Metallomics

Exploring the alternatives of biological nitrogen fixation

Journal:	*Metallomics*
Manuscript ID	MT-CRV-02-2018-000038.R1
Article Type:	Critical Review
Date Submitted by the Author:	28-Mar-2018
Complete List of Authors:	Mus, Florence; Washington State University, Institute of Biological Chemistry Alleman, Alexander; Washington State University, Institute of Biological Chemistry Pence, Natasha; Montana State University System, Chemistry and Biochemistry Seefeldt, Lance C.; Utah State University, Chemistry and Biochemistry Peters, John; Washington State University, Institute of Biological Chemistry

Exploring the alternatives of biological nitrogen fixation

Florence Mus[a], Alexender B. Alleman[a‡], Natasha Pence[b‡], Lance C. Seefeldt[c], John W. Peters[a*]

[a] Institute of Biological Chemistry, Washington State University, Pullman, WA 99163, USA

[b]Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717, USA

[c]Department of Chemistry and Biochemistry, Utah State University, Logan, UT 84322, USA

Key words: Nitrogen fixation, Mo-nitrogenase, V-nitrogenase, Fe-only nitrogenase, alternative nitrogenases, metals availability, biochemistry, genetics, evolution

‡ Both authors contributed equally to this work

*Address correspondence to John W. Peters, (jw.peters@wsu.edu)

Abstract

Most biological nitrogen fixation (BNF) results from the activity of the molybdenum nitrogenase (Mo-nitrogenase, Nif), an oxygen-sensitive metalloenzyme complex found in all known diazotrophs. Two alternative forms of nitrogenase, the vanadium nitrogenase (V-nitrogenase, Vnf) and the iron-only nitrogenase (Fe-only nitrogenase, Anf) have also been identified in the genome of some organisms that encode for Nif. It has been suggested that alternative nitrogenases were responsible for N_2-fixation on early Earth because oceans were depleted of bioavailable Mo. Results of recent phylogenetic- and structure-based studies suggest, however, that such an evolutionary path is unlikely, and favor a new model for a stepwise evolution of nitrogenase where the V-nitrogenase and the Fe-only nitrogenase are not the ancestor of the Mo-nitrogenase. Rather, Mo-nitrogenase emerged within the methanogenic archaea and then gave rise to the alternative forms suggesting they arose later in response to the availability of fixed N_2 and local environmental factors that influenced metal availability. This review summarizes the current state of knowledge on (1) the biochemistry of these complex systems highlighting the common and specific structural features and catalytic activities of the enzymes, (2) the recent progress in defining the discrete set of genes associated to N_2-fixation and the regulatory features that coordinate the differential expression of genes in response to metal availability, and (3) the diverse taxonomic and phylogenic distribution of nitrogenase enzymes and the evolutionary history of BNF from the perspective of metal content and metal availability.

Introduction

Biological nitrogen fixation (BNF), the reduction of dinitrogen (N_2) to ammonia, is an essential reaction in the global nitrogen cycle. BNF accounts for roughly two-thirds of the fixed nitrogen produced on Earth and is catalyzed by the nitrogenase complex.[1] Although N_2-fixation is not found in eukaryotes, it is widely distributed among the bacteria and archaea, revealing considerable biodiversity among diazotrophs. The ability to fix N_2 is found in most bacterial phylogenetic groups and is compatible with a wide range of physiologies.[2] Three distinct types of nitrogenase have been demonstrated to exist in diazotrophs: the Mo-nitrogenase, the V-nitrogenase, and the Fe-only nitrogenase.[3] Despite differences in their metal content, these enzymes are structurally, mechanistically, and phylogenetically related. The most widespread and intensively characterized system is the canonical Mo-nitrogenase. Mo-nitrogenase is a complex two component metalloenzyme comprised of two proteins. The smaller dimeric component, known as the iron protein or dinitrogenase reductase (NifH protein), functions as an ATP-dependent electron donor to the larger heterotetrameric component, known as the molybdenum-iron protein or dinitrogenase (comprising the NifD and NifK component proteins), which contains the enzyme catalytic site. Both of the nitrogenase-component proteins are extremely oxygen sensitive. The two other nitrogenases, V-nitrogenase and Fe-only nitrogenase, known as alternative nitrogenases, are enzyme homologs, with the exception of an additional subunit (VnfG or AnfG) in the dinitrogenase component, and the absence of the heteroatom Mo.

The structural components of the Mo-nitrogenase are encoded by *nifH*, *nifD*, and *nifK* genes. The V-nitrogenase is encoded by *vnfH*, *vnfD*, *vnfG*, and *vnfK*, and the Fe-only nitrogenase components are products of *anfH*, *anfD*, *anfG*, and *anfK* genes. Apart from the catalytic

components, additional gene products are required to produce fully functional nitrogenase enzymes. Although the number of proteins involved in the activation of nitrogenases seems to be species-specific and varies according to the physiology of the organism and environmental niche,[4,5] extensive biochemical and genetic characterization has implicated at least 82 genes products in the formation and regulation of the three forms of nitrogenases.[6–8] Despite variations in the precise inventory of proteins required for N_2-fixation, it is well acknowledged that the separate expression of the catalytic components is not enough to sustain N_2-fixation. A strong selective pressure to efficiently integrate N_2-fixation with other cellular metabolic processes is due to the high metabolic costs of N_2-fixation, in conjunction with the demands for fixed nitrogen in growing cells and the requirement to protect nitrogenase systems from oxidation inactivation.[4] Genome-wide transcription profiling studies were valuable for defining the full extent of the *nif*, *vnf*, and *anf* gene clusters and for identifying genes that are coordinately expressed with particular nitrogenase but not directly involved in N_2-fixation. Transcriptomic data has helped to refine understanding of the complex metabolic and regulation interactions that control N_2-fixation in diazotrophs. As well as uncover several interesting surprises regarding the evolutionary relationships between Mo-nitrogenase and the alternative nitrogenases.

This review will encompass the biochemistry, the gene expression and regulation context, and the phylogenetic relationships of the different metal containing forms of nitrogenase enzymes from the perspective of metal content and metal availability.

1-Discovery of nitrogenase systems

1-1 Biological nitrogen fixation is catalyzed by Mo-nitrogenase

Beijerinck first observed BNF in 1901 when a free-living soil bacterium, designated *Azotobacter chroococcum*, was able to grow rapidly in a medium free of any fixed nitrogen source.[9] The search for trace elements that would stimulate N_2-fixation identified molybdenum (Mo) in the form of Na_2MoO_4, and vanadium (V), in the form of $NaVO_3$, as playing key roles in nitrogen fixation.[9,10] Mo had been found to stimulate N_2-dependent growth over the control by 3-fold, V also stimulated N_2-dependent growth however it was not as effective in *Azotobacter*.[9,10] This led to the conclusion that Mo was essential for N_2-fixation.[9,10] Extracts from other microorganisms such as *Clostridium pasteurianum*, *Azotobacter vinelandii*, and *Rhodosporillum rubrum* further determined that BNF was a MgATP-dependent process that involved the production of H_2.[9,11] Studies focused on Mo-dependent N_2-fixation enabled the purification of a two-component enzyme system responsible for BNF known as the Mo-nitrogenase.[11] For many years, the Mo-nitrogenase was initially considered the sole enzyme system responsible for BNF.

1-2 Discovery of the alternative nitrogenases: the V- and Fe-only nitrogenases

Even though Mo was considered crucial for BNF, strains of *Azotobacter* had been observed to grow under N_2-fixing conditions in the absence of Mo.[9] To eliminate issues of Mo contamination, a study reexamined the trace metal requirement for BNF in *A. vinelandii* using metal purification techniques.[9] This study found that the wild type strain was able to fix N_2 in the absence of Mo, and adding V to the media slightly increased the amount of fixed nitrogen.[9] Despite these observations, the study concluded Mo was still essential for N_2-fixation.[9]

Growth studies using tungstate (W) instead of Mo enabled the isolation of W-tolerant mutant, which fixed nitrogen in the presence of Na_2WO_4. Importantly, these mutants expressed four proteins, which were different from the nitrogenase components usually expressed under

Mo. These mutants had mutations in the genes that encoded the Mo-nitrogenase. These same proteins could also be observed in the wild type when it was grown in the Mo-deficient media.[12,13] This observation gave the first clues to the existence of an alternative N_2-fixation system, which is expressed under Mo-deficient conditions.

Early attempts to purify nitrogenase from *A. vinelandii* cells grown in the presence of V yielded a less stable nitrogenase with a decreased affinity for substrates that favored H_2 production.[9] These attributes were not considered the product of a new enzyme system as it was thought that V could simply substitute for Mo in the Mo-nitrogenase.[9] The existence of alternative N_2-fixing enzyme systems was not confirmed until mutant strains of *A. vinelandii* that were unable to fix nitrogen due to mutational lesions in the nitrogenase structural genes reverted to fixing N_2 in Mo-deficient, nitrogen-free media.[14,15] The existence of an alternative N_2-fixation system was confirmed when deletion strains for the structural genes of the Mo-nitrogenase grew on Mo-deficient N_2 free media, were able to incorporate ^{15}N-labeled molecular N_2, and exhibited measurable nitrogenase activity.[16] The observation that BNF by microorganisms such as *A. vinelandii* and *A. chroococcum* can be catalyzed in the presence or absence of Mo or V in the media led to the purification of two alternative nitrogenases: the V-nitrogenase and the Fe-only nitrogenase.[12,17,18]

2-Biochemistry of nitrogenases

2-1 Structural features of nitrogenase systems: common and specific structural features

The Mo-, V- and Fe-only nitrogenases are all two-component enzyme systems where the catalytic component (dinitrogenase) is known as either the MoFe, VFe, and FeFe protein and the component involved in electron delivery (dinitrogenase reductase) is the Fe protein.[19–21] The Mo-, V- and Fe-only nitrogenases are mechanistically similar but genetically distinct as they are encoded by separate gene clusters designated, *nif*, *vnf*, and *anf*, respectively.[19,22,23] There is little sequence similarity between the MoFe, VFe and FeFe proteins, however residues that coordinate the FeMo-, FeV-, and FeFe-cofactors are conserved across all three enzyme systems (Figure 1a-c).[22–24]

The Mo-nitrogenase is the most widely studied of the nitrogenase enzymes and has been isolated from *A. vinelandii*, *Klebsiella pneumoniae*, and *C. pasteurianum.*[24] The crystal structure of the MoFe protein is a $\alpha_2\beta_2$ heterotetramer where the α-subunit is encoded by *nifD* and the β-subunit is encoded by *nifK* (Figure 1a).[25] The MoFe protein consists of a [7Fe-8S] P-cluster involved in mediating electrons to the [7Fe-9S-Mo-C-homocitrate] FeMo-cofactor, the site for substrate reduction (Figure 1d).[26–28] The Fe protein is a homodimer where a single [4Fe-4S] cluster bridges the two subunits and contains a nucleotide-binding site in each subunit.[29] During catalysis, the MoFe and Fe protein transiently associate, where the Fe protein delivers an electron to the MoFe protein, coupled to the hydrolysis of MgATP.[30–32] Complex formation between the MoFe protein and the Fe protein is followed by a conformationally-gated one electron transfer from the P-cluster to the FeMo-cofactor (Figure 1d).[33] The Fe protein then transfers one electron from its [4Fe-4S] cluster to the oxidized P-cluster (P^{+1}) in what has been described as a "deficit-spending" electron transfer process (Figure 1d).[34]

The V-nitrogenase has been purified from *A. vinelandii* and *A. chroococcum.*[24,35] The α- and β-subunits of the VFe protein is encoded by *vnfD* and *vnfK* respectively.[35,36] The VFe protein

has a FeV-cofactor at the active site and a P-cluster, which is involved in electron transfer to the active site.[24,35,36] The Fe protein component is encoded by *vnfH* and is distinct, although it is thought to be structurally similar to the Fe protein from the Mo-nitrogenase.[22,24] The recent crystal structure of the VFe protein shows that the VFe protein is a $\alpha_2\beta_2\gamma_2$ heterohexamer (Figure 1b).[35] It retains the structural core of the MoFe protein except for the gamma subunit, encoded by *vnf*G, which is attached to the α-subunit (Figure 1b).[35] The exact role of the gamma subunit is unknown, however it may play an indirect catalytic role or be involved in transferring the FeV-cofactor to the apoenzyme.[22,35,37] The P-cluster in the VFe protein resembles the P-cluster in the MoFe protein. There are differences between the FeMo-cofactor and the FeV-cofactor at the respective active sites (Figure 1a and 1b).[35] The FeV-cofactor resembles the FeMo-cofactor except the homocitrate moiety serves as a ligand to V instead of Mo (Figure 1a and 1b). The FeV-cofactor is slightly elongated due to the V-Fe distances, however the octahedral ligand environment gives distances that are almost indistinguishable from the FeMo-cofactor. The largest difference between the two active sites is the replacement of one of the belt sulfides with a ligand that is proposed to be carbonate (Figure 1b). The harder carbonate oxygen replacing the weaker bridging sulfide ligand enables extension of the cluster and disturbs the electronic structure, which could have an effect on the catalytic activity of the cluster.[35]

