

UNIVERSIDAD POLITÉCNICA DE MADRID

Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de
Biosistemas



**Engineering functional nitrogenase
cofactor biosynthesis protein NifEN in
mitochondria of *Saccharomyces
cerevisiae***

DOCTORAL THESIS

Submitted for the degree of Doctor by:

Katarzyna Dobrzyńska

BSc in Biology

MSc in Molecular Biology

Madrid, 2024



UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería Agronómica,
Alimentaria y de Biosistemas

**Doctoral Degree in Biotechnology and Genetic Resources of
Plants and Associated Microorganisms**

**Engineering functional nitrogenase
cofactor biosynthesis protein NifEN in
mitochondria of *Saccharomyces
cerevisiae***

DOCTORAL THESIS

Submitted for the degree of Doctor by:

Katarzyna Dobrzyńska

BSc in Biology
MSc in Molecular Biology

Under the supervision of:
Dr. Luis Manuel Rubio Herrero
Dr. Stefan Burén

Madrid, 2024

Title: Engineering functional nitrogenase cofactor biosynthesis protein NifEN in mitochondria of *Saccharomyces cerevisiae*

Author: Katarzyna Dobrzyńska

Doctoral Programme: Doctoral Degree in Biotechnology and Genetic Resources of Plants and Associated Microorganisms

Thesis Supervision:

Dr. Luis Manuel Rubio Herrero, Investigador Científicos, CSIC-INIA, Centro de Biotecnología y Genómica de Plantas UPM-INIA/CSIC, (Supervisor)

Dr. Stefan Burén, Associate Professor at Universidad Politécnica de Madrid, Centro de Biotecnología y Genómica de Plantas UPM-INIA/CSIC, (Co-supervisor)

External Reviewers:

Thesis Defense Committee:

Thesis Defense Date:

This thesis has been supported by FPI predoctoral fellowship [PRE2018-084951]

Kochającym dziadkom

Acknowledgement

I am deeply thankful to everyone who has played a part in bringing this thesis to completion, whether through direct assistance or indirect support. Though this journey has been challenging, it has also been incredibly fulfilling. Especially I want to express my sincerest gratitude to the people and institutions listed below:

First and foremost, I owe the completion of my degree to the opportunity provided by my thesis director, Luis Rubio. Without his acceptance into the project and inclusion in his research group, my thesis would have been incomplete. Working in your lab presented significant challenges due to its high standards. However, I have gained invaluable lessons throughout the past four-year. I am grateful for the guidance and advices you provided throughout the project. Additionally, I would like to thank you for the opportunity to undertake a stay abroad during my PhD thesis, which add a lot to my academic and personal growth.

I am deeply grateful to my thesis co-director, Stefan Buren, whose significant contribution was turning this project into a reality. Your guidance, advice, and insights have been incredibly helpful, and I am sincerely thankful for the opportunity to learn from you. Your words have always held a special significance for me, and I have always truly admired your professionalism. I deeply appreciate the care you took in mentoring me throughout the thesis process. You are not just a great scientist but also a person whom others should look up to.

I want to express my gratitude to the wonderful people who have been by my side throughout my PhD journey. Carlos, thank you for always helping me with tech stuff and checking in on how my thesis was going. I want to give a huge thanks to Diana, Cris, and Shandy. Whenever I was overwhelmed or needed help with experiments, I knew I could always count on you three. Your assistance, especially before I went abroad, was incredibly valuable and really helped me. Ana and Alvaro, your support during my time abroad played a big role in the success of my project. I am grateful for the knowledge and experience I gained from working alongside you. Natalia, I'll always remember your kindness and willingness to help from the moment we met. I'm grateful that we could always chat and share our experiences together. Manu, I appreciate our chats and your kindness. Your readiness to help out or just listen has meant a lot to me. German, your positive words, help with writing in Spanish, and readiness to support whenever I needed have always made the days in CBGP better. Thank you for being such a supportive friend! As well, a lot of gratitude to my more experienced friends Laura, Hind, and Lucas. Your chats, advice, and words of encouragement during the struggles of writing the thesis manuscript have meant the world to me. Thank you for being there when I needed support the most. Lucia, thank you for your huge help during the three years we worked together, particularly in teaching protocols. Your professionalism is

extremally impressed for me. I am grateful for your friendship and the assurance that I can always count on you. Jiang, my dear friend, without your company and support, this thesis might never have been finished. Thank you for showing me how to be a hardworking and tough person. I am incredibly grateful to have had you as a soulmate at CBGP.

Each person I've had the pleasure of working with during these four years has played an integral role in shaping my thesis experience, and for that, I am deeply grateful. Your support, guidance, and friendship have made this thesis journey more enriching and memorable.

I would also like to express my gratitude to Professor Dennis Dean for hosting me in his laboratory during my stay abroad. It was an incredible opportunity to work alongside you, and I truly appreciated the experience of collaborating with such an esteemed scientist. I am also thankful to the lab members, Valerie and Julia, for their kindness and assistance. Additionally, I want to extend my thanks to the people I met during my time there, who not only became my friends but also provided huge support there. Thank you to Larkin, Marin, Bernardo, Maria Helena, Rodrigo, and Marconi. Your presence made my stay in the US much more enjoyable.

Ogromne podziękowania kieruję do moich polskich przyjaciół, których poznałam w Madrycie. Zawsze byliście dla mnie w dobrych i trudnych chwilach, dając mi wsparcie i uczucie, że jestem otoczona rodziną. Wasza obecność była i nadal jest dla mnie niezwykle ważna. Aga, Twoje wsparcie od pierwszego dnia mojego pobytu było dla mnie ogromnie pomocne, a Twoja obecność sprawiała, że czułam się bardziej pewnie. Rozmowy z Tobą były zawsze pełne ciepła i zrozumienia, i cenię każdą chwilę spędzoną razem. Piotr, Twoja gotowość do pomocy i wsparcia sprawiła, że moje doświadczenia w Madrycie były jeszcze bardziej wyjątkowe. Dziękuję za wszystkie rady, rozmowy i za to, że zawsze mogłam na Ciebie liczyć. Kasia, Twoja przyjaźń i chęć do rozmów o każdej porze dnia i nocy były dla mnie bezcenne. Dziękuję Ci za wszystkie chwile spędzone razem. Jestem ogromnie wdzięczna za to, że mogłam Was poznać i mieć Was u boku. As well I want to express my gratitude to one of the most caring and amazing person I have met here, Dafne. I am truly thankful that we developed such a strong friendship. Even though you are no longer in Madrid, I still feel as close to you as ever before.

Niezmiernie jestem wdzięczna moim rodzicom, Mamie i Tacie, którzy byli nieocenionym wsparciem na każdym etapie mojej drogi doktorskiej. To dzięki ich wierze w moje umiejętności i stałemu wsparciu mogłam pokonywać kolejne wyzwania. Jestem przekonana, że bez ich pomocy nie zdołałabym osiągnąć tego, co udało mi się dzisiaj. Dzięki ich wsparciu i wychowaniu stałam się osobą ambitną, wytrwałą i pewną siebie. Kocham ich bardzo i doceniam każdą poświęconą mi chwilę.

Również chciałabym serdecznie podziękować mojemu bratu Gruszardowi za jego niezastąpione wsparcie i pomoc. Jego determinacja i poświęcenie stanowiły dla mnie

inspirację i wzór do naśladowania. Dziękuję, że zawsze mnie motywowałeś do wysiłku, pokazując, że ciężka praca przynosi efekty. Nie mogę również zapomnieć o moich dziadkach, cioci Izie, wujku Pawle i Oli, którzy zawsze interesowali się moimi postępami i wspierali mnie w trudnych chwilach, dodając otuchy i podnosząc na duchu. Dziękuję Wam za Waszą nieustającą miłość i wsparcie, które jest dla mnie niezwykle cenne.

Moim przyjaciółom w Polsce kieruję szczególne podziękowania. Przed wyjazdem do Hiszpanii obawiałam się, że nasze drogi się rozejdą i że odległość będzie trudna do pokonania. Jednak dzięki silnym więziom udało nam się utrzymać naszą przyjaźń, co sprawia, że teraz czuję się jeszcze bliżej Was. Zawsze mogłam liczyć na Waszą gotowość do rozmowy w trudnych momentach, Wasze odwiedziny w Hiszpanii i to, że za każdym razem, gdy wracałam do Polski, byliście gotowi się ze mną spotkać, wesprzeć i podnieść na duchu. Spotkania z Wami były niezwykle pomocne i zawsze dodawały mi nowej energii przed powrotem do Madrytu. Dziękuję Wam za wsparcie i za to, że mimo dystansu jesteście ze mną. Szczególne podziękowania dla kilku niezastąpionych osób, które były ze mną przez cały czas mojej drogi doktorskiej. Zosia, moja przyjaciółka od lat, zawsze stała przy mnie i potrafiła znaleźć właściwe słowa wsparcia w najtrudniejszych chwilach. Roza, znająca dokładnie szczegóły mojego doktoratu, zawsze była gotowa pomóc i wesprzeć mnie swoją wiedzą i doświadczeniem. Paulina, która nie tylko wierzyła we mnie, ale także nieustannie pytała o postępy i stawiała się motywacją do dalszych działań. Monika, swoją życzliwością i spokojem zawsze dawała mi siłę, by realizować moje cele. Ola, zawsze gotowa do wysłuchania i zainteresowania moimi postępami, była dla mnie wielkim wsparciem. Dziękuję również wszystkim innym, którzy wierzyli we mnie i dopingowali mnie na każdym etapie mojej drogi doktorskiej. Wasza obecność i wsparcie były dla mnie nieocenione i stanowiły niezwykle cenną część mojego życia.

Abstract

Synthetic fertilizers have unquestionably played a significant role in increasing crop productivity, which has facilitated a dramatic rise in human population growth over the past 100 years. However, with increasing evidence of climate change and the vast inequalities between farming in developed and developing countries, there is a drastic need for changes in how we can source the nitrogen needed for efficient agriculture. Biological nitrogen fixation (BNF), the conversion of inert dinitrogen (N_2) into metabolically tractable ammonia (NH_3) is a remarkable natural process performed by a select group of microorganisms called diazotrophs. This process, made possible by the enzyme nitrogenase, plays a crucial role in replenishing essential nitrogen compounds in the Earth's biosphere and offers a sustainable alternative to the industrial production of nitrogen fertilizers. If plants were engineered to also synthesize their own NH_3 like diazotrophs, the need for external nitrogen fertilizers would be reduced or eliminated. One way to enhance BNF in crops is to transfer and incorporate the prokaryotic genes required for nitrogenase expression into the plant's genome.

The nitrogenase enzyme complex, responsible for catalyzing BNF, is formed by the structural proteins NifH and NifDK and harbors three types of oxygen-sensitive metalloclusters: the [4Fe-4S] cluster at NifH and the P-cluster and FeMo-co at NifDK, each of which is essential for the activity of nitrogenase. Of these three FeMo-co is the most complex cluster and provides the catalytic site for N_2 reduction. Its synthesis is a multi-step process that involves the stepwise activity of several proteins, with NifB and NifEN playing pivotal roles. The central player is the scaffold protein NifEN that converts NifB-co, an [8Fe-9S-C] precursor produced by NifB, into FeMo-co by substituting Mo for an apical Fe atom within NifB-co and by attaching homocitrate to the Mo atom. The complexity of nitrogenase, involving a large number of genes required for its maturation and function, coupled with the cofactor's sensitivity towards oxygen, makes nitrogenase formation in plants a significant challenge. For this reason, most studies have focused on producing Nif components in the simpler eukaryotic model organism *Saccharomyces cerevisiae* (yeast) by targeting the proteins to O_2 -deprived compartments within the yeast cell such as mitochondria.

Previous work in our laboratory has demonstrated the expression of functional NifB variants and the *in vivo* formation of NifB-co in yeast mitochondria. As NifEN is the protein that accepts NifB-co and processes it into FeMo-co, and because our initial attempts to express NifEN from the diazotrophic model organism *Azotobacter vinelandii* only generated insoluble protein in yeast mitochondria, the identification of a NifEN variant that can be produced in functional form is a necessary next step for engineering eukaryotic FeMo-co biosynthesis.

To address this challenge, this thesis generated a library of seventeen NifEN protein variants originating from diverse bacterial and archaeal organisms to enhance the likelihood of identifying soluble and functional NifEN proteins. These variants were expressed in *S. cerevisiae*, targeted to the yeast mitochondria, and evaluated for their ability to form soluble complexes. In parallel, the NifEN variants were heterologously produced in *A. vinelandii* to determine their compatibility with the rest of the *A. vinelandii* nitrogenase machinery and their ability to support FeMo-co formation *in vivo*.

The conducted experiments produced three NifEN variants that, in addition to being active in *A. vinelandii*, accumulated as soluble and functional proteins in the mitochondria of aerobically cultured *S. cerevisiae*. By combining NifB and NifH proteins of different origins with these NifEN variants, all of which were produced and purified from *S. cerevisiae* mitochondria, and by proving *in vitro* FeMo-co synthesis, a pathway for eukaryotic FeMo-co biosynthesis is demonstrated in this doctoral thesis.

Resumen

Los fertilizantes sintéticos han desempeñado un papel significativo en el aumento de la productividad de los cultivos, lo que ha facilitado un aumento en el crecimiento de la población mundial en los últimos 100 años. Sin embargo, con la creciente evidencia del cambio climático y las grandes desigualdades entre la agricultura en países desarrollados y en desarrollo, existe una necesidad drástica de cambios en cómo podemos obtener el nitrógeno necesario para una agricultura eficiente. La fijación biológica de nitrógeno (FBN), la conversión del dinitrógeno inerte (N_2) en amoníaco (NH_3) metabólicamente manejable, es un proceso natural realizado por un grupo de microorganismos llamados diazotrofos. Este proceso, posible gracias a la enzima nitrogenasa, desempeña un papel crucial en la reposición de compuestos de nitrógeno esenciales en la biosfera de la Tierra y ofrece una alternativa sostenible a la producción industrial de fertilizantes de nitrógeno. Si las plantas fueran modificadas genéticamente para sintetizar su propio NH_3 al igual que los diazotrofos, se reduciría o eliminaría la necesidad de fertilizantes de nitrógeno externos. Una manera de mejorar la FBN en los cultivos es transferir e incorporar los genes procarióticos necesarios para la expresión de la nitrogenasa en el genoma de la planta.

El complejo enzimático de la nitrogenasa, responsable de catalizar la FBN, está formado por las proteínas estructurales NifH y NifDK y alberga tres tipos de metaloclústeres sensibles al oxígeno: el clúster [4Fe-4S] en NifH y el clúster P y FeMo-co en NifDK, cada uno de los cuales es esencial para la actividad de la nitrogenasa. De estos tres, FeMo-co es el clúster más complejo y proporciona el sitio catalítico para la reducción de N_2 . Su síntesis es un proceso multietapa que involucra la actividad gradual de varias proteínas, siendo NifB y NifEN de importancia crucial. El protagonista central es la proteína andamio NifEN que convierte el precursor NifB-co, un precursor [8Fe-9S-C] producido por NifB, en FeMo-co sustituyendo Mo por un átomo de Fe apical dentro de NifB-co y adjuntando homocitrato al átomo de Mo. La complejidad de la nitrogenasa, que involucra un gran número de genes necesarios para su maduración y función, junto con la sensibilidad del cofactor al oxígeno, hace que la formación de la nitrogenasa en las plantas sea un desafío significativo. Por esta razón, la mayoría de los estudios se han centrado en producir componentes de Nif en el organismo modelo eucariótico más simple, *Saccharomyces cerevisiae* (levadura), dirigiendo las proteínas a compartimentos privados de oxígeno dentro de la célula de levadura, como las mitocondrias.

Trabajos anteriores en nuestro laboratorio han demostrado la expresión de variantes funcionales de NifB y la formación in vivo de NifB-co en las mitocondrias de la levadura. Dado que NifEN es la proteína que acepta NifB-co y la procesa en FeMo-co, y porque nuestros

intentos iniciales de expresar NifEN del organismo modelo diazotrófico *Azotobacter vinelandii* solo generaron proteína insoluble en las mitocondrias de la levadura, la identificación de una variante de NifEN que se pueda producir en forma funcional es un próximo paso necesario para la ingeniería de la biosíntesis eucariótica de FeMo-co.

Para abordar este desafío, este estudio generó una biblioteca de diecisiete variantes de la proteína NifEN originarias de diversos organismos bacterianos y arqueas para aumentar la probabilidad de identificar proteínas NifEN solubles y funcionales. Estas variantes se expresaron en *S. cerevisiae*, se dirigieron a las mitocondrias de la levadura y se evaluaron por su capacidad para formar complejos solubles. Simultáneamente, las variantes de NifEN se produjeron heterológicamente en *A. vinelandii* para determinar su compatibilidad con el resto de la maquinaria de la nitrogenasa de *A. vinelandii* y su capacidad para respaldar la formación de FeMo-co *in vivo*.

Los experimentos realizados produjeron tres variantes de NifEN que, además de ser activas en *A. vinelandii*, se acumularon como proteínas solubles y funcionales en las mitocondrias de *S. cerevisiae* cultivadas aeróbicamente. Al combinar las proteínas NifB y NifH de diferentes orígenes con estas variantes de NifEN, todas las cuales se produjeron y purificaron a partir de las mitocondrias de *S. cerevisiae*, y al demostrar la síntesis *in vitro* de FeMo-co, se demuestra un camino para la biosíntesis eucariótica de FeMo-co en esta tesis doctoral.

Table of Contents

Acknowledgement	v
Abstract.....	viii
Resumen.....	x
List of Figures	xv
List of Tables.....	xvii
Abbreviations and Acronyms.....	xviii
1. INTRODUCTION.....	1
1.1 Nitrogen Fixation.....	1
1.1.1 Importance of nitrogen	1
1.1.2 Plant-available nitrogen.....	1
1.1.3 Biological nitrogen fixation	2
1.1.4 Engineering nitrogen fixation in plants.....	3
1.2 Nitrogenase.....	4
1.2.1 Genetic complexity of the molybdenum nitrogenase.....	5
1.3 Biosynthesis of nitrogenase clusters	8
1.3.1 Biosynthesis of [4Fe-4S] clusters by NifU and NifS.....	9
1.3.2 P-cluster formation	10
1.3.3 FeMo-co synthesis	11
1.3.4 NifB forms FeMo-co precursor	12
1.3.5 NifX carriers NifB-co to NifEN	15
1.3.6 NifH in FeMo-co synthesis	15
1.3.7 Role of NifQ in FeMo-co synthesis	16
1.3.8 NifV delivers R- homocitrate to form FeMo-co	16
1.3.9 NifEN scaffold as a central player in FeMo-co assembly.....	16
1.3.10 NafY: a FeMo-co carrier protein.....	18
1.4 Engineering nitrogenase into eukaryotic cells.....	19
1.5 Experimental design to engineer nitrogenase cofactor biosynthesis in the mitochondria of <i>S. cerevisiae</i>	21
2. OBJECTIVES.....	24
3. MATERIAL AND METHODS.....	25
3.1 Chemical and Reagents	25
3.2 Organisms.....	25
3.2.1 Strains of <i>E. coli</i>	25
3.2.2 Strains of <i>A. vinelandii</i>	25
3.2.3 Strains of <i>S. cerevisiae</i>	27
3.3 Microbial culture media.....	30
3.3.1 <i>E. coli</i> lysogenic broth	30

3.3.2	<i>A. vinelandii</i> Burk's modified medium.....	30
3.3.3	<i>S. cerevisiae</i> medium.....	30
3.4	Transformation methods.....	31
3.4.1	Preparation of chemically competent <i>E. coli</i> cells and heat-shock-mediated transformation	31
3.4.2	Preparation of competent <i>A. vinelandii</i> cells and transformation methods ...	31
3.4.3	Transformation of <i>S. cerevisiae</i>	32
3.5	DNA techniques.....	32
3.5.1	DNA amplification by Polymerase Chain Reaction (PCR)	32
3.5.1.1	High-fidelity DNA amplification for cloning purposes.....	33
3.5.1.2	Validation of DNA fragments by PCR	34
3.5.1.3	Gene Splicing by Overlap Extension (OE-PCR).....	35
3.5.2	Traditional cloning techniques.....	36
3.5.3	Exonuclease and Ligation-Independent Cloning (ELIC)	36
3.5.4	Plasmid DNA extraction	36
3.5.5	DNA gel electrophoresis.....	37
3.5.6	Extraction of DNA from agarose gels.....	37
3.5.7	Purification of PCR-amplified or restriction enzyme-digested DNA from solutions	38
3.6	Plasmid Construction.....	38
3.6.1	Plasmids generated for expression of NifEN proteins in <i>E. coli</i>	38
3.6.2	Generation of plasmids to overexpress NifEN library in <i>S. cerevisiae</i>	39
3.6.3	Plasmids generated to overexpress NifEN library and isolate NifEN proteins from <i>A. vinelandii</i>	43
3.7	Biomass production for Nif protein expression and purification.....	44
3.7.1	Nif protein overexpression in aerobically grown yeast cells	44
3.7.2	Nif protein overexpression in aerobic <i>E. coli</i> cultures.....	45
3.7.3	Nitrogenase derepression in small-scale <i>A. vinelandii</i> cultures	46
3.7.4	Nitrogenase derepression in large-scale <i>A. vinelandii</i> cultures.....	46
3.8	Protein expression analysis and purification.....	47
3.8.1	Preparation of <i>S. cerevisiae</i> protein extracts.....	47
3.8.2	Preparation of cell-free extracts	48
3.8.2.1	Preparation of <i>E. coli</i> cell-free extracts	48
3.8.2.2	Preparation of <i>S. cerevisiae</i> cell-free extracts	48
3.8.2.3	Preparation of <i>A. vinelandii</i> cell-free extracts	49
3.8.3	Purification of His-tagged NifEN from <i>E. coli</i>	49
3.8.4	Purification of SS- and TS-tagged NifEN from <i>E. coli</i>	49
3.8.5	Purification of SS- or TS-tagged NifEN from <i>S. cerevisiae</i>	50
3.8.6	Purification of SS-tagged NifEN from <i>A. vinelandii</i>	50
3.9	Protein analysis techniques.....	50
3.9.1	Protein quantification.....	50
3.9.2	Iron quantification	51
3.9.3	Protein Mass Spectrometry.....	51
3.9.4	Denaturing gel electrophoresis	51
3.9.5	Western blotting.....	52
3.10	The acetylene reduction assays for nitrogenase activity	53

