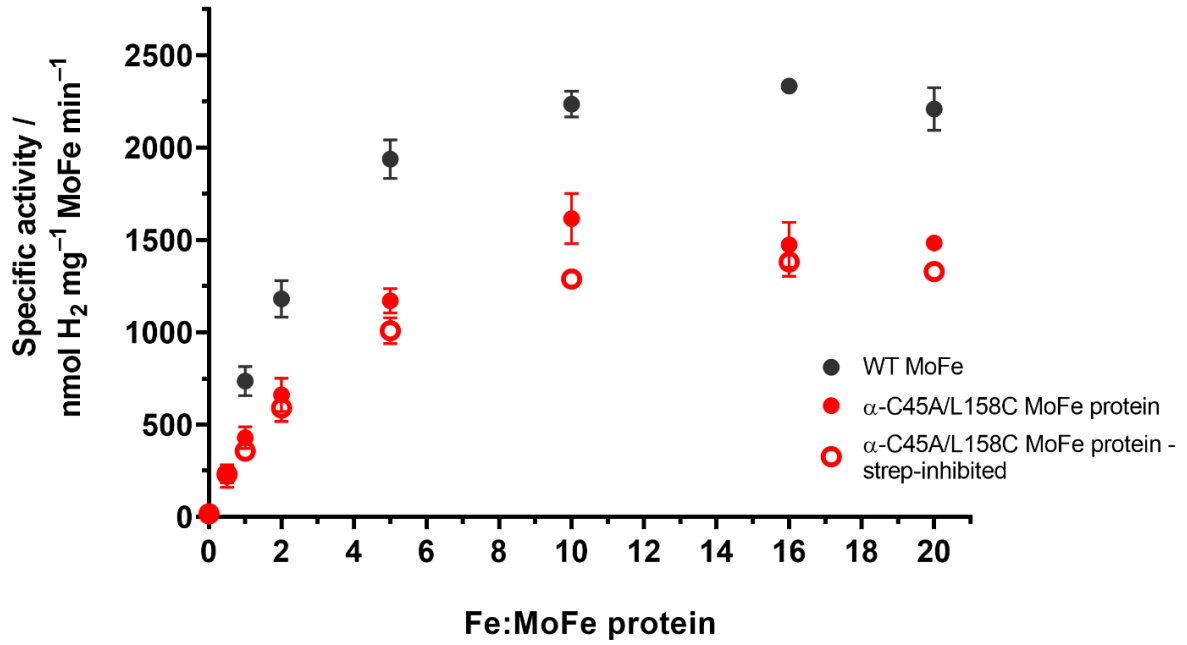


Figure S17: Fe protein titration. Specific activities for H_2 evolution under 1 atm Ar. The x axis indicates the molar equivalents of Fe protein per MoFe protein. Black dots = WT MoFe protein, red filled dots = α-C45A/L158C MoFe and red empty dots = strep-inhibited α-C45A/L158C MoFe. *n* = 3, error bars report standard deviation. All assays contained 0.1 mg mL^{-1} MoFe protein and were performed for 8 minutes at 30 °C. Note = the specific activities of the α-C45A/L158C MoFe proteins (native and strep-inhibited) were performed using protein isolated from a large-scale *A. vinelandii* culture (single biological repeat).

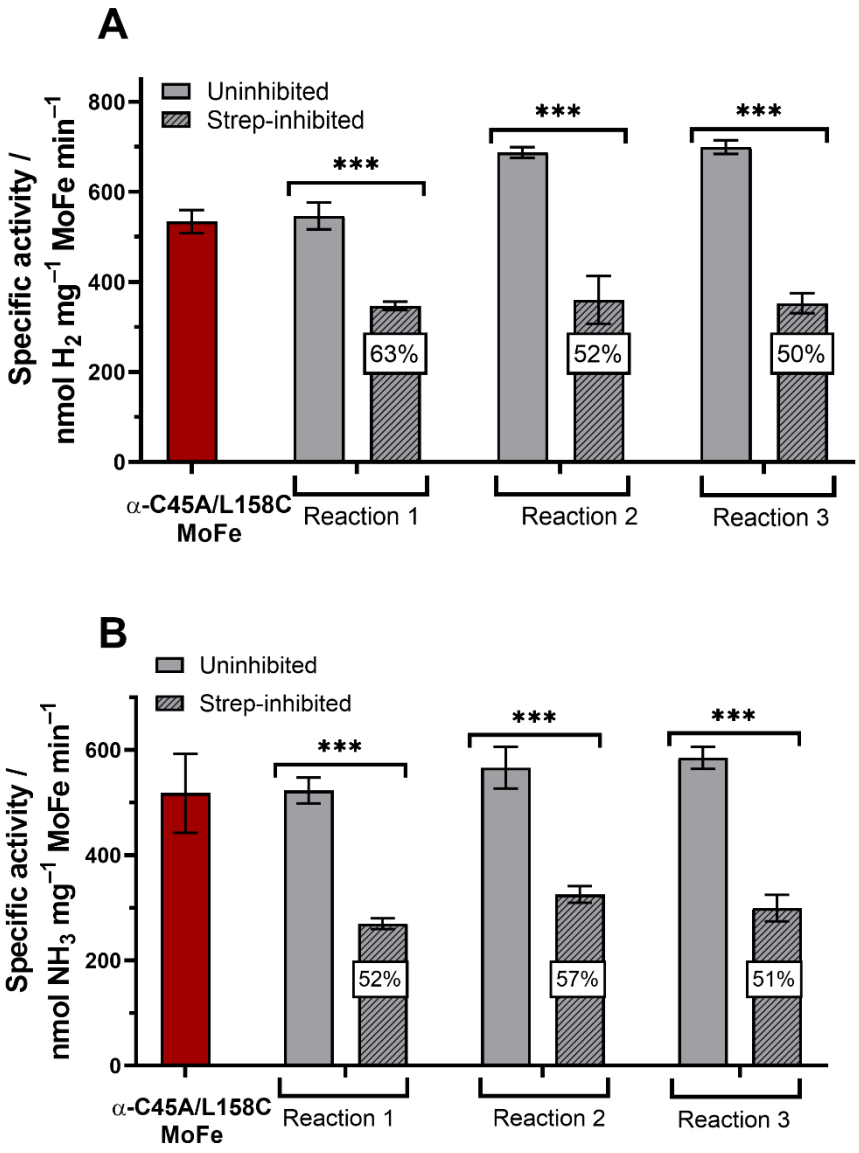


Figure S18: (**A**) Specific activity for H_2 evolution under 1 atm N_2 atmosphere. Red is the standard α-C45A/L158C MoFe purified in the absence of DT. Light grey represents the flow-through (uninhibited) MoFe proteins from three functionalization reactions, grey cross-hatched bars represent the strep-inhibited α-C45A/L158C MoFe. Percentage value in white square corresponds to the remaining specific activity of each strep-inhibited MoFe divided in relation to its unreacted MoFe (flow-through) fraction. (**B**) Specific activity for NH_3 evolution under 1 atm N_2 atmosphere. All activity assays were performed for 8 minutes at 30 °C with 0.1 mg mL^{-1} MoFe protein and 16.6 molar equivalents of Fe protein (0.48 mg). (**A-B**) $n = 3$ and error bars report standard deviation (propagated where necessary).

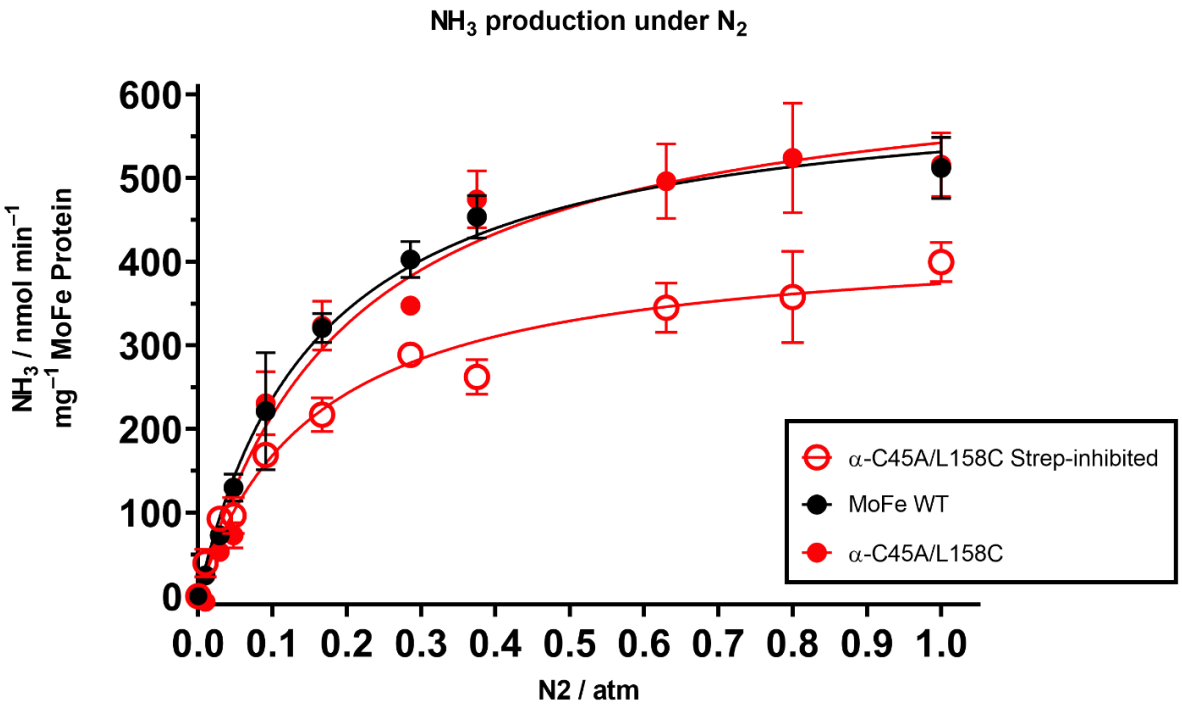


Figure S19: Specific activity of MoFe proteins *vs*. N_2 partial pressure (nmol NH_3 min^{-1} mg^{-1} MoFe). Black data: WT MoFe; red data: uninhibited α-C45A/L158C MoFe; grey data: strep-inhibited α-C45A/L158C MoFe. n = 3, error bars report standard deviation. All assays contained 0.1 mg mL^{-1} MoFe protein and were performed for 8 minutes at 30 °C.

Table S5: Michaelis-Menten kinetic parameters for WT MoFe protein.

Michaelis-Menten parameter	**WT MoFe**
K_M^{app} (atm)	0.16 ± 0.03
V_{max}^{app} (nmol NH_3 $min^{-1}.mg^{-1}$)	616.8 ± 41.6

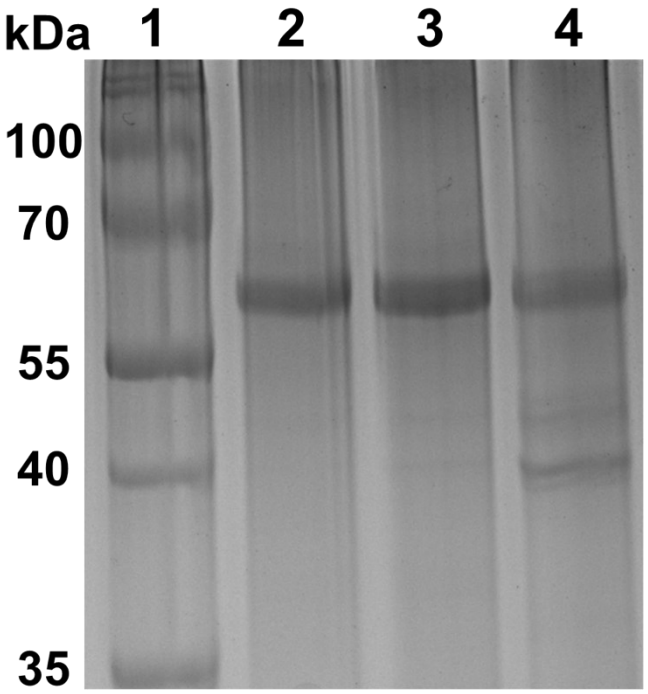


Figure S20: SDS-PAGE (10%, Coomassie blue stained), 180 V. Lanes: 1 = molecular weight markers, 2 = α-C45S MoFe protein, 3 = unreacted α-C45S MoFe protein, 4 = Strep-inhibited α-C45S MoFe. 400 ng of protein per well.

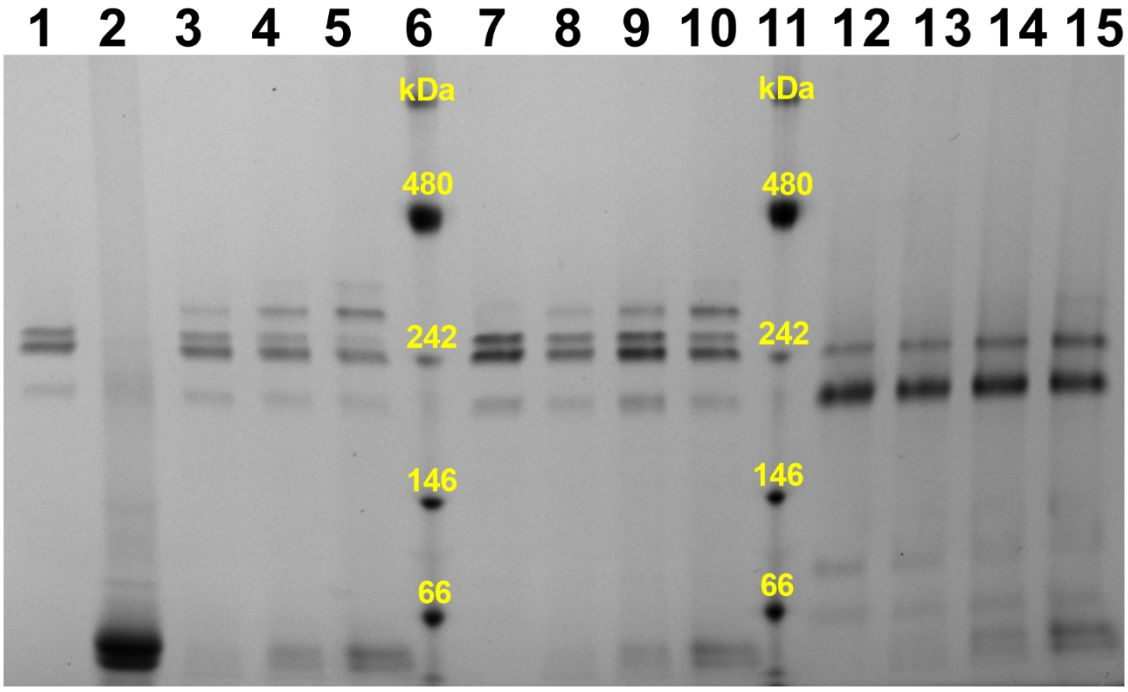


Figure S21: Anoxic Native PAGE 3-12%, 150 V. Lanes: 1 = WT MoFe protein, 2 = L127Δ Fe protein, 3 = WT MoFe + L127Δ Fe (1:1), 4 = WT MoFe + L127Δ Fe (1:2), 5 = WT MoFe + L127Δ Fe (1:5), 6,11 = molecular weight markers, 7 = α-C45S MoFe, 8 = α-C45S MoFe + L127Δ Fe (1:1), 9 = α-C45S MoFe + L127Δ Fe (1:2), 10 = α-C45S MoFe + L127Δ Fe (1:5), 12 = Strep-inhibited α-C45S MoFe, 13 = Strep-inhibited α-C45S MoFe + L127Δ Fe (1:1), 14 = Strep-inhibited α-C45S MoFe + L127Δ Fe (1:2), 15 = Strep-inhibited α-C45S MoFe + L127Δ Fe (1:5).