The Fe-only nitrogenase has been isolated from *A. vinelandii*, *R. rubrum*, and *Rhodobacter capsulatus*.[23,24] The α- and β-subunits of the FeFe protein is encoded by *anfD* and *anfK* respectively.[22,23] The FeFe protein has a proposed FeFe-cofactor at the active site and a P-cluster, which is involved in electron transfer to the active site. [23,24] The Fe protein component is encoded by *anfH* and is also distinct, and thought to be structurally similar to the Fe protein from the Mo-nitrogenase from spectroscopic studies.[22–24] The FeFe protein also includes the additional

gamma subunit, encoded by *anfG*, and is proposed to form a $\alpha_2\beta_2\gamma_2$ heterohexamer as seen in the VFe protein (Figure 1c).[23] There is no structural information for the FeFe protein, although spectroscopic studies propose the structures for the P-cluster and FeFe-cofactor to be analogous to the MoFe and VFe proteins respectively (Figure 1c).[22,23,38]

2-2 Catalytic activities of nitrogenase systems

Despite the similar features between all three nitrogenases, they differ greatly in their catalytic activities.[23,39] The Mo-nitrogenase reduces N_2 to ammonia following a limited stoichiometry of: $N_2 + 8H^+ + 16MgATP + 8e^- \rightarrow 2NH_3 + H_2 + 16MgADP + 16P_i$ as determined by Lowe and Thorneley.[40] Even at high N_2 pressures of 50 atm, the ratio of molecules of N_2 fixed to molecules of H_2 evolved remains 1:1 and approximately 27% of the total electron flux is directed towards H_2 evolution. [41] The alternative nitrogenases are less efficient at reducing N_2 than the Mo-nitrogenase.[17,19,36,39] More electron flux is diverted to H_2 production for the V-nitrogenase from *A. chroococcum*, and even under high N_2 pressures, the minimum stoichiometry for N_2 reduction is: $N_2 + 14H^+ + 40MgATP + 12\ e^- \rightarrow 2NH_4^+ + 3H_2 + 40MgADP + 40P_i$. [17,42] The Fe-only nitrogenase also exhibits a higher H_2 production activity in enzymes isolated from *A. vinelandii* and *R. capsulatus.*[23,43] The ratio of N_2 fixed to H_2 evolved is 9:1 for the enzyme from *R. capsulatus*, which is close to the reported ratio of 7:1 from *A. vinelandii.* [23,43] This difference in the ratio of N_2 fixed to H_2 evolved is accounted by the five-fold higher K_m for the Fe-only nitrogenase compared to the Mo-nitrogenase from *A. vinelandii* and therefore, higher N_2 pressures are required to outcompete H_2 production.[23]

N_2 activation at the FeMo-cofactor has been reported to follow a reversible reductive elimination/oxidative addition mechanism for both the Mo- and Fe-only nitrogenases with the formation of two Fe-H-Fe bridging hydrides as being key for producing H_2 and allowing N_2 to

bind.[23] While the V-nitrogenase does reduce N_2, the catalytic mechanism for N_2 reduction is unknown.[19,22,23] A key feature of the mechanism for N_2 reduction is the ability of Mo-nitrogenase to capture a substrate for multiple rounds of reduction with no observed intermediates.[39] This is important for understanding how the Mo-nitrogenase can reduce N_2 and other carbon-containing compounds. Early studies found that the Mo-nitrogenase could also reduce acetylene, C_2H_2, by two protons and two electrons to produce ethylene, C_2H_4, which contributed to the discovery that nitrogenase could reduce other alternative substrates besides N_2 and protons.[24,39] The V- and Fe-only nitrogenase can also reduce acetylene to ethylene. However, the alternative nitrogenases have been shown to release four electron-reduced products such as ethane during acetylene reduction and hydrazine during N_2 reduction for the V-nitrogenase.[24,39] Substitution of amino acids, Gln191 and His195 around the FeMo-cofactor have produced MoFe protein variants that can produce measurable ethane from acetylene reduction.[39,44] These results indicate that the amino acid environment around the cofactors, as well as the identity of the heterometal, can control product distribution during substrate reduction.[39,44] Table 1 compares the specific activities for N_2, H_2, C_2H_2, between all three enzyme systems isolated from *A. vinelandii*.

Known alternative substrates for nitrogenase have common themes in that they are generally small C-containing molecules with multiple bonds.[39] A known substrate for all three nitrogenases which has gained considerable interest for nitrogenase reduction is CO_2.[31,39,45–47] Recently *in vitro* and *in vivo* studies demonstrated that the Fe-only nitrogenase from *Rhodopseudomonas palustris* has the highest CO_2 to CH_4 reduction of the three enzyme systems.[46,48] The ability to carry out multiple electron reduction reactions of relatively inert molecules such as the reduction of N_2 to NH_3 or CO_2 to CH_4 is unique to nitrogenase.[39] The thermodynamics for the expected reaction pathways for N_2 and CO_2 reduction are strikingly

similar, where the standard Gibb's free energy change for first step is the largest and is followed by subsequent exergonic steps.[39] The ability to reduce other substrates is a product of a redox enzyme that is able to overcome the largest activation barrier in biology (+251 kJ/mol for N_2 compared to +43 kJ/mol for CO_2 reduction).[39,49]

3- Genetics of nitrogen fixation

3-1-Characterization of nitrogen fixation genes

The inventory of genes required for N_2-fixation catalyzed by Mo-nitrogenase varies amongst species, and is dependent upon the environmental niche and physiology of the host. In the genomes of N_2-fixing bacteria, including free-living (*K. pneumoniae*, *A. vinelandii*, *Methanococcus maripaludis*, *Paenibacillus sp.*), facultative- (*Frankia*) or obligated- symbionts (*Rhizobium meliloti, Bradyrhizobium japonicum*, and *Azorhizobium caulinodans*) symbionts, *nif* genes are found as several clusters. For example, in *K. pneumoniae* twenty *nif* genes are co-located within a 24 kb cluster,[50] whereas in *A. vinelandii* the *nif* genes are more dispersed and distributed as two clusters in the genome (Figure 2).[51] The major *nif* cluster contains the *nifHDKTYENX*, *iscA*^{nif}, *nifUSV*, *cysE*, *nifWZM*, *clpX*, and *nifF* genes and other interspersed open reading frames (Figure 2). The minor *nif* cluster contains the *nifLAB*, *fdxN*, *nifOQ*, *rhdN*, and *grx5* genes in one transcriptional direction (Figure 2). Other N_2-fixing organisms possess a more restricted *nif* gene set, for example, *Frankia* sp. contains 12 *nif* genes (*nifHnifDnifKnifEnifNnifXnifWnifZnifBnifUnifSnifN*)[52]; *R. meliloti*, *B. japonicum*, and *A. caulinodans* contain at least 9 *nif* genes (*nifHnifDnifKnifnifEnifNnifBnifSnifWnifXnifA*)[53]; and *M. maripaludis* and *Paenibacillus sp.*, contain 9 *nif* genes

(*nifBnifHnifDnifKnifEnifNnifXhesAnif*V).[54] Analysis of the distribution of *nif* genes sequences within microbial genomes indicates that nearly all diazotrophs have a minimal set consisting of six conserved genes *nifH, nifD, nifK, nifE, nifN*, and *nifB* (Figure 2). This concurs with the minimal catalytic core required to assemble the FeMo-cofactor *in vitro*.[1,55–57] Apart from the structural subunits encoded by *nifHDK*, several genes are required for the biosynthesis of the metalloclusters, in addition to other gene products necessary to produce a fully functional enzyme. the *nifE*, *nifN*, and *nifB* genes are required for synthesis of FeMo-cofactor, an essential component of the enzyme for substrate binding and reduction. It is also well established from genetic and biochemical analysis that *nifX*, *nifQ*, *nifV*, and *nifY* contribute to the synthesis and insertion of the FeMo-cofactor into the MoFe protein. The genes *nifU nifS* and *nifZ* play an important role in synthesis of metalloclusters, and that *nifM* and *nifW* are required for proper folding of the nitrogenase Fe protein (Figure 2).[1,55–57] The majority of the remaining genes of the *nif* clusters encode proteins important for FeMo-cofactor and P cluster biosynthesis and maturation of the nitrogenase or gene products necessary for optimal diazotrophic growth in *A. vinelandii* such as *iscA*^{nif}*, orf8*, *nifZ*, *clpX, nifF* (Figure 2).[84] The products of the *nifLA* operon are involved in the regulation of expression of all the other *nif* genes[58], in *K. pneumoniae* as well as in *A. vinelandii*.

The structural genes encoding dinitrogenase and dinitrogenase reductase of the V-nitrogenase have been described for *A. chroococcum*, *A. vinelandii*, and *R. palustris*.[17,59–61] In contrast to the single operon *nifHDK* encoding the subunits for Mo-nitrogenase, the genes encoding V-nitrogenase proteins are split between two operons (Figure 2). The *vnfH* gene encodes the dinitrogenase reductase subunits and is part of a two-gene operon. The *vnfDGK* operon, located downstream of the *vnfH* containing-operon, encodes the subunits for

dinitrogenase. The *vnfD* gene encodes the α-subunit, *vnfK* encodes the β-subunit and the *vnfG* encodes the γ-subunit. This gene does not have a counterpart in the Mo-containing nitrogenase system; however it does share some sequence similarity to the predicted translation product of the *anfG* gene. The structural genes encoding Fe-only nitrogenase in *A. vinelandii* and *R. palustris*, are organized as a single operon *anfHDGK* (Figure 2).[61–63] The subunits of dinitrogenase reductase are encoded by *anfH*, whereas the α- and β-subunits of dinitrogenase are encoded by *anfD* and *anfK*, respectively. The *anfG* gene encodes a third subunit (γ) for dinitrogenase, and is required for growth under Mo- and V- deficient conditions.[37] Other gene products that are required for active V-nitrogenase and Fe-only nitrogenase are encoded by the operon *vnfENX*, which is located downstream of *vnfH*. These genes product are homologous to the *nifENX* gene products that are required for FeMo-cofactor synthesis.[64] The regulatory *vnfA* and *anfA* genes encode transcriptional activators which are required for the transcription of the *vnf* and *anf* operons, respectively. The *vnfA* gene is located immediately upstream from *vnfENX* operon, and the *anfA* gene is located immediately upstream *anfHDGK*. In addition, *vnfY* gene located downstream from *vnfK* in *A. vinelandii* genome, is required for effective V-dependent diazotrophic growth and appears to play a role in the maturation of the FeV-cofactor and/or its insertion into apodinitrogenase of V-nitrogenase.[65] The *anf* operon from *A. vinelandii* also contains two other genes *anfO* and *anfR*. The predicted protein products of the *anfO* and *anfR* genes do not show overall similarity to any *nif* gene product; however, both genes are required for Fe-only nitrogenase dependent diazotrophic growth.[63] Some of the cofactor synthesis and assembly proteins encoded by the *nif* gene cluster (*nifB, nifU, nifS, nifV, and nifM*) participate in alternative nitrogenase cofactor synthesis and assembly.[9,66,67] Their common requirement suggests that the three enzyme systems share several common features.

3-2- Regulation of nitrogenase systems: metal availability, regulation by NH_4^+, transcriptional activators

Synthesis and activity of all types of nitrogenases are regulated in response to (1) ammonium ions, (2) the C/N ratio, (3) oxygen conditions, (4) the presence or the absence of Mo and V, and (5) the darkness and light in the case of phototrophic bacteria.