3.10.1	<i>In vivo</i> acetylene reduction assays in <i>A. vinelandii</i>	53
3.10.2	<i>In vitro</i> acetylene reduction assays.....	54
3.10.2.1	NifEN-dependent <i>in vitro</i> FeMo-co synthesis and insertion assays.....	56
3.10.2.2	NifB-co-dependent <i>in vitro</i> FeMo-co synthesis assays	56
3.10.2.3	NifB-dependent <i>in vitro</i> FeMo-co synthesis and insertion assay	56
3.11	Bioinformatic methods	57
3.11.1	Protein sequence and structure analysis.....	57
3.11.2	Statistical analysis	57
4.	RESULTS.....	58
4.1	Identification of a suitable NifEN protein for expression in <i>S. cerevisiae</i> mitochondria	58
4.1.1	Selection of tags for NifEN protein complex detection and purification	58
4.1.2	Design of the <i>nifEN</i> library	60
4.1.3	<i>In vivo</i> functionality of NifEN variants in <i>A. vinelandii</i>	61
4.1.4	Identification of soluble NifEN protein complex in the yeast mitochondria ..	63
4.1.5	Analysis of NifEN protein structures	66
4.1.6	<i>In vitro</i> FeMo-co synthesis using NifEN protein variants isolated from <i>A. vinelandii</i>	70
4.1.7	Cross-compatibility of selected NifEN variants with NifH proteins isolated from yeast mitochondria	70
4.1.7.1	Purification of <i>EcNifEN^{Ss}</i> , <i>EcNifEN^{Ht}</i> , and <i>EcNifEN^{Gm}</i> variants from <i>E. coli</i> cells	70
4.1.7.2	<i>In vitro</i> FeMo-co biosynthesis combined distinct NifH variants with three NifEN variants isolated from <i>E. coli</i>	72
4.1.8	FeMo-co biosynthesis using NifEN expressed in <i>S. cerevisiae</i>	73
4.1.8.1	Purification of <i>ScNifEN^{Ss}</i> , <i>ScNifEN^{Ht}</i> , and <i>ScNifEN^{Gm}</i>	73
4.1.8.2	Acetylene reduction assays using NifEN variants isolated from <i>S. cerevisiae</i>	75
4.1.8.3	FeMo-co biosynthesis pathway using proteins expressed in <i>S. cerevisiae</i> ..	76
5.	DISCUSSION	78
6.	CONCLUSIONS.....	87
	REFERENCES	88
	ANNEX.....	97

List of Figures

Figure 1. Structure of Mo-nitrogenase from <i>A. vinelandii</i> with associated metalloclusters.	5
Figure 2. Comparison of nif gene clusters from representative diazotrophic organisms.	8
Figure 3. Schematic illustration of NifS-mediated assembly of [4Fe-4S] clusters on NifU.	10
Figure 4. Structure of FeMo-co.	11
Figure 5. Simplified representation of FeMo-co synthesis centered at the transient complex formed between NifEN and NifH protein.	12
Figure 6. Proposed mechanism of NifB-co synthesis by NifB scaffold protein.	14
Figure 7. Comparison of NifEN and NifDK protein structures.	17
Figure 8. Experimental design followed in the presented thesis work to identify NifEN variants capable of forming soluble and functional protein complexes in <i>S. cerevisiae</i> mitochondria.	22
Figure 9. Simplified representation of parental plasmids pN2KD64 (A), pN2KD65 (B), and pN2KD66 (C) used to insert nifEN genes.	41
Figure 10. Typical batch fermentation process of the <i>S. cerevisiae</i> strain expressing ScNifEN ^{Gm} (KD37Y).	45
Figure 11. Simplified representation of FeMo-co synthesis and its insertion into the apo-form of NifDK.	54
Figure 12. In vitro FeMo-co synthesis and insertion assay procedure.	55
Figure 13. Schematic diagram of the strategy for tagging NifEN proteins expressed in <i>E. coli</i> cells.	58
Figure 14. Solubility and functionality of epitope-tagged NifEN proteins isolated from <i>E. coli</i>	59
Figure 15. Alignment of NifE and NifN proteins depicting conserved cluster-binding Cys residues.	61
Figure 16. In vivo functionality of NifEN candidates.	62
Figure 17. In vivo acetylene reduction activities of strains with restored Nif ⁺ phenotype.	63
Figure 18. Expression of NifEN variants in <i>S. cerevisiae</i>	64
Figure 19. The workflow for small-scale NifEN purification from cell-free extracts using magnetic beads.	65
Figure 20. Co-purification and purification experiments to identify soluble NifEN variants accumulating in the mitochondria of aerobically cultured <i>S. cerevisiae</i>	66
Figure 21. Sequence and structural divergence analysis of NifEN proteins in the library.	67
Figure 22. Isolation of NifEN variants from <i>A. vinelandii</i>	69
Figure 23. NifEN-dependent FeMo-co synthesis and apo-NifDK activation using AvNifEN ^{Ss} , AvNifEN ^{Ht} , and AvNifEN ^{Gm}	70
Figure 24. STAC purification of NifEN variants expressed in <i>E. coli</i>	71
Figure 25. Compatibility of NifEN candidates with distinct NifH variants.	73
Figure 26. STAC purification of NifEN variants expressed in aerobic cultures of <i>S. cerevisiae</i>	74

Figure 27. FeMo-co biosynthesis using NifEN isolated from <i>S. cerevisiae</i>	76
Figure 28. NifB-dependent FeMo-co synthesis and apo-NifDK activation using NifB, NifEN, and NifH proteins isolated from yeast.	77

List of Tables

Table 1. The genes involved in the synthesis and regulation of the Mo-nitrogenase.....	7
Table 2. E. coli-type strains used in this thesis.....	25
Table 3. A. vinelandii strains used and generated in this thesis.	26
Table 4. S. cerevisiae type-strain used in this thesis.	27
Table 5. S. cerevisiae W303-1a strains generated in this thesis.	28
Table 6. Primers used in this thesis.....	33
Table 7. PCR component concentrations.	34
Table 8. Standard PCR conditions.	34
Table 9. Colony PCR components.	34
Table 10. Colony PCR conditions.	35
Table 11. Scheme for a general overlapping PCR.....	35
Table 12. Plasmids to overexpress NifEN proteins in E. coli.....	39
Table 13. The list of plasmids to overexpress NifEN proteins in yeast.	42
Table 14. Generated plasmids to overexpress NifEN protein in A. vinelandii strain DJ2898.....	43
Table 15. Plasmids generated to isolate NifEN protein from Δ nifHDKTYnafAB A. vinelandii. ...	44
Table 16. Antibodies employed in this thesis.....	53
Table 17. Fe atom per NifEN protein and yield of NifEN from A. vinelandii cells.	69
Table 18. Fe atoms per protein and yield of NifEN from E. coli cells.....	72
Table 19. The yield of NifEN from S. cerevisiae cells.	75

Abbreviations and Acronyms

[Fe-S]	Sulfoferric cluster
5'dA·	5'deoxydadenosyl radical
A	Absorbance
AccB	Biotin carboxyl carrier protein of acetyl-CoA carboxylase
Acetyl-CoA	Acetyl Coenzyme A
Amp	Ampicillin
ARA	Acetylene Reduction Assay
ATP/ADP	Adenosine triphosphate / Adenosine diphosphate
<i>Av</i>	<i>Azotobacter vinelandii</i>
BNF	Biological Nitrogen Fixation
BSA	Bovine Serum Albumin
CFE	Cell-free protein extract
CV	Column Volume
Cys	Cysteine
<i>De</i>	<i>Dehalococcoides ethenogenes</i>
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphate
DTH	Sodium dithionite
DTT	1,4-dithiothreitol
E	Elution fraction
<i>e.g.</i>	Exempli gratia (“for example”)
<i>Ec</i>	<i>Escherichia coli</i>
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ELIC	Exonuclease and ligation-independent cloning
Fe	Iron
FeMo-co	Iron molybdenum cofactor [7Fe-9S-C-Mo-R-homocitrate]
FPLC	Fast Protein Liquid Chromatography
FT	Flow through fraction

g/mg/μg/	Grams/Milligrams/Micrograms/
GC	Gas Chromatography
GC-2014	Gas Chromatograph 2014
<i>Gm</i>	<i>Geobacter metallireducens</i>
<i>Gs</i>	<i>Geobacter sulfurreducens</i>
h	Hour
H ₂	Molecular hydrogen
HA	Human influenza hemagglutinin tag
His-tag	Histidine tag
HRP	Horseradish peroxidase
<i>Ht</i>	<i>Hydrogenobacter thermophilus</i>
<i>i.e.</i>	Id est (“that is”)
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
IMAC	Immobilized Metal Affinity Chromatography
ISC	Iron-Sulfur Cluster
Kan	Kanamycin
kb/bp	kilobase pair/base pair
kDa	kilodalton
l/ml/μl	liter/milliliter/microliter
LB	Lysogenic Broth
LiAc	Lithium acetate
LN ₂	Liquid nitrogen
M/mM/μM/nM	Molar/millimolar/micromolar/nanomolar
MCS	Multicloning sites
<i>Mi</i>	<i>Methanocaldococcus infernus</i>
min	Minute
Mo	Molybdenum
MOPS	3-(N-Morpholino)propane sulfonic acid
MPP	Mitochondrial processing peptidase
<i>Mt</i>	<i>Methanothermobacter thermautotrophicus</i>
MTP	Mitochondrial Targeting Peptides

MUSCLE	MULTiple Sequence Comparison by Log- Expectation
N	Nitrogen
N ₂	Molecular nitrogen
NH ₃	Ammonia
NH ₄ ⁺	Ammonium
NO ₂ ⁻	Nitrite
NO ₃ ⁻	Nitrate
O ₂	Molecular oxygen
OD _{600nm}	Optical density at 600 nm
<i>Para</i>	Arabinose promoter
PMSF	Phenylmethanesulfonyl fluoride
PycB	Pyruvate carboxylase subunit B
RMSD	Root-mean-square deviation
rpm	Revolutions per minute
<i>Rs</i>	<i>Roseiflexus</i> sp.
RT	Room Temperature
S	Sulfur
s	Second
SAH	S-Adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
<i>Sc</i>	<i>Saccharomyces cerevisiae</i>
SD	Standard deviation
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
<i>Ss</i>	<i>Synechococcus</i> sp. JA-2–3B'a
SS-tag	Strep-tag affinity tag
STAC	Strep-tag Affinity Chromatography
Strep-HRP	Streptactin-HRP
SUF	Sulfur mobilization pathway
TE	Total cell extract
TIM	Translocase of the Inner Membrane

TOM	Translocase of the Outer Membrane
TS/ts	TwinStrep-tag
V	Vanadium
v/v	Volume/volume
W	Wash fraction
w/v	Weight/volume
β-ME	β-Mercaptoethanol

1. INTRODUCTION

1.1 Nitrogen Fixation

1.1.1 Importance of nitrogen

Nitrogen is the seventh most abundant element in the universe, and it holds a pivotal role in sustaining life on Earth. Its versatility is evident from being a building block for an array of molecules essential for living organisms, e.g., amino acids (proteins), nucleic acids (such as DNA and RNA), adenosine triphosphate (ATP), and nicotinamide adenine dinucleotide (NAD⁺).

Approximately 67% of all nitrogen present in the biosphere is in the form of atmospheric dinitrogen gas (N₂) (Johnson and Goldblatt, 2015). Despite its abundance in the atmosphere, N₂ is largely inaccessible to most organisms, limiting primary productivity in many ecosystems. This limitation is particularly evident in agriculture because plants are unable to utilize atmospheric N₂ as a nitrogen source, leading to reduced crop productivity. Instead, plants rely on more accessible forms of nitrogen, such as ammonium (NH₄⁺) and nitrate (NO₃⁻), for their metabolic needs. These forms of nitrogen are either assimilated from the soil or can be obtained in the process called nitrogen fixation conducted by a group of prokaryotes capable of converting N₂ into these biologically active N species.

1.1.2 Plant-available nitrogen

In the soil, nitrogen exists in various inorganic forms such as ammonia (NH₃), NH₄⁺, nitrite (NO₂⁻), and NO₃⁻, which are directly absorbed by plants. Unfortunately, the largest portion of this essential nutrient is stored in the form of organic nitrogen, comprising urea, free amino acids, and short peptides, that are not readily accessible to plants. Nevertheless, certain plants (e.g., grass species found in the grasslands of North China) have the remarkable ability to directly take up low molecular weight organic nitrogen, including amino acids, peptides, and proteins (Tian et al., 2020). While plants cannot directly utilize most of the organic nitrogen, a gradual conversion process, known as mineralization, driven by soil microorganisms, transforms it into inorganic forms (NH₄⁺ or NO₃⁻), which are readily absorbed by plant roots (Liu et al., 2014). Despite this conversion process, the available nitrogen faces significant competition between plants and microorganisms for utilization. Additionally, nitrogen is consistently lost from the soil through various processes, including erosion, denitrification, leaching, volatilization of NH₃, and notably, the harvesting of crop residues containing nitrogen from the land. The intricate dynamics of nitrogen in the soil underscore the challenges and complexities associated with ensuring optimal nutrient availability for plant growth.

The limited availability of biologically usable nitrogen has been overcome by humans

through industrial processes, most notably the energy-intensive Haber-Bosch process. This technique involves the production of NH_3 from N_2 and hydrogen gas (H_2), utilizing iron (Fe) as a catalyst under high pressure and temperature conditions (Smil, 2004). The Haber-Bosch process has facilitated the large-scale production of ammonia, which serves as a critical component in the production of fertilizers, thereby supporting modern agriculture and enhancing crop productivity. However, this process, while instrumental in averting widespread starvation and ensuring food security for billions of people, also contributes to substantial environmental degradation. Notably, the Haber-Bosch process contributes significantly to greenhouse gas emissions. Firstly, hydrogen is generated through the steam-reformation of natural gas, which is subsequently combined with nitrogen gas from the atmosphere under extremely high temperatures ranging from 450°C to 600°C . Moreover, the production of nitrogenous fertilizers requires massive amounts of fossil fuels, accounting for approximately 2% of the world's energy consumption. Regrettably, a significant portion of the nitrogen applied to the soil does not get absorbed by crops. Nearly 50% of the nitrogen provided through fertilizers is lost due to leaching and other factors during various agricultural processes (Malinas et al., 2022). These cumulative effects result in increased levels of waste and pollution, leading to detrimental effects on soil health and the overall environment (Sammauria et al., 2020). In contrast to fossil fuel-dependent nitrogen fertilizers, biological nitrogen fixation offers a sustainable and environmentally friendly source of nitrogen for crop production.

1.1.3 Biological nitrogen fixation

The naturally occurring process of Biological Nitrogen Fixation (BNF), converting N_2 into NH_4^+ , serves as a major nitrogen source for plants and a pivotal step in nourishing the ecosystem. BNF is performed by certain microorganisms called diazotrophs that belong to the Bacteria or Archaea domains, as no eukaryotes can process the conversion of N_2 into a biologically active nitrogen species. Despite the relatively small number of species endowed with nitrogen-fixing abilities, diazotrophs exhibit impressive adaptability, thriving in diverse ecological environments across the globe. They represent a wide range of phylogenetic and physiological types and showcase different metabolic types, as they can be classified as phototrophs or chemotrophs, autotrophs, or heterotrophs. In nature, diazotrophs can exist as free-living organisms, and engage in non-symbiotic or symbiotic associations with living plants (Sprent and Sprent, 1990; Giller, 2001).

Nitrogen fixation is catalyzed by nitrogenase, an enzyme that is remarkably similar across most nitrogen-fixing bacteria. However, nitrogenase is rendered inactive in aerobic environments due to its extreme sensitivity to oxygen (O_2). Consequently, diazotrophs can only perform this process under anaerobic or O_2 -protective conditions and can be classified based on their sensitivity to O_2 . Obligate anaerobic N_2 -fixing organisms (such as *Clostridium*) can

only perform BNF in an O₂-free environment. Facultative anaerobes, like *Klebsiella oxytoca* and *Paenibacillus sabinae*, can grow in both anaerobic and aerobic conditions, but they can only fix nitrogen in the absence of O₂. Notably, the extensively studied diazotrophic species *Azotobacter vinelandii* belongs to the group of obligate aerobes and has developed special mechanisms to facilitate BNF under O₂ protective conditions (Burris and Roberts, 1993). In the case of *A. vinelandii*, rapid O₂ removal achieved from increased respiration can protect the enzyme from O₂ damage (Gallon, 1992), while the mechanism of extracellular alginate production is slowing down the diffusion of O₂ within the cell (Sabra et al., 2000). Furthermore, the *A. vinelandii* FeSII protein (called Shetna) binds to the nitrogenase enzyme, providing further protection against O₂-induced damage (Maier and Moshiri, 2000, Schlesier et al., 2016). Interestingly, in cyanobacteria, nitrogen fixation occurs in specialized thick-walled cells, called heterocysts. Heterocyst formation does not occur spontaneously; instead, it follows a complex and regulated mechanism known as temporal separation. This separation allows cyanobacteria to efficiently allocate their nitrogen-fixing machinery, ensuring that heterocysts are only produced when they are truly needed. The heterocyst formation involves nitrogen deficiency sensing, and heterocyst inception, development, and maturation. By carefully regulating the timing of heterocyst differentiation, cyanobacteria protect their nitrogen-fixing machinery and minimize energy consumption (Wolk, 1996).

1.1.4 Engineering nitrogen fixation in plants

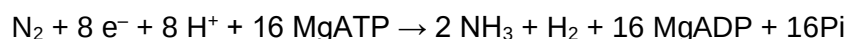
Considering the ecological significance and readily accessible nitrogen sources provided by free-living, non-symbiotic, and plant symbiotic diazotrophs, the ability to engineer BNF into plants holds immense promise for revolutionizing agriculture. This endeavor aims to empower plants with the ability to directly fix N₂ from the atmosphere, eliminating the reliance on chemical fertilizers and reducing environmental impact. Remarkable advancements in our understanding of the BNF process, coupled with the advancements in plant genetic tools and synthetic biology, have outlined three biotechnological approaches to enhance BNF in crop plants. One approach to enhance nitrogen fixation in cereal crops involves modifying naturally occurring bacteria that are already associated with these plants. By enhancing the bacteria's N₂-fixing capabilities, and ability to release nitrogenous compounds to the plant, crop nitrogen uptake and yield could be improved. The advantage of this approach is that it does not require genetic modification of the plant, making it more acceptable in countries with strict regulations regarding the growth of transgenic crops. The second approach aims at engineering symbiosis between N₂-fixing bacteria and non-legume crops such as cereals. This involves the engineering of cereals to recognize and interact with N₂-fixing bacteria, and to form nodules or nodule-like structures that provide the low-O₂ environment required for nitrogen fixation. The third strategy aims at transferring the prokaryotic genes encoding the nitrogenase enzyme into

the plant genome. This would allow the plant to synthesize its own nitrogenase and fix nitrogen independently of bacterial interactions. However, this approach is technically challenging due to the sensitivity of nitrogenase towards O₂, and the large collection of accessory proteins and metal clusters needed for maturation and functionality of nitrogenase.

The following chapters will explore the complex mechanisms of nitrogenase function, examining the formation of metalloclusters, with particular emphasis on the biosynthesis of the cofactor at the active site.

1.2 Nitrogenase

Nitrogenase catalyzes biological fixation of N₂ in a high energy-demanding reaction that requires eight electrons and a minimal of 16 ATP equivalents to fix one molecule of N₂. In aerobic respiratory metabolism, these requirements equal eleven molecules of NADH per molecule of N₂, while 26 molecules of NADH are required from anaerobic respiratory metabolism.



All diazotrophs harbor a molybdenum–iron nitrogenase (Mo-nitrogenase) system that contains a two-component O₂-sensitive protein complex: the dinitrogenase (NifDK) and dinitrogenase reductase (NifH). The NifH protein, a γ₂ homodimer of ~62 kDa dimer, harbors a [4Fe-4S] cluster that is located at the interface of the two γ subunits. Each γ subunit additionally contains a MgATP binding and hydrolysis site located at the C-terminus of each NifH polypeptide. The NifH protein serves as an ATP-dependent electron donor, transferring electrons to the larger 230 kDa heterotetramer NifDK protein composed of two asymmetric αβ dimers. The NifDK protein harbors two different types of complex metalloclusters: a [7Fe-9S-C-Mo-*R*-homocitrate] cluster called FeMo-co (for iron-molybdenum cofactor) and an [8Fe-7S] cluster called the P-cluster. The P-cluster is involved in intercomponent electron transfer, while the FeMo-co provides the site for N₂ binding and reduction (Figure 1).

Mo-nitrogenase reduces N₂ through a series of steps: transfer of electrons from NifH to the NifDK protein via inter-component electron transfer, leading to substrate activation and reduction. The process initiates with the formation of a transient complex between the reduced and MgATP-bound NifH protein and one-half of the NifDK protein (Hageman and Burris, 1978). Subsequently, this complex triggers a series of structural changes facilitating the efficient transfer of electrons from the reduced NifH to the P-cluster and ultimately to FeMo-co. This transfer is powered by the breakdown of ATP, releasing a Pi as a byproduct, and culminates in the dissociation of the complex. This cycle is repeated eight times to reduce one N₂ molecule

into two molecules of NH_3 (Hageman and Burris, 1978, Burgess and Lowe, 1996, Duval et al., 2013, Hoffman et al., 2014).

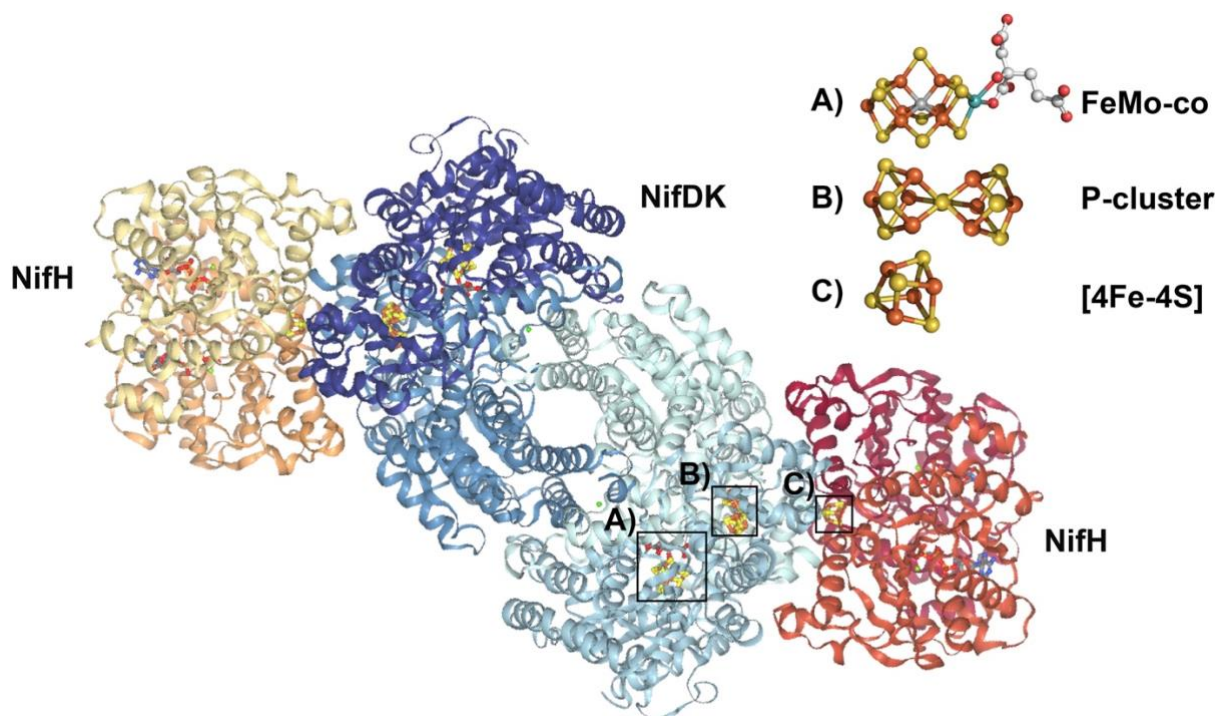


Figure 1. Structure of Mo-nitrogenase from *A. vinelandii* with associated metalloclusters. The NifH protein subunits are shown in yellow and red and the NifDK protein is shown in blue (PDB: 1G21). The right part shows the metal cofactors within the nitrogenase enzymatic complex where (A) indicates FeMo-co, (B) P-cluster, and (C) [4Fe-4S] cluster.