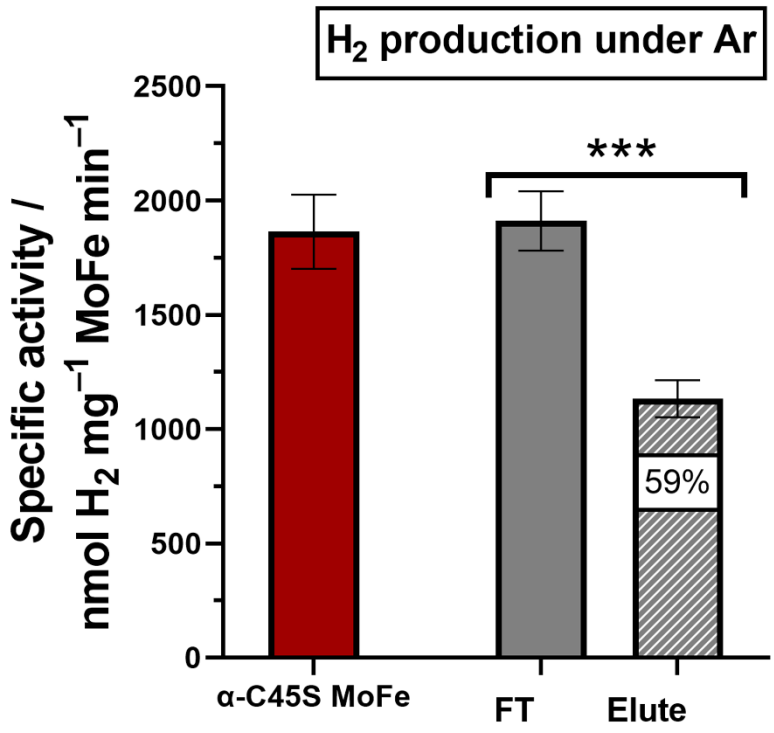


Figure S22: Specific activity for H_2 evolution under 1 atm Ar atmosphere. Red is the standard α-C45S MoFe purified in the absence of DT. Light grey represents the flow-through (uninhibited) MoFe proteins from functionalization reaction, grey cross-hatched bars represent the strep-inhibited α-C45S MoFe. Percentage value in white square corresponds to the remaining specific activity of strep-inhibited MoFe divided in relation to its unreacted MoFe (flow-through) fraction. All activity assays were performed for 8 minutes at 30 °C with 0.1 mg mL^{-1} MoFe protein and 16.6 molar equivalents of Fe protein (0.48 mg). n = 3 and error bars report standard deviation (propagated where necessary).

References

(1) Milton, R. D.; Cai, R.; Abdellaoui, S.; Leech, D.; De Lacey, A. L.; Pita, M.; Minteer, S. D. Bioelectrochemical Haber–Bosch Process: An Ammonia-Producing H2/N2 Fuel Cell. *Angew. Chemie - Int. Ed.* **2017**, *56* (10), 2680–2683. https://doi.org/10.1002/anie.201612500.

(2) Roth, L. E.; Tezcan, F. A. ATP-Uncoupled, Six-Electron Photoreduction of Hydrogen Cyanide to Methane by the Molybdenum–Iron Protein. *J. Am. Chem. Soc.* **2012**, *134* (20), 8416–8419. https://doi.org/10.1021/ja303265m.

(3) Puri, A. W.; Owen, S.; Chu, F.; Chavkin, T.; Beck, D. A. C.; Kalyuzhnaya, M. G.; Lidstrom, M. E. Genetic Tools for the Industrially Promising Methanotroph Methylomicrobium Buryatense. *Appl. Environ. Microbiol.* **2015**, *81* (5), 1775–1781. https://doi.org/10.1128/AEM.03795-14.

(4) Gu, W.; Ul Haque, M. F.; Baral, B. S.; Turpin, E. A.; Bandow, N. L.; Kremmer, E.; Flatley, A.; Zischka, H.; DiSpirito, A. A.; Semrau, J. D. A TonB-Dependent Transporter Is Responsible for Methanobactin Uptake by Methylosinus Trichosporium

OB3b. *Appl. Environ. Microbiol.* **2016**, *82* (6), 1917–1923. https://doi.org/10.1128/AEM.03884-15.

(5) Schäfer, A.; Tauch, A.; Jäger, W.; Kalinowski, J.; Thierbach, G.; Pühler, A. Small Mobilizable Multi-Purpose Cloning Vectors Derived from the Escherichia Coli Plasmids PK18 and PK19: Selection of Defined Deletions in the Chromosome of Corynebacterium Glutamicum. *Gene* **1994**, *145* (1), 69–73. https://doi.org/10.1016/0378-1119(94)90324-7.

(6) Christiansen, J.; Goodwin, P. J.; Lanzilotta, W. N.; Seefeldt, L. C.; Dean, D. R. Catalytic and Biophysical Properties of a Nitrogenase Apo-MoFe Protein Produced by a NifB-Deletion Mutant of Azotobacter Vinelandii. *Biochemistry* **1998**, *37* (36), 12611–12623. https://doi.org/10.1021/bi981165b.

(7) Ryle, M. J.; Seefeldt, L. C. Elucidation of a MgATP Signal Transduction Pathway in the Nitrogenase Iron Protein: Formation of a Conformation Resembling the MgATP-Bound State by Protein Engineering. *Biochemistry* **1996**, *35* (15), 4766–4775. https://doi.org/10.1021/bi960026w.

(8) Lanzilotta, W. N.; Fisher, K.; Seefeldt, L. C. Evidence for Electron Transfer from the Nitrogenase Iron Protein to the Molybdenum-Iron Protein without MgATP Hydrolysis: Characterization of a Tight Protein-Protein Complex. *Biochemistry* **1996**, *35* (22), 7188–7196. https://doi.org/10.1021/bi9603985.

(9) Jimenez-Vicente, E.; Yang, Z. Y.; Keith Ray, W.; Echavarri-Erasun, C.; Cash, V. L.; Rubio, L. M.; Seefeldt, L. C.; Dean, D. R. Sequential and Differential Interaction of Assembly Factors during Nitrogenase MoFe Protein Maturation. *J. Biol. Chem.* **2018**, *293* (25), 9812–9823. https://doi.org/10.1074/jbc.RA118.002994.

(10) Milton, R. D.; Abdellaoui, S.; Khadka, N.; Dean, D. R.; Leech, D.; Seefeldt, L. C.; Minteer, S. D. Nitrogenase Bioelectrocatalysis: Heterogeneous Ammonia and Hydrogen Production by MoFe Protein. *Energy Environ. Sci.* **2016**, *9* (8), 2550–2554. https://doi.org/10.1039/c6ee01432a.

(11) Baldwin, A. D.; Kiick, K. L. Tunable Degradation of Maleimide-Thiol Adducts in Reducing Environments. *Bioconjug. Chem.* **2011**, *22* (10), 1946–1953. https://doi.org/10.1021/bc200148v.

(12) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G. Data Processing and Analysis with the AutoPROC Toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 293–302. https://doi.org/10.1107/S0907444911007773.

(13) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, 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. https://doi.org/10.1107/S2059798319011471.

(14) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501. https://doi.org/10.1107/S0907444910007493.

(15) Bricogne, G.; Blanc, E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.; Sharff, A.; Smart, O. S.; Vonrhein, C.; Womack, T. O. Buster. *Buster version 2.10.4* **2017**, No. Cambridge, United Kingdom:, Global Phasing Ltd.

(16) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21. https://doi.org/10.1107/S0907444909042073.

CHAPTER 7: Discussion

At the start of my thesis, the available structural insights on archaeal nitrogenases were limited to a single structure of recombinantly expressed NifH from *M. acetivorans* (152). In contrast, there was an extensive data collection of structural information on bacterial nitrogenases encompassing various turnover states and their complexes with different partners, ligands and inhibitors. In the meantime, this knowledge base has expanded to include a comprehensive characterisation of all three isoforms of bacterial nitrogenases (58, 59, 68, 69).

The primary goal of this thesis was to address the existing knowledge gap by utilising (hyper)thermophilic members of *Methanococcales* as model organisms. This approach unveiled the first-ever structure of an archaeal nitrogenase (**Chapter 3, Chapter 4**). These insights were placed in a physiological context and complemented by in-depth transcriptomic analyses (**Chapter 2**) with a specific emphasis on regulatory mechanisms (**Chapter 4**). Additionally, the power of crystallography and structural biology as methods for providing crucial information about the intricate details of nitrogenase structural features was demonstrated once again. This was accomplished by employing these methods to help answer two long-standing questions in the field: the functionality of the mononuclear metal binding (MMB) site (**Chapter 5**) and nitrogenase cooperativity (**Chapter 6**) using the genetically tractable *A. vinelandii* model.

In this chapter, I will discuss the outcomes and contributions of the research outlined in this thesis, highlighting the significance of the findings and their implications for our understanding of N_2-fixation in methanogens and mechanistic details of nitrogenases.

7.1. Seeing the big picture: transcriptomics lessons from an environmental point of view.

This thesis was initiated by improving the growth conditions of several *Methanococcales* (*i.e.*, *M. thermolithotrophicus*, *M. vulcanius* and *M. infernus*) and adapting them to diazotrophic conditions to establish a strong foundation for the overall projects. Only then we started our molecular investigation. A deep understanding of *M. thermolithotrophicus* physiology, under diazotrophic conditions, was initiated by physiological experiments complemented with the first-ever comparative transcriptomic analysis between diazotrophic and non-diazotrophic *M. thermolithotrophicus* cells (**Chapter 2**).

The differential transcriptomics revealed the presence of at least three distinct strategies that the organism can employ to acquire essential nitrogen: *(i)* nitrogen fixation by nitrogenase (*nif* operon), *(ii)* ammonia uptake through *amtB*$_2$ transporter and assimilation through glutamine synthetase (*glnA*, GS)/glutamate synthase (GOGAT), *(iii)* nitrate import through *narK* transporter and reduction through *narB* and F_{420}-dependent sulfite reductase (*fsr*) (**Chapter 2**) (213). The data mining from this study confirmed the presence of a single molybdenum-dependent nitrogenase which is highly expressed under diazotrophic conditions, as well as a clue why *M. thermolithotrophicus* has lost its ability to feed on nitrate as the sole nitrogen source. Indeed, in contrast to the first report from Belay *et al.* in 1990, showing that *M. thermolithotrophicus* uses nitrate for nitrogen assimilation (214), nitrate-dependent growth could not be achieved under our laboratory conditions, even when reproducing the exact same medium (data obtained by Ramona Appel). We found that the encoded putative nitrate reductase gene of the nitrate-utilizing operon has an internal stop codon that could have disrupted its biological activity, explaining why this archaeon cultivated multiple times under laboratory conditions would have lost this ability compared to the wild type organism.

While laboratory conditions allow for observation of diazotrophic growth under precisely controlled, well-defined parameters, as described and discussed in **Chapter 2**, understanding of N_2-fixation in natural ecosystems such as deep-sea vents or shallow geothermal sediments where *M. thermolithotrophicus* comes from holds great significance from multiple standpoints. Ecologically, mapping the nutrient availability, N_2-fixation, and methane-production rates is crucial to quantify element cycling and budgets. From a bioenergetics standpoint, the ability of these organisms to perform one of the most energy-demanding catalytic processes under suboptimal conditions is truly remarkable and highlights their exceptional adaptation strategies. Nitrate is scarce or completely absent in hydrothermal vent fluids, making it an unfavourable nitrogen source (171). Consequently, it is unlikely that diazotrophic methanogens can depend on nitrate as a nitrogen source in such environments. Ammonium concentrations in vent fluids, deep-sea water and organic-rich anoxic sediments are exceedingly low ($\leq$1 µM), making dissolved N_2 the predominant reservoir of available nitrogen and the most likely nitrogen source for diazotrophic methanogens (171). Typically, organisms resort to nitrogen fixation only when reactive nitrogen species are absent. However, the occurrence of N_2-fixation in the presence of substantial ammonium concentrations has been observed in multiple environmental samples (215, 216), possibly due to the existence of ammonium depleted microniches within these habitats or other determinant factors (217). For example, the availability of carbon and electrons, could play a role in

whether organisms can rely on nitrogen fixation, as hypothesised in the case of ANME-2 isolated from Costa Rican deep-sea methane seep sediments (217). Another reason for N_2-fixation in the presence of ammonium might be that, in specific cases, this metabolic pathway fulfils another role, such as hydrogen production (217).

The availability of trace metals, which are essential components of the nitrogenase enzyme cofactors, represents another crucial ecological determinant of nitrogen fixation (39, 67, 218) and has played a significant role in driving the evolution of the three isoforms of nitrogenase (177-179, 219). In conjunction with this, diazotrophs have developed a sophisticated network of trafficking proteins involved to specifically recognize metals, various intermediates during synthesis, and the metallocofactors themselves (40). The roles, expression, and regulation of Mod, Tup, and Wtp transporters for metal selection during the diazotrophic growth of *M. thermolithotrophicus* are discussed in detail in **Chapter 2** (213). However, it can be speculated whether *nifX*, the sole gene in the *nif* operon of *M. thermolithotrophicus*, whose function remains unknown, might have a similar role. Although *Mt*NifX shares no homology with the homonymous proteins from *M. maripaludis* or *A. vinelandii*, it has a 37.76% identity with one of the two hypothetical proteins present in the *nif* operon of *M. infernus*. Based on AlphaFold models, *Mi*NifX seems to bind on the nitrogenase close to the FeMoco site and might play a role in the cofactor delivery without blocking NifH and $NifI_{1,2}$ binding sites (information obtained from recent analysis done by Tristan Wagner).

The transcriptomic analysis also revealed differences in transcript levels of two separate regions containing putative virus-like and mobile genetic elements. In **Chapter 2**, we hypothesised that respective upregulation and downregulation (213) of these regions under N_2-fixation is part of the systemic stress response triggered by energy and nutrient depletion. Alternatively, it is conceivable that these two regions might be actively involved in mediating environmental stressors and growth promotion and serve as regulatory switches for gene expression since prophages have been described to take on all of these roles in bacteria (220, 221). While partially conversed in other methanogens, further functional characterisation and transcriptomic studies can confirm whether they can be considered a part of the global transcriptional programme for environmental stress response in *Archaea* (222) and thus crucial for their survival of the N_2-fixation imposed stress in nature.

7.2. Zooming in: Revealing the structural details of nitrogenase from a hyperthermophilic methanogen. During diazotrophic growth under laboratory conditions, *M. thermolithotrophicus* fixes N_2 via a molybdenum-dependent nitrogenase, which is

significantly upregulated based on transcriptomic analysis (**Chapter 2**) (213). Since *M. thermolithotrophicus* harbours only one nitrogenase copy in its genome, it can be safely assumed that the same is happening in nature. Moreover, due to the identical metabolism and phylogenetic relation, the closely related *M. infernus* is supposed to react similarly under N_2-fixation conditions.

Until this thesis, knowledge regarding hyperthermophilic nitrogenases was predominantly derived from *in silico* analysis. However, extracting definitive conclusions on distinctive structural or biochemical characteristics from such analysis proved challenging, as can be beautifully illustrated in the example of the three dinitrogenase isoforms from *A. vinelandii*. Indeed, despite the high percentage of sequence similarity and the same physiological function of N_2-fixation (**Chapter 3**), they are not completely interchangeable in terms of biosynthetic machinery and differ in terms of their catalytic capacity (e.g. generation of hydrocarbons) (69). While the dinitrogenase reductase isoforms from *A. vinelandii* and *Methanosarcina acetivorans* have been shown to cross-react with the different dinitrogenase isoforms from *A. vinelandii*, none of these chimeric associations can match the activity of the natural couple, suggesting their co-adaptation during evolution, and raising the question of how the phylogenetically distant reductases from *Methanococcales* would cross-react with the bacterial systems.