Regulation of the nitrogenase systems by metal ion availability is a particularly interesting phenomenon. Although the Mo-nitrogenase component proteins are not regulated by Mo in *K. pneumoniae* and, hence, are formed constitutively under N_2-fixation conditions,[68] in *A. vinelandii*, Mo is absolutely required for the expression of the structural genes of the Mo-nitrogenase system.[69,70] Simultaneously, Mo represses the synthesis of both alternative nitrogenases found in *A. vinelandii*, the V-nitrogenase and the Fe-only nitrogenase (Figure 3).[70] The Fe-only nitrogenase system is additionally repressed by V when Mo is absent (Figure 3). In *R. capsulatus*, the alternative Fe-only nitrogenase is solely repressed by Mo.[71–73] The expression of alternative nitrogenases in *R. palustris* is not repressed by transition metals. Strains that were unable to express a functional Mo-nitrogenase due to mutations in Mo-nitrogenase structural genes synthesized functional V- and Fe-only nitrogenases, and expressed *vnf* and *anf* genes in N_2-fixing growth media that contained Mo and V at concentrations far in excess of those that repress alternative nitrogenase genes expression in other bacteria.[61] Synthesis of the Fe-only nitrogenase of *R. rubrum* required completely different growth conditions and does not appear to be metal regulated. This enzyme system is not repressed by Mo but is expressed whenever a strain lacks an active Mo-nitrogenase as a consequence of either physiological or genetic inactivation.[74]

The synthesis of all three nitrogenases is repressed by NH_4^+ in the free-living diazotroph *A. vinelandii*. Ammonium regulates V-nitrogenase at the post-transcriptional level, whereas both Mo- and Fe-only nitrogenases are repressed at the transcriptional level.[75] The regulatory genes *nifA*, *vnfA*, *anfA*, encode transcriptional activators, which are required for the transcription of the *nif*, *vnf*, and *anf* operons, respectively.[62,76] *A. vinelandii* requires NifA activator and the presence of Mo for the synthesis of Mo-nitrogenase (Figure 3). The *vnf* and *anf* operons are regulated by V and Mo and appears to be mediated through the effect of metal concentration on the transcriptional expression of the *vnfA* and *anfA*. Transcription of *vnfA* is not repressed by NH_4^+ but is repressed by Mo, whereas NH_4^+, V, and Mo repress expression of *anfA*. Mo acts by regulating transcriptional regulators of the alternative nitrogenases and does so by repressing the transcription of *vnfA*, as well as *anfA*. Additionally, it has been shown that a $VnfA^-$ strain can synthesize the Mo-nitrogenase in Mo-deficient conditions. Repression of Mo-nitrogenase structural genes in the absence of Mo is therefore mediated by VnfA, which under the same condition activates V-nitrogenase gene transcription. VnfA seems also to be involved, albeit indirectly, in mediating repression of Fe-only nitrogenase by V,[62,70,77] and surprisingly, NifA is required for full transcriptional expression of the *anfHDGKOR* operon.[77] Regulation of the alternative nitrogenases by V is not, however, a matter of switching between synthesis of V-nitrogenase and Fe-only nitrogenase. The genes encoding V-nitrogenase system are split into two operons, *vnfHFd* and *vnfDGK*, dinitrogenase reductase and dinitrogenase can be then expressed independently. Unlike dinitrogenase of V-nitrogenase, dinitrogenase reductase of V-nitrogenase is not only expressed under diazotrophic conditions in the presence of V but also in the absence of Mo and V, conditions under which Fe-only nitrogenase is expressed.[76,78] V regulates Fe-only nitrogenase synthesis by repressing transcription of its structural genes. However, this repression

is impaired in cells lacking the structural genes for the VFe protein of the V-nitrogenase.[70,75] This suggests that *A. vinelandii* represses Fe-only nitrogenase synthesis by sensing the presence of an active V-nitrogenase reductase, using this unique vanado-enzyme end-product, rather than V itself, as a regulatory signal. Also, mutations in the genes that encode the high affinity Mo-transport system (*modABCD* and *modE*) permit expression of the Mo-independent nitrogenases in the presence of low, but not high, concentrations of Mo.[79,80] ModE, repressor of the *modABCD* operon, appears to play a role in the repression of *anfA* transcription by Mo.[81] Thus, the genes involved in transport may also play a role in the ability to sense intracellular concentrations of Mo.

By contrast, expression of N_2-fixation genes in symbiotic diazotrophs is regulated predominantly in response to the cellular oxygen conditions. The lack of a strict nitrogen control mechanism probably relates to the physiological conditions under which symbiotic N_2-fixation takes place. *R. meliloti*, *B. japonicum*, and *A. caulinodans*, use largely identical regulatory elements (FixL, FixJ, FixK, NifA, RpoN); however these are integrated into different, species-specific networks. Two regulatory cascades which are present in all three species control different groups of target genes. The cascade including NifA controls expression of the nitrogenase structural genes and genes encoding accessory functions by means of their RpoN-dependent -24/-12 promoters, analogous to the situation in *K. pneumoniae*. On the other hand, the FixLJ and FixK proteins constitute a distinct regulatory cascade which may be specific for symbiotic diazotrophs since no functionally homologous proteins have been found in free-living diazotrophs.[53] The facultative symbiotic diazotroph, *Frankia*, fixes N_2 under both free-living and symbiotic conditions. The expression of *nif* genes in *Frankia* species is believed to be induced in response to low concentrations of fixed nitrogen and low oxygen pressure within the nodule

environment as well as those in soil under non-symbiotic conditions. Interestingly, Frankia genomes[82] do not contain many of the key regulators of the *nif* genes found in *Proteobacteria* (*nifA*, *nifL*, and *rpoN*)[2], indicating that the mechanisms responsible for NH_3-responsive transcriptional regulation differ from these well-studied bacteria.

3-3-Transcriptome profiles of nitrogen fixation

Global transcriptional profiling studies of N_2-fixation have been reported in different diazotrophic bacteria.[83–86] Several gene clusters were found to exhibit a significant change in expression under N_2-fixation conditions and revealed the discrete set of genes associated with each of the nitrogenase systems. These genes can be divided into two broad categories; genes directly related to N_2-fixation, and genes indirectly related to N_2-fixation.

Specific N_2-fixation genes: Up-regulation of most *nif* genes, which are directly involved in the synthesis, maturation, and function of nitrogenase, were found to be up-regulated in *Azoarcus* sp. strain BH72 and *P. stutzeri* during Mo-dependent diazotrophy.[5,84] In *Paenibacillus s*p. WLY78, the expression of the nine genes of the *nif* cluster (*nifBnifHnifDnifKnifEnifNnifXhesAnif*V) exhibited a significant transcript increase ranging from 150- to 1039-fold under the N_2-fixation condition. Furthermore, the transcript abundances varied greatly among these nine *nif* genes although all of them were co-transcribed from a common promoter. The different abundances of these co-transcribed genes suggest that these transcripts have different processing and segmental stabilities or that their synthesis may be regulated by transcriptional attenuation or processing, similar to the results found in *A. vinelandii*. However, the transcriptional profile of *Paenibacillus s*p. WLY78 is a little different from that of *A. vinelandii* where the transcript level of *nifH* (143-fold) were much higher than those of *nifD*, *nifK* (54-fold), and the other remaining genes of the

major *nif* cluster (2 to 24-fold) (Figure 2).[86] The genes of the minor *nif* cluster are co-transcribed (Figure 2); a large increase in the number of transcripts mapping *nifB*, *fdx*, *nifO* and *nifQ* was detected (12- to 18-fold), and a smaller increase in those for *rhd* and *grx5* was noted (4- to 6-fold). Low *nifL* and *nifA* transcript levels were detected under N_2-fixation and non-N_2-fixation conditions.[84]

Under V-dependent diazotrophy, large increases in the *vnfHF*, *vnfDGK*, *vnfEN*, *vnfX* and *vnfY* transcript levels were observed in *A. vinelandii*.[86] Transcripts of the *nif*-encoded cluster biosynthesis machinery, such as *nifUSV*, *nifM*, and *nifB*, also increased during V-dependent diazotrophy. Transcripts mapping to the *anfHDGK* structural genes increased in cultures dependent on the Fe-only nitrogenase for diazotrophic growth.[86] Large increases in *vnf*-encoded *H*, *E*, *N*, *X*, and *Y* transcripts were observed during heterometal-independent diazotrophy. In addition, an increase in *vnfH* transcripts and *nifH* transcripts was also noted. Interestingly, the transcript levels of *nifA*, *vnfA* and *anfA* increased under both V-dependent and heterometal-independent diazotrophic growth, and decreased in the presence of Mo.[86]

Non-specific N_2-fixation genes: In addition to the specific N_2-fixation genes, several of the gene clusters that were differentially regulated under N_2-fixation conditions encoded proteins that were not obviously related to N_2-metabolism but with completely different functions (Figure 4).

*Transport and secretion***:** Non-*nif* genes specifically required during Mo-dependent diazotrophy, such as *mod*, *feoAB* and *cys* encoding transporters of Mo, Fe, and S atoms were coordinately transcribed with *nif* genes in N_2-fixing conditions in *Paenibacillus* sp. WLY78.[85] Similarly, transcript levels of Mo transport system increased during Mo-dependent diazotrophy in *A. vinelandii*.[86] Similar increases in the level of transcripts mapping to Mod transport systems were

also noted in cultures expressing the alternative nitrogenases.[86] In *Azoarcus* sp. strain BH72, genes encoding for proteins related to the molybdopterin biosynthesis exhibited enhanced expression under N_2-fixation,[84] and gene clusters encoding for components of a type II and type VI secretion systems were strongly upregulated as well.

Iron-sulfur cluster assembly system: In *A. vinelandii*, the stimulation of *nifU* and *nifS* gene expression has been reported to be accompanied by a decrease in the expression of the iron sulfur cluster (ISC) system during Mo-dependent diazotrophy.[86] Interestingly, a decrease in the ISC machinery was also noted during both V-dependent and heterometal-independent diazotrophic growth. The genome of *Paenibacillus* sp. WLY78 does not have *nifUS*, but contains a complete *suf* (*sufCBSUD*) operon, a partial *suf* (*sufABC*) operon, a partial *isc* system (*iscR* and *fdx*), and two *nifS*-like genes. The transcript abundances of *sufCBSUD* were high under N_2-fixing condition, and the expression of the partial *suf* (*sufABC*) operon, partial *isc* system (*iscSR* and *fdx*), and two *nifS*-like genes were also induced by N_2-fixing condition.[85]

Respiratory and energy metabolism: In *A. vinelandii*, transcripts mapping to the genes coding for cytochrome *bd* oxidase I and the type II NADH-dependent oxidoreductase Ndh increased during Mo-dependent, V-dependent, and heterometal-independent diazotrophic growth. Small increases in transcript abundance for two other terminal oxidases, cytochrome cbb_3 oxidase and cytochrome *c* oxidase, were also observed during Mo-dependent diazotrophy.[86] In *Paenibacillus* sp. WLY78, 20 cytochrome oxidase genes (*qoxABCD*, *ctaCDEF*, *cydAB* and others) have been found to be induced under N_2-limitation conditions. The three *ndh* genes encoding NADH dehydrogenase, and the *hcaD* gene encoding NAD(FAD)-dependent dehydrogenase were also

up-regulated.[85] The gene encoding cytochrome *cc42* was found up-regulated in *Azoarcus* sp. strain BH72 as well.[84] Transcripts levels of ATP synthetases exhibited enhanced expression under N_2-fixation conditions in *A. vinelandii*, *Paenibacillus* sp. WLY78 and *Azoarcus* sp. strain BH72. In addition, genes encoding assimilatory nitrite reductases (*nar* genes) and NAD(P)H-nitrite-reductases (*nas* genes) were also found up-regulated in *Paenibacillus* sp. WLY78.

Electron transport: In *A. vinelandii*, the *nif* system encodes a specific flavodoxin encoded by the *nifF* gene. Although *nifF* is not required for Mo-dependent diazotrophy, the transcript level of *nifF* increased significantly during Mo-dependent diazotrophy. Three clusters of genes, *rnf1*, *rnf2*, and *fix*, also encode electron transport systems that are thought to provide reducing equivalents to nitrogenase in *A. vinelandii*. Both *rnf1* and *fix* transcript levels increased markedly during diazotrophic growth. The increases in transcripts from *rnf1* and *fix* operons were similar in magnitude during Mo-dependent diazotrophy (40-fold), and levels of transcripts from *rnf2* operon were relatively unchanged. In contrast, during V-dependent or heterometal-independent diazotrophy, *fix* operon transcript levels were higher than *rnf1* operon transcript levels (231-fold versus 18-fold).[86] In *Paenibacillus* sp. WLY78, the *fer* encoding ferredoxin, the *fldA* encoding flavodoxin and the *nfrA* encoding NAD(P)H-flavin oxidoreductase, were highly transcribed under N_2-fixing conditions.[85]

Nitrogen and carbon metabolism: In response to N_2-fixation, the transcription of *glnA* (glutamine synthetase), *glnP* and *glnH* (glutamine transporters) were enhanced in *Azoarcus* sp. strain. In contrast, the transcription of denitrification genes (*napD1* and *napE)*, as well as a gene for NO_2^- assimilation (*nirB*) were strongly repressed.[84] Similar results also were found in

Paenibacillus sp. WLY78.[83] The transcript level of the *glnRA* operon was up-regulated in N_2-fixing conditions and the other two *glnA* genes (*glnA*1, *glnA2*), *amtB*, *gltABD* involved in N_2 metabolism was significantly up-regulated.[85] Interestingly, expression of *hoxB* gene encoding for the small subunit of hydrogenase was enhanced in *Azoarcus* sp. strain BH72 under N_2-fixation condition.[84] Transcript levels also increased for the well characterized membrane-bound [NiFe] hydrogenase (*hoxVTRQOLMZGK*) in *A. vinelandii* during Mo-dependent diazotrophy.[86] Notable was an increase in the levels of transcripts for the genes encoding an uncharacterized soluble [NiFe] hydrogenase and the genes of the *hut* operon under diazotrophic conditions in the absence of Mo. Finally, N_2-fixation in *Azoarcus* sp. strain BH72 was accompanied by down-regulation of the expression of several of the Embden-Meyerhof pathway and tricarboxylic acid cycle enzymes. The transcription of the components of PDH complex as well as components of ethanol oxidation pathway producing acetyl-coA was reduced. Expression of some enzymes of the β-oxidation of the fatty acid pathway generating acetyl-CoA as end product, were also found to be transcriptionally repressed under N_2-fixation conditions.[84]

Secondary metabolism: In *A. vinelandii*, one of the largest increases in transcript abundance observed under Mo-dependent diazotrophic growth occurred for genes that encode type IV pili. Interestingly, the large increase in transcripts for pili formation was not observed during Mo-independent diazotrophic growth.[86] In *Azoarcus* sp. strain BH72, enhancement of expression of flagella and pili genes under diazotrophic growth was reported.[84]

Translation and transcription: Genes coding for the small subunit of ribosomal protein as well as transcription factors, translation initiation factors, and elongation factors, and DNA directed

RNA polymerase subunits were found modulated in their expression in response to N_2-fixation conditions in *Azoarcus* sp. strain BH72, *P. stutzeri*, *A. vinelandii, and Paenibacillus* sp. WLY78.[5,84–86]

4-Taxonomic distribution of nitrogenase genes in sequenced genomes and environmental diversity of nitrogenase systems

Nitrogenase is the only source of biotic ammonia in the environment but is only present in a subset of prokaryotes. This might be due to the complex and intricate process of BNF which leads to low gene accruement success rate in prokaryotes. By studying these required genes and their alternative metal counterparts, the taxonomic distribution of *nif*, *vnf*, and *anf* can be distinguished.