In addition to the Mo-nitrogenase, two other nitrogenases exist in Nature that are mechanistically similar but genetically distinct. These alternative nitrogenases are classified according to their active site cofactor: the vanadium-type nitrogenase (V-nitrogenase) containing FeV-co [V-7Fe-8S-C-R-homocitrate] and the Fe-only nitrogenase containing FeFe-co [8Fe-9S-C-R-homocitrate]. Among the three nitrogenases, the Mo-type is the most efficient in N_2 reduction and the most widely studied. The Mo nitrogenase is the subject of this work.

1.2.1 Genetic complexity of the molybdenum nitrogenase

Due to the complex biochemical nature of BNF, Mo-nitrogenase apart from NifH and NifDK proteins, requires multiple additional gene products to fulfill its functional capacity. The genes responsible for Mo-nitrogenase catalytic activity, assembly, and regulation are known as *nitrogen fixation (nif)* genes. The *nif* gene regulon consists of a cluster of genes that are typically organized into operons. Genomic studies have shown that the most basic *nif* operon comprises six *nif* genes: *nifH*, *nifD*, *nifK*, *nifE*, *nifN*, and *nifB*. These genes are present in all known diazotrophs and are essential to form a functional nitrogenase (Dos Santos et al., 2012). The *nifH*, *nifD*, and *nifK* genes encode the catalytic nitrogenase components, while *nifB*, *nifE*,

and *nifN* encode proteins required for the biosynthesis of the nitrogenase active site cofactor FeMo-co.

NifH, NifDK, NifB, and NifEN all require [Fe-S] clusters for their functionality. Often these clusters originate (in precursor form or as the final clusters) from NifU and NifS. Interestingly, not all known diazotrophs have *nifU* and *nifS* and these organisms instead employ other Fe-S assembly pathways, e.g., Isc or Suf (Ayala-Castro et al., 2008).

The following section describes the organization of *nif* genes from the model organism *A. vinelandii*. This diazotrophic organism, with its well-annotated *nif* gene clusters is an ideal model for understanding the function of nitrogenase genes. In *A. vinelandii*, *nif* genes are clustered in two distinct regions of its genome. These regions are named according to their size as major and minor *nif* gene clusters. Some of the *A. vinelandii* genes located within the major or minor *nif* clusters have the designation “*naf*”, as they encode nitrogenase-associated factors (Rubio et al., 2002). Nevertheless, all *nif* genes can be grouped based on their function (Buren et al., 2020): 1) genes encoding structural proteins of nitrogenase (*nifD* and *nifK* genes for the NifDK component, also called the MoFe protein, and *nifH* gene for the NifH component also called the Fe protein), 2) genes for FeMo-co biosynthesis and transfer (*nifB*, *fdxN*, *nifE*, *nifN*, *nifX*, *nifQ*, *nifV*) and insertion (*nafY* also involved in stabilization of NifDK protein with P-cluster but lacking FeMo-co), 3) genes encoding proteins for assembly of simpler Nif-related [Fe-S] clusters (*nifU* and *nifS*), 4) genes involved in electron transfer to the nitrogenase enzyme (*nifF* and *rnfABCDEFG*), 5) genes encoding proteins for maturation of NifH (*nifM*) and NifDK (*nafH*, *nifW* and *nifZ*), and 6) regulatory genes controlling expression of other *nif* genes in response to environmental signals (*nifA* and *nifL*) (Table 1).

Besides the *nif* gene clusters, *A. vinelandii* also possess genes for the synthesis and regulation of the alternative iron-only nitrogenase (*anf* genes) and vanadium-nitrogenase (*vnf* genes). The genes of these nitrogenase systems are in their own gene clusters. In addition, the production of functional Anf and Vnf nitrogenase isozymes requires five *nif* genes (*nifB*, *nifU*, *nifS*, *nifV*, and *nifM*) that do not have paralogous in the alternative gene clusters (Kennedy et al., 1986, Kennedy and Dean, 1992).

Table 1. The genes involved in the synthesis and regulation of the Mo-nitrogenase.

Gene	Roles of the gene products.
<i>nifD</i>	α -subunit of the NifDK protein; FeMo-co binding subunit.
<i>nifK</i>	β -subunit of the NifDK protein. Forms an $\alpha_2\beta_2$ tetramer with NifD, the catalytic component of nitrogenase.
<i>nifH</i>	Transfers electrons to NifDK. Inserts FeMo-co into the NifDK. Stabilizes NifDK protein.
<i>nifE</i>	Assembles FeMo-co. FeMo-co precursor binding subunit. Forms an $\alpha_2\beta_2$ tetramer with NifN.
<i>nifN</i>	Assembles FeMo-co. Forms an $\alpha_2\beta_2$ tetramer with NifE.
<i>nifB</i>	Catalyzes the FeMo-co precursor formation.
<i>nifU</i>	Provides [Fe-S] cluster precursors for the Nif proteins and their cofactors.
<i>nifS</i>	Provides S for [Fe-S] cluster assembly on NifU.
<i>nifX</i>	Involves in the transfer of FeMo-co precursor from NifB to NifE.
<i>nifV</i>	Homocitrate synthase involved in the attachment of R-homocitrate in the FeMo-co formation process.
<i>nifQ</i>	Supplies Mo prior to its incorporation into FeMo-co.
<i>nafY</i>	Stabilizes NifDK protein. Facilitates the transfer of FeMo-co from NifE to NifDK.
<i>nifF</i>	Flavodoxin. Participates in the nucleotide-dependent electron transfer to NifH.
<i>rnfABCDEF</i>	Membrane protein complex that generates electrons for nitrogenase.
<i>fdxN</i>	Supports FeMo-co precursor synthesis by delivering electrons to NifB during catalysis.
<i>nifM</i>	Ensures stabilization and maturation of NifH.
<i>nafH</i>	Binds to NifDK prior to its maturation process.
<i>nifW</i>	Interacts with the NifDK during NifDK maturation.
<i>nifZ</i>	Together with NifH involved in the maturation of NifDK.
<i>nifA</i>	Transcriptional activator of the <i>nif</i> gene cluster. Regulates and control the transcription of nitrogenase together with NifL.
<i>nifL</i>	Inhibits the activity of NifA by forming a complex with NifA.

The minimum set of six conserved *nif* genes (Dos Santos et al., 2012) is used for *in-silico* analysis when identifying diazotrophic species. A comparison of *nif* clusters among different diazotrophs shows that *A. vinelandii* exhibits extensive cluster complexity, with numerous auxiliary genes encoding proteins likely associated with the presence of alternative nitrogenase compartments or strict aerobic metabolism. *K. oxytoca*, another well-studied model diazotroph, has 20 *nif* genes while *Paenibacillus* sp. WLY78 only has nine (*nifBHDKENXhesAnifV*) (Figure 2). It has been shown that these clusters, when transferred to the non-diazotroph *Escherichia coli*, were sufficient to assemble an active nitrogenase (Wang et al., 2013, Dixon and Postgate, 1972). As *K. oxytoca* and *Paenibacillus* sp. WLY78 are facultative anaerobes and only fix nitrogen under anaerobic conditions, it has been proposed

that such an environment requires fewer genes for nitrogen fixation (Boyd et al., 2015). Some archaea contain the simplest *nif* gene organization. For example, in the case of *Methanococcus maripaludis*, six *nif* genes (*nifH*, *nifD*, *nifK*, *nifE*, *nifN*, and *nifX*) are located within a single operon. The *nifB* gene, required for FeMo-co synthesis, is located elsewhere in the genome of this organism, outside of the *nif* gene cluster (Kessler et al., 1998). In certain diazotrophs, *nif* genes are “fused”, so that two Nif proteins are expressed as a single polypeptide, e.g., *nifEN* in *Anabaena variabilis* (Thiel and Pratte, 2014) and *Geobacter sulfurreducens* (Ueki and Lovley, 2010), or *nifNB* in *Clostridia* (Chen et al., 2001).

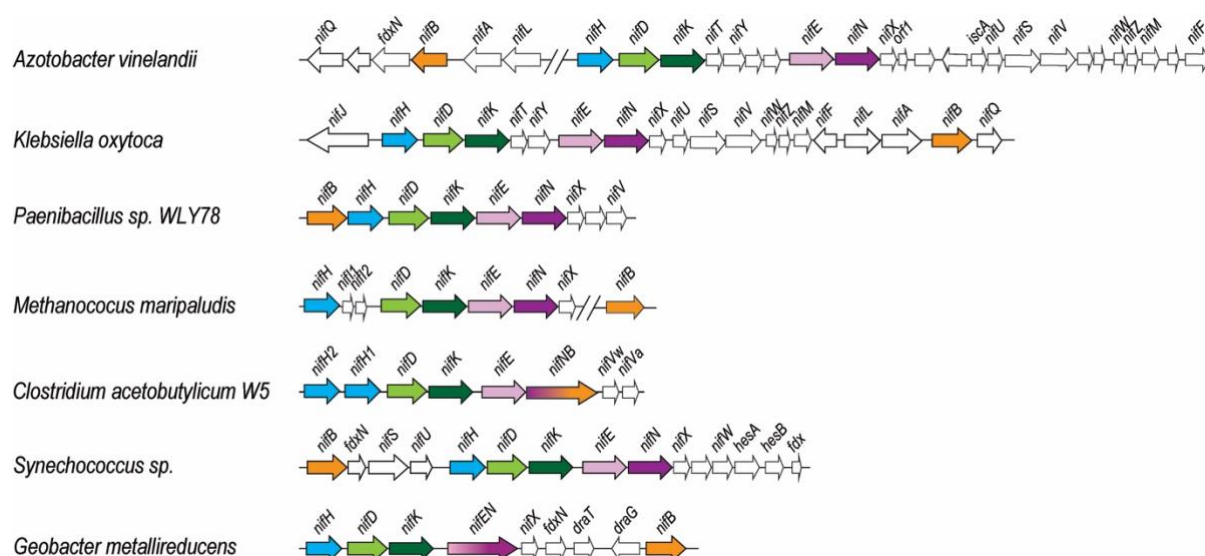


Figure 2. Comparison of nif gene clusters from representative diazotrophic organisms. Colored segments indicate six conserved nif genes essential for nitrogenase functionality (orange, nifB; blue, nifH; green, nifDK; and purple, nifEN).

1.3 Biosynthesis of nitrogenase clusters

The assembly and maturation of nitrogenase cofactors are crucial processes for Mo-nitrogenase functioning. The maturation of NifH necessitates the formation and incorporation of pre-formed [4Fe-4S] clusters, while the more complex metalloclusters at NifDK are assembled via different pathways. As mentioned before, in *A. vinelandii*, the biosynthesis of Fe-S clusters is achieved by the activity of the NifU and NifS proteins. These simple clusters are required for the biosynthesis of nitrogenase cofactors and are necessary for the maturation of nitrogenase components. The P-cluster is synthesized from two [4Fe-4S] precursors situated within the NifDK protein. At the same time, the biosynthesis of FeMo-co takes place outside of NifDK and involves NifB and NifEN proteins.

More details about nitrogenase cofactor biosynthesis and maturation of nitrogenase components into an active enzyme are provided in the following sections.

1.3.1 Biosynthesis of [4Fe-4S] clusters by NifU and NifS

To ensure the complete functionality and activation of Mo-nitrogenase, the catalytic components of nitrogenase and proteins involved in the biosynthesis of its essential cofactors necessitate the biosynthesis of [Fe-S] cluster building blocks. In this process of Fe-S cluster formation, NifU and NifS are involved (Yuvaniyama et al., 2000).

Previous studies have shown that *A. vinelandii* mutants lacking *nifU* did not grow under diazotrophic conditions, whereas very slow growth was observed in the *nifS* mutant strain. Moreover, when both the *nifU* and *nifS* genes were deleted, the activities of NifH and NifDK were significantly interrupted (about 90% lower activity), even when the NifDK protein was *in vitro* reconstituted with the addition of FeMo-co and a high amount of NifH. This proves that *nifUS* mutant strains produced a non-functional NifDK protein, likely due to the absence of the P-cluster (Jacobson et al., 1989b).

Furthermore, the non-diazotrophic growth of *A. vinelandii* strains with disrupted *nifU* and *nifS* genes results in non-functional Mo-dependent nitrogenase and impairs the accumulation of functional alternative nitrogenases (Kennedy and Dean, 1992).

Additionally, it has been shown that the phenotype observed in mutants with disrupted *isc* genes (iron-sulfur cluster genes required for maturation of [Fe-S] proteins involved in general metabolic processes) could be rescued by NifU (for *iscU*) but not NifS (for *iscS*). Also, *A. vinelandii* lacking *nifU* exhibited only partial diazotrophic growth when *iscU* was overexpressed, indicating that the NifU and NifS proteins are essential for nitrogenase [Fe-S] cluster maturation *in vivo* (Dos Santos et al., 2007, Johnson et al., 2005).

As mentioned previously, both diazotrophs and non-diazotrophs possess homologs of *nifU* and *nifS* involved in the assembly or restoration of [4Fe-4S] clusters. These homologous systems, including *isc* (Zheng et al., 1998) and *suf* (sulfur utilization factor system) (Takahashi and Tokumoto, 2002) appear to operate similarly to NifUS. Cysteine desulfurases (IscS/SufS) deliver inorganic sulfur for the biosynthesis of [2Fe-2S] and/or [4Fe-4S] clusters, which are then transferred to scaffold proteins (IscU/SufU or IscA/SufA) (Zheng et al., 1998, Takahashi and Tokumoto, 2002, Sendra et al., 2007).

NifS is a cysteine desulfurase homodimer of 88 kDa that contains a pyridoxal-5'-phosphate (PLP) cofactor to catalyze the reaction of desulfurization of L-cysteine into L-alanine and elemental sulfur (S) (Zheng et al., 1993, Zheng and Dean, 1994). This process involves the generation of NifS-bound persulfide intermediate that is transferred to NifU (Yuvaniyama et al., 2000).

NifU is a protein dimer made from two identical 33 kDa monomers (Fu et al., 1994). Analysis of NifU primary sequences shows nine conserved cysteines, located within three

domains in the NifU polypeptide, that independently coordinate [Fe-S] cluster assembly (Figure 3).

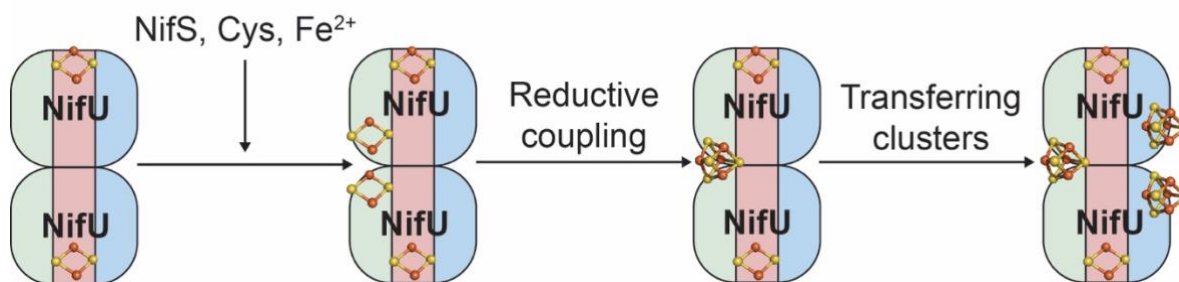


Figure 3. Schematic illustration of NifS-mediated assembly of [4Fe-4S] clusters on NifU. The [4Fe-4S] cluster assembly process starts at the N-terminal domain of NifU (area in green), where [2Fe-2S] clusters are initially formed. These clusters are subsequently united through reductive coupling to form the [4Fe-4S] cluster, which is transferred from the N-terminal domain to the C-terminal domain of NifU. The central domain (area in red) binds the permanently associated [2Fe-2S] cluster. The C-terminus of NifU (area in blue) facilitates the rapid assembly of [4Fe-4S] without requiring [2Fe-2S] intermediates. Adapted from (Buren et al., 2020).

The N-terminal domain has three Cys residues (Cys³⁵, Cys⁶², and Cys¹⁰⁶) involved in the assembly of two [2Fe-2S] clusters to form a transient [4Fe-4S] cluster, while the central domain contains a permanent [2Fe-2S] cluster ligated with four Cys residues (Cys¹³⁷, Cys¹³⁹, Cys¹⁷² and Cys¹⁷⁵). The two Cys residues (Cys²⁷² and Cys²⁷⁵) located at the C-terminal domain of each monomer coordinate the formation of a [4Fe-4S] cluster. Both N- and C-terminal [4Fe-4S] clusters are transient and destined for insertion into apo-Nif proteins, while the central domain binds a permanent and redox-active [2Fe-2S]^{2+/1+} cluster. This central cluster is not a precursor cluster but is necessary for NifU activity (Yuvaniyama et al., 2000, Dos Santos et al., 2004, Smith et al., 2005). Following their synthesis on NifU, assembled [4Fe-4S] clusters are delivered to NifH, NifB or NifDK, where they undergo further processing into more complex clusters.

1.3.2 P-cluster formation

For its functionality, NifDK requires the formation of the matured P-clusters before the insertion of the active site FeMo-co (Hu et al., 2007). The formation of the P-cluster relies on the delivery of 4Fe-4S clusters by NifU. It is hypothesized that in this process, two pairs of [4Fe-4S] clusters bind to NifDK (one cluster to each α and β subunit) and are then fused into two P-clusters by reductive coupling requiring the participation of NifH and MgATP (Hu et al., 2007, Cotton et al., 2009). The formation of the P-cluster in *A. vinelandii*, apart from the NifH protein, involves three accessory factors: NifH, NifW, and NifZ. Interestingly, the deletion of their respective gene does not render nitrogenase inactive, suggesting that they are not

essential for the maturation of the P-cluster (Hu et al., 2004, Jimenez-Vicente et al., 2019, Jacobson et al., 1989b). While the specific functions of NifH, NifW, and NifZ are not known yet, they are likely assisting in the binding of [4Fe-4S] clusters to the α and β subunits of NifDK, or associate with apo-NifDK to stabilize the conformation of the immature NifDK protein (Van Stappen et al., 2022, Jiménez-Vicente et al., 2018).

Activation of NifDK that contains mature P-clusters but not FeMo-co (often and from now on in this thesis designated as apo-NifDK) requires the separate formation and subsequent insertion of FeMo-co (Christiansen et al., 1998, Allen et al., 1993, Paustian et al., 1990). The FeMo-co biosynthesis process is mediated by the scaffold protein NifEN, encoded by *nifE* and *nifN* genes, which are paralogs of the *nifD* and *nifK* genes. The following sections will cover the mechanism of FeMo-co synthesis by NifEN.

1.3.3 FeMo-co synthesis

FeMo-co, a unique [7Fe-9S-C-Mo-*R*-homocitrate] cofactor, is composed of a single carbide atom situated in the cavity (central cage) formed by the six central Fe atoms. This inorganic structure is completed by one terminal Fe atom and one Mo atom located at the ends of the cofactor. The Mo atom is additionally coordinated to *R*-homocitrate (Figure 4).

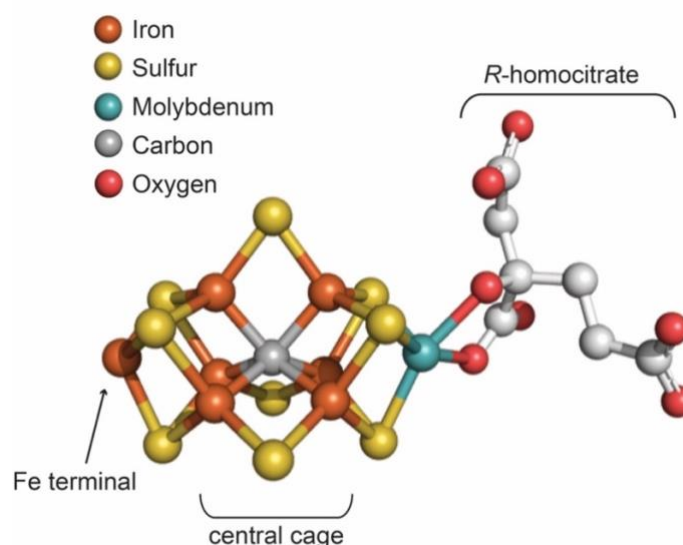


Figure 4. Structure of FeMo-co.

Assembly of FeMo-co is a complex, multi-step process involving seven Nif proteins categorized into three groups based on their specific roles during FeMo-co synthesis: 1) NifU, NifB, and NifEN act as a scaffold for the assembly of intermediate metalloclusters for FeMo-co biosynthesis, 2) NifS, NifV and NifQ deliver S, *R*-homocitrate and Mo, respectively, and 3) NifX facilitates the transport of the FeMo-co precursor metallocluster, NifB-co, from NifB to NifEN (Figure 5) (Buren et al., 2020).

The final assembly of FeMo-co occurs on the NifEN scaffold protein and requires a transient association between the NifEN and NifH protein (Hu et al., 2006, Hernandez et al., 2008). The NifEN/NifH complex serves as an acceptor for the essential substrates that are delivered by three distinct delivery branches to form FeMo-co (Figure 5): 1) the Fe-S cluster delivery branch involving NifU, NifS, and NifB to synthesize the FeMo-co precursor, called NifB-co, which is subsequently delivered to NifEN, 2) the molybdenum delivery pathway involving the action of NifQ, responsible for the delivery of the Mo atom to the NifEN/NifH complex, and 3) the homocitrate delivery pathway, involving NifV (homocitrate synthase) that produces the organic constituent *R*-homocitrate (McLean and Dixon, 1981, Hoover et al., 1989). Notably, under laboratory conditions, NifV and NifQ can be replaced by the addition of homocitrate and Mo (Curatti et al., 2007).