Chapter 3 investigated that question by studying the dinitrogenase reductases from *M. maripaludis*, *M. thermolithotrophicus*, and *M. infernus*. With a dimer size of approximately 60 kDa, NifH is larger compared to many other electron transfer proteins that typically consist of a single domain (65). NifH is also interesting in terms of reactivity as it is the only electron carrier known to reach the $[4Fe4S]^0$ redox state harbouring two electrons at a time (65) and differs in terms of reactivity as it has been shown to catalyse the reduction of CO and the generation of hydrocarbons (95, 96). The three structures obtained in this work highlight the structural conservation observed among nitrogenase reductases, particularly the conformation in the MgADP bound state, which holds great importance for studying the turnover state. The P-loop, switch I, and switch II motifs, which undergo nucleotide-state-dependent structural changes and trigger the necessary structural rearrangements enabling NifH to adopt the appropriate conformation and accurately position the [4Fe-4S] cluster for efficient electron transfer to the dinitrogenase (51, 60, 62, 63, 65) are conserved in all three examined archaeal models. While the recombinant NifH construct from *M. maripaludis* still contains a tag that might interfere with the ATP/ADP binding dynamic, the two native NifH structures from *M.*

thermolithotrophicus and *M. infernus* (the first ones from archaea) corroborate and confirm this structural conservation.

In order to fully comprehend intramolecular electron transfer in nitrogenase, it is important to examine NifH in complex with NifDK. The first reported complex of the two was the MgADP-AlF_4^- stabilised complex from *A. vinelandii* (223). Subsequently, additional complexes were characterised crystallographically, reflecting the three different docking geometries: DG1 (in the absence of ligands), DG2 (in the presence of AMPPCP or MgADP-AlF_4), and DG3 (in the presence of MgADP) (51). Based on the exposure and distance of the [4Fe-4S] cluster from the P cluster, only DG2 is considered "active" and relevant for the electron transfer to the dinitrogenase. According to the model proposed by Katz *et al.*, the NifH-NifDK association occurs initially based on electrostatic compatibility (224). Subsequently, a temporal sequence is proposed in which the three docking geometries are populated during turnover, corresponding to the different stages of the NifH cycle (225). In **Chapter 3**, we demonstrated that NifHs from methanogens could form a MgADP-AlF_4^- stabilised complex with *Av*NifDK by utilising migration profiles on native PAGE. Furthermore, our analysis revealed that specific key residues involved in contacting NifDK in different docking geometries are conserved among the NifHs from methanogens. Still, some substitutions might affect the appropriate binding to the bacterial system (**Chapter 3**). From this analysis, it was evident that binding of all three NifHs from methanogens to *Av*NifDK is possible and the electron transfer should happen, a hypothesis that was not corroborated by the cross-compatibility assay, where none of the archaeal NifHs allowed NH_3 generation when combined with the bacterial dinitrogenase. The observed incompatibility under these conditions may be attributed to two potential factors: the assay temperature, since the assay temperature does not accurately reflect the physiological temperature at which hyperthermophilic nitrogenases operate, or the subtle localised surface charge changes. Small variations in local surface charges could lead to repulsive interactions and affect the binding affinity and stability of the complex even though the surface residues are mainly conserved. However, since the observed (in)activities for mesophilic *Mm*Nif did not differ from the (hyper)thermophilic *Mt*NifH and *Mi*NifH, it is plausible that there are other factors, most likely a combination of them, at play.

To investigate the association and interaction of NifHs from methanogens with their natural partner, we focused on the nitrogenase from *M. infernus* since the tentative crystallisation of nitrogenase from *M. thermolithotrophicus* did not yield exploitable crystals. While it is difficult to speculate why the nitrogenase from *M. infernus* was successful in crystallising

compared to the one from *M. thermolithotrophicus*, despite a very similar purification protocol, one might assume that the robustness conferred by the thermostability of *M. infernus* nitrogenase might have played a favourable role in our crystallisation screenings. The obtained MgADP-AlF_4^- stabilised NifH-NifDK complex (**Chapter 4**) exhibits a conformation that corresponds to the DG2 docking geometry (51), with a distance of 17.6 Å between the [4Fe-4S] cluster and the P cluster conducive for electron transfer. Furthermore, this structure has provided confirmation of the involvement of conserved residues that were predicted to form the binding interface based on the analysis presented in **Chapter 3**.

Due to its ATPase function, NifH stands apart from traditional electron shuttles such as ferredoxins, flavodoxins, or cytochromes (65). On the other hand, NifDK also transcends the conventional role of passive electron acceptor observed in most redox catalytic enzymes. Through its intricate surface architecture, NifDK actively dictates the movement and ATPase activity of NifH and assumes an effector role. The transitions from DG1 to DG2 and from DG2 to DG3 of NifH are hypothesised to happen following its unidirectional motion along the asymmetrical NifDK surface, which functions as a guiding scaffold (51, 65). This phenomenon also provides a potential explanation for the relatively stable conformation of NifDK throughout these events. Rather than undergoing significant overall conformational changes, NifDK exhibits localised alterations in residues that connect the three clusters along the electron flow pathway, thus representing conformational gating. The same observation holds true for *Mi*NifDK, as supported by the root mean square deviation value of 0.299 Å for all atoms between the resting state model and the complex formed by *Mi*NifDK with NifH (**Chapter 4)**.

Chapter 4 presents a significant breakthrough as it unveils the first-ever structure of an archaeal nitrogenase that expands our knowledge about the architecture of this enzyme into the third domain of Life. The ($NifD_2K_2$) heterotetrameric organisation mirrors the three bacterial isoforms (70, 75). When compared to them, the deviation is in the same range for all three of them, thus our structural analysis cannot conclusively state which system would be the closest structural homologue. However, the V-nitrogenase is as compact as *Mi*NifDK (68). Albeit similar in size, the V-nitrogenase is still different from *Mi*NifDK as it carries the additional subunit VnfG and the sulphide-substituted carbonate on its metallocofactor that cannot be accommodated due to sterical encumbrance in *Mi*NifDK. The nitrogenase-defining metal clusters, the P cluster and FeMoco, occupy positions consistent with the three characterised bacterial isoforms, while the MMB site is occupied by Mg^{2+} (confirmed by the absence of an anomalous signal at the Fe K edge). Interestingly, this is the second structural

characteristic that bears more resemblance to *Av*VnfDK (68), where the MMB site is also occupied by Mg^{2+} (**Chapter 5**).

An anomalous scattering experiment at the Mo K edge confirmed that the metallocofactor present in *Mi*NifDK indeed contains molybdenum, corroborating the physiological results from *M. thermolithotrophicus* from **Chapter 2**. The snapshot of the archaeal nitrogenase obtained with an alternative conformation of the resting and turnover state might be one of the most compelling results for the field of N_2-fixation. It supports a common mode of reductive catalysis for all nitrogenases involving the opening of the binuclear substrate binding site at Fe2 and Fe6 upon sulfide S2B removal, regardless of their origin (69), which was still under debate for the Mo-nitrogenases (226). First observed in the structure of CO-bound molybdenum nitrogenase from *A. vinelandii*, in which CO replaced the bridging S2B sulfide between Fe2 and Fe6 (227), the S2B displacement was subsequently detected during SeCN reduction by Mo-nitrogenase (228) and during turnover in both V- (71) and Fe- only (69) nitrogenases. In all these examples, a conserved Gln^D residue ($Gln191^D$ in *Av*NifD, $Gln176^D$ in *Av*VnfD, $Gln176^D$ in *Av*AnfD) can assume two alternative conformations, corresponding to the resting and turnover states, respectively. In the resting state conformation, this residue points away from sulfide S2B with its side chain pointing towards the homocitrate ligand occupying the nearby cavity. In the second conformation, the side chain of glutamine is rotated so that S2B can be displaced to the otherwise occupied cavity, with the substrate taking its bridging position. These two conformations of *Mi*$Gln171^D$ are captured in the high-resolution model of *Mi*NifDK, with an N modelled as a bridging ligand in the turnover state (**Chapter 4**). The ligand identity must be confirmed by other technics as quantum mechanics rather argued for a hydroxyl group (229, 230).

During the turnover process, the P cluster, as the central electron transfer relay, undergoes substantial structural and coordination modifications (65, 78, 79) and oscillates between the all ferrous P^N resting state and P^{1+} one electron oxidised state, adding another layer of complexity to the nitrogenase catalytic cycle. Until recently, the P^N and P^{2+} oxidised states were mostly observed in structural models since the direct electron transfer step from the reduced [4Fe-4S] cluster of NifH to the P^{1+}-cluster is extremely fast (78, 79). Only one model of the P^{1+} state was tentatively obtained, but its modelling was difficult due to the low data quality (assumption based on the Rfree of 26.3%, the P cluster geometry and a suspicious mixture of P^{1+} and P^{2+} states (80). In the model of *Mi*NifDK in the mixed state, the P cluster exists as a fully occupied P^{1+} state modelled with two alternative conformations with a respective occupancy of 70 and 30% (**Chapter 4**). An anomalous dataset at the Fe K edge has

been measured very recently to confirm the exact location of Fe, which is already in high confidence from the anomalous signal extracted in the 12.7 keV dataset. In the first P^{1+} conformation, Fe6 is coordinated by $Ser143^{K}$ but lacks the amide nitrogen coordination of Fe5. Conversely, the second form is characterised by breaking the serine coordination at Fe6 while retaining the amide coordination at Fe5 (**Chapter 4**). The major populated state fits nicely with the bioinformatics prediction, favouring the Ser-bound state rather than amide-bound (80). In comparison, the full resting state model of *Mi*NifDK (referred to as monoclinic form in the manuscript) presents the P^{2+} state, characteristic for natively isolated nitrogenases, in which Fe5 is coordinated by a main chain amide nitrogen of $Cys72^{D}$ and Fe6 is coordinated by $Ser143^{K}$ (**Chapter 4**).

All mentioned results not only support existing hypotheses regarding the enzyme's mechanism but also establish a valuable framework for the future collection and integration of spectroscopic experiments. The high-quality electron density maps of our models will be particularly valuable for quantum mechanics calculations to depict with an even clearer view the states and geometry of nitrogenases´ metallocofactors during catalytic turnover. Understanding the electron transfer mechanism between nitrogenase subunits is vital in elucidating the overall catalytic process. The structure of the P^{1+} state holds particular significance as it is the link between the observed structural changes of NifH and the proposed conformationally gated electron transfer model (65, 79). The transition of NifH from DG1 to DG2 could trigger changes within NifDK around the P cluster cofactor, "priming" it for accepting electrons during the catalytic cycle. Thus, the capture of the P^{1+} state in the first identified archaeal nitrogenase is of exceptional importance.

Since both the [4Fe-4S] cluster of *Mi*NifH and the P cluster of *Mi*NifDK can exist as two different redox couples (**Chapter 4**), this raises an intriguing question regarding the feasibility of two-electron transfer *in vivo*. This plausible phenomenon holds particular significance, especially in the context of methanogenic diazotrophs, which face severe energy limitations. The ability of nitrogenase to facilitate two-electron transfer would be highly advantageous for these organisms as it would enhance their energy efficiency and enable them to optimise their metabolic processes despite the energy constraints they encounter.

7.3. Thriving against the odds: Regulation as the key to nitrogen fixation under energy limitation. Transcriptomic analysis of diazotrophic *M. thermolithotrophicus* revealed that its main energy-saving strategy during the transition to diazotrophy involves a targeted overall reduction of the transcription and translation processes (**Chapter 2**) (213). This is caused by

the downregulation of RNA polymerase subunits, the sigma factor 70, ribosomal proteins, rRNA, and genes with a putative function in tRNA/ribosome biogenesis and biosynthesis of nucleotide and amino acid precursors (**Chapter 2**) (213). Given that in growing cells, protein synthesis and ribosome production account for most nutrient and energy consumption, it becomes evident that these processes are prime targets for energy-saving measures (231). Nevertheless, similar changes in transcript levels of transcriptional and translational genes in *A. vinelandii* (138) and *M. mazei* (232) were not observed, indicating that the energy budget in *M. thermolithotrophicus* and other hyperthermophilic methanogens may be exceptionally constrained due to the hydrogenotrophic methanogenesis lifestyle.

To ensure efficient allocation of resources, prevent wasteful energy expenditure, and effectively balance nitrogen acquisition with energy conservation, intricate regulatory mechanisms tightly control the expression and activity of nitrogenase. As in many other parts of microbial metabolism, the nitrogenase is controlled by various regulatory layers, including transcriptional and post-translational regulation (40, 189). *A. vinelandii*, *M. mazei, and M. maripaludis* all prominently rely on the regulation of nitrogen metabolism at the transcriptional level (40, 200-202, 205, 206, 233), as described in detail in **Chapters 1 and 2**. However, the transcriptional regulation of metabolic enzymes is often driven by global physiological parameters (for example, the sensing of cellular carbon and oxygen status by NifAL in *A. vinelandii*), and the inherent stochasticity of transcription and translation events does not allow for precise fine-tuning of enzymatic activities (234). Consequently, emerging evidence suggests that cells do not rely on transcriptional regulation to intricately modulate enzymes´ activity but rather (234) use directly post-translational regulation, which offers greater flexibility and advantages.

In the case of *M. thermolithotrophicus*, the transition to diazotrophy did not show any upregulation of *nrpA*, the *nif* promotor-specific activator, known to be upregulated upon nitrogen limitation in methanogen (**Chapter 2**) (213). One plausible explanation could be a constitutive expression, as proposed in **Chapter 2**. Alternatively, given the stringent energy constraints in these organisms, post-translational mechanisms would offer an immediate response to synchronizing nitrogenase activity to the physiological needs, and more importantly, a reversible regulation of the existing pool of nitrogenase would provide substantial energy savings compared to *de novo* synthesis of such complex machinery.

In **Chapter 4**, a molecular description of the post-translational regulation systems, originally proposed by Leigh´s group (189, 197, 198, 206), is presented. The reversible inhibition

mechanism builds upon insights derived from the crystallographic analysis of the *Mi*NifDK-$NifI_{1,2}$ complex and the structures of *Mt*$NifI_{1,2}$ complexes in both apo and holo states. The high efficiency of nitrogenase inhibition mediated by $NifI_{1,2}$ is achieved through the formation of a circular closed complex in which three $NifI_{1,2}$ heterohexamers serve as bridging elements connecting three molecules of nitrogenase (**Chapter 4**). The relative simplicity of the circular closed complex achieves the inaccessibility of NifH docking surfaces on both sites of the dinitrogenase using fewer $NifI_{1,2}$ complexes. Unlike systems that rely on post-translational modifications (e.g. ribosylation), which require multiple enzymes for the addition and removal of covalently bound groups (235), here, the presence of activated $NifI_{1,2}$ alone is sufficient to physically block the nitrogenase complex formation. This switch off mechanism through NifI1,2, which responds to 2-OG and ATP levels integrating the signals about cellular nutrient and energy status is as simple and elegant, as it is efficient.

Despite the regulatory networks that control cellular responses to nutrient availability being extensive and complicated, even in microorganisms with a relatively small genome (e.g., below 1.8 megabase pair) (231), the differential transcriptomic did not reveal any other potential partners that could take on an additional regulatory role in *M. thermolithotrophicus* (**Chapter 2**) (213). In addition, the expression levels of other P_{II} proteins, except for $GlnK_2$ during growth on NH_4Cl, remained unchanged. Although an equivalent analysis of *M. infernus* is not available, the archaeon possesses additional genes within its *nif* operon with unknown functions. It is plausible that these proteins may play roles in fine-tuning the regulation of nitrogenase activity or other processes, such as the biosynthesis of the MoFe-cofactor, as *nifN* is missing.

7.4. Through the evolutionary lens: hydrogenotrophic methanogens as a window to ancestral nitrogenase systems. Considering the compelling evidence supporting the notion that nitrogenase evolved in hydrogenotrophic methanogens (137, 178, 179, 219) and that these organisms possess the most “primitive” and reduced genetic sets of *nif* operons (40), studying them holds significant value in terms of representing the ancestral system before the expansion and diversification of nitrogenase over time.