In all the prokaryote species with full genome sequences, 92 species contain coding sequences similar to NifD and NifH, 67 of which met the criteria for the minimum gene set (*nifHDKENB*).[87] The lack of abundance and distribution in the ecosystem could be due to the underrepresentation of full sequences in the database. The distribution of the alternative nitrogenase genes is infrequent throughout genome databases. Most N_2-fixing prokaryotes encode only a Mo-nitrogenase and prokaryotes that contain alternative nitrogenase(s), V-nitrogenase and/or Fe-only nitrogenase, always contain Mo-nitrogenase (*nif* operon).[87–89] In addition, some bacteria, including the purple non-sulfur phototrophs *R. capsulatus* and *R. rubrum* as well as various cyanobacteria and *C. pasteurianum* encode either Fe-only nitrogenase or V-nitrogenase. *A. vinelandii*, *Methanosarcina acetovorans*, and *R. palustris* has been shown to have genes and activity for all three different nitrogenases.[61] The alternative nitrogenases have been harder to distinguish in environmental studies due to the target of these surveys being NifH by

PCR amplicon sequencing. NifH does not contain metal cofactor differentiation and shows high similarity to *vnf*/*anfH*, rather than *nif*/*vnf*/*anf D* or *K*, which would show more diversity in genes. Even with this bias in environmental nitrogenase surveys, multiple other techniques show that alternative nitrogenase plays a role in the environment. Alternative nitrogenases have been found in a wide variety of diazotrophs within the phylums of Bacteriodetes, Chlorobi, Cyanobacteria, Firmicutes, Proteobacteria, and Verrucomicrobia. In addition, organisms with alternative nitrogenases have been isolated from lichen symbionts,[90] termite gut microbiota,[91] and woodchip, soil and mangrove marsh samples.[92]

With the onset of next-generation amplicon sequencing and metagenomic sequencing of environmental samples a better picture of the distribution of nitrogenase in the bacterial kingdom. There has been evidence of distinct distribution of diazotrophs across differing habitats allowing multiple bacterial groups to fill the ammonia-producing niche in the environment.[93,94] The trace metal content of Mo, V, and Fe influences the usage and presence of alternative nitrogenases. Environments with higher Mo concentration have lower *vnfD* and *anfD* amplicon counts and alternatively where Mo concentrations are low *vnfD* and *anfD* amplicons counts are higher.[93] With the onset of better sequencing technology, more diversity and complexity will begin to appear in the diazotrophic communities.[95,96]

5- Origin and evolution of biological nitrogen fixation

Evidence in geological and biological records establishes BNF starting as early as 3.2-3.5 Gya and is seen as a requirement to support the growing biomass in the Archean sea.[97,98] The evolutionary pressure for the development of nitrogenase comes from the ever expanding biosphere and the steady state production of reduced nitrogen from abiotic sources. The majority of the reduced abiotic nitrogen in the Archean oceans was produced from lighting and volcanic

sources on the order of 10^{10}-10^{11} mol of N/yr.[99] This lead to the onset of the nitrogen crisis in which abiotic nitrogen fluxes cannot account for the large and diverse biomass by 3.5 Gya.[100,101] In the reducing environment of Archean oceans, Mo-availability was decreased compared to modern levels, as it was captured in sedimentary sulfides.[102–104] This anoxic reducing ocean environment was also very rich in more soluble ferrous iron (Fe^{2+}) compared to the less bioavailable ferric iron (Fe^{3+}).[89,104] The reducing conditions of the Archean sea leads to low Mo bioavailability and high Fe bioavailability for Archean life. From these conditions, a logical theory has been presented that the alternative nitrogenases preceded the Mo-nitrogenase in the evolutionary lineage.[89,105,106] In recent years, this theory has shown little traction, and new phylogenetic and isotopic data further indicates that Mo-nitrogenase is ancestral to the alternatives.

Phylogenetic evidence shows the origin for nitrogenases in anaerobic and hydrogenotrophic methanogens and shows the vertical and lateral distribution in archaea and bacterial species.[97] The transfer of nitrogenase genes from archaea to the bacteria most likely occurred between a methanogen and an anaerobic firmicute which would be in proximity in the environment.[97,107] Earlier phylogenetic studies supported the theory that *anf* and *vnf* resulted in the dispersal and evolution of the more efficient *nif*, through gene duplication and gene transfer.[89] Though these early phylogenetic studies have proved the diversity of nitrogenase is a product of lateral gene transfer (LGT) and gene loss, they have failed to provide convincing evidence of the specific metal content of the "protonitrognease." [88,89,108] The dissimilarity between studies is most likely due to the significant homology difference between alternative nitrogenases and Mo-nitrogenase (~30% amino acid sequence identity).[89] By concatenating the core of nitrogenase (*nif*HDK) and performing a phylogentic reconstruction, a more clear and

well supported lineage agrees with observed metal content.[108] Within the concatenated phylogeny, two distinct lineages appear, one derived solely from hydrogenotrophic methanogens that branched at the base of the tree and the second comprised of more recently evolved NifHDK homologs from both bacterial and methanogen genomes.[97,108] Nested within the two lineages is a single lineage comprised of *vnf* and *anf* nitrogenase. Conjointly within the alternative nitrogenase lineage, *anf* is nested inside *vnf* suggesting that *anf* is derived from *vnf* (Figure 5).[108] A separate lineage comprised of biochemically uncharacterized nitrogenases appears in between the *nif* and *anf* lineages. Theses uncharacterized nitrogenase homologs are present in the genomes of anaerobic methanogens, anaerobic methanotrophic archaea, and anaerobic firmicutes.[109] By comparing conserved residues within the cofactor binding sites, the data suggest that the uncharacterized *nif* bind the FeMo-cofactor. These uncharacterized nitrogenase homologs have not been identified in aerobes or facultative anaerobes and therefore supports the concept that the alternative nitrogenase most likely diversified in an anoxic environment.[97,109]

All the organisms that contain the alternative nitrogenases also contain the Mo-nitrogenase genes. In addition, the gene clusters of the alternative nitrogenases lack the genes necessary for the biosynthesis of a complete co-factor (Figure 5). The transcriptional data show that *vnf* and *anf* require expression of multiple biosynthetic genes encoded in the *nif* operon.[86] The vital biosynthetic genes involved in cofactor maturation, *nifB* and *nifEN,* are missing in generally all alternative nitrogenase operons.[49] With the exception in *A. vinelandii* and *R. palustris* in which the copies of *vnfEN* are a product of a recent duplication event.[97] The reliance of the alternative nitrogenase on the biosynthetic genes of *nif* is shown in the duplication of the *vnf*/*anfHDK* from an original Mo-nitrogenase. The apparent lack of the biosynthetic genes is due

to the complexity of the metal-cofactors and the improbability for the alternative nitrogenase to develop a similar suite of genes independently.

The evolution of nitrogenase is a combined shaping of the gene's operon, protein structure, and maturation of the metal cofactors. A couple of theories have been presented to elucidate how evolution has overcome the highest activation barrier in biology. A theory suggests that proto-nitrogenase could have served a role as a detoxyase due to the lack of evolutionary pressure for fixed biotic N_2 before the nitrogen crisis.[100] This would mean that the development of the proto-metal active sites in the enzyme was mainly as a general reductase.[109] This theory has a grounding in that modern nitrogenases can reduce substrates such as cyanide, acetylene, and carbon monoxide.[57,100] But it seems that other detoxyase enzymes are more ancient than nitrogenase as well as nitrogenase's association with a more ancient methanogen porphyrin F_{430} biosynthesis enzyme (NflD). Nitrogenase and NflD share similar structural active sites, and presumably, the common ancestors were capable of binding similar porphyrins or metal clusters.[108,109] These metal clusters could have started as simple Fe-S clusters that acted as the electron donor during an oxidoreductase reaction. The specificity of the enzyme was then considerably pushed toward reduction of nitrogenase as the world biomass increased around 3.0-2.5 Gya.[49] The multiple substrates seen in modern nitrogenase is a not an artifact of ancient promiscuity but more likely just a side effect of evolving a redox enzyme that can overcome the triple bond of N_2.

The natural selection for the most efficient and specific nitrogenase involves the fine-tuning of its metal cofactors. The evolutionary path for a more specific cofactor can be seen in the FeMo-cofactor biosynthesis pathway in the *nif* operon (Figure 6C). NifB creates the early precursor of the FeMo-cofactor called NifB-co which has been proposed as the Fe6-S9-X core of

the FeMo-cofactor.[110,111] Without the scaffolding genes of *nifEN* the FeMo-cofactor is not able to associate with *nifDK*, and there is a build-up of the NifB-co.[112,113]The NifB-co has been shown to be active in the apo-nifDK enzyme and able to reduce acetylene and hydrogen but not nitrogen.[88] NifEN with the FeMo-cofactor has also been shown to be catalytically active but at a much lower rate.[114] This suggests that that metal cofactor maturation over time lead to a more specific and efficient nitrogenase. The complexity of the metal cofactor can only happen with the addition of biosynthetic genes. The operon duplication of *nifDK* to *nifEN* was required for the maturation of the FeMo-cofactor due to the high selection pressure of a more efficient nitrogen fixer. The operon duplication events for the alternative nitrogenase lead to the evolution of differentiated metal loading in the cofactor due to the environmental pressure of metal availability. The co-evolution of the *nifENB* with the nitrogenase enzyme reinstates the requirement the *nif* operon being the source of the alternative nitrogenase.

Recently isotopic differences in discrimination of alternative nitrogenase were used to provide geological data to support the theory that *anf* and *vnf* derived from *nif*.[98] Nitrogenases discriminate against heavy nitrogen isotopes leaving biomass isotopically lighter than the environment. Nitrogen isotopes ratios compare the ^{15}N ratio in the atmosphere and the ^{15}N ratio of the biological or geological sample [$\delta^{15}N‰ = (((^{15}N/^{14}N)_{sample}/(^{15}N/^{14}N)_{atmosphericN2} - 1)]$. The rate of discrimination of Mo-nitrogenase is down to -4‰ in laboratory experiments but on average is environmentally observed around -1‰.[106,115] The alternative nitrogenases discriminate even more against heavier N^{15} with isotope ratio differences from -6‰ to -8‰.[106] The discrimination of proto-nitrogenase cannot be known, but it is safe to assume their low affinity and promiscuity would favor small discrimination. The data showed a mean ratio of 0.0 +/- 1.2 $^{15}N‰$ from marine and fluvial sediment between 3.2 and 2.75 billion years ago. These ratios

cannot be due to abiotic sources and cannot support the discrimination differences of the alternative nitrogenases.[98,99] Though the data is clear that the discrimination in the reducing ocean ~3.5 billion years ago is due to Mo-nitrogenase, the N_2-fixation rate is still dependent on the Mo availability. The rate of N_2-fixation drastically decreases with the decrease of Mo in media, though when the concentration of Mo is below 5 nM nitrogenase activity is still measurable.[115] Even though Mo availability is limited during nitrogenase evolution in the archean sea local oxidative weathering or other unknown organically complexed Mo in the form of siderophores could have allowed for the uptake of Mo.[97,98] The limited amount of Mo still provides for the formation and evolution of the Mo-nitrogenase during the nitrogen crisis. The sedimentary data show that alternative nitrogenases most likely evolved after the nitrogen crisis and stayed prevalent due to the lack of Mo in microenvironments.

Conclusion

Our understanding of the BNF and nitrogen cycle has been expanded by extensive biochemical, genetic, environmental, phylogenetic, and structure-based characterizations of the three known nitrogenase systems. Mo-nitrogenase appears to give rise to the alternative nitrogenases, V- and Fe-only nitrogenases. The nature of the common ancestor of the nitrogenase enzymes and its associated bound metal cluster were likely controlled by the selective pressure imposed by fixed nitrogen limitation in combination with local environmental metal availability. Today, the majority of global N_2-fixation is attributed to the more efficient canonical Mo-nitrogenase, whereas V- and Fe-only nitrogenases are often considered “backup” enzymes, used when Mo is limiting. However, the environmental distribution and diversity of alternative nitrogenases suggest that these enzymes can make an important contribution to N_2-fixation in

response to constrains across global ecosystems, and that the trace metal controls on N_2-fixation may be more complex than previously thought.