The following section will describe the proteins involved in the delivery branches to the NifEN protein, contributing to the formation of mature FeMo-co.

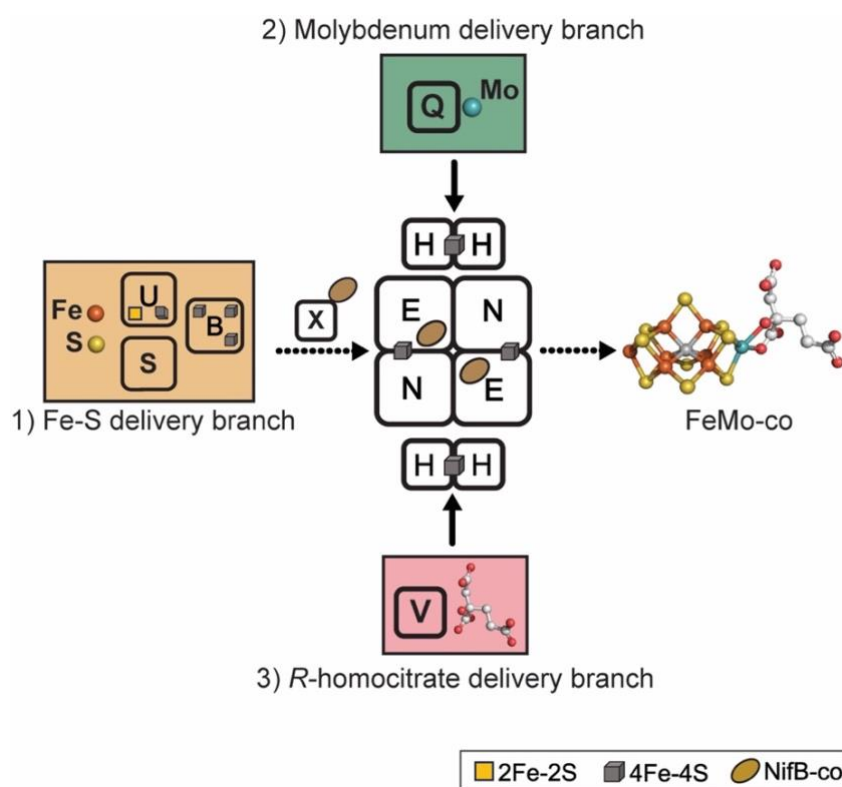


Figure 5. Simplified representation of FeMo-co synthesis centered at the transient complex formed between NifEN and NifH protein. The three delivery branches, annotated with the involved proteins, provide intermediate metallocluster (NifB-co), Mo, and *R*-homocitrate to the NifEN scaffold protein. The indicated NifX delivers NifB-co (assembled on NifB) to NifEN protein.

1.3.4 NifB forms FeMo-co precursor

NifB protein belongs to an S-adenosyl-L-methionine (SAM) radical enzyme group. This group of enzymes participates in a variety of biochemical reactions such as the synthesis of

cofactors or peptide modifications (Broderick et al., 2014, Nicolet, 2020). The radical SAM enzymes generate the radical species 5'deoxyadenosyl radical (5'dA $\cdot$) through the reductive cleavage of SAM using a [4Fe-4S] cluster as reductant. This radical intermediate abstracts the H atom from the substrate and initiates the subsequent catalytic reaction (Frey et al., 2008).

The majority of NifB proteins, originating from bacterial organisms (such as those found in *A. vinelandii* and *K. oxytoca*), consist of an N-terminal SAM-radical domain with a CxxxCxxC motif that binds a [4Fe-4S]-SAM cluster, and a C-terminal NifX-like domain. Later studies on nitrogen-fixing archaea identified NifB proteins that lack the C-terminal NifX-like domain, suggesting that it is not essential for NifB activity (Wilcoxon et al., 2016, Fay et al., 2015). These proteins contain only the SAM-radical domain and were identified as monomers of about 30-35 kDa (Echavarri-Erasun et al., 2014, Fay et al., 2015, Wilcoxon et al., 2016), in contrast to the NifB homodimers with 50-55 kDa subunits isolated from *A. vinelandii* or *K. oxytoca* (Sibold et al., 1983, Joerger and Bishop, 1988).

NifB is known to harbor three [4Fe-4S] clusters. One cluster is coordinated by a SAM molecule, named RS-cluster, and is required for the catalytic activity and insertion of the central carbide during NifB-co synthesis. The two other [4Fe-4S] clusters, called K1 and K2, are precursors that form the core of NifB-co (Wiig et al., 2011, Wilcoxon et al., 2016, Rettberg et al., 2018). Recent structural analysis of NifB suggests the presence of an [8Fe-8S] intermediate to NifB-co, called K-cluster, that is similar to the P-cluster (at NifDK) in structure (Jenner et al., 2021). The identification of this [8Fe-8S] intermediate led researchers to propose a model for the assembly of NifB-co (Jenner et al., 2021).

The identification of this [8Fe-8S] intermediate led researchers to propose a model for the assembly of NifB-co (Jenner et al., 2021). In this model (Figure 6), the K1 and K2 substrate clusters are initially delivered to NifB by the NifU scaffold protein. Subsequently, K1 and K2 undergo reductive coupling, generating a P-cluster-like species called the K-cluster ([8Fe-8S]). This event promotes the binding of SAM to the RS-cluster and is followed by the transfer of a methyl group to the external bridging sulfide of the K-cluster, forming S-Adenosyl-L-homocysteine (SAH) as subproduct. The SAH formed is then replaced by a second molecule of SAM, which participates in radical SAM reaction producing the 5'dA $\cdot$ radical. This then results in the abstraction of one H atom from the methyl group and facilitates the core rearrangement and two subsequent deprotonation reactions, generating the central carbide within the [Fe₈S₈C] cluster intermediate. To complete the NifB-co assembly process, an additional 9th sulfur atom is incorporated, likely by NifS or a sulfur transferase, resulting in the formation of [8Fe-9S-C] NifB-co.

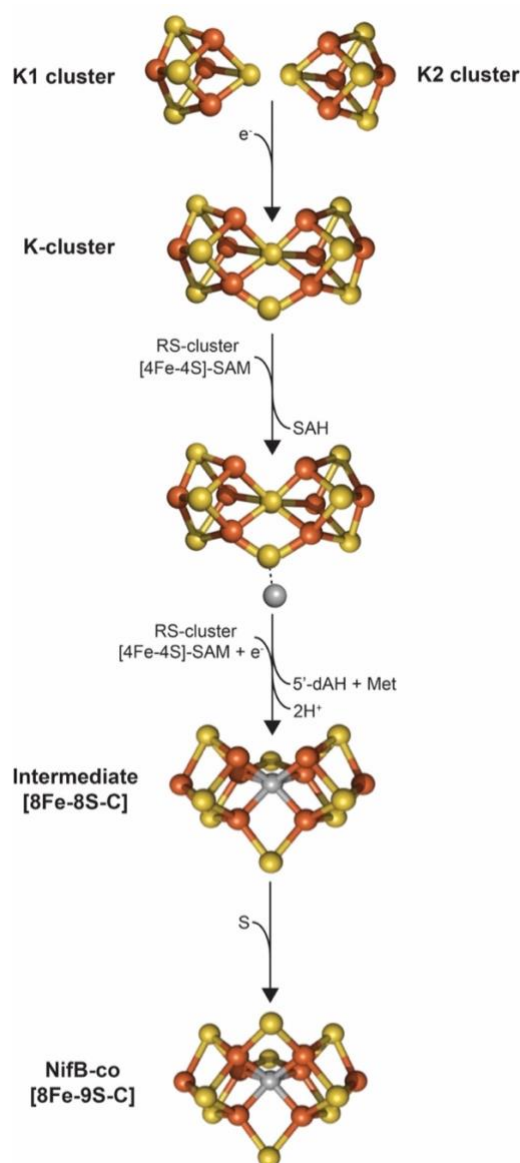


Figure 6. Proposed mechanism of NifB-co synthesis by NifB scaffold protein. Substrate K1 and K2 sub-clusters generate $[8Fe-8S]$ K-cluster. In the next step, the RS-cluster promotes the transfer of a methyl group from the RS-cluster-bound SAM to the K-cluster, producing SAH that is then exchanged by a second SAM molecule. This reaction produces 5'dA and leads to the abstraction of one H atom from the methyl group. The subsequent deprotonation event generates the carbide in the $[8Fe-8S-C]$ Intermediate. In the final step, NifB-co formation is completed by the incorporation of a sulfur atom, likely delivered by either NifS or a sulfur transferase enzyme. Adapted from (Martin Del Campo et al., 2022).

In the minor operon of *A. vinelandii*, *nifB* is located next to a ferredoxin gene (*fdxN*). While the precise function of FdxN is not yet fully known, analysis of an *A. vinelandii* $\Delta fdxN$ mutant strain showed a significantly reduced nitrogenase activity (Rodríguez-Quñones et al., 1993). Later, it was proposed that this decline was linked to impaired NifB-co biosynthesis, as $\Delta fdxN$ and $\Delta nifENX$ double mutant strains showed 20% lower cofactor activity compared to strains expressing FdxN (Jiménez-Vicente et al., 2014). It has been proposed that FdxN could fulfill a role as electron donor, as NifB requires reductant for the formation of the K cluster and to trigger the cleavage of SAM.

1.3.5 NifX carriers NifB-co to NifEN

NifX functions as a metallocluster transfer protein, facilitating the delivery of FeMo-co precursor synthesized on NifB (NifB-co) to NifEN, ultimately leading to FeMo-co synthesis (Hernandez et al., 2007). It has also been proposed that NifX acts as a NifB-co storage protein, regulating the flow of FeMo-co precursors to the NifEN protein (Hernandez et al., 2007).

Interestingly, the presence of NifX has not been reported in many N₂-fixing organisms, indicating that it is not strictly required *in vivo* (Jacobson et al., 1989a, Rubio et al., 2002). Notably, certain nitrogen-fixing *Clostridium* species possess the natural fusion of the *nifB* and *nifN* genes, potentially eliminating the need for NifX due to the direct transfer of NifB-co from NifB to NifEN (Chen et al., 2001). Also, although NifX is not required in *in vitro* FeMo-co synthesis assays (Curatti et al., 2007), its addition to the reaction mixture can enhance the activity (Shah et al., 1999, Curatti et al., 2007).

Importantly, in the context of this thesis, *in vitro* FeMo-co synthesis assays were performed in the presence of NifX, or NifB-co bound to NifX, to mimic *in vivo* FeMo-co biosynthesis in *A. vinelandii*.

1.3.6 NifH in FeMo-co synthesis

NifH has three functions. In addition to donating electrons to NifDK during N₂ reduction, and for P-cluster maturation, it plays an essential role during the synthesis of FeMo-co. Although being fundamentally important for FeMo-co synthesis, the exact function of NifH is still a subject of debate (Hu et al., 2005, Soboh et al., 2006, Curatti et al., 2007, Rubio and Ludden, 2008).

It was found that interaction between NifH and NifEN is required for the FeMo-co maturation process, wherein the NifQ protein, harboring a [3Fe-4S-Mo] cluster, transfers Mo to NifEN containing FeMo-co synthesis precursor. If NifH was absent from the FeMo-co synthesis assays *in vitro*, no FeMo-co was formed as neither Mo, nor homocitrate, was incorporated (Hernandez et al., 2008). According to this study, Ludden and collaborators proposed that NifH docks onto the surface of NifEN, and that this interaction induces conformational changes in the complex that enables the subsequent transfer of Mo, by NifQ, and the incorporation of homocitrate, delivered by NifQ, ultimately leading to the formation of FeMo-co.

However, some studies argue that Mo and homocitrate is delivered by NifH to NifEN, which replaces the Fe atom in the FeMo-co precursor to enable the FeMo-co maturation (Hu et al., 2006).

1.3.7 Role of NifQ in FeMo-co synthesis

A. vinelandii NifQ is a monomeric 20 kDa protein that is sensitive to O₂. It has been isolated with 3.1 Fe atoms and 0.3 Mo atoms per monomer. NifQ harbors a conserved Cx₄Cx₂Cx₅C motif at its C-terminus, which could be involved in metal cluster coordination. EPR analysis has shown that NifQ may carry either a redox-responsive [Mo-3Fe-4S] cluster, or a Mo-free [3Fe-4S] cluster (Hernandez et al., 2008). Moreover, the presence of independent MoO₂S₂ species bound to NifQ was observed. Intriguingly, both the [Mo-3Fe-4S] cluster and MoO₂S₂ have been found to be crucial for FeMo-co synthesis. Elimination of either of these Mo donors resulted in the abolishment of Mo donation by NifQ to the NifEN/NifH complex. Notably, when NifQ was incubated with NifEN or NifH separately, Mo transfer was not observed. However, co-incubation of all three protein components led to Mo transfer from NifQ to both NifEN and NifH (Hernandez et al., 2008). These findings indicate that the mobilization of Mo from NifQ requires the simultaneous presence of NifH and NifEN.

1.3.8 NifV delivers R- homocitrate to form FeMo-co

Homocitrate for FeMo-co formation is synthesized by the *nifV* gene product (homocitrate synthase). NifV is a homodimer protein that synthesizes R-homocitrate by the condensation of acetyl coenzyme A (acetyl-CoA) and α-ketoglutarate (Hoover et al., 1989, Zheng et al., 1997). The exact mechanism of how R-homocitrate is delivered to form FeMo-co is still not known. It is hypothesized that homocitrate incorporation occurs at NifEN (previously bound to NifH), presumably after Mo has been incorporated (Rangaraj and Ludden, 2002).

1.3.9 NifEN scaffold as a central player in FeMo-co assembly

The NifEN complex protein is the product of *nifE* and *nifN*. The polypeptides have mutual requirements for their stability. The deletion of either *nifE* or *nifN* disrupts the activity of NifDK but not NifH. However, NifDK could be activated *in vitro* by addition of purified FeMo-co to the reaction, suggesting that NifEN is critical for FeMo-co synthesis (Roberts et al., 1978). The first NifEN purification from *A. vinelandii* showed that the protein is a 200 kDa complex consisting of α₂β₂ heterotetramer structure, similar to the homologous NifDK protein (Paustian et al., 1989). NifD and NifE originating from diverse archaeal and bacterial organisms exhibit around 35% sequence identity and 60% sequence similarity, while NifK and NifN exhibit around 30% identity and 55% similarity (Fani et al., 2000) (Figure 7A-B). Additional studies have shown that NifEN contains two permanent [4Fe-4S] clusters coordinated by Cys³⁷, Cys⁶², and Cys¹²⁴ at NifE, and Cys⁴⁴ at NifN (Figure 7B). These permanent [4Fe-4S] clusters are located at positions equivalent to the P-clusters of the NifDK protein. Additionally, NifEN harbors FeMo-co precursors delivered by NifB, which are converted to VK-clusters (described

below) and ligated by Cys²⁵ from the NifE subunits (Figure 7C) (Kaiser et al., 2011). The NifEN and NifDK proteins share similarities at many levels. In addition to the amino acid sequence similarity and location of metalloclusters within the proteins, NifEN has demonstrated the ability to reduce acetylene (C₂H₂) and azide (N₃⁻), however at significantly low levels when compared to NifDK (Hu et al., 2009).

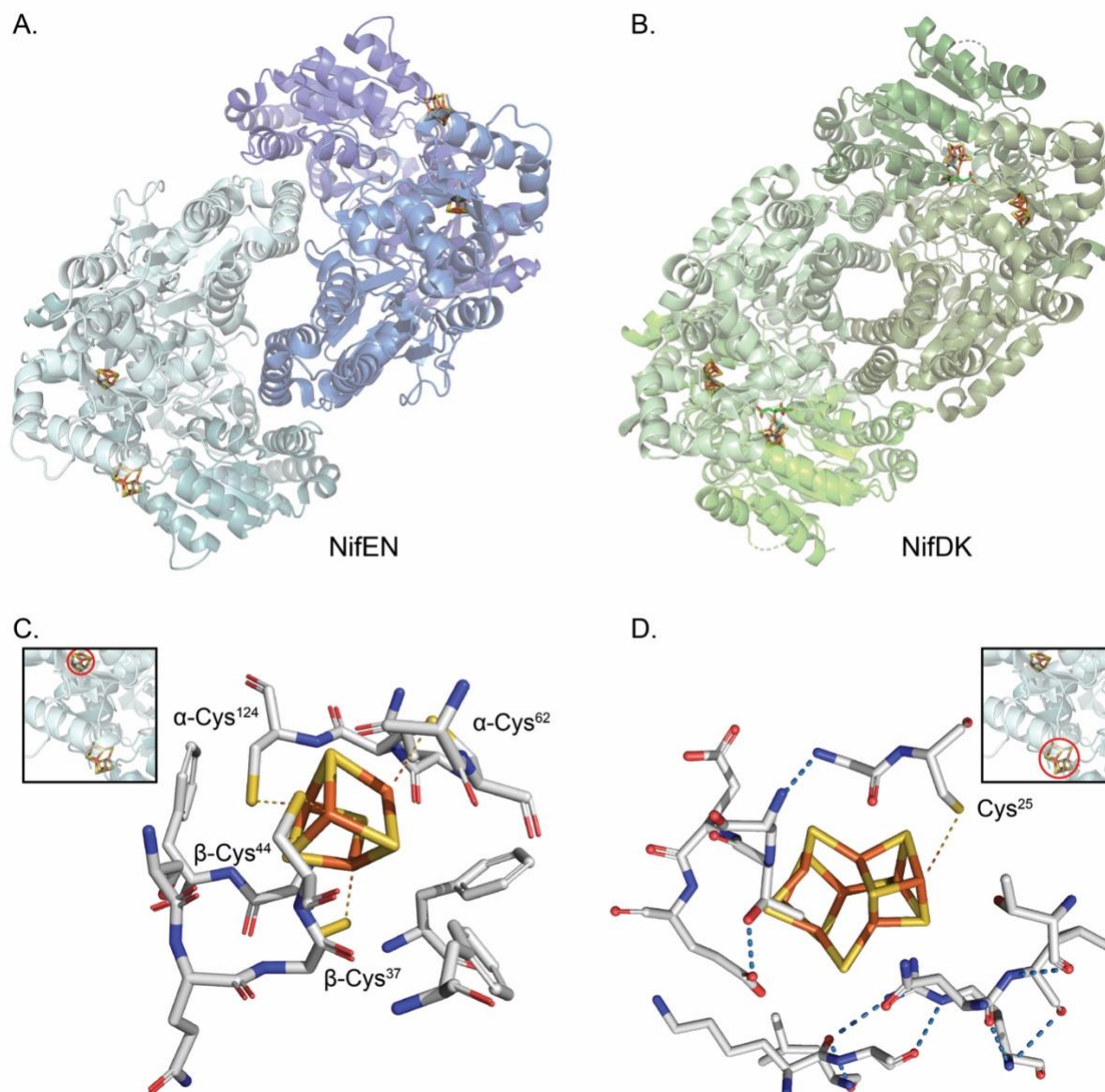


Figure 7. Comparison of NifEN and NifDK protein structures. The location of the metalloclusters within the NifEN (A) and NifDK (B). Ligating Cys residues for [4Fe-4S] (C) cluster and VK-cluster (D). Adapted from (Buren et al., 2020).

NifEN has a scaffold function that assembles and maturates FeMo-co before it is transferred to the apo-NifDK (Brigle et al., 1987). The first direct evidence for NifEN acting as an acceptor of NifB-co came from a study in which NifEN, incubated with ⁵⁵Fe-labeled NifB-co in an incomplete FeMo-co synthesis mixture, was demonstrated to accumulate ⁵⁵Fe.

Importantly, this mixture lacked essential components like MgATP, Mo, homocitrate, or the NifH protein, which are necessary for a complete FeMo-co synthesis (Allen et al., 1995).

The FeMo-co precursor that is delivered to NifEN was initially identified as NifB-co. However, later studies have shown that NifEN converts NifB-co into the next FeMo-co biosynthetic intermediate called VK-cluster. The studies performed by Hernandez and colleagues have shown that NifB-co and VK-cluster share the same atomic composition and that they contain neither Mo nor homocitrate (Hernandez et al., 2007). However, the VK-cluster exhibits EPR signals in both the dithionite-reduced and the thionine-oxidized states and it is therefore different from the EPR-silent NifB-co. This suggested that NifB-co and the VK-cluster are different entities or at least metalloclusters with distinct electromagnetic signatures. Unfortunately, X-ray crystallography of the NifEN structure has not allowed for the determination of the VK-cluster structure at sufficient resolution, nor the involvement of additional amino acid residues in its ligation.

As the VK-cluster does not contain Mo, it was hypothesized that NifEN could have a Mo binding site. This speculation arose from the observation that NifEN, when purified from the *A. vinelandii* strain DJ1041 ($\Delta nifHDKTYnafAB$), contained Mo at a level sufficient to enable *in vitro* FeMo-co formation without further addition of Mo. The analysis of the Mo environment in the purified protein suggested that NifEN could bind a transient [Mo-3Fe-4S] cluster acting as a Mo-donor (Soboh et al., 2006). The homocitrate content of the purified NifEN remains undetermined, although it has been suggested that NifEN could bind and supply *R*-homocitrate as *in vitro* reaction mixtures devoid of Mo and homocitrate have been shown to exhibit marginal apo-NifDK protein activation.

1.3.10 NafY: a FeMo-co carrier protein

In *A. vinelandii*, mature FeMo-co is delivered to apo-NifDK by NafY, a protein encoded by the *nafY* gene located in the minor *nif* cluster and regulated independently of the *nif* genes. NafY is a small monomeric protein (ca. 26 kDa) with two key domains: an N-terminal domain responsible for apo-NifDK binding, and a C-terminal core domain that has been shown to bind mature FeMo-co formed on NifEN (Dyer et al., 2003). Further work has demonstrated that NafY bound to apo-NifDK can protect Cys²⁷⁵ at NifDK. This amino acid is, together with His⁴⁴², crucial for the ligation of FeMo-co to NifDK (Jiménez-Vicente et al., 2018). These findings show the dual role of NafY, both as a FeMo-co carrier protein, facilitating insertion of the cofactor into apo-NifDK, but also as a protector of the FeMo-co ligand α -Cys²⁷⁵ from oxidation.