A particularly striking observation is the singularity of the known biological fixation pathway (236). While at least seven distinct pathways are known for carbon fixation (237), the modern nitrogen fixation pathway stands out as the sole recognised mechanism. This discrepancy raises intriguing questions regarding the mechanisms underlying the recruitment of structural

domains and regulatory networks, the selective pressures that have influenced the metal dependence of nitrogenases, and the establishment of the enzymatic mechanism (236, 237).

The recent study in which ancestral nitrogenase was resurrected within *A. vinelandii* suggests that the core strategy for biological nitrogen fixation was conserved across the investigated evolutionary timeline and that life's available strategies for N_2-fixation may be fundamentally limited (236). Simultaneously, this constraint imposed by the imperative preservation of the nitrogenase mechanism may have served as the primary limitation for the evolution of nitrogenase despite varying selective pressures, including environmental factors, ecological diversification, and even the occurrence of V- and Fe-only isoforms (236). Alternatively, it is plausible to consider that alternative strategies for biological nitrogen fixation may have emerged early in the evolutionary history of life but were subsequently outcompeted or rendered obsolete, although the evidence for the existence of these hypothetical alternative mechanisms is not described in any known microbial genome. Nevertheless, exploring the potential existence and subsequent decline of these alternative nitrogen fixation strategies could provide valuable insights into the evolutionary dynamics and the factors that lead to the dominance and preservation of modern nitrogenase as the prevailing mechanism for biological nitrogen fixation. The remarkable conservation of the N_2 reduction mechanism throughout nitrogenase evolutionary history and diversification of diazotrophs suggests that life indeed may have been constrained in optimising the most energetically challenging catalysis in nature.

This diversification was followed by *nif*'s expansion operons to accommodate the increased productivity under aerobic conditions (139). The study, which aimed to correlate the different metabolic strategies with overall *nif* gene content and the phylogenetic distribution of individual *nif* genes in different taxa, revealed that the distributions of *nifA* and *nifL* genes are positively correlated with the ability to use O_2 in metabolism. On the other hand, the distributions of $nifI_1$ and $nifI_2$ genes were inversely correlated with the ability to use O_2 in metabolism suggesting a shift in the primary strategies by which Nif is regulated during the transition from anaerobic ($NifI_{1,2}$) to aerobic (NifAL) metabolism (139). This theory is further supported by the observation that nearly all organisms that lack $nifI_1I_2$ have *nifA*. However, there are examples of taxa in which both (anaerobic Chlorobi) or neither (aerobic cyanobacteria and anaerobic *Geobacter spp.*) $nifI_1I_2$ and *nifA* are found (139).

Contemporary diazotrophs, however, despite their diverse lifestyles and ecological niches, share a fundamental principle of sensing the metabolic state and coordinating their response

to nitrogen limitation (187-189). This common framework with the utilisation of 2-oxoglutarate (2-OG) as a central signalling molecule highlights once more the importance of stringent regulation of nitrogenase activity (**Chapter 4**). Modern nitrogenases have to be discussed and studied in a broader context of the co-evolved network of accessory proteins responsible for biosynthesis, metal trafficking, and oxygen protection among other things. Studies of diazotrophy in hydrogenotrophic methanogens, in which the nitrogenase system might have initially evolved, offer a unique perspective that can yield valuable insights into ancestral nitrogenase systems.

M. infernus presents itself as a captivating model from this perspective, as it showcases an even further reduction of an already reduced *nif* operon found in hydrogenotrophic methanogens. This reduction, also observed in the genomes of other ancient groups, such as basally branching Chloroflexi and Firmicutes, is reflected in the absence of the complete set of *nifBEN* biosynthetic genes and involves the absence of one or both *nifEN* genes (137). *M. infernus* lacks the *nifN* gene, which has traditionally been regarded as essential for FeMoco synthesis (40), both *in vivo* and *in vitro*, leading to questioning its capabilities to express a functional nitrogenase (137). However, in this thesis we demonstrated that *M. infernus* could express a functional nitrogenase supporting its growth under diazotrophic condition, breaking this dogma and indicating the usage of an alternative and/or simplified pathway for cofactor assembly, as proposed by Garcia *et al* (137). They also argue that it is more probable that the organism in which the last common nitrogenase ancestor was present did not possess *nifEN* genes and had such pathway, than *nifEN* genes being later lost in multiple lineages (137). In this case, the ancient nitrogenases might have relied on a FeMoco precursor synthesized only by NifB for catalysis. This type of cofactor synthesis pathway, representing a transition to the canonical FeMoco biosynthetic pathway may still be present in organisms such as *M. infernus*, offering us a window into the evolutionary past of nitrogenases. Moreover, although today all three nitrogenase isoforms are encoded by separate operons (67), it is possible that ancient nitrogenases were generalists concerning their metal usage (177), and preliminary results point that nitrogenase from *M. infernus* might behave in the similar way. Namely, despite encoding only one molybdenum-dependent nitrogenase, it seems that *M. infernus* can grow diazotrophically in the absence of molybdenum, but with vanadium ad iron in the medium (personal communication and preliminary results obtained from Ramona Appel and Mustafa Törer). By investigating organisms such as *M. infernus*, we can gain a better understanding of the origins and early evolutionary stages of nitrogenases and better

understand the intertwining of nitrogen fixation with energy conservation and environmental adaptation.

In this context, experimental validation of various features described for the bacterial molybdenum-nitrogenase in phylogenetically distant archaeal homologues is particularly important. For example, while we demonstrated that the MMB site is not involved in the P cluster and FeMoco insertion, catalytic turnover or overall stability of *A. vinelandii* molybdenum-nitrogenase (**Chapter 5**), we hypothesised that it might affect the intracellular stability of the nitrogenase under fluctuating conditions. Confirming whether this site has a more prominent role in stabilizing nitrogenase in hyperthermophilic *Methanococcales* with different intracellular composition, and thermotolerance would therefore be very interasting in the context of their environmental fitness. The MMB site might further affect N_2-fixation by having a role in subtle structural changes influencing NifH-NifDK dynamics or cooperativity between the subunits (**Chapter 5**, **Chapter 6** (238)). In **Chapter 6** (238), our study revealed the presence of negative cooperativity in N_2 fixation by *A. vinelandii* molybdenum-nitrogenase. This finding represents a substantial contribution to our understanding of the catalytic mechanism of the enzyme. However, as discussed there, further investigations are required to determine the significance and extent of this negative cooperativity *in vivo*, particularly in the presence of different NifH/NifDK ratios and interacting partners. Additionally, exploring the existence of negative cooperativity in extremely energy-limited hyperthermophilic *Methanococcales* could provide valuable insights into the adaptation strategies and energy saving mechanisms of these organisms.

7.5. The journey does not end here: future directions in the characterisation of archaeal nitrogenases. Just as the initial elucidation of nitrogenase structures from *A. vinelandii* (57, 75) provided the framework for subsequent investigations of the system (65), a comprehensive biochemical and biophysical characterization of the archaeal counterpart, requiring multidisciplinary approach, awaits following the structures obtained in this thesis. By providing the first-ever structural insights into the fascinating system of archaeal nitrogenases (**Chapter 3, Chapter 4**) and the physiological shifts occurring during diazotrophy in thermophilic *Methanococcales* (**Chapter 2**), this thesis established foundations and paves the road for further advances in understanding different aspects of this system, but also raises many questions that appeared over the successive discoveries.

The transcriptomic analysis conducted in **Chapter 2** (213), provided invaluable insights into the specific metabolic alterations occurring in *M. thermolithotrophicus* as a result of change to

diazotrophic lifestyle, but did not yield a definitive answer regarding potential candidates for an electron donor, thereby leaving this question open for further investigation. In **Chapter 2** (213), we hypothesised that the NifH in *M. thermolithotrophicus* could be reduced by a ferredoxin, and that the ferredoxin itself could be reduced by coupling of H_2 oxidation with the sodium ion pumping by Ech complex. The proteomic analysis conducted on nitrogen-limited *M. maripaludis* highlighted the disparity between changes observed at the transcriptome level and those at the proteome level (239). A similar analysis on diazotrophic *M. thermolithotrophicus* could provide a more comprehensive understanding of the protein-level changes and has the potential to pinpoint putative electron donors.

Although all three isoforms of nitrogenase likely share a common mechanism of N_2 reduction (69), they exhibit varying efficiencies in catalytic turnover for both N_2 and other substrates (52, 67). The underlying reasons for these differences have not been definitively established. Consequently, it is plausible to hypothesize that the thermostable archaeal molybdenum nitrogenase may demonstrate different reactivity or utilize different substrates in comparison to its bacterial counterpart. Thus, probing the activity, reactivity towards different substrates and distribution of reducing equivalents between H_2 production and N_2 reduction during turnover is one of the most exciting future goals in characterization of nitrogenases from hyperthemophilic methanogens, as it holds great promise for uncovering new functional diversity and potential applications of these enzymes.

One of the challenges towards this goal would be the setup of a high-temperature activity assay that would allow the testing of specific activities for NH_3 evolution and electron distribution at physiological temperatures of 65°-85°C. The traditional activity assay used for testing the activities of nitrogenases uses mesophilic creatine-phosphokinase from rabbit muscle for ATP regeneration, which cannot withstand higher temperatures. While a preliminary probing of the activity of the three NifH proteins from methanogens was conducted through a cross-reactivity assay with *A. vinelandii* nitrogenase (**Chapter 3**), the measurement of activity rates for the hyperthermophilic nitrogenase reductase and nitrogenase proved to be beyond the scope of this thesis, as the establishment of a thermostable ATP regeneration system is required as a prerequisite before successful measurement of activity rates can be achieved. An alternative thermostable system to regenerate ATP could be: *(i)* pyruvate kinase (phosphoenol-pyruvate + ADP → pyruvate + ATP), *(ii)* polyphosphate kinase (polyphosphate + ADP → polyphosphate-1 + ATP) or *(iii)* acetate kinase (acetyl-phosphate + ADP → acetate + ATP). A thermostable pyruvate kinase could be obtained either from *M. infernus* itself or *Aeropyrum pernix* (240), which is stable at temperatures as high as 95° C.

Polyphosphate kinase from both *Thermosynechococcus elongatus* BP-1 (241) and *Thermus thermophiles* (242) can be recombinantly expressed and could be used for this purpose because they are stable at temperatures up to 70° C. Finally, thermostable acetate kinase from *Thermotoga maritima* was successfully recombinantly purified and showed stable activity for one hundred eighty minutes at 80° C (243). Once established, these thermostable ATP regeneration systems can be integrated into the already implemented robust protocols for the NH_3 production and quantification assays (**Chapter 3**) at increased temperatures suitable for probing the activity of *Mt*/*Mi*NifHDK.

Undoubtedly, the trapping and characterization of various substrate- and intermediate-bound nitrogenase species plays a pivotal role in acquiring further insights. Therefore, having the enzyme readily available in both solution and crystal forms (**Chapter 3**, **Chapter 4**) proves to be of immense value. For instance, employing spectroscopic techniques already used in the study of bacterial nitrogenases (such as ^{57}Fe nuclear resonance vibrational spectroscopy, Fe-edge extended X-ray absorption fine structure spectroscopy, and density functional theory calculations) would allow for the assessment of different oxidation states and redox thermodynamics of the cofactors (65), while their reduction potential(s) could be determined in the electrochemical experiments (244).

The establishment of robust crystallization protocols for nitrogenase from hyperthermophilic methanogens, as detailed in **Chapter 3** and **Chapter 4** of this thesis, presents a significant opportunity for co-crystallization studies. By co-crystallizing nitrogenase with various potential substrates and intermediates, valuable insights can be gained regarding their binding sites and the underlying catalytic mechanism. Furthermore, soaking the crystals with different redox agents allows for the identification and characterization of distinct redox states of the enzyme and its cofactors as well as redox-triggered conformational changes, all of which are crucial for understanding the electron transfer processes and the catalytic cycle of nitrogenase.

The rapid advancement of crystallographic techniques also presents unparalleled opportunities for further acquirement of structural data. Microcrystals, like ones obtained during our crystallization trials, can be exploited by using hard X-ray free-electron lasers (XFELs) (245), which facilitate the data collection from a stream of small crystals. This approach is particularly promising in the case of large/flexible protein complexes, which tend to crystallize poorly yielding small crystals (246, 247) and can be applied for crystals of nitrogenase itself or nitrogenase in complex with different partners. Meanwhile, the new generation of time-resolved X-ray crystallographic techniques permits measuring the

dynamics of proteins and visualization of conformational changes at high temporal and spatial resolutions, adding a temporal dimension to X-ray crystallography experiments (248), a prompt particularly attractive for following for example *Mi*NifDK during turnover.

Furthermore, the high-resolution structure of *Mi*NifDK in mixed state obtained in **Chapter 4** is a promising starting point for the charactization of its FeMoco electronic structure, as successfully done for the FeVco through computational modelling starting from the 1.35 Å crystal structure (249).

All structures presented and discussed in this thesis represent the Mo-dependent nitrogenase isoform and related partners. There are several archaeal species that harbor two or even all three nitrogenase isoforms, yet their three-dimensional structures remain undescribed. The crystallisation of the remaining two nitrogenase isoforms holds great interest as the obtained structures can reveal additional unexpected differences and surprises like the ones that accompanied the structure of bacterial vanadium nitrogenase (68). The pipeline for anaerobic native nitrogenase purification and obtaining high-quality samples for crystallisation, established here (**Chapter 3, Chapter 4**), can be readily adapted to other diazotrophic archaeal model organisms. Crystallisation of native nitrogenases requires substantial protein yield, which in turn calls for large-scale cultivation of the selected organism beforehand. However, for certain archaeal diazotrophic model organisms such as *Methanosarcina*, this task is feasible as they are already being successfully cultivated in laboratory settings.

Having a robust genetic toolbox similar to the ones already available for *M. maripaludis* and *Methanosarcinales* (191) could have facilitated the previous task, as it would enable not only the homologous expression of the required proteins but also provide opportunities for metabolic engineering. Limited knowledge about plasmid replication mechanisms and endogenous DNA defence systems in thermophilic methanogens poses another barrier to progress in this area. However, the already developed methods for genetic manipulation, for example in *M. jannaschii* (250), show promise in their potential translation to closely related methanogens.The availability of genetic tool systems for *M. jannaschii* holds immense potential and can be used to open exciting new avenues of investigation. For instance, heterologous expression of the *nif* operon from *M. infernus* to further validate its sufficiency for diazotrophic capabilities. Moreover, the targeted mutagenesis of genes for proteins involved in cofactor assembly pathway could shed more light on the potential ancestral features of this pathway in *M. infernus*.

Finally, collecting and integrating the environmental data from the natural habitats of (hyper)thermophilic methanopgens with available physiological data is essential, since although multiple environmental studies have detected *nifH* from thermophilic methanogens in various environments (171-174), indicating their active nitrogen fixation in nature, there is a lack of metadata on the different organic and inorganic nitrogen sources available *in situ*. Quantifying the available nutrients that condition the utilisation of different uptake strategies, such as ammonia import versus nitrogen fixation, would greatly enhance our understanding of the role, activity, and fitness of these methanogens in their respective niches. Despite the known presence of methanogens capable of N_2 fixation in diverse environments, the contribution and ecological significance of methanogenic N_2 fixation for the global N cycle is not very well characterized and requires further study.

Understanding the nitrogen-fixing capabilities and determining the nitrogen fixation rates of *M. infernus* and *M. thermolithotrophicus* within their natural habitats is of utmost importance in estimating their contribution to overall nitrogen dynamics in their ecosystems. In this thesis, both organisms were cultivated under the conditions favouring high cell yields and robust growth under diazotrophic conditions, but not reflecting the environmental conditions. A laboratory set up which would mimic these conditions (e.g. in a bioreactor) could allow for such measurments and enhance our understanding of their adaptation and survival strategies in nature.