Our understanding of this complex reaction mechanism for N_2 reduction, from the view-point of evolutionary and system complexity, can be exploited to develop new strategies for agricultural soil fertilization. The heavy use of chemical N fertilizers in agriculture, and the rising concerns over the possible harmful environmental effect of the fertilizers, as well as their cost for small-scale farmers in developing countries, have highlighted the importance of expanding the use of agricultural production methods that are both agronomically and economically sustainable. In this context, there is a significant interest in research on BNF and prospects for increasing its importance in an agricultural setting. Symbiotic N_2-fixing systems offer an economically attractive and ecologically sound means of reducing external inputs and improving internal resources. Symbiotic N_2-fixation is, however, restricted mainly to legumes in agricultural systems, and there is considerable interest in exploring whether similar symbioses can be developed in nonlegumes, which produce the bulk of human food.

Acknowledgements

J.W.P. thanks the National Science Foundation [grant number NSF-1331098] for supporting the work on gene expression and physiological studies. A portion of this work (structure function studies: principal investigators J.W.P. and L.C.S.) was supported as part of the Biological Electron Transfer and Catalysis Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0012518.

References

1 L. M. Rubio and P. W. Ludden, Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase, *Annu. Rev. Microbiol.*, 2008, **62**, 93–111.

2 R. Dixon and D. Kahn, Genetic regulation of biological nitrogen fixation, *Nat. Rev. Microbiol.*, 2004, **2**, 621–631.

3 L. S. Hartmann and S. R. Barnum, Inferring the evolutionary history of Mo-dependent nitrogen fixation from phylogenetic studies of *nifK* and *nifDK*, *J. Mol. Evol.*, 2010, **71**, 70–85.

4 T. L. Hamilton, E. S. Boyd and J. W. Peters, Environmental Constraints Underpin the Distribution and Phylogenetic Diversity of *nifH* in the Yellowstone Geothermal Complex, *Microb. Ecol.*, 2011, **61**(4), 860-70.

5 W. Lu, Y. Yan, S. Ping, J. Peng, Y. Han, L. Li, J. Yang, Y. Dou, Y. Li and H. Fan, Global transcriptional analysis of nitrogen fixation and ammonium repression in root- associated *Pseudomonas stutzeri* A1501 Global transcriptional analysis of nitrogen fixation and ammonium repression in root-associated *Pseudomonas stutzeri* A1501, *BMC Genomics*, 2010, **11**, 1–13.

6 J. W. Peters, K. Fisher and D. R. Dean, Nitrogenase structure and function: a biochemical-genetic perspective., *Annu. Rev. Microbiol.*, 1995, **49**, 335–366.

7 L. M. Rubio and P. W. Ludden, Maturation of Nitrogenase: a Biochemical Puzzle, *J. Bacteriol.*, 2005, **187**, 405–414.

8 L. C. Seefeldt, B. M. Hoffman and D. R. Dean, Mechanism of Mo-Dependent Nitrogenase, *Annu. Rev. Biochem.*, 2009, **78**, 701–22.

9 R. D. Joerger, P. E. Bishop and H. J. Evans, Bacterial Alternative Nitrogen Fixation Systems, *CRC Crit. Rev. Microbiol.*, 1988, **16**, 1–14.

10 H. Bortels, Further studies on the significance of molybdenum, vanadium, tungsten and other soil materials on nitrogen-fixing and other microorganisms, *Zentr. Bakt. Parasitenk*, 1936, **95**, 193–218.

11 W. A. Bulen and J. R. LeComte, The nitrogenase system from *Azotobacter*: two-enzyme requirement for N_2 reduction, ATP-dependent H_2 evolution, and ATP hydrolysis., *Proc. Natl. Acad. Sci. U. S. A.*, 1966, **56**, 979–86.

12 J. R. Chisnell, R. Premakumar and P. E. Bishop, Purification of a second alternative nitrogenase from a *nifHDK* deletion strain of *Azotobacter vinelandii*., *J. Bacteriol.*, 1988, **170**, 27–33.

13 G. D. Riddle, J. G. Simonson, B. J. Hales and H. D. Braymer, Nitrogen fixation system of tungsten-resistant mutants of *Azotobacter vinelandii*, *J. Bacteriol.*, 1982, **152**, 72–80.

14 P. E. Bishop, D. M. Jarlenski and D. R. Hetherington, Evidence for an alternative nitrogen fixation system in *Azotobacter vinelandii*., *Proc. Natl. Acad. Sci.*, 1980, **77**, 7342–7346.

15 P. E. Bishop, D. M. Jarlenski and D. R. Hetherington, Expression of an Alternative Nitrogen Fixation System in *Azotobacter vinelandii*, *J. Bacteriol.*, 1982, **150**, 1244–1251.

16 P. E. Bishop, R. Premakumar, D. R. Dean, M. R. Jacobson, J. R. Chisnell, T. M. Rizzo and J. Kopczynski, Nitrogen Fixation by *Azotobacter vinelandii* Strains Having Deletions in Structural Genes for Nitrogenase, *Science,* 1986, **232**(4746), 92-4.

17 R. L. Robson, R. R. Eady, T. H. Richardson, R. W. Miller, M. Hawkins and J. R. Postgate, The alternative nitrogenase of *Azotobacter chroococcum* is a vanadium enzyme, *Nature*, 1986, **322**, 388–390.

18 B. J. Hales, E. E. Case, J. E. Morningstar, M. F. Dzeda and L. A. Mauterer, Isolation of a New Vanadium-Containing Nitrogenase from *Azotobacter vinelandii*, *Biochemistry*, 1986, **25**, 7251–7255.

19 R. R. Eady and D. N. Prevent, Structure-function relationships of alternative nitrogenases, *Chem. Rev.*, 1996, **96**, 3013–3030.

20 B. K. Burgess and D. J. Lowe, Mechanism of Molybdenum Nitrogenase, *Chem. Rev.*, 1996, **96**, 2983–3012.

21 Y. Hu, C. C. Lee and M. W. Ribbe, Vanadium nitrogenase: A two-hit wonder?, *Dalt. Trans.*, 2012, **41**, 1118–1127.

22 M. A. Schneider K., Catalysts for Nitrogen Fixation. Nitrogen Fixation: Origins, Applications, and Research Progress (Iron-Only Nitrogenase: Exceptional Catalytic, Structural and Spectroscopic Features), Springer, Dordrecht, 1st edn., 2004.

23 D. Harris, D. Lukoyanov, S. Shaw, P. D. Compton, M. Tokmina-Lukaszewska, B. Bothner, N. L. Kelleher, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, The Mechanism of N_2 Reduction Catalyzed by Fe-Nitrogenase Involves Reductive Elimination of H_2, *Biochemistry*, 2017, acs.biochem.7b01142.

24 Y. Hu and M. W. Ribbe, Nitrogenase and homologs, *J. Biol. Inorg. Chem.*, 2015, **20**, 435–445.

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

26 M. K. Chan, J. Kim and D. C. Rees, The nitrogenase FeMo-cofactor and P-cluster pair: 2.2 A resolution structures, *Science (80-.).*, 1993, **260**, 792 LP-794.

27 O. Einsle, A. F. Tezcan, S. L. A. Andrade, B. Schmid, M. Yoshida, J. B. Howard and D. C. Rees, Nitrogenase MoFe-Protein at 1.16 Å Resolution: A Central Ligand in the FeMo-Cofactor, 2002, **297**, 1696–1700.

28 T. Spatzal, M. Aksoyoglu, L. Zhang, S. L. A. Andrade, E. Schleicher, S. Weber, D. C. Rees and O. Einsle, Evidence for interstitial carbon in nitrogenase FeMo cofactor, *Science (80-.).*, 2011, **334**, 940.

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

30 R. N. F. Thorneley, R. R. Eady and D. J. Lowe, Biological nitrogen fixation by way of an enzyme-bound dinitrogen-hydride intermediate, *Nature*, 1978, 272, 557–558.

31 Z.-Y. Yang, V. R. Moure, D. R. Dean and L. C. Seefeldt, Carbon dioxide reduction to methane and coupling with acetylene to form propylene catalyzed by remodeled nitrogenase, *Proc. Natl. Acad. Sci.*, 2012, **109**, 19644–19648.

32 R. V. Hageman and R. H. Burris, in *Current Topics in Bioenergetics*, Elsevier BV, 1980, pp. 279–291.

33 K. Danyal, B. S. Inglet, K. A. Vincent, B. M. Barney, B. M. Hoffman, F. A. Armstrong, D. R. Dean and L. C. Seefeldt, Uncoupling nitrogenase: Catalytic reduction of hydrazine to ammonia by a MoFe protein in the absence of Fe protein-ATP, *J. Am. Chem. Soc.*, 2010, **132**, 13197–13199.

34 K. Danyal, D. R. Dean, B. M. Hoffman and L. C. Seefeldt, Electron transfer within nitrogenase: Evidence for a deficit-spending mechanism, *Biochemistry*, 2011, **50**, 9255–9263.

35 D. Sippel and O. Einsle, The structure of vanadium nitrogenase reveals an unusual bridging ligand, *Nat. Chem. Biol.*, 2017, **13**, 956–960.

36 C. C. Lee, Y. Hu and M. W. Ribbe, Unique features of the nitrogenase VFe protein from *Azotobacter vinelandii*, *Proc. Natl. Acad. Sci.*, 2009, **106**, 9209–9214.

37 S. I. Waugh, D. M. Paulsen, P. V. Mylona, R. H. Maynard, R. Premakumar and P. E. Bishop, The genes encoding the delta subunits of dinitrogenases 2 and 3 are required for Mo-independent diazotrophic growth by *Azotobacter vinelandii*, *J. Bacteriol.*, 1995, **177**, 1505–1510.

38 E. Krahn, B. J. R. Weiss, M. Kröckel, J. Groppe, G. Henkel, S. P. Cramer, A. X. Trautwein, K. Schneider and A. Müller, 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, *J. Biol. Inorg. Chem.*, 2002, **7**, 37–45.

39 L. C. Seefeldt, Z. Y. Yang, S. Duval and D. R. Dean, Nitrogenase reduction of carbon-containing

compounds, *Biochim. Biophys. Acta - Bioenerg.*, 2013, **1827**, 1102–1111.

40 R. N. Thorneley and D. J. Lowe, Nitrogenase of *Klebsiella pneumoniae*. Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: identification of the rate-limiting step for substrate reduction., *Biochem. J.*, 1983, **215**, 393–403.

41 F. B. Simpson and R. H. Burris, A nitrogen pressure of 50 atmospheres does not prevent evolution of hydrogen by nitrogenase, *Science (80-.).*, 1984, **224**, 1095 LP-1097.

42 J. A. Rees, R. Bjornsson, J. K. Kowalska, F. A. Lima, J. Schlesier, D. Sippel, T. Weyhermüller, O. Einsle, J. A. Kovacs and S. DeBeer, Comparative electronic structures of nitrogenase FeMoco and FeVco, *Dalt. Trans.*, 2017, **46**, 2445–2455.

43 K. Schneider, U. Gollan, M. Dröttboom, S. Selsemeier-Voigt and A. Müller, Comparative biochemical characterization of the iron-only nitrogenase and the molybdenum nitrogenase from *Rhodobacter capsulatus*., *Eur. J. Biochem.*, 1997, **244**, 789–800.

44 K. Fisher, M. J. Dilworth, C.-H. Kim and W. E. Newton, *Azotobacter vinelandii* Nitrogenases with Substitutions in the FeMo-Cofactor Environment of the MoFe Protein: Effects of Acetylene or Ethylene on Interactions with H^+, HCN, and CN^-, *Biochemistry*, 2000, **39**, 10855–10865.

45 N. Khadka, D. R. Dean, D. Smith, B. M. Hoffman, S. Raugei and L. C. Seefeldt, CO_2 Reduction Catalyzed by Nitrogenase: Pathways to Formate, Carbon Monoxide, and Methane, *Inorg. Chem.*, 2016, **55**, 8321–8330.

46 B. Hu, D. F. Harris, D. R. Dean, T. L. Liu, Z.-Y. Yang and L. C. Seefeldt, Electrocatalytic CO_2 reduction catalyzed by nitrogenase MoFe and FeFe proteins, *Bioelectrochemistry*, 2018, **120**, 104–109.

47 N. S. Sickerman, K. Tanifuji, Y. Hu and M. W. Ribbe, Synthetic Analogues of Nitrogenase Metallocofactors: Challenges and Developments., *Chemistry*, 2017, **23**, 12425–12432.

48 Y. Zheng, D. F. Harris, Z. Yu, Y. Fu, S. Poudel, R. N. Ledbetter, K. R. Fixen, Z.-Y. Yang, E. S. Boyd, M. E. Lidstrom, L. C. Seefeldt and C. S. Harwood, A pathway for biological methane production using bacterial iron-only nitrogenase, *Nat. Microbiol.*, , DOI:10.1038/s41564-017-0091-5.

49 E. S. Boyd and J. W. Peters, New insights into the evolutionary history of biological nitrogen fixation, *Front. Microbiol.*, 2013, **4**, 1–12.