1.4 Engineering nitrogenase into eukaryotic cells

The strategy of introducing prokaryotic *nif* genes into plants, enabling the plant cell to synthesize its own N₂-fixing machinery independently of bacterial interactions, holds immense potential to revolutionize agriculture. Plants capable of fixing their own nitrogen would enhance sustainability and resilience to fertilizer shortages, with the potential to improve food security and reduce the environmental footprint of agriculture. However, the approach of introducing prokaryotic *nif* genes into plants faces several challenges, including:

- The complex and delicate nature of the nitrogenase components
- The sensitivity of nitrogenase and its supporting accessory proteins and metal clusters to O₂
- The difficulty of coupling the plant's metabolism to the energy and reducing power needed for nitrogen fixation

Despite the challenges presented, significant progress has been made in recent years. Researchers have identified the minimal set of *nif* genes essential for nitrogenase activity and have developed strategies to tackle the obstacles mentioned above. For instance, heterologous promoters can be employed to regulate the level and timing of *nif* gene expression, while plant regulators can be manipulated to optimize nitrogenase function. Additionally, in addressing the demanding energy requirements of nitrogenase, one promising approach involves relocating nitrogenase synthesis to mitochondria and chloroplasts, the energy powerhouses of plant cells. These compartments provide an environment metabolically suitable for hosting nitrogenase, as they are abundant in energy (ATP) and reducing power (NADH). Mitochondria is additionally a site with high O₂ consumption. Moreover, as nitrogenase maturation can be compartmentalized into independent modules, *nif* genes required for different blocks can be engineered and studied individually, without the need to assemble the complete nitrogenase complex directly in plants. The distinct modules can be engineered simultaneously, and the compatibility between them can be assessed once each individual module demonstrates functional competence.

The laborious process of transforming plants and plant cells demands alternative approaches for rapid and efficient investigations. *Saccharomyces cerevisiae* (baker's yeast), a rapidly growing and easily transformed eukaryotic model organism, offers a promising platform for initial assessments of eukaryotic Nif protein functionality. The presence of [Fe-S] cluster assembly machinery in mitochondria, highly similar to that found in diazotrophic organisms, suggests its potential as a good candidate for the transfer of *nif* genes to the plant cell (Allen et al., 2017; Liu et al., 2018). The combination of [Fe-S] cluster machinery together

with high levels of energy and reducing power, and presumably a low-O₂ environment within mitochondria, could help in hosting heterologous O₂-labile nitrogenase proteins and metalloclusters. No easy and well-established methods have been described for transforming the plant mitochondria genome. For this reason, nitrogenase components that are tested in yeast should be expressed in the cytosol, either from *nif* genes integrated into the yeast nuclear genome or from expression plasmids, and then targeted to the mitochondria using targeting peptides.

The import of Nif protein into mitochondria is a multi-step process involving the use of mitochondrial targeting peptides (MTPs) derived from naturally occurring mitochondrial proteins, that are added to the N-terminus of the expressed Nif protein. These MTPs will guide the Nif proteins toward the mitochondrial import machinery. The transport process begins with cytosolic chaperones, proteins that bind to the MTP-labeled Nif proteins and escort them to the mitochondrial outer membrane. At this point, the proteins encounter the translocase of the outer membrane (TOM), a complex that facilitates Nif protein entry into the mitochondrial intermembrane space. Next, the proteins must cross the mitochondrial inner membrane and this process requires the assistance of the translocase of the inner membrane (TIM) complex. This sophisticated machinery recognizes the MTPs and guides the preproteins into the mitochondrial matrix, the innermost compartment of the organelle. Within the mitochondrial matrix, proteolytic cleavage of the MTPs by the mitochondrial processing peptidase (MPP) removes the MTPs, leaving behind the mature Nif protein, and often with a residual amino acid sequence at the N-terminus of the protein. These N-terminal modifications could impact the function of the Nif proteins, causing improper protein folding and solubility issues.

Several Nif proteins exhibit poor solubility when targeted to yeast mitochondria. This solubility issue is likely due to a combination of factors, including pH, metabolite concentration, incompatibility of chaperones with Nif proteins, protein-membrane interactions, and improper folding. Several Nif proteins originating from model diazotrophic organisms like *K. oxytoca* or *A. vinelandii* have been found to be insoluble or poorly soluble when targeted to yeast mitochondria (Buren et al., 2017, López-Torrejón et al., 2016). To address these solubility issues, researchers are exploring different approaches, including the construction of polyproteins by fusing multiple proteins into single polypeptide chains or exploiting biodiversity by using other bacterial or archaeal variants of the nitrogen fixation pathway.

Previous studies have shown that Nif proteins from less characterized diazotrophs can be more soluble in eukaryotic cells and therefore better suited for the plant engineering (Jiang et al., 2021, Paya-Tormo et al., 2022, Burén et al., 2019). For example, NifH from the thermophilic archaea *Hydrogenobacter thermophilus* supported P-cluster maturation, FeMo-co synthesis, and ATP-dependent electron transfer to NifDK *in vitro* when isolated from yeast mitochondria. Similarly, NifB proteins from two other archaea (*Methanocaldococcus infernus*

and *Methanothermobacter thermautotrophicus*) were shown to support FeMo-co synthesis *in vitro*.

1.5 Experimental design to engineer nitrogenase cofactor biosynthesis in the mitochondria of *S. cerevisiae*

After NifH and NifB, NifEN is the next protein in the FeMo-biosynthetic pathway that must be engineered. Previous works had already shown that NifEN from the model organisms *A. vinelandii* and *K. oxytoca* were challenging to express in soluble forms in mitochondria of *S. cerevisiae* and *N. benthamiana* (Buren et al., 2017, Okada et al., 2020b). The encouraging outcomes of the NifH and NifB library screening studies were the basis for the design and experimental approach applied in this thesis (Figure 8). Therefore, in this thesis work, a library was designed comprising homologous gene variants originating from diverse organisms that could increase the prospects of identifying NifEN variants that were favorably accepted by the eukaryotic cell and efficiently transported to the yeast mitochondria. Previous studies have already shown that *A. vinelandii* NifU and NifS are active when produced in *S. cerevisiae*. Therefore, and because NifEN requires [4Fe-4S] clusters for its functionality, the library screening was conducted in *S. cerevisiae* cells simultaneously expressing *A. vinelandii* NifU and NifS. Notably, the initial NifEN library screening was in parallel also conducted in *A. vinelandii*. The rationale behind this strategy was to ensure compatibility between selected library NifEN variants and the *A. vinelandii* nitrogenase machinery, prior to more time-consuming purification experiments and complex biochemical analyses.

NifEN functionality not only requires that the NifE and NifN polypeptides are soluble, but also that they form a complex. Therefore, to enable the study of NifEN complex formation, two distinct epitope tags were attached to the NifE and NifN subunits, respectively. Using pulldown experiments, the co-purification of the respective polypeptide partner, indicative of complex formation, could be performed.

Once functional (in *A. vinelandii*), soluble, and complex forming (in *S. cerevisiae* mitochondria) NifEN candidates were identified, these candidates were expressed and isolated from *E. coli*. *E. coli* is a well-characterized and efficient host for heterologous protein expression and offers a robust system for abundant production of proteins with less effort. This approach allowed for performing extensive compatibility assays to identify the optimal NifH variant for each NifEN, something that would have been challenging to conduct using NifEN isolated from yeast mitochondria (as the yield of NifEN from yeast is much lower). Finally, the selected NifEN proteins were isolated from yeast mitochondria, combined with other yeast-produced Nif components, and used to assay *in vitro* FeMo-co biosynthesis.

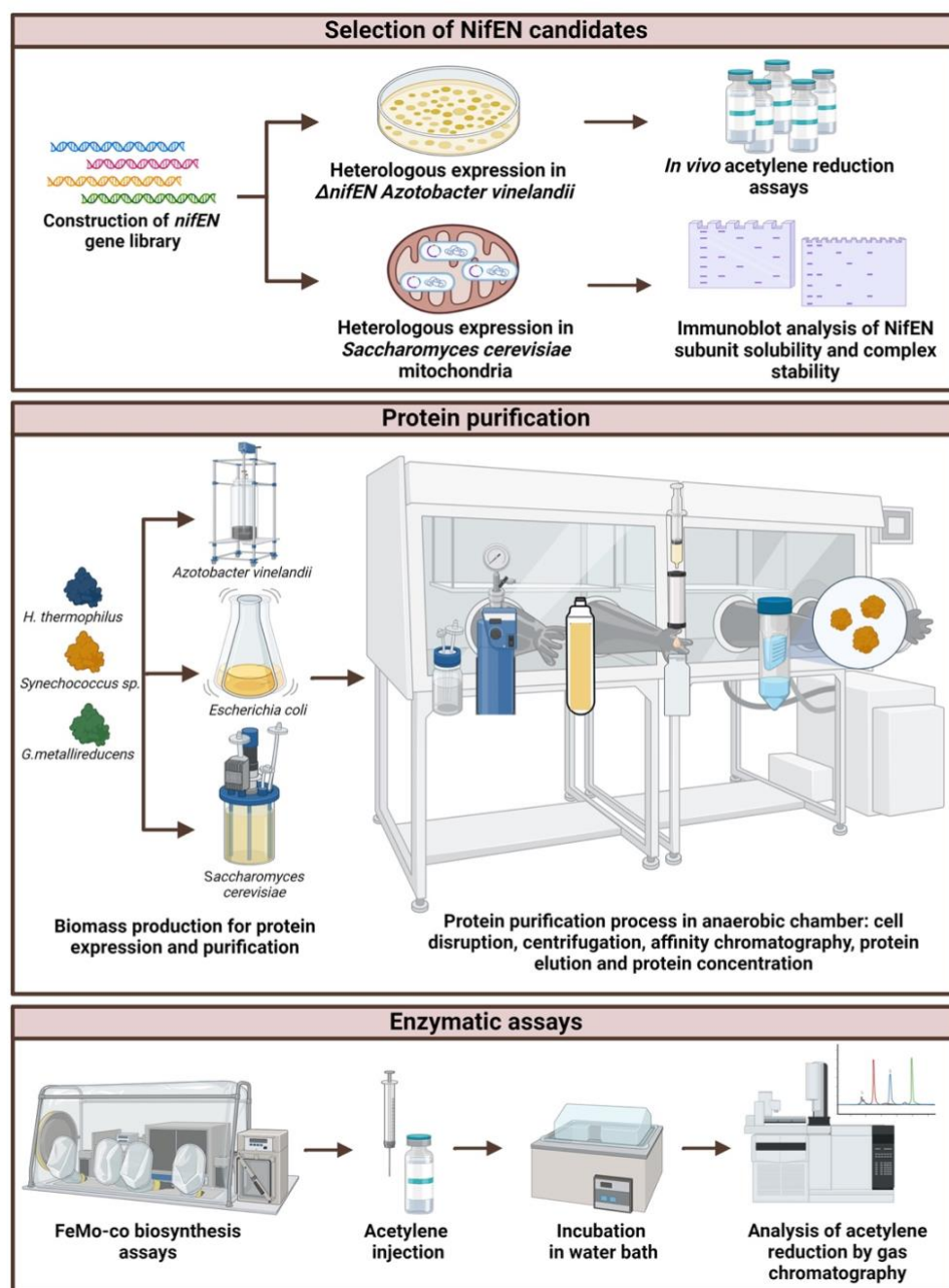


Figure 8. Experimental design followed in the presented thesis work to identify NifEN variants capable of forming soluble and functional protein complexes in *S. cerevisiae* mitochondria.

The approach used in this study demonstrated that evaluating the functionality of individual assembly proteins, such as NifEN, in a yeast is a promising strategy. By expressing individual modules separately, we can better control their expression and ensure that they are fully functional before attempting to assemble a complete nitrogenase pathway.

Directing multiple genes into the mitochondria of yeast for simultaneous expression can be challenging and inefficient, as the expression of one gene may affect the expression

and functionality of others. Therefore, the modular approach enables a more concentrated examination of soluble proteins and the systematic identification of functional components, subsequently facilitating the assembly of a functional complex pathway. The ability to independently evaluate module functionality also provides a significant advantage when applying this approach to plant cells, whose transformation process is notably more intricate. By expressing individual modules separately, we can potentially circumvent these challenges and simplify the process of engineering nitrogen fixation into plant cells.

2. OBJECTIVES

The assembly of the nitrogenase FeMo-cofactor stands as a pivotal process in ensuring the functionality of nitrogenase, as it provides the catalytic site for N₂ reduction in the Mo-dependent nitrogenase. This complex biosynthetic pathway requires the cooperative interactions of various nitrogen fixation (*nif*) genes. At the center of this process is the scaffold protein NifEN, which accepts the precursor cluster NifB-co produced by NifB. NifB-co is converted to FeMo-co by the incorporation of Mo and homocitrate, in a reaction that involves NifH.

Prior work established that NifB-co can be produced in *S. cerevisiae* mitochondria. The subsequent step in FeMo-co formation involves NifEN. However, NifEN from *A. vinelandii* resulted in the generation of insoluble protein when targeted to yeast mitochondria, underscoring the need to identify better NifEN variants that are capable of forming soluble and functional protein complexes in eukaryotic cells.

The specific objectives of the presented thesis are:

1. To generate a library of *nifE* and *nifN* genes by systematically exploring diverse diazotrophic organisms across the bacterial and archaeal domains using relevant databases.
2. To identify NifEN variants that form soluble and stable protein complexes within the mitochondria of *S. cerevisiae*.
3. To assess the *in vivo* functionality and compatibility of NifEN protein candidates with other Nif components of *A. vinelandii*.
4. To examine the ability of selected NifEN variants, when isolated from *A. vinelandii*, to support the formation of FeMo-co *in vitro*.
5. To identify a NifH protein that is compatible with each selected NifEN variant, enabling efficient FeMo-co assembly and activation of apo-NifDK protein.
6. To purify the selected NifEN variants from aerobically cultured *S. cerevisiae* to perform *in vitro* FeMo-co synthesis together with yeast-isolated NifH and NifB proteins.

3. MATERIAL AND METHODS

3.1 Chemical and Reagents

All chemicals were purchased from Bio-Rad, Difco, IBA Lifesciences, Invitrogen, Merck, New England Biolabs, Promega, Sigma-Aldrich, ThermoFisher Scientific, GeneScript, VWR. All reagents were of appropriate purity and were used following the manufacturer's guidelines.

3.2 Organisms

3.2.1 Strains of *E. coli*

E. coli strains used in this work are summarized in Table 2. DH5 α and TOP10 strains were used to construct plasmids for the expression of *nif* genes in *E. coli*, *S. cerevisiae*, and *A. vinelandii*. *E. coli* cells were cultured at 37°C overnight in an oven when plated or in a shaking incubator at 150 rpm when grown in liquid cultures. Strain BL21 (DE3) was used for heterologous overexpression of Nif proteins in *E. coli* and fresh transformations were performed each time. The generated strains were preserved in 40% glycerol solution (v/v) in cryovials and stored at -80°C.

Table 2. E. coli-type strains used in this thesis.

Strain	Genetic background	Source
DH5 α	<i>F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR phoA</i> <i>Φ80dlacZΔM15 Δ(lacZYA-argF) U169 hsdR17(rK- mK+) λ-</i>	(Sambrook and Russell, 2001)
TOP10	<i>F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15</i> <i>Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL</i> (StrR) <i>endA1 nupG</i>	ThermoFisher Scientific
BL21 (DE3)	<i>F- ompT gal dcm lon hsdS_B(r_B- m_B-) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5] [malB⁺]_{K-12}(λ^S))</i>	Promega

3.2.2 Strains of *A. vinelandii*

A. vinelandii DJ (wild-type strain) was obtained from the laboratory of Professor Dennis Dean (Virginia Tech, USA) and used for both transformation and Nif protein overexpression. A list of *A. vinelandii* strains used for this thesis work is shown in Table 3. When *A. vinelandii* was grown on a solid medium, cells were incubated at 30°C. For liquid medium, a maximum of 20% of the total flask volume was filled with Burk medium,

and cells were incubated with shaking at 200 rpm at 28°C using baffled flasks. The *A. vinelandii* strains were preserved in 9% DMSO (v/v) in cryovials and stored at -80°C.

Table 3. *A. vinelandii* strains used and generated in this thesis.

Strain	Genetic background	Source	Purpose
DJ	Wild type	Dennis Dean	Positive control for diazotrophic growth. Nif protein source
DJ33	$\Delta nifDK$ (<i>nifD376-nifK524</i>)	Dennis Dean	NifH isolation
DJ1118	$\Delta nifB::KanR$ <i>PnifH::polyHis::nifEN</i>	Dennis Dean	Apo-NifEN isolation (devoid of FeMo-co precursor)
DJ1143	$\Delta nifB$ <i>polyHis::NifDK</i>	Dennis Dean	Apo-NifDK isolation (devoid of FeMo-co)
DJ2898	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$, <i>StrR</i> in <i>scr</i> region	This work	<i>In vivo</i> complementation experiments
DJ2917	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Azotobacter vinelandii</i>	This work	<i>In vivo</i> complementation experiments
DJ2918	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Fisherella thermalis</i>	This work	<i>In vivo</i> complementation experiments
DJ2919	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Synechococcus</i> sp.	This work	<i>In vivo</i> complementation experiments
DJ2920	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Hydrogenobacter thermophilus</i>	This work	<i>In vivo</i> complementation experiments
DJ2921	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Nostoc</i>	This work	<i>In vivo</i> complementation experiments
DJ2922	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Paenibacillus sabinae</i>	This work	<i>In vivo</i> complementation experiments
DJ2923	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Arcobacter nitrofigilis</i>	This work	<i>In vivo</i> complementation experiments
DJ2924	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Rhodobacter capsulatus</i>	This work	<i>In vivo</i> complementation experiments
DJ2825	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Geobacter metallireducens</i>	This work	<i>In vivo</i> complementation experiments
DJ2926	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifENX</i> <i>Azotobacter vinelandii</i>	This work	<i>In vivo</i> complementation experiments
DJ2927	$\Delta vnfDGK$, $\Delta anfHDGK::GentR$, $\Delta nifEN::KanR$ <i>Para::Strep nifEN</i> <i>Methanothermobacter marburgensi</i>	This work	<i>In vivo</i> complementation experiments

Strain	Genetic background	Source	Purpose
DJ2928	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Klebsiella oxytoca</i>	This work	<i>In vivo</i> complementation experiments
DJ2929	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Metanosarcina acetivorans</i>	This work	<i>In vivo</i> complementation experiments
DJ2930	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Methanococcus maripaludis</i>	This work	<i>In vivo</i> complementation experiments
DJ2931	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Methanothermobacter thermoautotrophicus</i>	This work	<i>In vivo</i> complementation experiments
DJ2932	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Chlorobaculum tepidum</i>	This work	<i>In vivo</i> complementation experiments
DJ2933	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Methanosarcina barkeri</i> CM1	This work	<i>In vivo</i> complementation experiments
DJ2934	$\Delta vnfDGK$, $\Delta anfHDGK:GentR$, $\Delta nifEN:KanR$ Para: <i>Strep nifEN</i> <i>Anabaena variabilis</i>	This work	<i>In vivo</i> complementation experiments
DJ2957	$\Delta nifHDKTYnafABnifEN$ <i>PnifH::Strep nifEN</i> <i>Synechococcus</i> sp. JA-2-3B'a(2-13)	This work	NifEN isolation
DJ2963	$\Delta nifHDKTYnafABnifEN$ <i>PnifH::Strep nifEN</i> <i>Hydrogenobacter thermophilus</i>	This work	NifEN isolation
DJ2965	$\Delta nifHDKTYnafABnifEN$ <i>PnifH::Strep nifEN</i> <i>Geobacter metallireducens</i>	This work	NifEN isolation

3.2.3 Strains of *S. cerevisiae*

S. cerevisiae strain W303-1a (Table 4) was used as a parental strain for Nif protein expression and isolation.

Table 4. *S. cerevisiae* type-strain used in this thesis.

Strain	Genetic background	Source	Purpose
W303-1a	leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15 [phi+]	EURO-SCARF (ATCC 208352)	Overexpression of Nif proteins

The generated *S. cerevisiae* strains used in this work are listed in Table 5. Growth of *S. cerevisiae* in the solid medium was performed in a growth chamber at 28°C and when grown in a liquid medium, flasks were incubated with shaking at 150 rpm at 28°C inside a shaking incubator. The *S. cerevisiae* strains were preserved in cryovials with 40% glycerol solution (v/v) and stored at -80°C.

Table 5. *S. cerevisiae* W303-1a strains generated in this thesis.