7.6. Concluding remarks. My thesis centred around the structural characterisation of nitrogen fixation machinery in (hyper)thermophilic methanogens and their ability to seamlessly integrate this highly demanding pathway with the rest of their metabolism under laboratory conditions. The transcriptomic analysis of *M. thermolithotrophicus* revealed simple, yet highly effective strategies employed by this organism to cope with energy constraints and nutrient limitations during diazotrophy. This analysis also facilitated the confirmation of crucial protein players within nitrogen fixation network. I described the first-ever structure of an archaeal molybdenum-containing nitrogenase, its reductase, and its regulatory partner $NifI_{1,2}$ as stand-alone and as complexes. Their structural examination revealed the molecular basis of this peculiar post-translational regulation organized as a sequestration mechanism. The atomic resolution structure of the nitrogenase from a hyperthermophile and its potential could be a game-changer model to decipher in more details how N_2-fixation operates. All further information retrieved from this model study could be transposed to mimicry biochemistry. While nitrogen was originally named *azote*, meaning “no life” due to its perceived unreactivity, not only is it an indispensable component of all

biomolecules, but it also significantly supports modern society through various nitrogen-based compounds that serve as fertilisers, biofuels, pharmaceuticals, industrial polymers, and catalysts, among other applications. As such, it is set to remain a key topic of very much lively research in chemistry, biology, ecology, and chemical and environmental engineering both in the present day and in the future.

The List of references for Introduction and Discussion

1. Galloway JN, Cowling EB. 2002. Reactive Nitrogen and The World: 200 Years of Change. AMBIO: A Journal of the Human Environment 31:64-71, 8.
2. Galloway JN, Townsend AR, Erisman JW, Bekunda M, Cai Z, Freney JR, Martinelli LA, Seitzinger SP, Sutton MA. 2008. Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. Science 320:889-892.
3. Erisman JW, Sutton MA, Galloway J, Klimont Z, Winiwarter W. 2008. How a century of ammonia synthesis changed the world. Nature Geoscience 1:636-639.
4. Stevens CJ. 2019. Nitrogen in the environment. Science 363:578-580.
5. Zhang X, Ward BB, Sigman DM. 2020. Global Nitrogen Cycle: Critical Enzymes, Organisms, and Processes for Nitrogen Budgets and Dynamics. Chemical Reviews 120:5308-5351.
6. Tamm CO. 2012. Nitrogen in terrestrial ecosystems: questions of productivity, vegetational changes, and ecosystem stability, vol 81. Springer Science & Business Media.
7. LeBauer DS, Treseder KK. 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology 89:371-379.
8. Matson P, Lohse KA, Hall SJ. 2002. The globalization of nitrogen deposition: consequences for terrestrial ecosystems. Ambio 31:113-119.
9. Kuypers MMM, Marchant HK, Kartal B. 2018. The microbial nitrogen-cycling network. Nature Reviews Microbiology 16:263-276.
10. Smil V. 2004. Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production. MIT press.
11. Galloway JN, Bleeker A, Erisman JW. 2021. The human creation and use of reactive nitrogen: a global and regional perspective. Annual Review of Environment and Resources 46:255-288.
12. Haber F, Van Oordt G. 1905. Über die bildung von ammoniak den elementen. Zeitschrift für anorganische Chemie 44:341-378.
13. Haber F. 1922. Über die darstellung des ammoniaks aus stickstoff und wasserstoff. Naturwissenschaften 10:1041-1049.
14. Haber F. 1923. The history of the ammonia process. Naturwissenschaften 11:339-340.
15. Smil V. 2002. Nitrogen and food production: proteins for human diets. AMBIO: A Journal of the Human Environment 31:126-131.

16. Lehnert N, Dong HT, Harland JB, Hunt AP, White CJ. 2018. Reversing nitrogen fixation. Nature Reviews Chemistry 2:278-289.
17. Beman JM, Arrigo KR, Matson PA. 2005. Agricultural runoff fuels large phytoplankton blooms in vulnerable areas of the ocean. Nature 434:211-214.
18. Diaz RJ, Rosenberg R. 2008. Spreading dead zones and consequences for marine ecosystems. Science 321:926-929.
19. Rabalais NN, Díaz RJ, Levin LA, Turner RE, Gilbert D, Zhang J. 2010. Dynamics and distribution of natural and human-caused hypoxia. Biogeosciences 7:585-619.
20. Van Breemen N, Burrough P, Velthorst E, Van Dobben H, de Wit T, Ridder T, Reijnders H. 1982. Soil acidification from atmospheric ammonium sulphate in forest canopy throughfall. Nature 299:548-550.
21. Driscoll CT, Lawrence GB, Bulger AJ, Butler TJ, Cronan CS, Eagar C, Lambert KF, Likens GE, Stoddard JL, Weathers KC. 2001. Acidic Deposition in the Northeastern United States: Sources and Inputs, Ecosystem Effects, and Management Strategies: The effects of acidic deposition in the northeastern United States include the acidification of soil and water, which stresses terrestrial and aquatic biota. BioScience 51:180-198.
22. Stevens CJ, Dise NB, Mountford JO, Gowing DJ. 2004. Impact of nitrogen deposition on the species richness of grasslands. Science 303:1876-1879.
23. Bobbink R, Hicks K, Galloway J, Spranger T, Alkemade R, Ashmore M, Bustamante M, Cinderby S, Davidson E, Dentener F, Emmett B, Erisman JW, Fenn M, Gilliam F, Nordin A, Pardo L, De Vries W. 2010. Global assessment of nitrogen deposition effects on terrestrial plant diversity: a synthesis. Ecological Applications 20:30-59.
24. Simkin SM, Allen EB, Bowman WD, Clark CM, Belnap J, Brooks ML, Cade BS, Collins SL, Geiser LH, Gilliam FS, Jovan SE, Pardo LH, Schulz BK, Stevens CJ, Suding KN, Throop HL, Waller DM. 2016. Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States. Proceedings of the National Academy of Sciences of the United States of America 113:4086-4091.
25. Sutton MA, Howard CM, Kanter DR, Lassaletta L, Móring A, Raghuram N, Read N. 2021. The nitrogen decade: mobilizing global action on nitrogen to 2030 and beyond. One Earth 4:10-14.
26. Anas M, Liao F, Verma KK, Sarwar MA, Mahmood A, Chen Z-L, Li Q, Zeng X-P, Liu Y, Li Y-R. 2020. Fate of nitrogen in agriculture and environment: agronomic, eco-

physiological and molecular approaches to improve nitrogen use efficiency. Biological Research 53:47.

27. Kirchman DL. 2020. The First “Dead Zone”. Limnology and Oceanography Bulletin 29:107-109.
28. Bowman WD, Ayyad A, Bueno de Mesquita CP, Fierer N, Potter TS, Sternagel S. 2018. Limited ecosystem recovery from simulated chronic nitrogen deposition. Ecological Applications 28:1762-1772.
29. Nayak-Luke R, Bañares-Alcántara R, Wilkinson I. 2018. “Green” Ammonia: Impact of Renewable Energy Intermittency on Plant Sizing and Levelized Cost of Ammonia. Industrial & Engineering Chemistry Research 57:14607-14616.
30. Sutton MA, Bleeker A, Howard C, Erisman JW, Abrol YP, Bekunda M, Datta A, Davidson E, Vries W, Zhang FS. 2013. Our nutrient world. The challenge to produce more food & energy with less pollution. Global Overview of Nutrient Management.
31. Tanabe Y, Nishibayashi Y. 2021. Comprehensive insights into synthetic nitrogen fixation assisted by molecular catalysts under ambient or mild conditions. Chemical Society Reviews 50:5201-5242.
32. Oldroyd GE, Dixon R. 2014. Biotechnological solutions to the nitrogen problem. Current Opinion in Biotechnology 26:19-24.
33. Buren S, Rubio LM. 2018. State of the art in eukaryotic nitrogenase engineering. FEMS Microbiology Letters 365:fnx274.
34. Bueno Batista M, Dixon R. 2019. Manipulating nitrogen regulation in diazotrophic bacteria for agronomic benefit. Biochemical Society Transactions 47:603-614.
35. Mus F, Crook MB, Garcia K, Garcia Costas A, Geddes BA, Kouri ED, Paramasivan P, Ryu M-H, Oldroyd GE, Poole PS. 2016. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. Applied and Environmental Microbiology 82:3698-3710.
36. Rapson TD, Gregg CM, Allen RS, Ju H, Doherty CM, Mulet X, Giddey S, Wood CC. 2020. Insights into nitrogenase bioelectrocatalysis for green ammonia production. ChemSusChem 13:4856-4865.
37. Ritter SK. 2008. The Haber-Bosch reaction: an early chemical impact on sustainability. Chemical & Engineering News 86.
38. Humphreys J, Lan R, Tao S. 2021. Development and Recent Progress on Ammonia Synthesis Catalysts for Haber–Bosch Process. Advanced Energy and Sustainability Research 2:2000043.

39. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. 2018. Exploring the alternatives of biological nitrogen fixation. Metallomics 10:523-538.
40. Martin del Campo JS, Rigsbee J, Bueno Batista M, Mus F, Rubio LM, Einsle O, Peters JW, Dixon R, Dean DR, Dos Santos PC. 2023. Overview of physiological, biochemical, and regulatory aspects of nitrogen fixation in *Azotobacter vinelandii*. Critical Reviews in Biochemistry and Molecular Biology 57:492-538.
41. Burgess BK, Lowe DJ. 1996. Mechanism of molybdenum nitrogenase. Chemical Reviews 96:2983-3012.
42. Hoffman BM, Lukoyanov D, Dean DR, Seefeldt LC. 2013. Nitrogenase: a draft mechanism. Accounts of Chemical Research 46:587-595.
43. Hoffman BM, Lukoyanov D, Yang ZY, Dean DR, Seefeldt LC. 2014. Mechanism of nitrogen fixation by nitrogenase: the next stage. Chemical reviews 114:4041-4062.
44. Mortenson LE, Thorneley RN. 1979. Structure and function of nitrogenase. Annual Review of Biochemistry 48:387-418.
45. Spiro TG. 1986. Molybdenum enzymes. John Wiley & Sons.
46. Howard JB, Rees DC. 2006. How many metals does it take to fix N_2? A mechanistic overview of biological nitrogen fixation. Proceedings of the National Academy of Sciences 103:17088-17093.
47. Seefeldt LC, Hoffman BM, Dean DR. 2009. Mechanism of Mo-dependent nitrogenase. Annual Review of Biochemistry 78:701-722.
48. Seefeldt LC, Dean DR. 1997. Role of nucleotides in nitrogenase catalysis. Accounts of Chemical Research 30:260-266.
49. Igarashi RY, Seefeldt LC. 2003. Nitrogen fixation: the mechanism of the Mo-dependent nitrogenase. Critical Reviews in Biochemistry and Molecular Biology 38:351-384.
50. Dos Santos PC, Igarashi RY, Lee H-I, Hoffman BM, Seefeldt LC, Dean DR. 2005. Substrate interactions with the nitrogenase active site. Accounts of Chemical Research 38:208-214.
51. Tezcan FA, Kaiser JT, Mustafi D, Walton MY, Howard JB, Rees DC. 2005. Nitrogenase complexes: multiple docking sites for a nucleotide switch protein. Science 309:1377-1380.
52. Seefeldt LC, Yang ZY, Lukoyanov DA, Harris DF, Dean DR, Raugei S, Hoffman BM. 2020. Reduction of Substrates by Nitrogenases. Chemical Reviews 120:5082-5106.

53. Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, Fitzhugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, Dearellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Craig Kuettner H, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, De Macario EC, Ferry JG, Jarrell KF, Jing H, Macario AJL, Paulsen I, Pritchett M, Sowers KR, et al. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Research 12:532-542.

54. Oda Y, Samanta SK, Rey FE, Wu L, Liu X, Yan T, Zhou J, Harwood CS. 2005. Functional Genomic Analysis of Three Nitrogenase Isozymes in the Photosynthetic Bacterium *Rhodopseudomonas palustris*. Journal of Bacteriology 187:7784-7794.

55. Schneider K, Gollan U, Dröttboom M, Selsemeier-Voigt S, Müller A. 1997. Comparative Biochemical Characterization of the Iron-Only Nitrogenase and the Molybdenum Nitrogenase from *Rhodobacter Capsulatus*. European Journal of Biochemistry 244:789-800.

56. Thiel T. 1993. Characterization of genes for an alternative nitrogenase in the cyanobacterium *Anabaena variabilis*. Journal of Bacteriology 175:6276-6286.

57. Georgiadis M, Komiya H, Chakrabarti P, Woo D, Kornuc J, Rees D. 1992. Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. Science 257:1653-1659.

58. Rohde M, Trncik C, Sippel D, Gerhardt S, Einsle O. 2018. Crystal structure of VnfH, the iron protein component of vanadium nitrogenase. Journal of Biological Inorganic Chemistry 23:1049-1056.

59. Trncik C, Müller T, Franke P, Einsle O. 2022. Structural analysis of the reductase component AnfH of iron-only nitrogenase from *Azotobacter vinelandii*. Journal of Inorganic Biochemistry 227:111690.

60. Einsle O, Rees DC. 2020. Structural Enzymology of Nitrogenase Enzymes. Chemical Reviews 120:4969-5004.

61. Robson R. 1984. Identification of possible adenine nucleotide-binding sites in nitrogenase Fe-and MoFe-proteins by amino acid sequence comparison. FEBS letters 173:394-398.

62. Koonin EV. 1993. A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. Journal of Molecular Biology 229:1165-1174.

63. Leipe DD, Wolf YI, Koonin EV, Aravind L. 2002. Classification and evolution of P-loop GTPases and related ATPases. Journal of molecular biology 317:41-72.
64. Bange G, Sinning I. 2013. SIMIBI twins in protein targeting and localization. Nature Structural & Molecular Biology 20:776-780.
65. Rutledge HL, Tezcan FA. 2020. Electron Transfer in Nitrogenase. Chemical Reviews 120:5158-5193.
66. Burén S, Jiménez-Vicente E, Echavarri-Erasun C, Rubio LM. 2020. Biosynthesis of Nitrogenase Cofactors. Chemical Reviews 120:4921-4968.
67. Jasniewski AJ, Lee CC, Ribbe MW, Ribbe MW, Hu Y. 2020. Reactivity, Mechanism, and Assembly of the Alternative Nitrogenases. Chemical Reviews 120:5107-5157.
68. Sippel D, Einsle O. 2017. The structure of vanadium nitrogenase reveals an unusual bridging ligand. Nature Chemical Biology 13:956-960.
69. Trncik C, Detemple F, Einsle O. 2023. Iron-only Fe-nitrogenase underscores common catalytic principles in biological nitrogen fixation. Nature Catalysis doi:10.1038/s41929-023-00952-1.
70. Einsle O, Tezcan FA, Andrade SL, Schmid B, Yoshida M, Howard JB, Rees DC. 2002. Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. Science 297:1696-1700.
71. Sippel D, Rohde M, Netzer J, Trncik C, Gies J, Grunau K, Djurdjevic I, Decamps L, Andrade SLA, Einsle O. 2018. A bound reaction intermediate sheds light on the mechanism of nitrogenase. Science 359:1484-1489.
72. Rohde M, Grunau K, Einsle O. 2020. CO Binding to the FeV Cofactor of CO-Reducing Vanadium Nitrogenase at Atomic Resolution. Angewandte Chemie International Edition 59:23626-23630.
73. Rohde M, Laun K, Zebger I, Stripp ST, Einsle O. 2021. Two ligand-binding sites in CO-reducing V nitrogenase reveal a general mechanistic principle. Science Advances 7:eabg4474.
74. Zhang L, Kaiser JT, Meloni G, Yang K-Y, Spatzal T, Andrade SLA, Einsle O, Howard JB, Rees DC. 2013. The Sixteenth Iron in the Nitrogenase MoFe Protein. Angewandte Chemie International Edition 52:10529-10532.
75. Kirn J, Rees D. 1992. Crystallographic structure and functional implications of the nitrogenase molybdenum–iron protein from *Azotobacter vinelandii*. Nature 360:553-560.