50 W. Arnold, A. Rump, W. Klipp, U. B. Priefer and A. Pühler, Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*, *J. Mol. Biol.*, 1988, **203**, 715–738.

51 J. C. Setubal, P. Dos Santos, B. S. Goldman, H. Ertesvåg, G. Espin, L. M. Rubio, S. Valla, N. F. Almeida, D. Balasubramanian, L. Cromes, L. Curatti, Z. Du, E. Godsy, B. Goodner, K. Hellner-Burris, J. A. Hernandez, K. Houmiel, J. Imperial, C. Kennedy, T. J. Larson, P. Latreille, L. S. Ligon, J. Lu, M. Mærk, N. M. Miller, S. Norton, I. P. O'Carroll, I. Paulsen, E. C. Raulfs, R. Roemer, J. Rosser, D. Segura, S. Slater, S. L. Stricklin, D. J. Studholme, J. Sun, C. J. Viana, E. Wallin, B. Wang, C. Wheeler, H. Zhu, D. R. Dean, R. Dixon and D. Wood, Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes, *J. Bacteriol.*, 2009, **191**, 4534–4545.

52 C. J. Oh, H. B. Kim, J. Kim, W. J. Kim, H. Lee and C. S. An, Organization of nif gene cluster in *Frankia* sp. EuIK1 strain, a symbiont of *Elaeagnus umbellata*, *Arch. Microbiol.*, 2012, **194**, 29–34.

53 H. M. Fischer, Genetic regulation of nitrogen fixation in Rhizobia., *Microbiol. Rev.*, 1994, **58**, 352–386.

54 L. Wang, L. Zhang, Z. Liu, D. Zhao, X. Liu, B. Zhang, J. Xie, Y. Hong, P. Li, S. Chen, R. Dixon and J. Li, A Minimal Nitrogen Fixation Gene Cluster from *Paenibacillus* sp. WLY78 Enables Expression of Active Nitrogenase in *Escherichia coli*, *PLoS Genet.*, 2013, **9**, 1–11.

55 G. P. Roberts, T. MacNeil, D. MacNeil and W. J. Brill, Regulation and characterization of protein products coded by the nif (nitrogen fixation) genes of *Klebsiella pneumoniae*, *J. Bacteriol.*, 1978, **136**, 267–279.

56 N. Gavini, S. Tungtur and L. Pulakat, Peptidyl-prolyl cis/trans isomerase-independent functional NifH mutant of *Azotobacter vinelandii*, *J. Bacteriol.*, 2006, **188**, 6020–6025.

57 Y. Hu, C. C. Lee and M. W. Ribbe, Extending the carbon chain: Hydrocarbon formation catalyzed by vanadium/molybdenum nitrogenases, *Science*, 2011, **333**, 753–755.

58 G. N. Gussin, C. W. Ronson and F. M. Ausubel, Regulation of nitrogen fixation genes, *Annu. Rev. Genet.*, 1986, **20**, 567–91.

59 R. D. Joerger, T. M. Loveless, R. N. Pau, L. A. Mitchenall, B. H. Simon and P. E. Bishop, Nucleotide sequences and mutational analysis of the structural genes for nitrogenase 2 of *Azotobacter vinelandii*, *J. Bacteriol.*, 1990, **172**, 3400–3408.

60 R. Raina, M. A. Reddy, D. Ghosal and H. K. Das, Characterization of the gene for the Fe-protein of the vanadium dependent alternative nitrogenase of *Azotobacter vinelandii* and construction of a Tn5 mutant, *Mol. Gen. Genet.*, 1988, **214**, 121–127.

61 Y. Oda, S. K. Samanta, F. E. Rey, T. Yan, J. Zhou and C. S. Harwood, Functional Genomic Analysis of Three Nitrogenase Isozymes in the Photosynthetic Bacterium *Rhodopseudomonas palustris*, *J. Bacteriol.*, 2005, **187**, 7784–7794.

62 R. D. Joerger, M. R. Jacobson, R. Premakumar, E. D. Wolfinger and P. E. Bishop, Nucleotide sequence and mutational analysis of the structural genes (*anfHDK*) for the second alternative nitrogenase from *Azotobacter vinelandii*, *J.Bacteriol.*, 1989, **171**, 1075–1086.

63 P. V. Mylona, R. Premakumar, R. N. Pau and P. E. Bishop, Characteristics of orf1 and orf2 in the *anfHDGK* genomic region encoding nitrogenase 3 of *Azotobacter vinelandii*, *J. Bacteriol.*, 1996, **178**, 204–208.

64 R. D. Joerger, E. D. Wolfinger and P. E. Bishop, The Gene Encoding Dinitrogenase Reductase-2 Is Required for Expression of the Second Alternative Nitrogenase from *Azotobacter Vinelandii*, *J. Bacteriol.*, 1991, **173**, 4440–4446.

65 C. Rüttimann-Johnson, L. M. Rubio, D. R. Dean and P. W. Ludden, VnfY is required for full activity of the vanadium-containing dinitrogenase in *Azotobacter vinelandii*, *J. Bacteriol.*, 2003, **185**, 2383–2386.

66 C. Kennedy, R. Gamal, R. Humphrey, J. Ramos, K. Brigle and D. Dean, The *nifH, nifM* and *nifN* genes of *Azotobacter vinelandii*: Characterisation by Tn5 mutagenesis and isolation from pLAFR1 gene banks, *Mol. Gen. Genet. MGG*, 1986, **205**, 318–325.

67 C. Kennedy and D. Dean, The *nifU, nifS* and *nifV* gene products are required for activity of all three nitrogenases of *Azotobacter vinelandii*, *Mol. Gen. Genet.*, 1992, **231**, 494–498.

68 R. Dixon, R. R. Eady, G. Espin, S. Hill, M. Iaccarino, D. Kahn and M. Merrick, Analysis of regulation of *Klebsiella pneumoniae* nitrogen fixation (*nif*) gene cluster with gene fusions, *Nature*, 1980, **286**, 128–132.

69 H. H. Nagatani, V. K. Shah and W. J. Brill, Activation of inactive nitrogenase by acid-treated component I, *J. Bacteriol.*, 1974, **120**, 697–701.

70 F. Luque and R. N. Pau, Transcriptional regulation by metals of structural genes for *Azotobacter vinelandii* nitrogenases, *Mol. Gen. Genet.*, 1991, **227**, 481–487.

71 K. Schneider, A. Muller, U. Schramm and W. Klipp, Demonstration of a molybdenum and vanadium independent nitrogenase in a *nifHDK* deletion mutant of *Rhodobacter capsulatus*, *Eur. J. Biochem.*, 1991, **195**, 653–661.

72 U. Gollan, K. Schneider, A. Muller, K. Schuddekopf and W. Klipp, Detection of the in vivo incorporation of a metal cluster into a protein: The FeMo cofactor is inserted into the FeFe protein of the alternative nitrogenase of *Rhodobacter capsulatus*, *Eur. J. Biochem.*, 1993, **215**, 25–35.

73 G. Wang, S. Angermuller and W. Klipp, Characterization of *Rhodobacter capsulatus* genes encoding a molybdenum transport system and putative molybdenum-pterin-binding proteins, *J.*

Bacteriol., 1993, **175**, 3031–3042.
74 L. J. Lehman and G. P. Roberts, Identification of an alternative nitrogenase system in *Rhodospirillum rubrum*, *J. Bacteriol.*, 1991, **173**, 5705–5711.
75 S. Jacobitz and P. E. Bishop, Regulation of nitrogenase-2 in *Azotobacter vinelandii* by ammonium, molybdenum, and vanadium, *J. Bacteriol.*, 1992, **174**, 3884–3888.
76 R. Premakumar, M. R. Jacobson, T. M. Loveless and P. E. Bishop, Characterization of transcripts expressed from nitrogenase-3 structural genes of *Azotobacter vinelandii*., *Can. J. Microbiol.*, 1992, **38**, 929–936.
77 J. Walmsley, A. Toukdarian and C. Kennedy, The role of regulatory genes *nifA, vnfA, anfA, nfrX, ntrC,* and *rpoN* in expression of genes encoding the three nitrogenases of *Azotobacter vinelandii*, *Arch. Microbiol.*, 1994, **162**, 422–429.
78 R. Premakumar and E. Bishop, A comparison of the three dinitrogenase reductases expressed by *Azotobacter vinelandii*, 1988, Can. J. Microbiol., 35, 344-348.
79 R. N. Pau, eds. R. Palacios, J. Mora and W. E. Newton, Springer Netherlands, Dordrecht, 1993, pp. 117–122.
80 N. J. Mouncey, L. A. Mitchenall and R. N. Pau, Mutational analysis of genes of the mod locus involved in molybdenum transport, homeostasis, and processing in *Azotobacter vinelandii*, *J. Bacteriol.*, 1995, **177**, 5294–5302.
81 R. Premakumar, R. N. Pau, L. A. Mitchenall, M. Easo and P. E. Bishop, Regulation of the transcriptional activators AnfA and VnfA by metals and ammonium in *Azotobacter vinelandii*, *FEMS Microbiol. Lett.*, 1998, **164**, 63–68.
82 P. Normand, P. Lapierre, L. S. Tisa, J. P. Gogarten, N. Alloisio, E. Bagnarol, C. A. Bassi, A. M. Berry, D. M. Bickhart, N. Choisne, A. Couloux, B. Cournoyer, S. Cruveiller, V. Daubin, N. Demange, M. P. Francino, E. Goltsman, Y. Huang, O. R. Kopp, L. Labarre, A. Lapidus, C. Lavire, J. Marechal, M. Martinez, J. E. Mastronunzio, B. C. Mullin, J. Niemann, P. Pujic, T. Rawnsley, Z. Rouy, C. Schenowitz, A. Sellstedt, F. Tavares, J. P. Tomkins, D. Vallenet, C. Valverde, L. G. Wall, Y. Wang, C. Medigue and D. R. Benson, Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography, *Genome Res.*, 2007, **17**, 7–15.
83 Y. Yan, J. Yang, Y. Dou, M. Chen, S. Ping, J. Peng, W. Lu, W. Zhang, Z. Yao, H. Li, W. Liu, S. He, L. Geng, X. Zhang, F. Yang, H. Yu, Y. Zhan, D. Li, Z. Lin, Y. Wang, C. Elmerich, M. Lin and Q. Jin, Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501, *Proc. Natl. Acad. Sci.*, 2008, **105**, 7564–7569.
84 A. Sarkar and B. Reinhold-Hurek, Transcriptional profiling of nitrogen fixation and the role of NifA in the diazotrophic endophyte *Azoarcus sp.* strain BH72, *PLoS One*, 2014, **9**, 1–12.
85 H. W. Shi, L. Y. Wang, X. X. Li, X. M. Liu, T. Y. Hao, X. J. He and S. F. Chen, Genome-wide transcriptome profiling of nitrogen fixation in *Paenibacillus* sp. WLY78, *BMC Microbiol.*, 2016, **16**, 1–10.
86 T. L. Hamilton, M. Ludwig, R. Dixon, E. S. Boyd, P. C. Dos Santos, J. C. Setubal, D. A. Bryant, D. R. Dean and J. W. Peters, Transcriptional profiling of nitrogen fixation in *Azotobacter vinelandii*, *J. Bacteriol.*, 2011, **193**, 4477–4486.
87 P. C. Dos Santos, Z. Fang, S. W. Mason, J. C. Setubal and R. Dixon, Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes, *BMC Genomics*, 2012, **13**, 1–12.
88 B. Soboh, E. S. Boyd, D. Zhao, J. W. Peters and L. M. Rubio, Substrate specificity and evolutionary implications of a NifDK enzyme carrying NifB-co at its active site, *FEBS Lett.*, 2010, **584**, 1487–1492.
89 J. Raymond, J. L. Siefert, C. R. Staples and R. E. Blankenship, The Natural History of Nitrogen Fixation, *Mol. Biol. Evol.*, 2004, **21**, 541–554.
90 B. P. Hodkinson, J. L. Allen, L. L. Forrest, B. Goffinet, E. Sérusiaux, Ó. S. Andrésson, V. Miao, J. P.

Bellenger and F. Lutzoni, Lichen-symbiotic cyanobacteria associated with Peltigera have an alternative vanadium-dependent nitrogen fixation system, *Eur. J. Phycol.*, 2014, **49**, 11–19.

91 S. Noda, M. Ohkuma, R. Usami, K. Horikoshi and T. Kudo, Culture-independent characterization of a gene responsible for nitrogen fixation in the symbiotic microbial community in the gut of the termite *Neotermes koshunensis*, *Appl. Environ. Microbiol.*, 1999, **65**, 4935–4942.

92 D. A. Betancourt, T. M. Loveless, J. W. Brown and P. E. Bishop, Characterization of diazotrophs containing Mo-independent nitrogenases, isolated from diverse natural environments, *Appl. Environ. Microbiol.*, 2008, **74**, 3471–3480.

93 D. L. McRose, X. Zhang, A. M. L. Kraepiel and F. M. M. Morel, Diversity and activity of alternative nitrogenases in sequenced genomes and coastal environments, *Front. Microbiol.*, 2017, **8**, 1–13.