Strain	Plasmids	Expressed proteins	Promoter	Full-length (kDa)	Processed (kDa)
SB320Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
KD38Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD104	Su9-TS-NifEN ^{<i>Anabaena variabilis</i>}	GAL10	109.3	102
KD49Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD79	SU9-HA-NifE ^{<i>Arcobacter nitrofigilis</i>}	GAL10	57.7	50.4
	pN2KD130	SU9-NifN ^{<i>Arcobacter nitrofigilis</i>} -SS	GAL10	56	48.7
KD39Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD79	SU9-HA-NifE ^{<i>Azotobacter vinelandii</i>}	GAL10	60.8	53.5
	pN2KD130	SU9-NifN ^{<i>Azotobacter vinelandii</i>} -SS	GAL10	58.3	51
KD51Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD81	SU9-HA-NifE ^{<i>Chlorobaculum tepidum</i>}	GAL10	59.1	51.8
	pN2KD132	SU9-NifN ^{<i>Chlorobaculum tepidum</i>} - SS	GAL10	59.1	51.8
KD41Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD69	SU9-HA-NifE ^{<i>Fischerella thermalis</i>}	GAL10	62.2	54.9
	pN2KD122	SU9-NifN ^{<i>Fischerella thermalis</i>} -SS	GAL10	56.9	49.6
KD37Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD103	Su9-TS-NifEN ^{<i>Geobacter metallireducens</i>}	GAL10	110.5	103.2
KD45Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD69	SU9-HA-NifE ^{<i>Hydrogenobacter thermophilus</i>}	GAL10	58.9	51.6
	pN2KD126	SU9-NifN ^{<i>Hydrogenobacter thermophilus</i>} -SS	GAL10	61	53.7
KD42Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD70	SU9-HA-NifE ^{<i>Klebsiella oxytoca</i>}	GAL10	59	51.7
	pN2KD123	SU9-NifN ^{<i>Klebsiella oxytoca</i>} -SS	GAL10	59.6	52.3
KD48Y	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD78	SU9-HA-NifE ^{<i>Methanococcus maripaludis</i>}	GAL10	62	54.7

Strain	Plasmids	Expressed proteins	Promoter	Full-length (kDa)	Processed (kDa)
KD43Y	pN2KD129	SU9-NifN ^{<i>Methanococcus maripaludis</i>} _SS	GAL10	59.2	51.9
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD71	SU9-HA-NifE ^{<i>Methanosarcina acetivorans</i>}	GAL10	65.7	58.4
KD52Y	pN2KD124	SU9-NifN ^{<i>Methanosarcina acetivorans</i>} _SS	GAL10	66.5	59.2
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD82	SU9-HA-NifE ^{<i>Methanosarcina barkeri</i>}	GAL10	66.1	58.8
KD40Y	pN2KD133	SU9-NifN ^{<i>Methanosarcina barkeri</i>} _SS	GAL10	61.5	54.2
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD68	SU9-HA-NifE ^{<i>Methanothermobacter marburgensis</i>}	GAL10	61.7	54.4
KD50Y	pN2KD121	SU9-NifN ^{<i>Methanothermobacter marburgensis</i>} _SS	GAL10	57.9	50.6
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD80	SU9-HA-NifE ^{<i>Methanothermobacter thermotrophicus</i>}	GAL10	61.8	54.5
KD46Y	pN2KD131	SU9-NifN ^{<i>Methanothermobacter thermotrophicus</i>} _SS	GAL10	58.1	50.8
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD76	SU9-HA-NifE ^{<i>Nostoc sp.</i>}	GAL10	61.7	54.4
KD47Y	pN2KD127	SU9-NifN ^{<i>Nostoc sp.</i>} _SS	GAL10	57.3	50
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD77	SU9-HA-NifE ^{<i>Paenibacillus sabinae</i>}	GAL10	58.6	51.3
KD53Y	pN2KD128	SU9-NifN ^{<i>Paenibacillus sabinae</i>} _SS	GAL10	57.3	50
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD83	SU9-HA-NifE ^{<i>Rhodobacter capsulatus</i>}	GAL10	64.8	57.5
KD44Y	pN2KD134	SU9-NifN ^{<i>Rhodobacter capsulatus</i>} _SS	GAL10	58.4	51.1
	pN2GLT4	Su9-NifU ^{<i>Azotobacter vinelandii</i>}	GAL1	40.8	33.6
		Su9-NifS ^{<i>Azotobacter vinelandii</i>}	GAL10	51.3	44
	pN2KD72	SU9-HA-NifE ^{<i>Synechococcus sp.</i>}	GAL10	59.6	52.3
	pN2KD125	SU9-NifN ^{<i>Synechococcus sp.</i>} _SS	GAL10	65.8	58.5

3.3 Microbial culture media

3.3.1 *E. coli* lysogenic broth

Liquid Lysogenic Broth (LB) was used to culture *E. coli* for molecular biology and protein isolation purposes (Bertani, 1951). LB contained 10 g tryptone, 5 g yeast extract, and 10 g NaCl per liter of distilled water. Solid LB for the preparation of Petri dishes additionally required 15 g of agar per liter. All media were autoclaved for 20 min at 121°C.

When necessary, culture media was supplemented with the following antibiotics (final concentrations): carbenicillin (100 µg/ml), ampicillin (100 µg/ml), and kanamycin (50 µg/ml).

3.3.2 *A. vinelandii* Burk's modified medium

For diazotrophic growth *A. vinelandii* cells were cultured in liquid or solid media of modified Burk's medium. The medium was autoclaved and stored in two separate solutions: phosphate buffer (0.2 g KH₂PO₄, 0.8 g K₂HPO₄, and distilled water up to 900 ml) and concentrated 10x salts prepared by resuspending 200 g of sucrose, 2 g MgSO₄·7H₂O, 0.9 g CaCl₂·2H₂O, 1 ml of 10mM Na₂MoO₄·2H₂O, 50 mg FeSO₄·7H₂O in one liter of distilled water. To prepare 1000 ml of Burk media 900 ml of phosphate buffer and 100 ml of 10x salt solution were mixed. The media were supplemented with 3 g/l arabinose to grow *A. vinelandii* cells previously transformed with arabinose inducible expression plasmid. The fixed-nitrogen component (13 mM NH₄AOc) was added to the media when *A. vinelandii* was cultured under non-diazotrophic conditions. For Petri dishes, 16 g/l of agar was added to Burk's media. All media components were autoclaved for 20 min at 121°C.

When necessary for selection purposes, kanamycin (0.5 µg/ml) was added.

3.3.3 *S. cerevisiae* medium

The *S. cerevisiae* strains generated in this study using W303-1a as parental strain were cultured in a synthetic defined (SD) medium (1.9 g/l yeast nitrogen base, 5 g/l ammonium sulfate, 20 g/l glucose, and drop-out mixture). The growth was carried out either in solid media (containing 1% agar) on plates or in liquid media in flasks, supplemented with appropriate amino acids (uracil 20 µg/ml, tryptophan 30 µg/ml, adenine 22 µg/ml, leucine 80 µg/ml, histidine 20 µg/ml) depending on autotrophic requirements.

3.4 Transformation methods

3.4.1 Preparation of chemically competent *E. coli* cells and heat-shock-mediated transformation

Transformation of *E. coli* was performed due to standard protocols (Sambrook and Russell, 2001).

The *E. coli* cells were grown on LB solid media overnight at 37°C and fresh biomass was used to inoculate a 5 ml liquid LB media and grown overnight at 37°C with shaking. Next day a 200 ml culture of LB was inoculated at OD₆₀₀ of 0.02-0.05 and incubated at 37°C with shaking at 150 rpm until reaching an OD₆₀₀ of 0.4-0.45. Then, all procedures were performed at 4°C. Growth cells were pelleted by centrifugation for 7 min at 3,000 x *g*, resuspended in 40 ml ice-cold MgCl₂ (0.1 mM), and incubated on ice for 15 min. Then the cells were pelleted again and resuspended in 8 ml of ice-cold CaCl₂ (0.1 M) and kept on ice for 1 h. For cryo-preservation, 3 ml of cold 60% (v/v) glycerol solution was added. Hundred µl aliquots were transferred to sterile Eppendorf tubes, snap-frozen in liquid N₂ (LN₂), and stored at -80°C.

Chemically competent cells were thawed on ice for 5 min and then used for heat-shock transformation. Initially, 50-100 ng of plasmid DNA was added to the cells, gently mixed, and incubated on ice for 20 min. Subsequently, the DNA/cell mixture was immersed in a water bath at 42°C for 1 min and then incubated on ice for 5 min. To facilitate cell recovery, 1 ml of LB medium was added, and the mixture was incubated at 37°C for 1 h with shaking at 200 rpm. Finally, the cells were pelleted by centrifugation at 13,000 x *g* for 2 min and spread onto a Petri dish containing solid LB medium supplemented with the appropriate antibiotic(s) for selection. The plates were then incubated overnight at 37°C.

3.4.2 Preparation of competent *A. vinelandii* cells and transformation methods

Gene transformation and manipulation of *A. vinelandii* were based on the procedures previously described (Page and von Tigerstrom, 1979, Dos Santos and Dean, 2011).

In short, competent cells of *A. vinelandii* were streaked out onto Burk's medium with a fixed nitrogen source (no Mo added) and incubated at 30°C for 2-3 days. A loop of cells was then resuspended in 50 ml Burk's medium (-Mo, -Fe), vortexed vigorously, and grown overnight with shaking (170 rpm) at 30°C. The cultures were incubated until exhibited a bright greenish color, indicating the excretion of the iron siderophore azotobactin, that facilitates DNA acquisition by *A. vinelandii*.

Transformation of *A. vinelandii* was performed by mixing 200 µl of fresh cell suspension with 200 µl of transformation buffer (20 mM MOPS pH 7.4 and 15 mM MgSO₄) and 5 µl of plasmid DNA. The mixture was kept for 20 minutes at RT and transferred to sterile test tubes to make 3-4 serial dilutions with 1x PO₄ solutions. Selection for transformed cells was performed by plating cells on solid Burk's medium with 3 g/l arabinose.

When needed, the transformation of *A. vinelandii* was carried out by adding 440 µl of the fresh cell suspension to 3 ml of Burk's media, followed by the addition of 20-40 µl of plasmid DNA. The mixture was then transferred to 50 ml flasks and incubated overnight at 30°C with shaking at 170 rpm. The overnight culture was then transferred to sterile test tubes for serial dilutions using 1x PO₄ solutions. The selected cells were finally plated onto solid Burk's medium containing 0.5 µg/mL of kanamycin to isolate transformed *A. vinelandii* cells.

3.4.3 Transformation of *S. cerevisiae*

S. cerevisiae cells were transformed according to the protocol described by (Gietz and Schiestl, 2007) with slight modifications.

SD medium supplemented with appropriate amino acid(s) was inoculated with *S. cerevisiae* cells and grown overnight with shaking (150 rpm) at 30°C. For each transformation, 1 ml culture was transferred to an Eppendorf tube and pelleted by centrifugation at 13.000 x g for 30 s. The cells were resuspended in 1 ml of ice-cold LiAc (lithium acetate) (100 mM) and incubated for 5 min at 30°C. The cells were then pelleted again, and the mixture of 240 µl PEG (50% v/v), 36 µl LiAc (1 M), 5 µl denatured salmon sperm DNA (10 mg/ml), 1-2 µg of each yeast plasmid and 55 µl of Milli-Q water was added to resuspend the yeast cell pellet. The tubes with the mixture were immediately placed in an incubator at 30°C for 30 min and then transferred to a water bath (42°C) for 3 h. After incubation, the cells were finally pelleted by centrifugation, resuspended in 100 µl of sterile water, and plated on minimal SD solid medium containing the appropriate amino acid composition (uracil (20 µg/ml), tryptophan (30 µg/ml), adenine (22 µg/ml), leucine (80 µg/ml), and histidine (20 µg/ml)) for selection of prototrophic revertants.

3.5 DNA techniques

3.5.1 DNA amplification by Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) DNA was used to perform DNA amplification, molecular cloning, and validation of DNA fragments. The choice of Taq polymerase for the reactions depended on the provider used in the laboratory at each given moment,

taking into account cost, experimental requirements, sequence fidelity requirements, and ease of use. Primers were ordered from ThermoFisher Scientific and Integrated DNA Technologies, the complete list of primers is shown in Table 6.

Table 6. Primers used in this thesis.

Number	Sequence
2910	ATACTTTAACGTCAAGGAGAAAAACCCCCCATGGGTAAT
2911	CCCGGGCCCTATAGTGAGTCGTATTACCCATGGGGGGTTT
2937	TCAACCCTCACTAAAGGGCGGCCGCATGGCCTCCACTCGT
2938	GAATTGTTAATTAAGAGCTCTTATCCGGATTTTTTGGATC
2943	CCTCACTAAAGGGCGGCCGCATGGCCTCCACTCGTGTCT
2944	AAGAATTGTTAATTAAGAGCTCTTATCCGGATTTTTTGGATC CTTTTCAAATTGTGGAT
3211	GGTTTCTTTACCAGACTCGAGTCATTTCTCAAAGTGAAGGA
3212	TATAAGAAGGAGATATACATATGGCTCAATTGATTCAAGA
3213	TATAAGAAGGAGATATACATATGGCTGTCTTGAGAAAGTC
3214	GAAGGAGATATACCATGGTATCAGCATGGAGTCATCCTCA
3215	GGTTTCTTTACCAGACTCGAGTCATCCGGAGTTGTGAGCC
3230	GAGCTCGAATTCGGATCCTCATCCGGAACCAACTTCTTCT
3231	GAAGGAGATATACCATGGTATATCCATATGATGTTCCAGA
3583	GAGCTCGAATTCGGATCCTCATCCGGAGTATTGCCATGGT

3.5.1.1 High-fidelity DNA amplification for cloning purposes

The enzyme used to perform high-fidelity DNA amplification for molecular cloning was Phusion Hot Start II High-Fidelity DNA Polymerase (ThermoFisher Scientific) which has been engineered to excel at amplifying large DNA fragments, having proofreading activity and the ability to generate blunt ends on the amplified DNA product. The concentration of each component used for a typical PCR reaction is shown in Table 7, while in Table 8 are shown the standard PCR conditions used in this thesis. The amounts of DNA templates were 50 ng (for plasmid DNA) or 100 ng (for genomic DNA).

The annealing temperatures for each primer pair were calculated using an online tool (available at the ThermoFisher Scientific [website software](#)). The calculated annealing temperature was set at 72°C (or higher), while a 2-step PCR protocol was carried out following manufacturer recommendations. DNA fragments amplified by PCR

for cloning purposes were purified from agarose gels using QIAquick Gel Extraction Kit (Qiagen) as detailed in section 3.5.6.

Table 7. PCR component concentrations.

PCR component	Final concentration
5 x Phusion buffer HF or GF	1X
dNTPs	200 µM each
Forward primer	0.5 µM
Reverse primer	0.5 µM
DMSO	3%
DNA template	50-100 ng
Phusion Hot Start Polymerase	0.02 U/µl

Table 8. Standard PCR conditions.

Cycle Step	2-step protocol		3-step protocol		N° cycles
	Temperature	Time	Temperature	Time	
Initial denaturation	98 °C	30 s	98 °C	30 s	1
Denaturation	98 °C	10 s	98 °C	10 s	34
Annealing	-	-	X °C	20 s	
Extension	72 °C	30 s/kb	72 °C	30 s/kb	
Final extension	72 °C	8 min	72 °C	8 min	1
	4 °C	Hold	4 °C	Hold	

3.5.1.2 Validation of DNA fragments by PCR

Identification of positive clones by colony PCR was tested using the KAPA2G Fast HotStart ReadyMix PCR Kit according to the guidelines provided by the manufacturer Sigma-Aldrich. The concentration of components and the conditions used for standard PCR are described in Table 9 and Table 10. The annealing temperatures of employed primers were calculated using the online tool available at the ThermoFisher Scientific [website](#). The size of the amplified DNA fragments was visualized by standard agarose gel electrophoresis (section 3.5.5).

Table 9. Colony PCR components.

PCR component	Final concentration
KAPA2G Fast HotStart ReadyMix	1X
Forward primer	0.5 µM
Reverse primer	0.5 µM
DNA template	10-100 ng

Table 10. Colony PCR conditions.

Step	Temperature	Time	N° cycles
Initial denaturation	95 °C	3 min	1
Denaturation	95 °C	10-15 s	30-40
Annealing	X °C	20 s	
Extension	72 °C	15 s/kb	
Final extension	72 °C	1 min/kb	1
	4 °C	Hold	

3.5.1.3 Gene Splicing by Overlap Extension (OE-PCR)

Two DNA sequences were combined due to Gene Splicing by Overlap Extension reaction (OE-PCR) (Horton et al., 1990). This cloning method fuses DNA fragments without relying on restriction endonucleases and it is carried out in two steps. First, DNA fragments are PCR-amplified with primers containing at the 5' ends share homologous sequences. Then, two distinct fragments create a linear insert due to the hybridization of homologous ends and are filled to completion by the DNA polymerase. OE-PCR was performed to generate fused proteins or to add detection/purification tags.

The scheme for the overlapping PCR procedure is presented in Table 11, where for each primer pair combination were used specific PCR conditions. DNA fragments generated in step one (PCR fragments A and B) were used as templates for the final (overlapping) PCR. The PCRs were carried out using Phusion™ High-Fidelity DNA Polymerase (ThermoFisher Scientific) according to manufacturer instructions Table 7 and Table 8. The annealing temperatures of each primer combination were calculated using the ThermoFisher Scientific [online tool](#).

Table 11. Scheme for a general overlapping PCR.

Step	PCR	Primer combination	Template
1	PCR A	Forward primer 1A +	Original gene sequence (plasmid)
		Reverse primer 2A	
1	PCR B	Forward primer 1B +	Original gene sequence (plasmid)
		Reverse primer 2B	
2	Overlapping PCR	Forward primer 1A +	PCR fragment A +
		Reverse primer 2B	PCR fragment B

3.5.2 Traditional cloning techniques

The traditional cloning method involves the restriction enzyme digestion of DNA fragments, followed by the subsequent ligation of these fragments into a targeted plasmid. For this reason, the commercial plasmids contain multicloning sites (MCS) with many restriction sites to enable DNA fragment insertion.

The restriction enzymes enable DNA ligase-mediated hybridization of the insert, and the backbone vector fragments with compatible overhang sequences. Therefore, primer sequences contain restriction enzyme sites that allow the generation of unique and complementary cutter sites in the amplified DNA fragment with those of the target vector. The type II restriction enzymes used in this thesis were purchased from NEB (New England Biolabs). Adhering to the manufacturer's recommendations and employing bioinformatic tools available on the NEB website (nebcloner.neb.com), the appropriate conditions for restriction enzyme digestion were determined. The digested DNA fragments were either extracted from agarose gels or purified using DNA binding columns (refer to sections 3.5.6 and 3.5.7, respectively).

DNA fragments of interest were inserted into the target vector using T4 DNA ligase (Promega). Ligation reactions were carried out following the manufacturer's instructions. Normally, a 3 to 1 ratio of insert to vector was used together with 1x ligation buffer and T4 DNA ligase (0.1-1 units). Reactions were kept overnight at 4°C. The complete ligation volume (usually 10 µl) was used for the transformation of chemically competent *E. coli* cells (section 3.4.1). The insertion of the DNA constructs was verified by colony PCR (section 3.5.1.2) and DNA sequencing. The sequencing data was analyzed using SnapGene software (GSL Biotech LLC).

3.5.3 Exonuclease and Ligation-Independent Cloning (ELIC)

This molecular cloning method originally described by (Koskela and Frey, 2015) is based on homologous recombination between purified and linear double-stranded DNA molecules using complementary ends. In the specific, DNA fragments were amplified with primers equipped with homologous overhangs to the ends of the enzymatically digested target vectors to enable directional cloning. In the reaction, vector and insert (ratio at 1:4) were mixed with water up to 10 µl of final volume and then the mixture was used to transform chemically competent *E. coli* cells.

3.5.4 Plasmid DNA extraction

Extraction of plasmid DNA from *E. coli* strains was performed by using the QIAGEN Plasmid Mini Kit (Qiagen) following the manufacturer's instructions.

For this purpose, it was used 10 ml of LB overnight-grown *E. coli* culture supplemented with antibiotics for the plasmid selection. The next day, the grown cells were pelleted for 5 min at 5,000 x *g* and resuspended in 250 µl buffer P1. The cells were lysed with P2 buffer (250 µl) followed by incubation for 5 min at room temperature. Additionally, to neutralize the lysis solution, 350 µl of N3 reagent was added. Then, cell debris and genomic DNA were removed by centrifugation at 13,000 x *g* for 15 min. The resulting supernatant was loaded onto a silica QIAprep column and passed through by centrifugation for 1 min at 13,000 x *g*. The extracted DNA was washed with 750 µl of PE buffer, and centrifuged twice for 1 min at 13,000 x *g*. Finally, onto the column was added 50 µl of Milli-Q water (preheated to 50°C), and after incubation for 2 min and centrifugation for 1 min at 13,000 x *g*, the plasmid DNA was eluted. The DNA concentration and quality of eluted plasmid were analyzed by absorbance ($A_{260/280}$) using a NanoDrop® ND-1000 Spectrophotometer.

3.5.5 DNA gel electrophoresis

DNA fragments after restriction enzyme digestion or from PCR amplifications were mixed with 6x loading dye buffer (30% (v/v) glycerol, 0.25% (w/v) bromophenol blue and 0.25% (w/v) xylene cyanol FF) and separated by loading samples on agarose gel. The percentage of agarose in the gels was adjusted depending on the expected size of the DNA fragment: 1-2% agarose for 100-200 bp fragments and 0.9% agarose for fragments above 200 bp. SYBR Safe (ThermoFisher Scientific) was used to visualize the separated DNA in the agarose gel (added at a 1:100,000 dilution before jellification). Gel electrophoresis was performed in TAE buffer (40 mM Tris-HCl pH 8, 20 mM acetic acid, 1 mM EDTA). The sizes of the DNA fragments were estimated according to DNA bands present in commercial ladders (Nippon Genetics).

3.5.6 Extraction of DNA from agarose gels

QIAEX® II Gel Extraction Kit (QIAGEN) was used for DNA extraction from agarose gel according to manufacturer instructions. The gel slices containing DNA fragments were excised from agarose gel under ultraviolet light using a clean scalpel and then transferred to a 1.5 ml Eppendorf tube. DNA fragments ranging from 100 bp to 4 kb were mixed with 3 volumes of QX1 buffer per 1 volume of gel. For fragments larger than 4 kb, 2 volumes of Milli-Q water and 1 volume of QX1 buffer were added per 1 volume of gel. Then, to each sample was added 15 µl QIAEX II resin and the mixture was incubated with shaking at 800 rpm 50°C for 10 min. The DNA bound to resin was pelleted by centrifugation at 13,000 x *g* for 30 s. After removing the supernatant, the pellet was washed by the addition of 500 µl QX1 buffer and then again centrifuged.

Finally, the resin was washed twice using 500 µl PE buffer and air-dried until it turned white. DNA was eluted by incubating the pellets at 50°C for 10 with 20 µl preheated to 50°C Milli-Q water. The eluted DNA was analyzed by absorbance ($A_{260/280}$) using a NanoDrop ND-1000 Spectrophotometer to estimate DNA concentration and ensure quality.

3.5.7 Purification of PCR-amplified or restriction enzyme-digested DNA from solutions

Apart from gel extraction and purification, DNA in solution, such as PCR amplifications or restriction enzyme digests, was occasionally purified using the QIAquick PCR purification Kit (Qiagen) following the provided user manual. In brief, five volumes of PB buffer were added to one volume of DNA solution. The pH of the mixture was checked using the provided colorimetric pH indicator, which turns orange/violet if the solution's pH is too alkaline. If this occurred, one µl of sodium acetate (3 M, pH 5.0) was added until the indicator turned yellow. The mixture was then transferred to a QIAquick spin column and centrifuged for 30 seconds at 13,000 x g. The flow-through was discarded, and the column was washed with 750 µl of PE buffer. The QIAquick column was centrifuged to remove the wash buffer and again to remove any remaining ethanol. Following a 5-minute incubation, the DNA was eluted with 30 µl of preheated Milli-Q water (50°C). The eluted DNA was analyzed using a NanoDrop ND-1000 Spectrophotometer to determine its concentration and ensure its purity by measuring its absorbance ($A_{260/280}$).

3.6 Plasmid Construction

3.6.1 Plasmids generated for expression of NifEN proteins in *E. coli*

For epitope tag screening, *nifE* and *nifN* genes originating from *A. vinelandii* were inserted behind a single T7 promoter in the pETDuet-1 vector (Novagen). To express genetic constructs of His-NifEN (pN2KD4), His-NifE/Strep-NifN (pN2KD8), and His-NifE/NifN-Strep (pN2KD10), the *nifE* and *nifN* genes were amplified by PCR from *A. vinelandii* DJ genomic DNA and inserted into NdeI (for *nifE*) or BamHI+HindIII (for *nifN*) digested pETDuet-1 vector using ELIC, described in section 3.5.3. The sequence encoding for Strep-Tag II was amplified and inserted with EcoRI and KpnI (for pN2KD8) or BglII and KpnI (for pN2KD10) using overlapping PCR. All generated plasmids are listed in Table 12.