76. Kim J, Rees D. 1992. Structural models for the metal centers in the nitrogenase molybdenum-iron protein. Science 257:1677-1682.
77. Peters JW, Stowell MHB, Soltis SM, Finnegan MG, Johnson MK, Rees DC. 1997. Redox-Dependent Structural Changes in the Nitrogenase P-Cluster. Biochemistry 36:1181-1187.
78. Chica B, Ruzicka J, Pellows LM, Kallas H, Kisgeropoulos E, Vansuch GE, Mulder DW, Brown KA, Svedruzic D, Peters JW, Dukovic G, Seefeldt LC, King PW. 2022. Dissecting Electronic-Structural Transitions in the Nitrogenase MoFe Protein P-Cluster during Reduction. Journal of the American Chemical Society 144:5708-5712.
79. Keable SM, Zadvornyy OA, Johnson LE, Ginovska B, Rasmussen AJ, Danyal K, Eilers BJ, Prussia GA, LeVan AX, Raugei S. 2018. Structural characterization of the P^{1+} intermediate state of the P-cluster of nitrogenase. Journal of Biological Chemistry 293:9629-9635.
80. Cao L, Borner MC, Bergmann J, Caldararu O, Ryde U. 2019. Geometry and electronic structure of the P-cluster in nitrogenase studied by combined quantum mechanical and molecular mechanical calculations and quantum refinement. Inorganic Chemistry 58:9672-9690.
81. Li Z, Guo S, Sun Q, Chan GK-L. 2019. Electronic landscape of the P-cluster of nitrogenase as revealed through many-electron quantum wavefunction simulations. Nature Chemistry 11:1026-1033.
82. Rutledge HL, Rittle J, Williamson LM, Xu WA, Gagnon DM, Tezcan FA. 2019. Redox-Dependent Metastability of the Nitrogenase P-Cluster. Journal of the American Chemical Society 141:10091-10098.
83. Pierik AJ, Wassink H, Haaker H, Hagen WR. 1993. Redox properties and EPR spectroscopy of the P clusters of *Azotobacter vinelandii* MoFe protein. European Journal of Biochemistry 212:51-61.
84. Spatzal T, Aksoyoglu M, Zhang L, Andrade SL, Schleicher E, Weber S, Rees DC, Einsle O. 2011. Evidence for interstitial carbon in nitrogenase FeMo cofactor. Science 334:940-940.
85. Lancaster KM, Roemelt M, Ettenhuber P, Hu Y, Ribbe MW, Neese F, Bergmann U, DeBeer S. 2011. X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. Science 334:974-977.
86. Henthorn JT, Arias RJ, Koroidov S, Kroll T, Sokaras D, Bergmann U, Rees DC, DeBeer S. 2019. Localized Electronic Structure of Nitrogenase FeMoco Revealed by

Selenium K-Edge High Resolution X-ray Absorption Spectroscopy. Journal of the American Chemical Society 141:13676-13688.

87. Grönberg KLC, Gormal CA, Durrant MC, Smith BE, Henderson RA. 1998. Why R-Homocitrate Is Essential to the Reactivity of FeMo-Cofactor of Nitrogenase: Studies on NifV--Extracted FeMo-Cofactor. Journal of the American Chemical Society 120:10613-10621.

88. Decamps L, Rice DB, DeBeer S. 2022. An Fe_6C Core in All Nitrogenase Cofactors. Angewandte Chemie International Edition 61:e202209190.

89. Eady RR. 1996. Structure-function relationships of alternative nitrogenases. Chemical Reviews 96:3013-3030.

90. Harris DF, Lukoyanov DA, Shaw S, Compton P, Tokmina-Lukaszewska M, Bothner B, Kelleher N, Dean DR, Hoffman BM, Seefeldt LC. 2018. Mechanism of N_2 Reduction Catalyzed by Fe-Nitrogenase Involves Reductive Elimination of H_2. Biochemistry 57:701-710.

91. Harris DF, Yang Z-Y, Dean DR, Seefeldt LC, Hoffman BM. 2018. Kinetic understanding of N_2 reduction versus H_2 evolution at the E_4 (4H) Janus state in the three nitrogenases. Biochemistry 57:5706-5714.

92. Hu B, Harris DF, Dean DR, Liu TL, Yang Z-Y, Seefeldt LC. 2018. Electrocatalytic CO_2 reduction catalyzed by nitrogenase MoFe and FeFe proteins. Bioelectrochemistry 120:104-109.

93. Hu Y, Lee CC, Ribbe MW. 2011. Extending the carbon chain: hydrocarbon formation catalyzed by vanadium/molybdenum nitrogenases. Science 333:753-755.

94. Zheng Y, Harris DF, Yu Z, Fu Y, Poudel S, Ledbetter RN, Fixen KR, Yang ZY, Boyd ES, Lidstrom ME, Seefeldt LC, Harwood CS. 2018. A pathway for biological methane production using bacterial iron-only nitrogenase. Nature Microbiology 3:281-286.

95. Rebelein JG, Stiebritz MT, Lee CC, Hu Y. 2017. Activation and reduction of carbon dioxide by nitrogenase iron proteins. Nature Chemical Biology 13:147-149.

96. Stiebritz MT, Hiller CJ, Sickerman NS, Lee CC, Tanifuji K, Ohki Y, Hu Y. 2018. Ambient conversion of CO_2 to hydrocarbons by biogenic and synthetic $[Fe_4S_4]$ clusters. Nature Catalysis 1:444-451.

97. Hales BJ. 1990. Alternative nitrogenase. Advances in Inorganic Biochemistry 8:165-198.

98. Schneider K, Müller A. 2004. Iron-Only Nitrogenase: Exceptional Catalytic, Structural and Spectroscopic Features, p 281-307. *In* Smith BE, Richards RL, Newton

WE (ed), Catalysts for Nitrogen Fixation: Nitrogenases, Relevant Chemical Models and Commercial Processes. Springer Netherlands.

99. Krahn E, Weiss R, Kröckel M, Groppe J, Henkel G, Cramer P, Trautwein X, Schneider K, Müller A. 2002. The Fe-only nitrogenase from *Rhodobacter capsulatus*: identification of the cofactor, an unusual, high-nuclearity iron-sulfur cluster, by Fe K-edge EXAFS and 57Fe Mössbauer spectroscopy. Journal of Biological Inorganic Chemistry 7:37-45.

100. Harris DF, Lukoyanov DA, Kallas H, Trncik C, Yang Z-Y, Compton P, Kelleher N, Einsle O, Dean DR, Hoffman BM, Seefeldt LC. 2019. Mo-, V-, and Fe-Nitrogenases Use a Universal Eight-Electron Reductive-Elimination Mechanism To Achieve N_2 Reduction. Biochemistry 58:3293-3301.

101. Fixen KR, Zheng Y, Harris DF, Shaw S, Yang Z-Y, Dean DR, Seefeldt LC, Harwood CS. 2016. Light-driven carbon dioxide reduction to methane by nitrogenase in a photosynthetic bacterium. Proceedings of the National Academy of Sciences 113:10163-10167.

102. Lee CC, Tanifuji K, Newcomb M, Liedtke J, Hu Y, Ribbe MW. 2018. A comparative analysis of the CO-reducing activities of MoFe proteins containing Mo-and V-Nitrogenase cofactors. ChemBioChem 19:649-653.

103. Rebelein JG, Lee CC, Newcomb M, Hu Y, Ribbe MW. 2018. Characterization of an M-cluster-substituted nitrogenase VFe protein. mBio 9:e00310-18.

104. Shin SH, Kim S, Kim JY, Lee S, Um Y, Oh M-K, Kim Y-R, Lee J, Yang K-S. 2012. Complete genome sequence of *Klebsiella oxytoca* KCTC 1686, used in production of 2, 3-Butanediol. Journal of Bacteriology 194:2371 - 2372.

105. Barney BM, Igarashi RY, Dos Santos PC, Dean DR, Seefeldt LC. 2004. Substrate interaction at an iron-sulfur face of the FeMo-cofactor during nitrogenase catalysis. Journal of Biological Chemistry 279:53621-53624.

106. Igarashi RY, Laryukhin M, Dos Santos PC, Lee H-I, Dean DR, Seefeldt LC, Hoffman BM. 2005. Trapping H-bound to the nitrogenase FeMo-cofactor active site during H_2 evolution: Characterization by ENDOR spectroscopy. Journal of the American Chemical Society 127:6231-6241.

107. Lukoyanov D, Khadka N, Yang Z-Y, Dean DR, Seefeldt LC, Hoffman BM. 2016. Reductive Elimination of H_2 Activates Nitrogenase to Reduce the N≡N Triple Bond: Characterization of the E_4 (4H) Janus Intermediate in Wild-Type Enzyme. Journal of the American Chemical Society 138:10674-10683.

108. Lukoyanov D, Khadka N, Yang Z-Y, Dean DR, Seefeldt LC, Hoffman BM. 2016. Reversible photoinduced reductive elimination of H_2 from the nitrogenase dihydride state, the E_4 (4H) Janus intermediate. Journal of the American Chemical Society 138:1320-1327.

109. Lukoyanov D, Yang Z-Y, Khadka N, Dean DR, Seefeldt LC, Hoffman BM. 2015. Identification of a key catalytic intermediate demonstrates that nitrogenase is activated by the reversible exchange of N_2 for H_2. Journal of the American Chemical Society 137:3610-3615.

110. Rees DC, Akif Tezcan F, Haynes CA, Walton MY, Andrade S, Einsle O, Howard JB. 2005. Structural basis of biological nitrogen fixation. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 363:971-984.

111. Danyal K, Dean DR, Hoffman BM, Seefeldt LC. 2011. Electron transfer within nitrogenase: evidence for a deficit-spending mechanism. Biochemistry 50:9255-9263.

112. Lanzilotta WN, Seefeldt LC. 1997. Changes in the midpoint potentials of the nitrogenase metal centers as a result of iron protein-molybdenum-iron protein complex formation. Biochemistry 36:12976-12983.

113. Howard JB, Rees DC. 1994. Nitrogenase: a nucleotide-dependent molecular switch. Annual Review of Biochemistry 63:235-264.

114. Rutledge HL, Cook BD, Nguyen HPM, Herzik MA, Jr., Tezcan FA. 2022. Structures of the nitrogenase complex prepared under catalytic turnover conditions. Science 377:865-869.

115. Eady RR, Lowe DJ, Thorneley RN. 1978. Nitrogenase of *Klebsiella pneumoniae*: A pre-steady state burst of ATP hydrolysis is coupled to electron transfer between the component proteins. FEBS letters 95:211-213.

116. Lanzilotta WN, Ryle MJ, Seefeldt LC. 1995. Nucleotide hydrolysis and protein conformational changes in *Azotobacter vinelandii* nitrogenase iron protein: defining the function of aspartate 129. Biochemistry 34:10713-10723.

117. Yang Z-Y, Ledbetter R, Shaw S, Pence N, Tokmina-Lukaszewska M, Eilers B, Guo Q, Pokhrel N, Cash VL, Dean DR. 2016. Evidence that the Pi release event is the rate-limiting step in the nitrogenase catalytic cycle. Biochemistry 55:3625-3635.

118. Hallenbeck PC. 1983. Nitrogenase reduction by electron carriers: Influence of redox potential on activity and the ATP2e− ratio. Archives of Biochemistry and Biophysics 220:657-660.

119. Hardy R, Knight Jr E. 1966. Reductant-dependent adenosine triphosphatase of nitrogen-fixing extracts of *Azotobacter vinelandii*. Biochimica et Biophysica Acta (BBA)-Enzymology and Biological Oxidation 122:520-531.
120. Mortenson LE. 1964. Ferredoxin and ATP, requirements for nitrogen fixation in cell-free extracts of *Clostridium pasteurianum*. Proceedings of the National Academy of Sciences 52:272-279.
121. Martin A, Burgess B, Iismaa S, Smartt C, Jacobson M, Dean D. 1989. Construction and characterization of an *Azotobacter vinelandii* strain with mutations in the genes encoding flavodoxin and ferredoxin I. Journal of bacteriology 171:3162-3167.
122. Bulen WA. 1965. Biological nitrogen fixation. Science 147:310-312.
123. Thorneley RN, Deistung J. 1988. Electron-transfer studies involving flavodoxin and a natural redox partner, the iron protein of nitrogenase. Conformational constraints on protein-protein interactions and the kinetics of electron transfer within the protein complex. Biochemical Journal 253:587-595.
124. Deistung J, Cannon FC, Cannon MC, Hill S, Thorneley RNF. 1985. Electron transfer to nitrogenase in *Klebsiella pneumoniae*. *nifF* gene cloned and the gene product, a flavodoxin, purified. Biochemical Journal 231:743-753.
125. Silverstein R, Bulen WA. 1970. Kinetic studies of the nitrogenase-catalyzed hydrogen evolution and nitrogen reduction reactions. Biochemistry 9:3809-3815.
126. Hageman RV, Burris R. 1980. Electron allocation to alternative substrates of *Azotobacter* nitrogenase is controlled by the electron flux through dinitrogenase. Biochimica et Biophysica Acta (BBA)-Bioenergetics 591:63-75.
127. Wherland S, Burgess BK, Stiefel EI, Newton WE. 1981. Nitrogenase reactivity: effects of component ratio on electron flow and distribution during nitrogen fixation. Biochemistry 20:5132-5140.
128. Badalyan A, Yang Z-Y, Seefeldt LC. 2019. A voltammetric study of nitrogenase catalysis using electron transfer mediators. ACS Catalysis 9:1366-1372.
129. Bush J, Wilson P. 1959. A non-gummy chromogenic strain of *Azotobacter vinelandii*. Nature 184:381-381.
130. Noar JD, Bruno-Bárcena JM. 2018. *Azotobacter vinelandii*: the source of 100 years of discoveries and many more to come. Microbiology 164:421-436.
131. Lineweaver H, Burk D. 1934. The determination of enzyme dissociation constants. Journal of the American Chemical Society 56:658-666.

132. Ortiz PJ, Ochoa S. 1959. Studies on Polynucleotides Synthesized by Polynucleotide Phosphorylase: IV. P^{32}-LABELED RIBONUCLEIC ACID. Journal of Biological Chemistry 234:1208-1212.
133. Joerger RD, Bishop PE, Evans HJ. 1988. Bacterial alternative nitrogen fixation systems. CRC Critical Reviews in Microbiology 16:1-14.
134. Joerger R, Jacobson M, Premakumar R, Wolfinger E, Bishop P. 1989. Nucleotide sequence and mutational analysis of the structural genes (*anfHDGK*) for the second alternative nitrogenase from *Azotobacter vinelandii*. Journal of Bacteriology 171:1075-1086.
135. Chisnell JR, Premakumar R, Bishop PE. 1988. Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii*. Journal of Bacteriology 170:27-33.
136. Dasgupta D, Panda AK, Mishra R, Mahanty A, De Mandal S, Bisht SS. 2021. Nif genes: tools for sustainable agriculture, p 413-434, Recent Advancement in Microbial Biotechnology. Elsevier.
137. Garcia AK, McShea H, Kolaczkowski B, Kaçar B. 2020. Reconstructing the evolutionary history of nitrogenases: Evidence for ancestral molybdenum-cofactor utilization. Geobiology 18:394-411.
138. Hamilton TL, Ludwig M, Dixon R, Boyd ES, Dos Santos PC, Setubal JC, Bryant DA, Dean DR, Peters JW. 2011. Transcriptional Profiling of Nitrogen Fixation in *Azotobacter vinelandii*. Journal of Bacteriology 193:4477-4486.
139. Boyd ES, Garcia Costas AM, Hamilton TL, Mus F, Peters JW. 2015. Evolution of Molybdenum Nitrogenase during the Transition from Anaerobic to Aerobic metabolism. Journal of Bacteriology 197:1690-1699.
140. Bulen WA, LeComte JR. 1966. The nitrogenase system from *Azotobacter*: two-enzyme requirement for N_2 reduction, ATP-dependent H_2 evolution, and ATP hydrolysis. Proceedings of the National Academy of Sciences 56:979-986.
141. Yates MG, Planque K. 1975. Nitrogenase from *Azotobacter chroococcum*: Purification and properties of the component proteins. European Journal of Biochemistry 60:467-476.
142. Eady R, Smith B, Cook K, Postgate J. 1972. Nitrogenase of *Klebsiella pneumoniae*. Purification and properties of the component proteins. Biochemical Journal 128:655-675.