94 J. P. Zehr, B. D. Jenkins, S. M. Short and G. F. Steward, Nitrogenase gene diversity and microbial community structure: a cross-system comparison, *Environ. Microbiol.*, 2003, **5**, 539–554.

95 Z. Dai, X. Guo, H. Yin, Y. Liang, J. Cong and X. Liu, Identification of nitrogen-fixing genes and gene clusters from metagenomic library of acid mine drainage, *PLoS One*, 2014, **9**, 1–9.

96 K. Yu and T. Zhang, Metagenomic and Metatranscriptomic Analysis of Microbial Community Structure and Gene Expression of Activated Sludge, *PLoS One*, 2012, **7**(5), 1-13. DOI:10.1371/journal.pone.0038183.

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

98 E. E. Stüeken, R. Buick, B. M. Guy and M. C. Koehler, Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr, *Nature*, 2015, **520**, 666–669.

99 S. M. Som, R. Buick, J. W. Hagadorn, T. S. Blake, J. M. Perreault, J. P. Harnmeijer and D. C. Catling, Earth's air pressure 2.7 billion years ago constrained to less than half of modern levels, *Nat. Geosci.*, 2016, **9**, 448–452.

100 R. Navarro-González, C. P. McKay and D. N. Mvondo, A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning, *Nature*, 2001, **412**, 61–64.

101 A. L. Zerkle and S. Mikhail, The geobiological nitrogen cycle: From microbes to the mantle, *Geobiology*, 2017, **15**, 343–352.

102 B. E. Erickson and G. R. Helz, Molybdenum(VI) speciation in sulfidic waters: Stability and lability of thiomolybdates, *Geochim. Cosmochim. Acta*, 2000, **64**, 1149–1158.

103 D. Wang, Redox chemistry of molybdenum in natural waters and its involvement in biological evolution, *Front. Microbiol.*, 2012, **3**, 1–7.

104 M. A. Saito, D. M. Sigman and F. M. M. Morel, The bioinorganic chemistry of the ancient ocean: The co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean-Proterozoic boundary?, *Inorganica Chim. Acta*, 2003, **356**, 308–318.

105 A. D. Anbar and A. H. Knoll, Proterozoic Ocean Chemistry and Evolution: A Bioionorganic Bridge?, *Science*, 2002, **297**, 1137–1142.

106 X. Zhang, D. M. Sigman, F. M. M. Morel and A. M. L. Kraepiel, Nitrogen isotope fractionation by alternative nitrogenases and past ocean anoxia, *Proc. Natl. Acad. Sci.*, 2014, **111**, 4782–4787.

107 A. J. M. Stams, Metabolic interactions between anaerobic bacteria in methanogenic environments, *Antonie Van Leeuwenhoek*, 1994, **66**, 271–294.

108 E. S. Boyd, T. L. Hamilton and J. W. Peters, An alternative path for the evolution of biological nitrogen fixation, *Front. Microbiol.*, 2011, **2**, 1–11.

109 S. E. McGlynn, E. S. Boyd, J. W. Peters and V. J. Orphan, Classifying the metal dependence of uncharacterized nitrogenases, *Front. Microbiol.*, 2012, **3**, 1–8.

110 S. J. George, R. Y. Igarashi, Y. Xiao, J. A. Hernandez, M. Demuez, D. Zhao, Y. Yoda, P. W. Ludden, L. M. Rubio and S. P. Cramer, Extended X-ray absorption fine structure and nuclear resonance vibrational spectroscopy reveal that NifB-co, a FeMo-co precursor, comprises a 6Fe core with an

interstitial light atom, *J. Am. Chem. Soc.*, 2008, **130**, 5673–5680.

111 L. Curatti, C. S. Brown, P. W. Ludden, L. M. Rubio and S. Kustu, Genes required for rapid expression of nitrogenase activity in *Azotobacter vinelandii*, *Proc. Natl. Acad. Sci.*, 2005, **102**, 6291–6296.

112 S. Tal, T. W. Chun, N. Gavini and B. K. Burgess, The delta-nifB (or delta-nifE) FeMo cofactor-deficient MoFe protein is different from the delta-nifH protein, *J. Biol. Chem.*, 1991, **266**, 10654–10657.

113 V. K. Shah, J. R. Allen, N. J. Spangler and P. W. Ludden, In vitro synthesis of the iron-molybdenum cofactor of nitrogenase. Purification and characterization of NifB cofactor, the product of NifB protein., *J. Biol. Chem.*, 1994, **269**, 1154–1158.

114 Y. Hu, J. M. Yoshizawa, A. W. Fay, C. C. Lee, J. A. Wiig and M. W. Ribbe, Catalytic activities of NifEN: Implications for nitrogenase evolution and mechanism, *Proc. Natl. Acad. Sci.*, 2009, **106**, 16962 LP-16966.

115 A. L. Zerkle, C. K. Junium, D. E. Canfield and C. H. House, Production of ^{15}N-depleted biomass during cyanobacterial N_2-fixation at high Fe concentrations, *J. Geophys. Res. Biogeosciences*, 2008, **113**, 1–9.

Figures

Figure 1: A. The structure of one catalytic half of the MoFe protein is a $\alpha_2\beta_2$ heterotetramer where the α-subunit (green) houses the FeMo-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). Residues His195, Gln191, Val70 and Arg96 coordinate the FeMo-cofactor, which is a [7Fe-9S-Mo] cluster with a central carbide and homocitrate moiety. B. The structure of one catalytic half of the VFe protein is a $\alpha_2\beta_2\gamma_2$ heterohexamer where the α-subunit (green) houses the FeV-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). The gamma subunit (magenta) contacts the α-subunit as seen in the VFe protein. Residues His180, Gln176, Val57 and Lys83 coordinate the FeV-cofactor, which is a [7Fe-9S-V] cluster with a central carbide and homocitrate moiety. The FeV-cofactor is slightly elongated and has a carbonate ligand, which differs from the bridging sulfide atom observed in the FeMo-cofactor. C. Homology model of the proposed structure of one catalytic half of the FeFe protein is a $\alpha_2\beta_2\gamma_2$ heterohexamer where the α-subunit (green) houses the FeV-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). The gamma subunit (magenta) presumably contacts the α-subunit as seen in the VFe protein. Residues His180, Gln176, Val57 and Lys83 reside at the proposed site of the FeFe-cofactor, which is analogous to the FeV- and FeMo-cofactors however, the current structure is unknown. D. The Fe protein (orange) is involved in electron transfer to the MoFe protein (blue and green). Upon association, (1) one electron is transferred from the P-cluster to the FeMo-cofactor where (2) the Fe protein can then transfer an electron from its [4Fe-4S] cluster to the P-cluster in what is called the "deficit spending model."

Table 1: Comparison of the specific activities for N_2, H_2 and C_2H_2 for the Mo-, V- and Fe-only nitrogenases isolated from *A. vinelandii*.

Figure 2: Organization of Mo, V, and Fe nitrogenase gene clusters in *A. vinelandii*.

Figure 3: Putative pathways for metal regulation of *A. vinelandii* nitrogenases. Gene encoding transcriptional activator proteins are designated by *nifA*, *vnfA*, *anfA*. Genes of Mo-, V-, and Fe-nitrogenase systems are indicated by *nif*, *vnf* and *anf* respectively. Transcriptional activation is shown by solid arrows pointed to promoters which are represented by purple chevron symbols. Transcriptional repression is shown by dashed arrows. Activator NifA and the presence of molybdenum (Mo) are required for the synthesis of Mo-nitrogenase. Repression of Mo-nitrogenase structural genes in the absence of Mo is mediated by VnfA, which under the same condition activates V-nitrogenase gene transcription. Mo acts by repressing transcription of the transcriptional regulators of the alternative nitrogenases, *vnfA* and *anfA*. Vanadium (V) regulates Fe-only nitrogenase by repressing the transcription of its structural genes. Fe protein of V-nitrogenase (VnfH) is always synthesized under Mo-deficient conditions. In absence of V, the activation of *AnfA* is mediated by VnfH.

Figure 4: Schematic illustration of cellular functions directly and indirectly involved in N_2-fixation. In addition of the specific N_2-fixation gene clusters, *nif*, *vnf* and *anf*, encoding for nitrogenase enzymes, there are other core genes that encode for major functional groups of proteins: transport (Mo, Fe, S, NH_4^+ transporters), respiratory and energy metabolism (cytochromes, NADH dehydrogenases, ATP synthetases), electron transport (RNF1, FIX, Fdx,

Flv), nitrogen and carbon metabolism (*gln*, *glt*, hydrogenase genes), secondary metabolism (pili, flagella). OM: outer membrane; P: periplasm; IM: inner membrane; H_2ase: hydrogenase; $Fdx_{ox/red}$: ferredoxin oxidized/reduced; Flv: flavodoxin.

Figure 5: Rendition of phylogenetic results from Boyd et al. 2011, concatenation of *nif*/*vnf*/*anf* and characterized HDK protein sequences placed in a maximum-likelihood reconstruction tree. The metal content of nitrogenase are colored, blue (*nif*), pink (uncharacterized), green (*anf*), red (*vnf*). Methanogen Mo-nitrogenase is labeled on the alternative nitrogenase branch. Uncharacterized nitrogenase show closer homology to *nif* than the alternative nitrogenases show in McGlynn et al. 2013. Distinct nesting shown in the tree leads to the theory of the alternative nitrogenases operon duplication in the before LUCA and *vfn* preceding *anf*.

Figure 6: The evolution of nitrogenase. A Protonitrogenase most likely existed before LCUA which could have possibly associated with metal cofactors due to the similarity of the active site cavity of modern. After a gene duplication event subunit differentiation occurred and created a heterotetramer with more specific activity. This gene duplication was then followed by operon duplication and created the scaffolding proteins of *nifEN*. These evolved to help with co-factor maturation and development of a more specific and efficient cofactor. This enzyme was very similar to the modern nitrogenase. This operon was then duplicated again and environmental pressures differentiated the cofactor biosynthesis to develop alternative metal nitrogenases.

Enzyme (Wild type)	Specific Activity for N_2 (nmol/min/mg)	Specific Activity for H_2 (nmol/min/mg)	Specific Activity for C_2H_2 (nmol/min/mg)	Reference
Mo-nitrogenase	605	2226	1819	Harris et al. 2017
V-nitrogenase	660	1400	220	Sickerman et al. 2017
Fe-only nitrogenase	181	1085	306	Harris et al. 2017

Table I: Comparison of the specific activities for N2, H2 and C2H2 for the Mo-, V- and Fe-only nitrogenases isolated from A. vinelandii.

190x107mm (300 x 300 DPI)

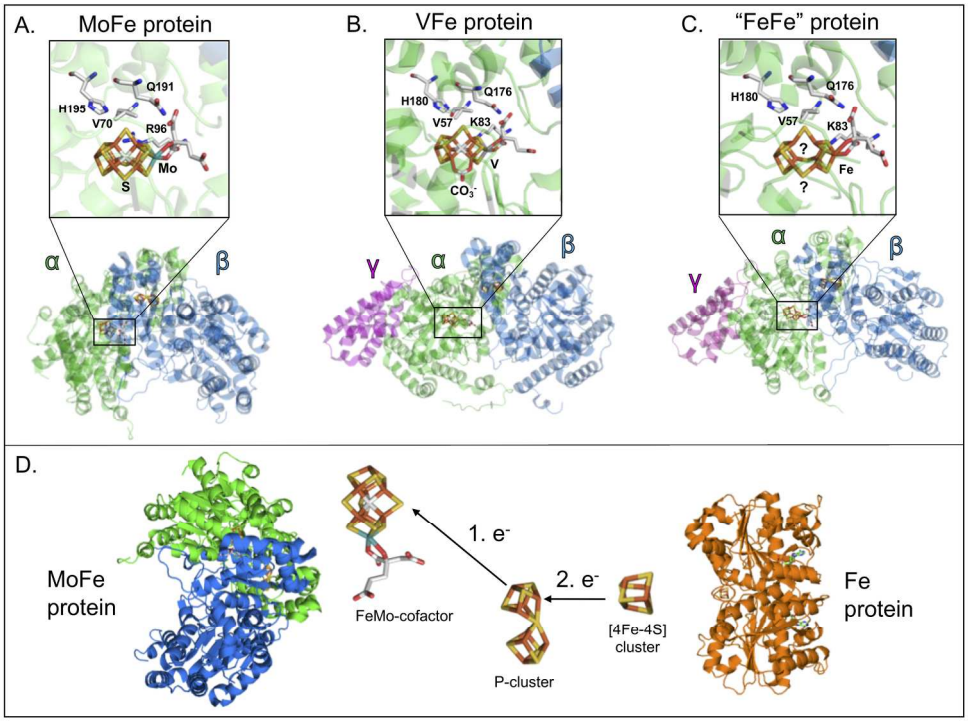

Figure 1: A. The structure of one catalytic half of the MoFe protein is a α2β2 heterotetramer where the α-subunit (green) houses the FeMo-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). Residues His195, Gln191, Val70 and Arg96 coordinate the FeMo-cofactor, which is a [7Fe-9S-Mo] cluster with a central carbide and homocitrate moiety. B. The structure of one catalytic half of the VFe protein is a α2β2γ2 heterohexamer where the α-subunit (green) houses the FeV-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). The gamma subunit (magenta) contacts the α-subunit as seen in the VFe protein. Residues His180, Gln176, Val57 and Lys83 coordinate the FeV-cofactor, which is a [7Fe-9S-V] cluster with a central carbide and homocitrate moiety. The FeV-cofactor is slightly elongated and has a carbonate ligand, which differs from the bridging sulfide atom observed in the FeMo-cofactor. C. Homology model of the proposed structure of one catalytic half of the FeFe protein is a α2β2γ2 heterohexamer where the α-subunit (green) houses the FeV-cofactor and P-cluster, which is at the interface of the α-subunit and β-subunit (blue). The gamma subunit (magenta) presumably contacts the α-subunit as seen in the VFe protein. Residues His180, Gln176, Val57 and Lys83 reside at the proposed site of the FeFe-cofactor, which is analogous to the FeV- and FeMo-cofactors however, the current structure is unknown. D. The Fe protein (orange) is involved in electron transfer to the MoFe protein (blue and green). Upon association, (1) one electron is transferred from the P-cluster to the FeMo-cofactor where (2) the Fe protein can then transfer an electron from its [4Fe-4S] cluster to the P-cluster in what is called the "deficit spending model."