Plasmids pN2KD140 (for expression of TS-NifEN from *Geobacter metallireducens*), pN2KD142 (for expression of HA-NifE/NifN-SS from *Synechococcus*

sp. JA-2-3B'a) and pN2KD143 (for expression of HA-NifE/NifN-SS from *H. thermophilus*) were generated to test compatibility of NifEN proteins with optimum NifH partner. The gene and tag sequences from yeast expression vectors were amplified by PCR with following primer combination #3214 and #3215 (for TS-NifEN from *G. metallireducens*), #3231 and #3230 (for HA-NifE from *Synechococcus* sp. JA-2-3B'a), #3212 and #3211 (for NifN-SS from *Synechococcus* sp. JA-2-3B'a), #3231 and #3583 (for HA-NifE from *H. thermophilus*) and #3213 and #3211, then inserted into NcoI+XhoI (for HA-NifE), NcoI+BamHI (for HA-NifE) or NdeI+XhoI (for NifN-SS) digested pETDuet-1 vectors (Novagen) using ELIC method. The generated plasmids are listed in Table 12.

Overexpression of NifEN proteins in *E. coli* required the additional expression of proteins involved in the [Fe-S] cluster biosynthesis (Wilcoxon et al., 2016, Li et al., 2016). For this reason, in this work, *E. coli* BL21(DE3) cells were transformed with generated plasmids listed in Table 12 together with pN2LP30 (pRSF*nifUSmetK*Duet-1), a plasmid for inducible expression of *nifU*, *nifS* and *metK* (Barahona et al., 2021).

Table 12. Plasmids to overexpress NifEN proteins in *E. coli*.

Plasmid backbone	Construct	Promoter	Plasmid name
pETDuet-1	His-NifE ^{<i>A. vinelandii</i>}	T7	pN2KD4
	NifN ^{<i>A. vinelandii</i>}	T7	
pETDuet-1	His-NifE ^{<i>A. vinelandii</i>}	T7	pN2KD8
	Strep- NifN ^{<i>A. vinelandii</i>}	T7	
pETDuet-1	His-NifE ^{<i>A. vinelandii</i>}	T7	pN2KD10
	NifN ^{<i>A. vinelandii</i>} -Strep	T7	
pETDuet-1	TS-NifEN ^{<i>Geobacter metallireducens</i>}	T7	pN2KD140
pETDuet-1	HA-NifE ^{<i>Hydrogenobacter thermophilus</i>}	T7	pN2KD143
	NifN-SS ^{<i>Hydrogenobacter thermophilus</i>}	T7	
pETDuet-1	HA-NifE ^{<i>Synechococcus</i> sp.}	T7	pN2KD142
	NifN-SS ^{<i>Synechococcus</i> sp.}	T7	

3.6.2 Generation of plasmids to overexpress NifEN library in *S. cerevisiae*

Protein sequences from UniProt (UniProt, 2023) were used to optimize by the GenSmart Codon Optimization tool (GenScript, NJ, USA), *nifE* and *nifN* genes for the expression in *S. cerevisiae*. All DNA fragments were cloned into pUC cloning vectors and synthesized by GenScript (AnnexI). The *nifE* and *nifN* genes, lacking the stop codon sequence, were flanked by BamHI and BspEI restriction sites at the 5' and 3' ends, respectively (*i. e.*, GGATCCATG...TCCGGA).

The pESC-HIS vector (Agilent Technologies) was used to express NifE proteins. First, the BamHI site within pESC-HIS was replaced with NcoI using PCR with primers #2910 and #2911, generating plasmid pN2KD13. Then, a fragment containing SU9-HA with a sequence for BamHI-BspEI-Stop added to the 3' end of SU9-HA was synthesized by GenScript, amplified with primers #2937 and #2938 and inserted using ELIC into NotI+SacI digested pN2KD13, generating pN2KD64 (pESC-HIS with SU9-HA-BamHI-BspEI-Stop) Figure 9A. The *nifE* genes were then inserted into BamHI+BspEI digested pN2KD64 by cloning, generating pESC-HIS variants for expression of SU9-HA-NifE.

The pESC-LEU vector (Agilent Technologies) was used for NifN expression. First, a fragment containing SU9- Strep-tag II (denoted here as SS-tag) with a sequence for BamHI-BspEI recognition sites inserted between the SU9 mitochondria signal and the SS-tag, and a Stop sequence added to the 3' end of SS, was synthesized by GenScript and inserted to vector digested with NotI and SacI. Then, the XhoI+PacI fragment from pN2KD40 was inserted by cloning into XhoI+PacI digested pESC-LEU, generating pN2KD65 (pESC-LEU with SU9-BamHI-BspEI-SS-Stop) Figure 9B. The *nifN* genes were then inserted into BamHI+BspEI digested pN2KD65 by cloning, generating pESC-LEU variants for expression of SU9-NifN-SS.

The pESC-HIS vector (Agilent Technologies) was used for NifEN fusion protein expression. First, the sequence encoding for SU9-TS was amplified with primers #2943 and #2944 by PCR from plasmid pN2SB103 and inserted using ELIC into NotI+SacI digested pN2KD13, generating pN2KD66 (pESC-HIS with SU9-TS-BamHI-BspEI-Stop) Figure 9C. The *nifEN* fusion genes were then inserted into BamHI+BspEI digested pN2KD66 by cloning, generating pESC-HIS variants for expression of SU9-HA-NifEN.

All generated yeast plasmids (Table 13) were introduced into parental strain SB320Y, which expressed NifU^{Av} and NifS^{Av} targeted to mitochondria.

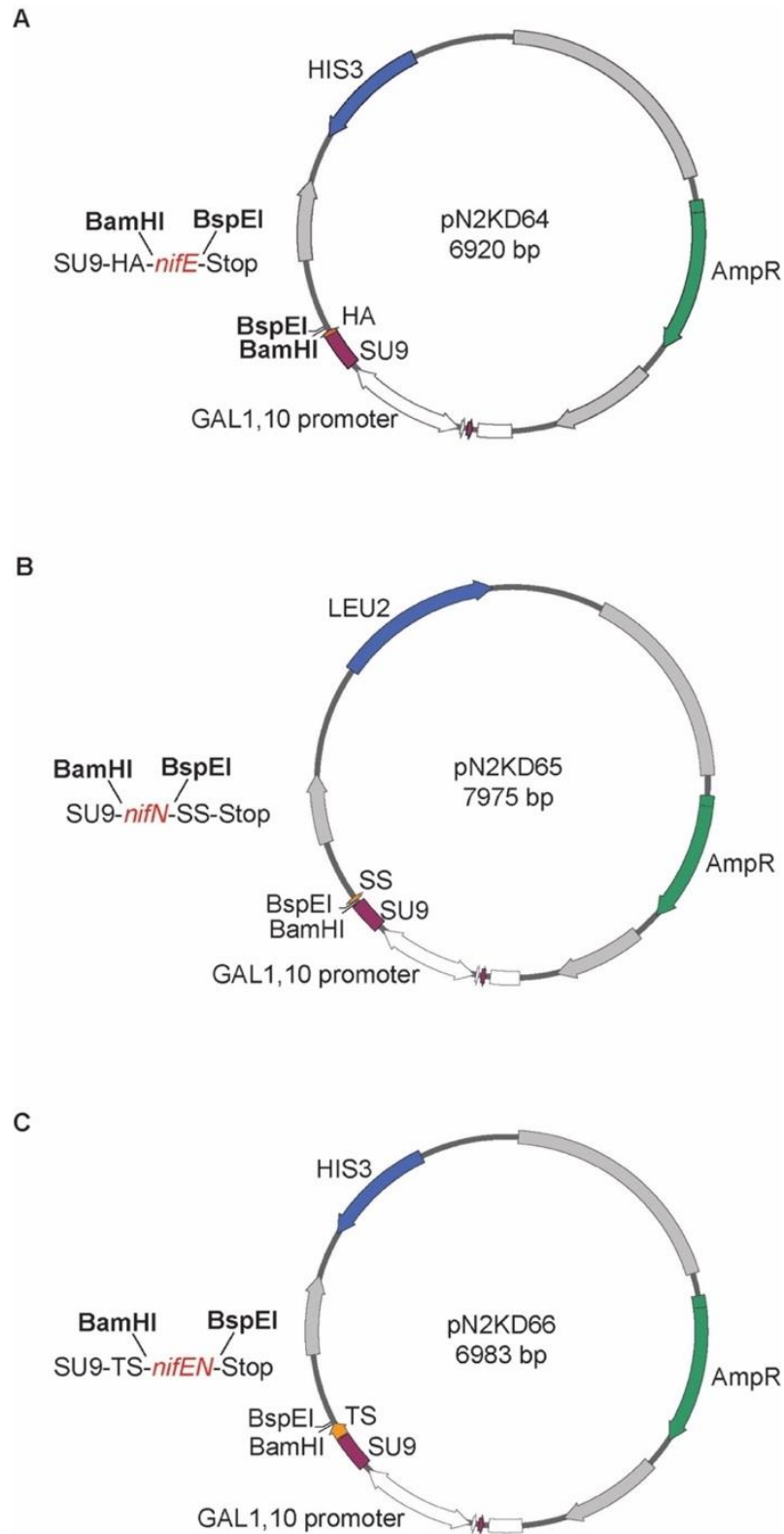


Figure 9. Simplified representation of parental plasmids pN2KD64 (A), pN2KD65 (B), and pN2KD66 (C) used to insert *nifEN* genes.

Table 13. The list of plasmids to overexpress NifEN proteins in yeast.

Backbone	Constructs	Promoter	Plasmids
pESC-His (pN2KD66)	SU9-TS-NifEN ^{Anabaena variabilis}	GAL10	pN2KD104
pESC-His (pN2KD64)	SU9-HA-NifE ^{Arcobacter nitrofigilis}	GAL10	pN2KD79
pESC-Leu (pN2KD65)	SU9-NifN ^{Arcobacter nitrofigilis} -SS	GAL10	pN2KD130
pESC-His (pN2KD64)	SU9-HA-NifE ^{Azotobacter vinelandii}	GAL10	pN2KD67
pESC-Leu (pN2KD65)	SU9-NifN ^{Azotobacter vinelandii} -SS	GAL10	pN2KD120
pESC-His (pN2KD64)	SU9-HA-NifE ^{Chlorobaculum tepidum}	GAL10	pN2KD81
pESC-Leu (pN2KD65)	SU9-NifN ^{Chlorobaculum tepidum} -SS	GAL10	pN2KD132
pESC-His (pN2KD64)	SU9-HA-NifE ^{Fischerella thermalis}	GAL10	pN2KD69
pESC-Leu (pN2KD65)	SU9-NifN ^{Fischerella thermalis} -SS	GAL10	pN2KD122
pESC-His (pN2KD66)	SU9-TS-NifEN ^{Geobacter metallireducens}	GAL10	pN2KD103
pESC-His (pN2KD64)	SU9-HA-NifE ^{Hydrogenobacter thermophilus}	GAL10	pN2KD74
pESC-Leu (pN2KD65)	SU9-NifN ^{Hydrogenobacter thermophilus} -SS	GAL10	pN2KD126
pESC-His (pN2KD64)	SU9-HA-NifE ^{Klebsiella oxytoca}	GAL10	pN2KD70
pESC-Leu (pN2KD65)	SU9-NifN ^{Klebsiella oxytoca} -SS	GAL10	pN2KD123
pESC-His (pN2KD64)	SU9-HA-NifE ^{Methanococcus maripaludis}	GAL10	pN2KD78
pESC-Leu (pN2KD65)	SU9-NifN ^{Methanococcus maripaludis} -SS	GAL10	pN2KD129
pESC-His (pN2KD64)	SU9-HA-NifE ^{Methanosarcina acetivorans}	GAL10	pN2KD71
pESC-Leu (pN2KD65)	SU9-NifN ^{Methanosarcina acetivorans} -SS	GAL10	pN2KD124
pESC-His (pN2KD64)	SU9-HA-NifE ^{Methanosarcina barkeri}	GAL10	pN2KD82
pESC-Leu (pN2KD65)	SU9-NifN ^{Methanosarcina barkeri} -SS	GAL10	pN2KD133
pESC-His (pN2KD64)	SU9-HA-NifE ^{Methanothermobacter marburgensis}	GAL10	pN2KD68
pESC-Leu (pN2KD65)	SU9-NifN ^{Methanothermobacter marburgensis} -SS	GAL10	pN2KD121
pESC-His (pN2KD64)	SU9-HA-NifE ^{Methanothermobacter thermotrophicus}	GAL10	pN2KD80
pESC-Leu (pN2KD65)	SU9-NifN ^{Methanothermobacter thermotrophicus} -SS	GAL10	pN2KD131
pESC-His (pN2KD64)	SU9-HA-NifE ^{Nostoc sp.}	GAL10	pN2KD76
pESC-Leu (pN2KD65)	SU9-NifN ^{Nostoc sp.} -SS	GAL10	pN2KD127
pESC-His (pN2KD64)	SU9-HA-NifE ^{Paenibacillus sabinae}	GAL10	pN2KD77
pESC-Leu (pN2KD65)	SU9-NifN ^{Paenibacillus sabinae} -SS	GAL10	pN2KD128
pESC-His (pN2KD64)	SU9-HA-NifE ^{Rhodobacter capsulatus}	GAL10	pN2KD83

Backbone	Constructs	Promoter	Plasmids
pESC-Leu (pN2KD65)	SU9-NifN ^{Rhodobacter capsulatus} -SS	GAL10	pN2KD134
pESC-His (pN2KD64)	SU9-HA-NifE ^{Synechococcus sp.}	GAL10	pN2KD72
pESC-Leu (pN2KD65)	SU9-NifN ^{Synechococcus sp.} -SS	GAL10	pN2KD125

3.6.3 Plasmids generated to overexpress NifEN library and isolate NifEN proteins from *A. vinelandii*

To conduct *in vivo* complementation experiments in the *A. vinelandii* strain DJ2898 ($\Delta vnfDGK$, $\Delta anfHDGK$: *Gent*^R, $\Delta nifEN$:*kan*^R, *str*^R in *scr* region), the 17 *nifEN* candidate gene cassettes were cloned into a plasmid suitable for *A. vinelandii* transformation. Two strategies were employed to avoid improper *nifEN* gene translation. First, the native intergenic sequence naturally found between *nifE* and *nifN* in *A. vinelandii* was inserted between the yeast codon-optimized *nifE* and *nifN* variants using overlapping PCR, except for the two *nifEN* natively fused (from *A. variabilis* and *G. metallireducens*). In addition, the last three codons of *nifE* and the first three codons of *nifN* were replaced by the corresponding codons from the *A. vinelandii* *nifE* and *nifN* genes by PCR. The *nifEN* cartridges were then subcloned into plasmid pN2KD163 using ELIC. Derivative from pDB1588, plasmid pN2KD163 contains recombination sites for integrating the genes into the sucrose (*scr*) metabolism gene region, under the regulation of the arabinose promoter (*Para*) (Dos Santos et al., 2007). A DNA sequence encoding an SS-tag was placed at the 5' of each yeast-optimized *nifE* gene to facilitate NifEN isolation. Strain DJ2898 was independently transformed with the 17 plasmids containing the *nifEN* candidate genes listed in Table 14.

Table 14. Generated plasmids to overexpress NifEN protein in *A. vinelandii* strain DJ2898.

Plasmid backbone	Constructs	Promoter	Plasmids
pN2KD163	Strep-NifEN ^{Anabaena variabilis}	<i>Para</i>	pN2KD161
pN2KD163	Strep-NifEN ^{Arcobacter nitrofigilis}	<i>Para</i>	pN2KD155
pN2KD163	Strep-NifEN ^{Azotobacter vinelandii}	<i>Para</i>	pN2KD144
pN2KD163	Strep-NifEN ^{Chlorobaculum tepidum}	<i>Para</i>	pN2KD157
pN2KD163	Strep-NifEN ^{Fischerella thermalis}	<i>Para</i>	pN2KD146
pN2KD163	Strep-NifEN ^{Geobacter metallireducens}	<i>Para</i>	pN2KD160
pN2KD163	Strep-NifEN ^{Hydrogenobacter thermophilus}	<i>Para</i>	pN2KD150
pN2KD163	Strep-NifEN ^{Klebsiella oxytoca}	<i>Para</i>	pN2KD147
pN2KD163	Strep-NifEN ^{Methanococcus maripaludis}	<i>Para</i>	pN2KD154

Plasmid backbone	Constructs	Promoter	Plasmids
pN2KD163	Strep-NifEN ^{<i>Methanosarcina acetivorans</i>}	Para	pN2KD148
pN2KD163	Strep-NifEN ^{<i>Methanosarcina barkeri</i>}	Para	pN2KD158
pN2KD163	Strep-NifEN ^{<i>Methanothermobacter marburgensis</i>}	Para	pN2KD145
pN2KD163	Strep-NifEN ^{<i>Methanothermobacter thermautotrophicus</i>}	Para	pN2KD156
pN2KD163	Strep-NifEN ^{<i>Nostoc sp.</i>}	Para	pN2KD152
pN2KD163	Strep-NifEN ^{<i>Paenibacillus sabiniae</i>}	Para	pN2KD153
pN2KD163	Strep-NifEN ^{<i>Rhodobacter capsulatus</i>}	Para	pN2KD159
pN2KD163	Strep-NifEN ^{<i>Synechococcus sp.</i>}	Para	pN2KD149

For NifEN protein isolation from *A. vinelandii*, the yeast codon-optimized *nifEN* genes were cloned into NheI and NsiI digested pDB2617 (placing an SS-tag at the 5′-end of the *nifE* gene) that allows for recombination between the *nifH* promoter and *nifX* in *A. vinelandii*. The resulting plasmids (Table 15) were used to transform *A. vinelandii* DJ, replacing *nifHDKTYnafABnifEN* genes with the respective *nifEN* gene.

Table 15. Plasmids generated to isolate NifEN protein from Δ nifHDKTYnafAB *A. vinelandii*.

Plasmid backbone	Constructs	Promoter	Plasmids
pDB2617	Strep-NifEN ^{<i>Geobacter metallireducens</i>}	<i>nifH</i>	pDB2630
pDB2617	Strep-NifEN ^{<i>Hydrogenobacter thermophilus</i>}	<i>nifH</i>	pDB2628
pDB2617	Strep-NifEN ^{<i>Synechococcus sp.</i>}	<i>nifH</i>	pDB2621

3.7 Biomass production for Nif protein expression and purification

3.7.1 Nif protein overexpression in aerobically grown yeast cells

Yeast strains were grown in SD plates supplemented with amino acids for two days in a 30°C incubator. After two days, single yeast colonies were used to inoculate a 500 ml flask with 100 ml of SD medium containing 2% glucose and appropriate amino acids. Then, the precultures were incubated overnight at 28°C with shaking at 200 rpm and used to inoculate a 4-liter bioreactor (Applikon) containing SD medium supplemented with the same components as before. After around 22 hours, the addition of 0.25% yeast extract, 0.25% bactopectone, 0.25% bactotryptone, 2.5% sucrose, 25

mg/l ammonium iron (III) citrate, 0.225 mM iron (II) sulfate, 1.25 mM magnesium sulfate, 1.5 mM calcium chloride and vitamin solution were used to enrich the medium. The next day, the culture was supplemented with 2.25% galactose, 0.25% bactopectone, and 0.25% yeast extract for galactose-induced expression of Nif proteins. The induction phase was maintained for around 20h, then the cells were collected by centrifugation at 5,000 x *g* for 5 min and the pellet was stored in LN₂. As an example, the typical batch fermentation process of *S. cerevisiae* is presented below in Figure 10.

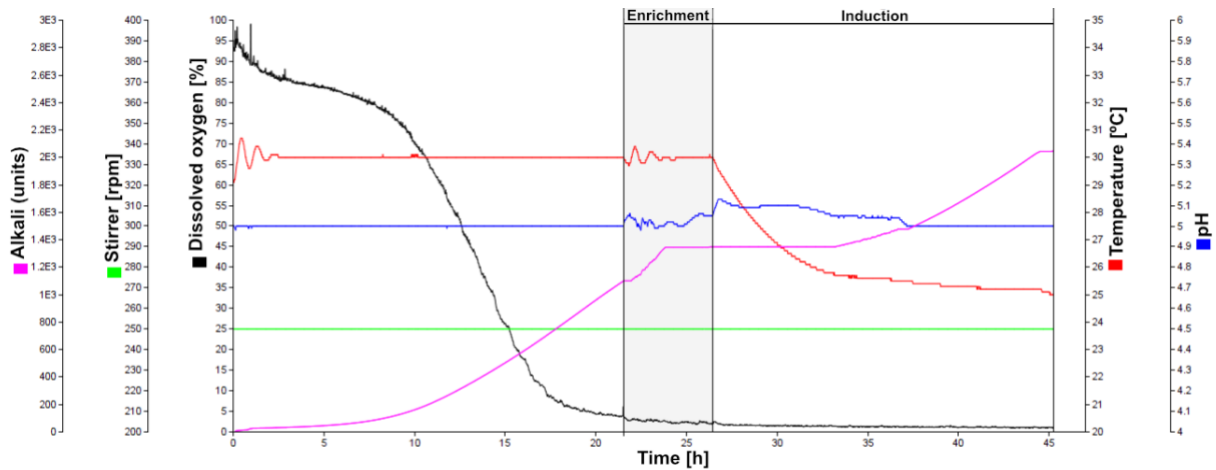


Figure 10. Typical batch fermentation process of the *S. cerevisiae* strain expressing ScNifEN^{Gm} (KD37Y). The following parameters were monitored during the process: dissolved oxygen (black), temperature (red), pH (blue), and alkali addition (pink) (to maintain constant pH). Stirring (green) was maintained at 250 rpm throughout the process. Culture turbidity (OD₆₀₀) was measured manually throughout the process. In the initial phase, the culture was grown in a synthetic drop-out (SD) medium supplemented with amino acids containing 3.3% glucose. When the dissolved oxygen dropped to below ~5%, the culture was enriched (indicated in grey) (around 22 h). Dissolved oxygen remained close to 0% during the rest of the process. The induction phase started at 27h (the white area) and was maintained until 45 h. The turbidity at the end of the process was 22.5 (OD₆₀₀) and the cells were harvested.