143. Vandecasteele J-P, Burris R. 1970. Purification and properties of the constituents of the nitrogenase complex from *Clostridium pasteurianum*. Journal of Bacteriology 101:794-801.

144. Fisher K, Newton WE. 2005. Nitrogenase proteins from *Gluconacetobacter diazotrophicus*, a sugarcane-colonizing bacterium. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1750:154-165.

145. Hales BJ, Case EE, Morningstar JE, Dzeda MF, Mauterer LA. 1986. Isolation of a new vanadium-containing nitrogenase from *Azotobacter vinelandii*. Biochemistry 25:7251-7255.

146. Eady R, Robson RL, Richardson TH, Miller RW, Hawkins M. 1987. The vanadium nitrogenase of *Azotobacter chroococcum*. Purification and properties of the VFe protein. Biochemical journal 244:197-207.

147. Chisnell J, Premakumar R, Bishop P. 1988. Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii*. Journal of Bacteriology 170:27-33.

148. Müller A, Schneider K, Gollan U, Krahn E, Dröttboom M. 1995. Characterization of the iron only nitrogenase from *Rhodobacter Capsulatus*. Journal of Inorganic Biochemistry 59:551-551.

149. Davis R, Lehman L, Petrovich R, Shah VK, Roberts GP, Ludden PW. 1996. Purification and characterization of the alternative nitrogenase from the photosynthetic bacterium *Rhodospirillum rubrum*. Journal of Bacteriology 178:1445-1450.

150. Zumft WG, Castillo F. 1978. Regulatory properties of the nitrogenase from *Rhodopseudomonas palustris*. Archives of Microbiology 117:53-60.

151. Schlessman JL, Woo D, Joshua-Tor L, Howard JB, Rees DC. 1998. Conformational variability in structures of the nitrogenase iron proteins from *Azotobacter vinelandii* and *Clostridium pasteurianum*. Journal of Molecular Biology 280:669-685.

152. Rettberg LA, Kang W, Stiebritz MT, Hiller CJ, Lee CC, Liedtke J, Ribbe MW, Hu Y. 2019. Structural Analysis of a Nitrogenase Iron Protein from *Methanosarcina acetivorans*: Implications for CO_2 Capture by a Surface-Exposed [Fe_4S_4] Cluster. mBio 10:e01497-19.

153. Mayer SM, Lawson DM, Gormal CA, Roe SM, Smith BE. 1999. New insights into structure-function relationships in nitrogenase: a 1.6 Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe-protein. Journal of Molecular Biology 292:871-891.

154. Kim J, Woo D, Rees DC. 1993. X-ray crystal structure of the nitrogenase molybdenum-iron protein from *Clostridium pasteurianum* at 3.0-.ANG. resolution. Biochemistry 32:7104-7115.
155. Owens CP, Katz FEH, Carter CH, Oswald VF, Tezcan FA. 2016. Tyrosine-Coordinated P-Cluster in *G. diazotrophicus* Nitrogenase: Evidence for the Importance of O-Based Ligands in Conformationally Gated Electron Transfer. Journal of the American Chemical Society 138:10124-10127.
156. Atalah J, Cáceres-Moreno P, Espina G, Blamey JM. 2019. Thermophiles and the applications of their enzymes as new biocatalysts. Bioresource Technology 280:478-488.
157. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. Nature Reviews Microbiology 6:579-591.
158. Enzmann F, Mayer F, Rother M, Holtmann D. 2018. Methanogens: biochemical background and biotechnological applications. AMB Express 8:1.
159. Sauterey B, Charnay B, Affholder A, Mazevet S, Ferrière R. 2020. Co-evolution of primitive methane-cycling ecosystems and early Earth's atmosphere and climate. Nature Communications 11:2705.
160. Shima S, Huang G, Wagner T, Ermler U. 2020. Structural Basis of Hydrogenotrophic Methanogenesis. Annu Rev Microbiol 74:713-733.
161. Offre P, Spang A, Schleper C. 2013. Archaea in Biogeochemical Cycles. Annu Rev Microbiol 67:437-457.
162. Aschenbach K, Conrad R, Reháková K, Doležal J, Janatková K, Angel R. 2013. Methanogens at the top of the world: occurrence and potential activity of methanogens in newly deglaciated soils in high-altitude cold deserts in the Western Himalayas. Front Microbiol 4:359.
163. Wen X, Yang S, Horn F, Winkel M, Wagner D, Liebner S. 2017. Global Biogeographic Analysis of Methanogenic Archaea Identifies Community-Shaping Environmental Factors of Natural Environments. Front Microbiol 8:1339.
164. Garcia JL, Patel BK, Ollivier B. 2000. Taxonomic, phylogenetic, and ecological diversity of methanogenic Archaea. Anaerobe 6:205-226.
165. Angenent LT, Usack JG, Sun T, Fink C, Molitor B, Labatut R, Posmanik R, Hörl M, Hafenbradl D. 2022. Upgrading anaerobic digestion within the energy economy–the methane platform. Resource Recovery from Water: Principles and Application Edited

by Pikaar I, Guest J, Ganigué R, Jensen P, Rabaey K, Seviour T, Trimmer J, van der Kolk O, Vaneeckhaute C, Verstraete W IWA Publishing.

166. Kassem N, Hockey J, Lopez C, Lardon L, Angenent LT, Tester JW. 2020. Integrating anaerobic digestion, hydrothermal liquefaction, and biomethanation within a power-to-gas framework for dairy waste management and grid decarbonization: a techno-economic assessment. Sustainable Energy & Fuels 4:4644-4661.
167. Kurth JM, Op den Camp HJM, Welte CU. 2020. Several ways one goal-methanogenesis from unconventional substrates. Applied Microbiology and Biotechnology 104:6839-6854.
168. Murray PA, Zinder SH. 1984. Nitrogen fixation by a methanogenic archaebacterium. Nature 312:284-286.
169. Belay N, Sparling R, Daniels L. 1984. Dinitrogen fixation by a thermophilic methanogenic bacterium. Nature 312:286-288.
170. Dekas AE, Poretsky RS, Orphan VJ. 2009. Deep-Sea Archaea Fix and Share Nitrogen in Methane-Consuming Microbial Consortia. Science 326:422-426.
171. Mehta MP, Butterfield DA, Baross JA. 2003. Phylogenetic Diversity of Nitrogenase (*nifH*) Genes in Deep-Sea and Hydrothermal Vent Environments of the Juan de Fuca Ridge. Applied and Environmental Microbiology 69:960-970.
172. Man-Aharonovich D, Kress N, Zeev EB, Berman-Frank I, Béjà O. 2007. Molecular ecology of *nifH* genes and transcripts in the eastern Mediterranean Sea. Environmental Microbiology 9:2354-2363.
173. Dang H, Luan X, Zhao J, Li J. 2009. Diverse and Novel *nifH* and *nifH*-Like Gene Sequences in the Deep-Sea Methane Seep Sediments of the Okhotsk Sea. Applied and Environmental Microbiology 75:2238-2245.
174. Bae HS, Morrison E, Chanton JP, Ogram A. 2018. Methanogens Are Major Contributors to Nitrogen Fixation in Soils of the Florida Everglades. Applied and Environmental Microbiology 84:1-16.
175. Mei R, Kaneko M, Imachi H, Nobu MK. 2023. The origin and evolution of methanogenesis and Archaea are intertwined. PNAS Nexus 2.
176. Stüeken EE, Buick R, Guy BM, Koehler MC. 2015. Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr. Nature 520:666-669.
177. Raymond J, Siefert JL, Staples CR, Blankenship RE. 2004. The Natural History of Nitrogen Fixation. Molecular Biology and Evolution 21:541-554.

178. Boyd ES, Anbar AD, Miller S, Hamilton TL, Lavin M, Peters JW. 2011. A late methanogen origin for molybdenum-dependent nitrogenase. Geobiology 9:221-232.
179. Boyd ES, Hamilton TL, Peters JW. 2011. An alternative path for the evolution of biological nitrogen fixation. Frontiers in Microbiology 2:1-11.
180. Poulton SW, Fralick PW, Canfield DE. 2010. Spatial variability in oceanic redox structure 1.8 billion years ago. Nature Geoscience 3:486-490.
181. Kendall B, Reinhard CT, Lyons TW, Kaufman AJ, Poulton SW, Anbar AD. 2010. Pervasive oxygenation along late Archaean ocean margins. Nature Geoscience 3:647-652.
182. Anbar AD, Knoll AH. 2002. Proterozoic ocean chemistry and evolution: A bioinorganic bridge? Science 297:1137-1142.
183. Lyons TW, Reinhard CT, Planavsky NJ. 2014. The rise of oxygen in Earth's early ocean and atmosphere. Nature 506:307-315.
184. Canfield DE, Glazer AN, Falkowski PG. 2010. The evolution and future of Earth's nitrogen cycle. Science 330:192-196.
185. Navarro-González R, McKay CP, Mvondo DN. 2001. A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. Nature 412:61-64.
186. Kasting J, Walker JC. 1981. Limits on oxygen concentration in the prebiological atmosphere and the rate of abiotic fixation of nitrogen. Journal of Geophysical Research: Oceans 86:1147-1158.
187. Forchhammer K, Selim KA, Huergo LF. 2022. New views on PII signaling: from nitrogen sensing to global metabolic control. Trends in Microbiology 30:722-735.
188. Huergo LF, Chandra G, Merrick M. 2013. P(II) signal transduction proteins: nitrogen regulation and beyond. FEMS Microbiology Reviews 37:251-83.
189. Leigh JA, Dodsworth JA. 2007. Nitrogen Regulation in Bacteria and Archaea. Annual Review of Microbiology 61:349-377.
190. Huergo LF, Dixon R. 2015. The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. Microbiol Mol Biol Rev 79:419-35.
191. Contreras G, Thomsen J, Pfitzer M, Hafenbradl D, Kostner D, Holtmann D, Schmitz RA, Rother M, Molitor B. 2022. New perspectives for biotechnological applications of methanogens. Current Research in Biotechnology 4:468-474.
192. Kessler PS, McLarnan J, Leigh JA. 1997. Nitrogenase Phylogeny and the Molybdenum Dependence of Nitrogen Fixation in *Methanococcus maripaludis*. Journal of Bacteriology 179:541-543.

193. Lobo AL, Zinder SH. 1990. Nitrogenase in the Archaebacterium *Methanosarcina barkeri* 227. Journal of Bacteriology 172:6789-6796.
194. Blank CE, Kessler PS, Leigh JA. 1995. Genetics in Methanogens: Transposon Insertion Mutagenesis of a *Methanococcus maripaludis nifH* Gene. Journal of Bacteriology 177:5773-5777.
195. Kessler PS, Blank C, Leigh JA. 1998. The *nif* Gene Operon of the Methanogenic Archaeon *Methanococcus maripaludis*. Journal of Bacteriology 180:1504-1511.
196. Chien Y-T, Zinder SH. 1996. Cloning, Functional Organization, Transcript Studies, and Phylogenetic Analysis of the Complete Nitrogenase Structural Genes (*nifHDK2*) and Associated Genes in the Archaeon *Methanosarcina barkeri* 227. Journal of Bacteriology 178:143-148.
197. Dodsworth JA, Cady NC, Leigh JA. 2005. 2-Oxoglutarate and the PII homologues $NifI_1$ and $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. Molecular Microbiology 56:1527-1538.
198. Dodsworth JA, Leigh JA. 2006. Regulation of nitrogenase by 2-oxoglutarate-reversible, direct binding of a PII-like nitrogen sensor protein to dinitrogenase. Proceedings of the National Academy of Sciences of the United States of America 103:9779-9784.
199. Kessler PS, Leigh JA. 1999. Genetics of Nitrogen Regulation in *Methanococcus maripaludis*. Genetics 152:1343-1351.
200. Weidenbach K, Ehlers C, Kock J, Ehrenreich A, Schmitz RA. 2008. Insights into the NrpR regulon in *Methanosarcina mazei* Gö1. Archives of Microbiology 190:319-332.
201. Weidenbach K, Ehlers C, Kock J, Schmitz RA. 2010. NrpRII mediates contacts between NrpRI and general transcription factors in the archaeon *Methanosarcina mazei* Gö1. FEBS Journal 277:4398-411.
202. Weidenbach K, Ehlers C, Schmitz RA. 2014. The transcriptional activator NrpA is crucial for inducing nitrogen fixation in *Methanosarcina mazei* Gö1 under nitrogen-limited conditions. FEBS Journal 281:3507-3522.
203. Prasse D, Förstner KU, Jäger D, Backofen R, Schmitz RA. 2017. $sRNA_{154}$ a newly identified regulator of nitrogen fixation in *Methanosarcina mazei* strain Gö1. RNA Biology 14:1544-1558.
204. Prasse D, Schmitz RA. 2018. Small RNAs Involved in Regulation of Nitrogen Metabolism. Microbiol Spectrum 6.

205. Lie TJ, Leigh JA. 2003. A novel repressor of *nif* and *glnA* expression in the methanogenic archaeon *Methanococcus maripaludis*. Molecular Microbiology 47:235-246.

206. Lie TJ, Wood GE, Leigh JA. 2005. Regulation of *nif* Expression in *Methanococcus maripaludis*: ROLES OF THE EURYARCHAEAL REPRESSOR NrpR, 2-OXOGLUTARATE, AND TWO OPERATORS. Journal of Biological Chemistry 280:5236-5241.

207. Jäger D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA. 2009. Deep sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. Proceedings of the National Academy of Sciences of the United States of America 106:21878-21882.

208. Dodsworth JA, Cady NC, Leigh JA. 2005. 2-Oxoglutarate and the PII homologues $NifI_1$ and $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. Molecular Microbiology 56:1527-1538.

209. Kessler PS, Daniel C, Leigh JA. 2001. Ammonia Switch-Off of Nitrogen Fixation in the Methanogenic Archaeon *Methanococcus maripaludis*: Mechanistic Features and Requirement for the novel GlnB homologues, $NifI_1$ and $NifI_2$. Journal of Bacteriology 183:882-889.

210. Mehta MP, Baross JA. 2006. Nitrogen Fixation at 92 °C by a Hydrothermal Vent Archaeon. Science 314:1783-1786.

211. Huber H, Thomm M, König H, Thies G, Stetter KO. 1982. *Methanococcus thermolithotrophicus*, a Novel Thermophilic Lithotrophic Methanogen. Archives of Microbiology 132:47-50.

212. Jeanthon C, L'Haridon S, Reysenbach AL, Vernet M, Messner P, Sleytr UB, Prieur D. 1998. *Methanococcus infernu*s sp. nov., a novel hyperthermophilic lithotrophic methanogen isolated from a deep-sea hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology 48 Pt 3:913-919.

213. Maslać N, Sidhu C, Teeling H, Wagner T. 2022. Comparative Transcriptomics Sheds Light on Remodeling of Gene Expression during Diazotrophy in the Thermophilic Methanogen *Methanothermococcus thermolithotrophicus*. mBio 13:e02443-22.

214. Belay N, Jung KY, Rajagopal BS, Kremer JD, Daniels L. 1990. Nitrate as a Sole Nitrogen Source for *Methanococcus thermolithotrophicus* and Its Effect on Growth of Several Methanogenic Bacteria. Curr Microbiol 21:193-198.