203x147mm (300 x 300 DPI)

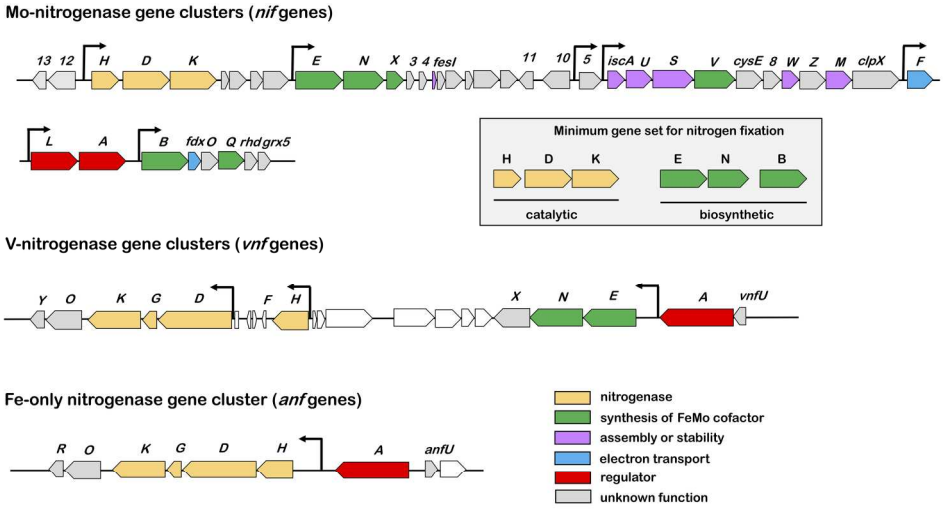

Figure 2: Organization of Mo, V, and Fe nitrogenase gene clusters in A. vinelandii.

190x107mm (300 x 300 DPI)

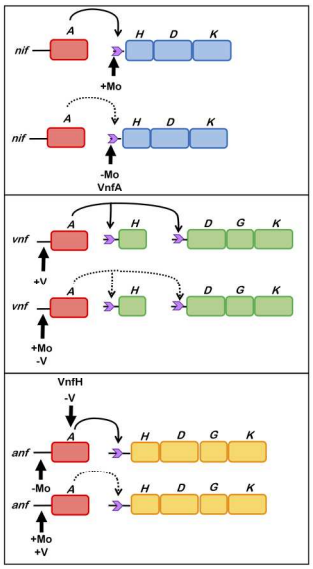

Figure 3: Putative pathways for metal regulation of A. vinelandii nitrogenases. Gene encoding transcriptional activator proteins are designated by nifA, vnfA, anfA. Genes of Mo-, V-, and Fe- nitrogenase systems are indicated by nif, vnf and anf respectively. Transcriptional activation is shown by solid arrows pointed to promoters which are represented by purple chevron symbols. Transcriptional repression is shown by dashed arrows. Activator NifA and the presence of molybdenum (Mo) are required for the synthesis of Mo-nitrogenase. Repression of Mo-nitrogenase structural genes in the absence of Mo is mediated by VnfA, which under the same condition activates V-nitrogenase gene transcription. Mo acts by repressing transcription of the transcriptional regulators of the alternative nitrogenases, vnfA and anfA. Vanadium (V) regulates Fe-only nitrogenase by repressing the transcription of its structural genes. Fe protein of V-nitrogenase (VnfH) is always synthesized under Mo-deficient conditions. In absence of V, the activation of AnfA is mediated by VnfH.

190x107mm (300 x 300 DPI)

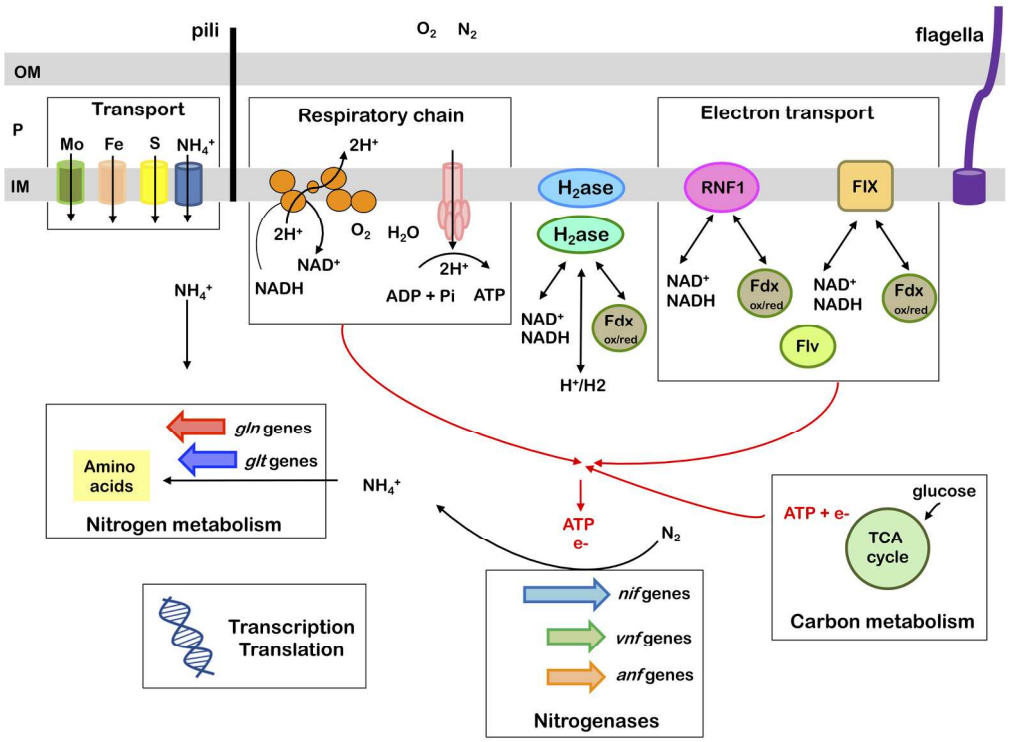

Figure 4: Schematic illustration of cellular functions directly and indirectly involved in N2-fixation. In addition of the specific N2-fixation gene clusters, nif, vnf and anf, encoding for nitrogenase enzymes, there are other core genes that encode for major functional groups of proteins: transport (Mo, Fe, S, NH4+ transporters), respiratory and energy metabolism (cytochromes, NADH dehydrogenases, ATP synthetases), electron transport (RNF1, FIX, Fdx, Flv), nitrogen and carbon metabolism (gln, glt, hydrogenase genes), secondary metabolism (pili, flagella). OM: outer membrane; P: periplasm; IM: inner membrane; H2ase: hydrogenase; Fdxox/red: ferredoxin oxidized/reduced; Flv: flavodoxin.

190x142mm (300 x 300 DPI)

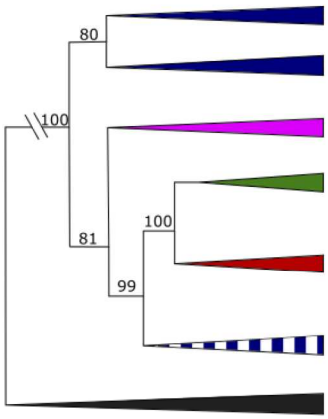

Figure 5: Rendition of phylogenetic results from Boyd et al. 2011, concatenation of nif/vnf/anf and characterized HDK protein sequences placed in a maximum-likelihood reconstruction tree. The metal content of nitrogenase are colored, blue (nif), pink (uncharacterized), green (anf), red (vnf). Methanogen Mo-nitrogenase is labeled on the alternative nitrogenase branch. Uncharacterized nitrogenase show closer homology to nif than the alternative nitrogenases show in McGlynn et al. 2013. Distinct nesting shown in the tree leads to the theory of the alternative nitrogenases operon duplication in the before LUCA and vfn preceding anf.

279x361mm (300 x 300 DPI)

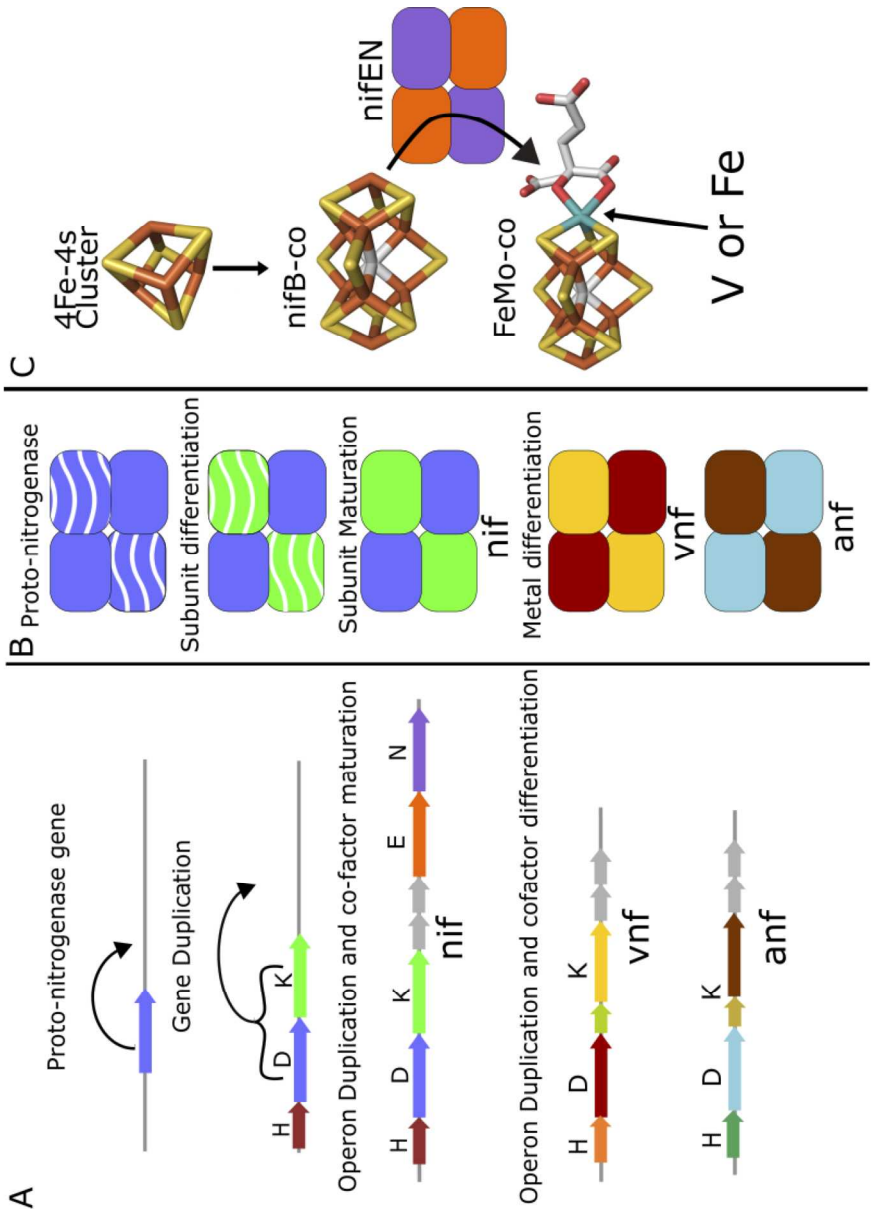

Figure 6: The evolution of nitrogenase. A Protonitrogenase most likely existed before LCUA which could have possibly associated with metal cofactors due to the similarity of the active site cavity of modern. After a gene duplication event subunit differentiation occurred and created a heterotetramer with more specific activity. This gene duplication was then followed by operon duplication and created the scaffolding proteins of nifEN. These evolved to help with co-factor maturation and development of a more specific and efficient cofactor. This enzyme was very similar to the modern nitrogenase. This operon was then duplicated again and environmental pressures differentiated the cofactor biosynthesis to develop alternative metal nitrogenases.

279x361mm (300 x 300 DPI)