215. Capone D. 1988. Nitrogen cycling in coastal marine environments. Blackburn, TH.

216. Bertics VJ, Löscher C, Salonen I, Dale AW, Gier J, Schmitz R, Treude T. 2013. Occurrence of benthic microbial nitrogen fixation coupled to sulfate reduction in the seasonally hypoxic Eckernförde Bay, Baltic Sea. Biogeosciences 10:1243-1258.
217. Dekas AE, Chadwick GL, Bowles MW, Joye SB, Orphan VJ. 2014. Spatial distribution of nitrogen fixation in methane seep sediment and the role of the ANME archaea. Environmental Microbiology 16:3012-3029.
218. Zhang X, McRose DL, Darnajoux R, Bellenger JP, Morel FMM, Kraepiel AML. 2016. Alternative nitrogenase activity in the environment and nitrogen cycle implications. Biogeochemistry 127:189-198.
219. Boyd ES, Peters JW. 2013. New insights into the evolutionary history of biological nitrogen fixation. Front Microbiol 4:1-12.
220. Wang X, Kim Y, Ma Q, Hong SH, Pokusaeva K, Sturino JM, Wood TK. 2010. Cryptic prophages help bacteria cope with adverse environments. Nature Communications 1:147.
221. Feiner R, Argov T, Rabinovich L, Sigal N, Borovok I, Herskovits AA. 2015. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. Nature Reviews Microbiology 13:641-650.
222. Hackley RK, Schmid AK. 2019. Global Transcriptional Programs in Archaea Share Features with the Eukaryotic Environmental Stress Response. Journal of Molecular Biology 431:4147-4166.
223. Schindelin H, Kisker C, Schlessman JL, Howard JB, Rees DC. 1997. Structure of ADP·AlF_4–-stabilized nitrogenase complex and its implications for signal transduction. Nature 387:370-376.
224. Katz FE, Owens CP, Tezcan F. 2016. Electron transfer reactions in biological nitrogen fixation. Israel Journal of Chemistry 56:682-692.
225. Owens CP, Katz FEH, Carter CH, Luca MA, Tezcan FA. 2015. Evidence for Functionally Relevant Encounter Complexes in Nitrogenase Catalysis. Journal of the American Chemical Society 137:12704-12712.
226. Kang W, Lee CC, Jasniewski AJ, Ribbe MW, Hu Y. 2021. Response to Comment on "Structural evidence for a dynamic metallocofactor during N_2 reduction by Mo-nitrogenase". Science 371:eabe5856.
227. Spatzal T, Perez KA, Einsle O, Howard JB, Rees DC. 2014. Ligand binding to the FeMo-cofactor: Structures of CO-bound and reactivated nitrogenase. Science 345:1620-1623.

228. Spatzal T, Perez KA, Howard JB, Rees DC. 2015. Catalysis-dependent selenium incorporation and migration in the nitrogenase active site iron-molybdenum cofactor. eLife 4:1-11.
229. Cao L, Caldararu O, Ryde U. 2020. Does the crystal structure of vanadium nitrogenase contain a reaction intermediate? Evidence from quantum refinement. JBIC Journal of Biological Inorganic Chemistry 25:847-861.
230. Benediktsson B, Thorhallsson AT, Bjornsson R. 2018. QM/MM calculations reveal a bridging hydroxo group in a vanadium nitrogenase crystal structure. Chemical Communications 54:7310-7313.
231. Chubukov V, Gerosa L, Kochanowski K, Sauer U. 2014. Coordination of microbial metabolism. Nature Reviews Microbiology 12:327-340.
232. Veit K, Ehlers C, Ehrenreich A, Salmon K, Hovey R, Gunsalus RP, Deppenmeier U, Schmitz RA. Global transcriptional analysis of *Methanosarcina mazei* strain Gö1 under different nitrogen availabilities. Mol Gen Genomics 276:41-55.
233. Lie TJ, Dodsworth JA, Nickle DC, Leigh JA. 2007. Diverse homologues of the archaeal repressor NrpR function similarly in nitrogen regulation. FEMS Microbiol Lett 271:281-288.
234. Kochanowski K, Sauer U, Noor E. 2015. Posttranslational regulation of microbial metabolism. Current Opinion in Microbiology 27:10-17.
235. Millar AH, Heazlewood JL, Giglione C, Holdsworth MJ, Bachmair A, Schulze WX. 2019. The Scope, Functions, and Dynamics of Posttranslational Protein Modifications. Annual Review of Plant Biology 70:119-151.
236. Garcia AK, Harris DF, Rivier AJ, Carruthers BM, Pinochet-Barros A, Seefeldt LC, Kaçar B. 2023. Nitrogenase resurrection and the evolution of a singular enzymatic mechanism. eLife 12:e85003.
237. Garcia AK, Cavanaugh CM, Kacar B. 2021. The curious consistency of carbon biosignatures over billions of years of Earth-life coevolution. The ISME Journal 15:2183-2194.
238. Cadoux C, Ratcliff D, Maslać N, Gu W, Tsakoumagkos I, Hoogendoorn S, Wagner T, Milton RD. 2023. Nitrogen Fixation and Hydrogen Evolution by Sterically Encumbered Mo-Nitrogenase. JACS Au 3:1521-1533.
239. Xia Q, Wang T, Hendrickson EL, Lie TJ, Hackett M, Leigh JA. 2009. Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen *Methanococcus maripaludis*. BMC Microbiol 9:149.

240. Johnsen U, Hansen T, Schönheit P. 2003. Comparative analysis of pyruvate kinases from the hyperthermophilic archaea *Archaeoglobus fulgidus*, *Aeropyrum pernix*, and *Pyrobaculum aerophilum* and the hyperthermophilic bacterium T*hermotoga maritima*: unusual regulatory properties in hyperthermophilic archaea. Journal of Biological Chemistry 278:25417-25427.

241. Sato M, Masuda Y, Kirimura K, Kino K. 2007. Thermostable ATP regeneration system using polyphosphate kinase from *Thermosynechococcus elongatus* BP-1 for D-amino acid dipeptide synthesis. Journal of bioscience and bioengineering 103:179-184.

242. Restiawaty E, Iwasa Y, Maya S, Honda K, Omasa T, Hirota R, Kuroda A, Ohtake H. 2011. Feasibility of thermophilic adenosine triphosphate-regeneration system using *Thermus thermophilus* polyphosphate kinase. Process Biochemistry 46:1747-1752.

243. Bock A-K, Glasemacher Jr, Schmidt R, Schönheit P. 1999. Purification and characterization of two extremely thermostable enzymes, phosphate acetyltransferase and acetate kinase, from the hyperthermophilic eubacterium *Thermotoga maritima*. Journal of bacteriology 181:1861-1867.

244. Cadoux C, Milton RD. 2020. Recent Enzymatic Electrochemistry for Reductive Reactions. ChemElectroChem 7:1974-1986.

245. Branlard J, Ayvazyan G, Ayvazyan V, Grecki M, Hoffmann M, Jezynski T, Kudla I, Lamb T, Ludwig F, Mavric U. 2012. The european XFEL LLRF system. IPAC 12:55-57.

246. Keedy DA, Kenner LR, Warkentin M, Woldeyes RA, Hopkins JB, Thompson MC, Brewster AS, Van Benschoten AH, Baxter EL, Uervirojnangkoorn M, McPhillips SE, Song J, Alonso-Mori R, Holton JM, Weis WI, Brunger AT, Soltis SM, Lemke H, Gonzalez A, Sauter NK, Cohen AE, van den Bedem H, Thorne RE, Fraser JS. 2015. Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. eLife 4:e07574.

247. Gisriel C, Coe J, Letrun R, Yefanov OM, Luna-Chavez C, Stander NE, Lisova S, Mariani V, Kuhn M, Aplin S, Grant TD, Dörner K, Sato T, Echelmeier A, Cruz Villarreal J, Hunter MS, Wiedorn MO, Knoska J, Mazalova V, Roy-Chowdhury S, Yang J-H, Jones A, Bean R, Bielecki J, Kim Y, Mills G, Weinhausen B, Meza JD, Al-Qudami N, Bajt S, Brehm G, Botha S, Boukhelef D, Brockhauser S, Bruce BD, Coleman MA, Danilevski C, Discianno E, Dobson Z, Fangohr H, Martin-Garcia JM, Gevorkov Y, Hauf S, Hosseinizadeh A, Januschek F, Ketawala GK, Kupitz C, Maia L,

Manetti M, Messerschmidt M, et al. 2019. Membrane protein megahertz crystallography at the European XFEL. Nature Communications 10:5021.

248. Wilson MA. 2022. Mapping Enzyme Landscapes by Time-Resolved Crystallography with Synchrotron and X-Ray Free Electron Laser Light. Annual Review of Biophysics 51:79-98.

249. Benediktsson B, Bjornsson R. 2020. Quantum Mechanics/Molecular Mechanics Study of Resting-State Vanadium Nitrogenase: Molecular and Electronic Structure of the Iron–Vanadium Cofactor. Inorganic Chemistry 59:11514-11527.

250. Susanti D, Frazier MC, Mukhopadhyay B. 2019. A Genetic System for *Methanocaldococcus jannaschii*: An Evolutionary Deeply Rooted Hyperthermophilic Methanarchaeon. Frontiers in Microbiology 10.

Acknowledgements

Without the constant support and advice throughout the last four years, this thesis would not have been possible. I would like to thank everyone who contributed to this thesis and was part of this journey.

Dr. Tristan Wagner, first, for the opportunity to do my PhD in the Microbial Metabolism Group and for everyday support, advice and discussion throughtout the years. Without your experience, guidance and scientific input, which greatly shaped my project, this thesis would not be possible, and I learned so much from the years of working with you. Thank you for sharing your passion and excitement for research and science and being a very dedicated supervisor. Thank you for being a reviewer of my thesis, a member of my Thesis Advisory Committee and my Thesis Defence Committee.

Prof. Dr. Ross D. Milton, for the successful collaboration and contributions towards the goals of this thesis, sharing your knowledge and advice and hosting me in the Milton Lab in Geneva. Thank you for being an active member of my Thesis Advisory Committee, the reviewer of my thesis and a member of my Thesis Defence Committee.

Dr. Marco Bellinzoni, for invaluable contributions of his crystallographic experience to the discussion about this thesis, being an external reviewer, an active member of my Thesis Advisory Committee and Thesis Defence Committee.

Prof. Dr. Jan-Hendrik Hehemann, for inspiring discussions during my Thesis Advisory Committee meetings and accepting the role of the Chair of the Thesis Defence Committee.

Prof. Dr. Rudolf Amann, for finding the time to participate in the final discussion of the thesis and being a member of my Thesis Defence Committee.

Thank you **Marie-Caroline Müller** and **Amelie Heidenreich** for joining my Thesis Defence Committee as the non-voting members.

Microbial Metabolism Group, thank you to all past and current members of MicroMet:

Ramona Appel and Christina Probian, technicians at MicroMet for all the technical assistance, constant support and advice during experimental phase of this thesis.

Marion Jespersen, for all the selfless help and support, in and outside the lab. Thank you for going through all the highs and lows of this adventure together, I cannot imagine the last four and a half years without you. Special thanks for translation of the Summary to German.

My office mates, **Dr. Olivier Lemaire and Melissa Belhamri** for always having an open ear for my questions and all the helpful suggestions.

All my collaborators and co-authors for their scientific input and involvement in the realization of this thesis:

Dr. Chandni Sidhu and **Dr. Hanno Teeling** (Molecular Ecology Department), for the successful collaboration on analyzing the transcriptomic data and sharing your experience with me. I learned a lot from you and I am very grateful for your patience and effort in teaching me.

Cécile Cadoux (Milton Group at University of Geneva), your work represents invaluable contribution to this thesis. Thank you for hosting me in your lab with a contagious enthusiasam and teaching me your tricks about the nitrogenase activity assays. Thank you for the successful colaborations so far, I really enjoy collaborating with you and hope that the colaborations will continue.

Max Planck Gesellschaft Facilities and their Staff: **Dr. Bruno Huettel** from MPG Genome Centre Cologne for RNA library preparation and sequencing and **Dr**. **Anna Shevchenko** from MPG Mass Spectrometry Facility Dresden for the mass spectrometry characterization of our samples.

Staff at Syncrotron Facilities and Beamlines: **PROXIMA-I** at **Synchrotron SOLEIL** (Paris, France), **PXI** and **PXIII** at **Swiss Light Source and the Paul Scherrer Institut** (Villingen, Switzerland) and **FIP2** at **European Synchrotron Radiation Facility** (Grenoble, France) for their advice, help and support during data collection, especially **Dr. Sylvain Engilberge.**

To all MarMic master students, interns and HiWis who contributed to this thesis throughout the years: **Fenja Murken**, **Anna Carlotta Kück**, **Pauline Bolte** and **Sia Shende.**

Max Planck Society which funded this thesis through the **Max Planck Institute for Marine Microbiology** and the **International Max Planck Research School MarMic.**

Dr. Christiane Glöckner, MarMic programme coordinator**, and Anita Tingberg** from MPI international office for help with all the administrative matters during the MarMic programme

Petar Čuček and **Znanstveno Edukacijski Centar Višnjan** for introducing me to marine microbiology and **Petnica Science Centre Biology Department** where I discovered my passion for biology.

I am incredibly thankful to have the support of great friends in my life without whom I cannot imagine achieving this: **Marko, Silvija, Stevan, Katarina, Nevena and Mila.**

Paloma, Darjan and Marion, thank you for being my rocks, making me feel understood, sharing the PhD experience and the mutual support for going through this part of our lives. Thank you for all the memories we created during these last years and I am looking forward to the future adventures.

Finally, I have been incredibly lucky to have the unwavering support of my family back in Serbia. I am thankful every day to have them in my life.

To my family, especially my parents and sister, thank you for unconditional support in all my scientific and non-scientific endeavours.

Hvala mojoj porodici, posebno mojim roditeljima i sestri, na bezuslovnoj podršci u svim mojim poduhvatima.

Declaration on the contribution of the candiate to the multi-author manuscripts included as chapters of the submitted doctoral thesis

	Contribution of the candidate in % of the total work				
	Ch. 2	Ch. 3	Ch. 4	Ch. 5	Ch. 6
Experimental concept and design	90	90	90	20	10
Experimental work/ Acquisition of experimental data	100	80	80	50	10
Data analysis and interpretation	70	80	80	70	10
Preparation of figures and tables	80	100	100	50	20
Drafting of the manuscript	80	80	80	50	30

Additional contributions: In addition to participating in the initial laboratory and equipment setup during the building phase (e.g. anaerobic tents), the candidate successfully implemented a robust protocol for large-scale cultivation of *M. thermolithotrophicus* and established a pipeline for anaerobic native nitrogenase purification yielding high-quality samples for crystallisation. Although these efforts are not explicitly documented in the thesis or published, the work presented in **Chapters 3** and **4** heavily relies on this groundwork.

We declare that all authors of these manuscripts have been informed on and did not object to the listed contributions of the candidate as author.

Date: _____

Signatures: _____ _____

(Candidate) (Supervisor)

Universität Bremen

Universität Bremen | Fachbereich 02 | Postfach 33 04 40,
28334 Bremen

Universität Bremen
Fachbereich 2
Prüfungsamt Biologie/Chemie
z. Hd. Frau Frauke Ernst
Leobener Straße
28359 Bremen
Deutschland

Prüfungsamt
Biologie/Chemie
Frauke Ernst
Geschäftsstelle
Fachbereich 02
Leobener Str. / NW2
D-28359 Bremen
Verwaltungspavillon 06
Tel. 0421 218-62802
Fax 0421 218-9862802
frauke.ernst@uni-bremen.de

Versicherung an Eides Statt

Name, Vorname: Nevena Maslać

Matrikel-Nr. 3118122

Ich, Maslać Nevena,

versichere an Eides Statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe. Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe. Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Ort, Datum / Unterschrift