3.7.2 Nif protein overexpression in aerobic *E. coli* cultures

Chemically competent *E. coli* BL21 (DE3) cells were freshly transformed with the plasmids of interest, spread on selective agar plates, and incubated overnight at 37°C. The biomass from the plates was used to inoculate 0.5-liter flasks containing 100 ml LB supplemented with appropriate antibiotics for selection and the prepared precultures was incubated at 37°C with shaking (200 rpm) for 2-3 h until the OD₆₀₀ reached 0.8-1. About 25 ml of growth cells were transferred to a 4-liter flask containing 1 liter of LB medium supplemented with 20 µM ammonium iron (III) citrate and antibiotics (resulting in an initial OD_{600nm} of about 0.02). First aerobic growth of cultures was performed at 37°C with constant agitation at 150 rpm until reaching an OD_{600nm} of 0.6-0.8, then after around 4h was added lactose (3 g/l final concentration) to induce protein expression. Additionally,

the medium was supplemented with 2 mM L-cysteine and 0.2 mM ammonium iron (III) citrate (final concentrations) to support cellular [Fe-S] cluster biosynthesis and insertion (Crack et al., 2014, Hoffman et al., 1995). Protein induction was performed overnight (typically about 16 h) at 30°C and 105 rpm. The cells were collected by centrifugation at 5,000 x g for 5 min at 4°C. The resulting cell biomass was snap-frozen in LN₂ and stored at -80°C until used.

3.7.3 Nitrogenase derepression in small-scale *A. vinelandii* cultures

For diazotrophic growth determinations, *A. vinelandii* strains were grown on solid Burk's medium for 2-3 days at 30°C. Fresh *A. vinelandii* biomass was then used to inoculate 125 ml flasks containing 25 ml of Burk medium (supplemented with 5 mM NH₄AOc) at an initial OD₆₀₀ of 0.00125. Flasks were cultured overnight at 30°C with constant agitation at 250 rpm. The next day, the presence of a fixed nitrogen source in the media was confirmed by the Nessler test. Hundred twenty-five ml flasks containing 25 ml of Burk's medium without a fixed nitrogen source but supplemented with 3 g/l arabinose to stimulate expression of Nif proteins were then inoculated at an initial OD₆₀₀ of 0.3 and incubated at 30°C and 250 rpm for several hours until cell cultures reached an OD₆₀₀ of about 1. These *A. vinelandii* cultures were used for *in vivo* acetylene reduction assays described below in section 3.10.1.

3.7.4 Nitrogenase derepression in large-scale *A. vinelandii* cultures

The cultures of *A. vinelandii* strains generated in this work were grown in 60-120 ml of Burk's medium using a 150-liter custom-built fermenter (W.B. Moore Inc.), to produce NifEN proteins for purifications and biochemical assays described in this thesis. A typical fermenter of *A. vinelandii* culture yields around 350-500 g of wet cell paste at this scale. Moreover, to perform *in vitro* nitrogenase activity assays were used proteins from following sources: NifH and holo-NifDK were isolated from *A. vinelandii* DJ wild-type, Strep-tagged apo-NifDK from $\Delta nifB$ DJ2107 strain (Jiménez-Vicente et al., 2018), His-tagged holo-NifEN isolated from DJ1041 strain ($\Delta nifHDKTYorf1orf2$ *his-nifEN*) (Goodwin et al., 1998), His-tagged apo-NifEN from DJ1118 strain ($\Delta nifB::kanR$ *PnifH::polyHis::nifEN*), which corresponds to the *A. vinelandii* strain DJ1041 (expresses His-tagged NifEN with the deleted *nifB* gene). As an example, the procedure for the growth of *A. vinelandii* DJ2965 (expresses Strep-tagged NifEN from *G. metallireducens*) is detailed below.

A. vinelandii cells were plated on solid Burk's medium and incubated at 30°C for 2-3 days. Cells were scraped and used to inoculate a 2-liter flask with 500 Burk's medium and incubated with vigorous shaking at 30°C overnight. Early the following day, three 2-

liter flasks, each containing 500 ml of Burk's medium, were inoculated with 50 ml of overnight culture. Meanwhile, the fermenter was sterilized and filled with 0.5 mM sucrose, 0.8 mM MgSO₄, 0.67 mM CaCl₂·2H₂O, 27 μM FeSO₄·7H₂O and 1 μM Na₂MoO₄ diluted in 120 l of distilled water. The fermenter was autoclaved for 5 min at 121.5°C and then cooled down and kept at 30°C. Prior to inoculation of the fermenter, 66 mM ammonium iron (III) citrate, 5.7 mM NH₄AOC, 2-3 drops antifoam 204 (Sigma-Aldrich), and 200 ml of autoclaved buffer containing 0.2 M KH₂PO₄ and 1.3 M K₂HPO₄ were added. Then, the previously prepared culture was used to inoculate the fermenter and the conditions were set to 30°C, 125 rpm, and 80 l air/min at 5-6 psi internal pressure. The cells were grown until the nitrogen-fixing source was depleted, and at this point, the cells were derepressed to induce nitrogenase expression. After 4 hours, the culture was concentrated using a hollow fiber system (120 cm/500-kDa pore diameter cartridge) from 120 l to 8-10 l and transferred into 1-liter centrifuge bottles, and the concentrated cell suspension was pelleted by centrifuging at 6,000 x g for 20 min and immediately stored at -80°C.

3.8 Protein expression analysis and purification

3.8.1 Preparation of *S. cerevisiae* protein extracts

Total and soluble yeast protein extracts were prepared by centrifuging yeast cells at 5,000 x g for 10 min at 4°C, and then incubating yeast pellets for 20 minutes at room temperature in YeastBuster protein extraction reagent (Novagen) with shaking at 800 rpm in an anaerobic chamber (Coy Laboratories). Fifty μl of the resulting total extract was mixed with 2x Laemmli buffer and heated for 10 min at 95°C. To prepare the soluble extract, the insoluble cell debris was removed by centrifuging the total extract at 16,000 x g for 20 min at 4°C. Fifty μl of the resulting supernatant was mixed with 2x Laemmli buffer and heated for 10 min at 95°C.

To test soluble complex formation, soluble protein extract (supernatant) was added to MagStrep "type3" XT beads (IBA Lifesciences) or anti-HA magnetic beads (MedChemExpress) and incubated in an anaerobic chamber for 2 h at room temperature with shaking. The beads were then washed three times using 500 μl of 100 mM Tris-HCl pH 8.0, 150 mM NaCl for MagStrep "type3" XT beads, or 137 mM NaCl, 2.7 mM KCl pH 7.4, 8 mM Na₂HPO₄, 2 mM KH₂PO₄, 0.05% Tween 20 for anti-HA magnetic beads. The soluble NifEN protein complex was eluted by the addition of 50 μl of 2x Laemmli buffer and subsequent heating for 10 min at 95°C.

NifEN expression, solubility, and complex formation were analyzed by SDS-PAGE and immunoblotting.

3.8.2 Preparation of cell-free extracts

In this section are described the procedures followed for Nif protein purifications. As Nif proteins are extremely sensitive to O₂, for this reason, all steps of the purification process were performed under strict anaerobic conditions. This includes storage of cell pellets in LN₂, transfer to an anaerobic chamber operating at <1 ppm O₂ (containing 5% of hydrogen and 95% of nitrogen) (Coy Laboratories), cell lysis, centrifugation, chromatographic separations, protein concentration, and storage of the final isolated protein. All buffers used in the procedures were first made anaerobic by bubbling N₂ inside a glovebox to remove dissolved O₂.

3.8.2.1 Preparation of *E. coli* cell-free extracts

The frozen cells were typically resuspended in the lysis buffer (100 mM Tris-HCl pH 8, 200 mM NaCl, 10% glycerol, 2 mM DTH) at a cell to buffer (w/v) ratio of 1:2. Protein inhibitors (1 mM PMSF, 1 µg/ml leupeptin) and 5 µg/ml DNase I was added during resuspension of the cells. The cells were lysed using an Emulsiflex-C5 homogenizer (Avestin Inc.) operating at 15,000 under strict anaerobic conditions to obtain the total cell extract (TE). The broken cell extract was transferred to centrifuge tubes equipped with sealing closures (Beckman Coulter) and centrifuged at 50,000 x *g* for 1 h at 4°C using Avanti J-26 XP (Beckman Coulter) with JA-25.50 rotor. The supernatant obtained after centrifugation was loaded onto the affinity chromatography column for protein purification purposes.

3.8.2.2 Preparation of *S. cerevisiae* cell-free extracts

Purifications of Nif proteins expressed in *S. cerevisiae* were based on Strep-tag Affinity Chromatography (STAC). For this, the mitochondria-targeted Nif proteins had a TS-tag or SS-tag fused to the N-terminus of NifEN or C-terminus of NifN protein, respectively.

Cells produced in the 4-liter fermenter and stored in LN₂ were transferred to an anaerobic glovebox and resuspended in anaerobic lysis buffer containing 100 mM Tris-HCl pH 8.6, 300 mM NaCl, 10% glycerol, and freshly added 2 mM DTH. The ratio of cells to buffer (w/v) was typically 1:2. During cell resuspension, 1 mM PMSF, 1 µg/ml leupeptin, 5 µg/ml DNase I, and 1:200 (w/v) BioLock (IBA Lifesciences) was added. The cells were lysed using an HPH 2000/4-DH5 high-pressure homogenizer (IKA) operating at 20,000 psi under anaerobic conditions to obtain the total extract (TE). The TE was centrifuged at 50,000 x *g* for at least 1,5 h at 4°C in centrifuge tubes equipped with sealing closures (Beckman Coulter) using an Avanti J-26 XP with JA-25.50 rotor. The

supernatant was passed through a 0.2 μm pore size filtering cup (ThermoFisher Scientific) to remove cell debris and precipitated material, resulting in the soluble cell-free extract (CFE). The resin used to purify the proteins was Streptactin XT (IBA Lifesciences). The capacity of the Streptactin resin to bind the TS or SS-tagged proteins requires the pH to be in the range of 7-10. At lower pH values, the binding capacity is affected. Therefore, the pH of the CFE was always determined and corrected to 8 if the pH was lower than 7 before loading on the column.

3.8.2.3 Preparation of *A. vinelandii* cell-free extracts

The purification of the Strep-tagged NifEN proteins from the *A. vinelandii* cell paste was performed using STAC following previously described procedures (Jimenez-Vicente et al., 2018), with some modifications.

The cells stored at $-80\text{ }^{\circ}\text{C}$, were transferred into an anaerobic chamber resuspend and disrupted in ratio of 1:1 (w/v) of lysis buffer (50 mM Tris-HCl pH 8, 2 mM DTH) using a nano-DeBee homogenizer (B.E.E. International, Inc.) maintaining a stream of Ar over the reservoir during the disruption process. Disrupted cells were supplemented with 1 ml of 100 mM PMSF, 10 mg of DNase I, and 5 mg of Pepstatin A. The TE was centrifuged at $167,000 \times g$ for 45 min at 4°C using a Type 45 Ti rotor (Beckman Coulter). The resulting supernatant was then filtered to remove cell debris and precipitated material (generating the CFE), prior to loading onto the affinity chromatography column.

3.8.3 Purification of His-tagged NifEN from *E. coli*

The CFE was obtained as described in section 3.8.2.1 and was loaded at 2 ml/min into a 5 ml HiTrap IMAC HP column (GE Healthcare) charged with Co^{2+} following the manufacturer's instructions. The column was washed with seven washes (10 column volumes each) of lysis buffer supplemented with 0, 10, 30, 50, 150, 300, and 500 mM of imidazole, respectively. Fractions containing NifEN were collected and pooled. The eluted protein was concentrated using 100-kDa cut-off centrifugal filters (Amicon Ultra-15, Millipore) and desalted to remove biotin using a PD-10 column. Finally, His-NifEN preparations were frozen and stored in LN_2 in beads for long-term storage.

3.8.4 Purification of SS- and TS-tagged NifEN from *E. coli*

The CFE was obtained as described previously in section 3.8.2.1. and transferred to the MBraun glovebox ($<1\text{ ppm}$ of O_2). The CFE was loaded into a pre-equilibrated 5 ml Streptactin XT column (IBA Lifesciences) affinity column attached to an ÄKTA FPLC (GE Healthcare). After loading, the column was washed with 10 column volumes of

washing buffer (100 mM Tris-HCl pH 8.0, 200 mM NaCl, 10% glycerol, and 2 mM DTH). The bound protein was eluted from the column by applying 20 ml of wash buffer supplemented with 50 mM biotin (IBA Lifesciences), concentrated using centrifugal filter units with 100-kDa cut-off (Amicon Ultra-15, Millipore) and then snap-frozen and stored cryovials in LN₂.

3.8.5 Purification of SS- or TS-tagged NifEN from *S. cerevisiae*

The soluble CFE was obtained as described in sections 3.8.2.2 and loaded onto a 5 ml Streptactin XT column (IBA Lifesciences) previously equilibrated with wash buffer (100 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol and supplemented with freshly added 2 mM DTH) following the manufacturer's instructions. The purification procedures were conducted using an AKTA Prime FPLC system (GE Healthcare) inside the MBraun glovebox (<0.1 ppm of O₂). The CFE was loaded to the column at 2 ml/min flow. Once loaded, the column was washed using 15 column volumes (CV) of wash buffer. Finally, the protein sample was eluted using D-biotin dissolved in wash buffer at a final concentration of 50 mM. Biotin was removed from the eluted protein using PD-10 (GE Healthcare) desalting columns previously equilibrated with wash buffer. The final sample was concentrated using a 100-kDa cut-off centrifugal filter (Amicon Ultra-15, Millipore) and stored in LN₂ as drops.

3.8.6 Purification of SS-tagged NifEN from *A. vinelandii*

The filtered cell-free extracts were loaded at 2 ml/min into a 5 ml Strep-Tactin column (IBA Lifesciences). The column was washed with buffer B (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM DTH). Protein bound to the column was eluted with buffer B supplemented with 50 mM biotin. Finally, purified NifEN proteins were snap-frozen and stored in LN₂ for further analysis.

3.9 Protein analysis techniques

3.9.1 Protein quantification

Protein concentration was determined using the BCA Protein Assay Kit (Pierce, ThermoFisher Scientific) following the manufacturer's instructions (Smith et al., 1985). Protein samples were first incubated with iodoacetamide (2 mg/ml) for 15 min at 37°C to prevent interference from thiol groups with the BCA reagent (Hill and Straka, 1988).

3.9.2 Iron quantification

The Fe content of purified protein samples was estimated using the bipyridyl assay (Moss and Mellon, 1942) and an iron standard solution (1000 mg/l Fe in nitric acid, Merck). A calibration curve was prepared using Milli-Q water ranging from 0 to 300 μM in a 50 μl volume. Protein samples were also prepared in a final volume of 50 μl . Proteins were denatured and Fe^{2+} was released by the addition of 25 μl of 0.1 N HCl solution and incubation for 10 min at 80°C. Then, 147 μl Milli-Q water, 12.5 μl of 20% hydroxylamine solution, and 15.5 μl of 20 mM bipyridyl solution were added to the samples at RT and then incubated for 1 h at RT. The absorbance at 520 nm was recorded using a SPECTROstar Nano spectrophotometer. The Fe^{2+} concentration in the sample was estimated by comparison to the calibration curve, and the number of Fe atoms per protein molecule could be determined by knowing the concentration and the molecular mass of the protein.

The Fe content of NifEN proteins purified from *A. vinelandii* was determined by inductively coupled plasma mass spectrometry (ICP-MS) at the Metal Analysis Service Center of Virginia Tech.

3.9.3 Protein Mass Spectrometry

Peptide mass fingerprinting was carried out at the Proteomics facility of the Facultad de Farmacia of Universidad Complutense de Madrid (Madrid, Spain) using a 4800 Proteomics Analyzer equipped with a TOF/TOF mass spectrometer. The purified proteins were separated by SDS-PAGE, stained with freshly prepared Coomassie R-250 (Sigma-Aldrich, 0.1% Coomassie in 40% ethanol and 10% glacial acetic acid), destained in a solution containing 40% ethanol and 10% glacial acetic acid, and then, protein bands were excised from the gel using a sterile scalpel blade and transferred to an Eppendorf tube containing Milli-Q water.

3.9.4 Denaturing gel electrophoresis

All SDS-PAGE work was performed using equipment, chemicals, and protocols from Bio-Rad unless otherwise indicated.

Protein samples were mixed with 2x Laemmli buffer (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% (w/v) SDS, and 0.005% (w/v) Bromophenol Blue, supplemented with freshly added 10% β -ME) and incubated for 10 min at 95°C. Before loading the polyacrylamide gels, the samples were centrifuged at maximum speed for 1 min in a tabletop centrifuge. Depending upon protein size, SDS-PAGE was performed using gels with 7-12% acrylamide. The running buffer consisted of 250 mM glycine, 25 mM Tris-

HCl pH 8.3, and 0.1% SDS. Proteins were separated by electrophoresis, normally first run at constant 150 V for 10 min, and then 35 min at constant 250 V. For proteins visualized by Coomassie, the gels were stained overnight in a solution of 0.1% Coomassie brilliant blue R-250, 40% ethanol and 10% glacial acetic acid and then extensively washed with destaining solution (40% ethanol and 10% glacial acetic acid). The sizes of the denatured polypeptides were estimated by comparison to genetics molecular weight marker (ThermoFisher Scientific).

3.9.5 Western blotting

Protein immunodetection was used to confirm protein expression and the identity of purified proteins. The proteins were resolved by SDS-PAGE and then transferred to a 0.45 µm nitrocellulose membrane using a TransBlot Semi-dry transfer device (Bio-Rad) after a 5 min equilibration of SDS gels, membranes, and Whatman papers in transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% methanol). Protein transfer was carried out at a constant 22 V for 40 min. The successful transfer was verified by incubating the membranes for 5 min at RT in Ponceau staining solution (0.2% (w/v) Ponceau S and 5% glacial acetic acid) and removing excess Ponceau by washing with Milli-Q water until the background was light.

After ensuring a correct transfer of the proteins, the Ponceau stain was completely removed by washing with TBS-T (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.02% Tween-20), and the membranes were blocked for 45 min at RT in TBS-T buffer supplemented with 5% (w/v) non-fat dried milk. The membrane was washed 3 times (10 min for each incubation) using TBS-T buffer and then incubated with the primary antibody (primary antibody dilutions are listed in Table 16) at 4°C overnight in an orbital shaker. The primary antibody solution was typically prepared in TBS-T buffer containing 5% BSA (w/v) and 0.1% NaN₃ (w/v). The following day the membranes were washed 3 times (10 min each incubation) using TBS-T buffer to remove excess NaN₃ and primary antibody, and then incubated with secondary antibody solution (dilutions of secondary antibodies conjugated to horseradish peroxidase (HRP) listed in Table 16) at RT for at least 2 h. The secondary antibody solution was freshly prepared in TBS-T buffer supplemented with 2% milk. The membranes were washed 3 times for 10 min with TBS-T buffer and then incubated with the enhanced chemiluminescence (ECL) reagent solutions 1 (2.5 mM Luminol, 0.4 mM p-coumaric acid in 100 mM Tris-HCl buffer pH 8.5) and 2 (0.036% of H₂O₂ in 100 mM Tris-HCl buffer pH 8.5) at 1:1 ratio for 1 min. The membranes were developed using an iBright™ CL1000 Imaging system (ThermoFisher Scientific).

Table 16. Antibodies employed in this thesis.

Target protein	Antibody type	Dilution	Reference number	Source
α -NifU ^{Av}	Primary	1:2000	-	Rubio lab
α -NifS ^{Av}	Primary	1:1000	-	Rubio lab
α -HA	Primary	1:500	A01244	Gene Script
α -His	Primary	1:5000	SAB1305538	Sigma Aldrich
StrepMAB	Primary	1:5000	A01732	Gene Script
α -Rabbit HRP	Secondary	1:50000	11859140	Invitrogen
α -Mouse HRP	Secondary	1:20000	sc-516102	Santa Cruz Biotechnology
α -Rat HRP	Secondary	1:20000	GENA935	Sigma-Aldrich

3.10 The acetylene reduction assays for nitrogenase activity

Nitrogenase activity can be measured in two ways: *in vitro* using pure protein components or *in vivo* in free-living cells. The most used method is the acetylene reduction assay, which is based on the ability of nitrogenase to reduce acetylene to ethylene. Acetylene is highly soluble in water, while ethylene is not. This difference in solubility allows for the easy separation and quantification of ethylene using gas chromatography (GC).

The acetylene reduction assays were conducted with either 6% (v/v) of acetylene in the gas phase for the *A. vinelandii* *in vivo* nitrogenase assay or 8% (v/v) of acetylene in the gas phase for *in vitro* nitrogenase assays. The gas phase composition varied depending on the assay type. For the *A. vinelandii* *in vivo* nitrogenase assays, 94% air was present, whereas *in vitro* nitrogenase assays were carried out in an atmosphere comprising 94% argon to prevent oxidation of Nif proteins. The assays were performed using 9 ml vials sealed with rubber stoppers or fitted with butyl rubber caps. Incubation for the acetylene reduction step was conducted at 30°C for 15-20 min in a water bath with shaking. The reactions were stopped by injecting 100 μ l of 8 M NaOH. The ethylene produced was measured by gas chromatography using a GC-2010 Plus, Shimadzu or GC-2014. The carrier gas was N₂ (25 ml/min), and a mixture of H₂ and synthetic air was used to flame the ionization detector. The GC separation column was a Shimadzu PORAPAK N 80/100 efficient for short-chain hydrocarbon compound chromatographic separation (Shah et al., 1972, David et al., 1980).

3.10.1 *In vivo* acetylene reduction assays in *A. vinelandii*

In vivo acetylene reduction assay was performed with 1 ml of *A. vinelandii* culture prepared according to the procedure described in section 3.7.3. The cell culture was

transferred to a glass vial as described above, then 500 μ l of acetylene was injected and overpressure was relieved. For this type of acetylene reduction assay, the vial was incubated at 30°C for 15 min. The reaction was then stopped and the concentration of ethylene in the vial headspace was measured using the GC method using a flame ionization detector (GC-2010 Plus, Shimadzu).

3.10.2 *In vitro* acetylene reduction assays

FeMo-co synthesis can be determined by measuring the amount of apo-NifDK activated by FeMo-co using *in vitro* acetylene reduction assays. The process of FeMo-co synthesis and insertion into apo-NifDK can be divided into three separate steps presented in (Figure 11). Step 1, synthesis and transfer of NifB-co to NifEN. Step 2, conversion of NifB-co to FeMo-co at NifEN in a process that incorporates Mo and homocitrate and that, in addition to NifEN, requires the action of NifH. Step 3, transfer of FeMo-co from NifEN to apo-NifDK (P-cluster containing but FeMo-co-deficient NifDK) to form active NifDK. To determine the activity of reconstituted NifDK, the reduction of acetylene into ethylene at NifDK in the presence of excess NifH and with DTH is measured by GC.

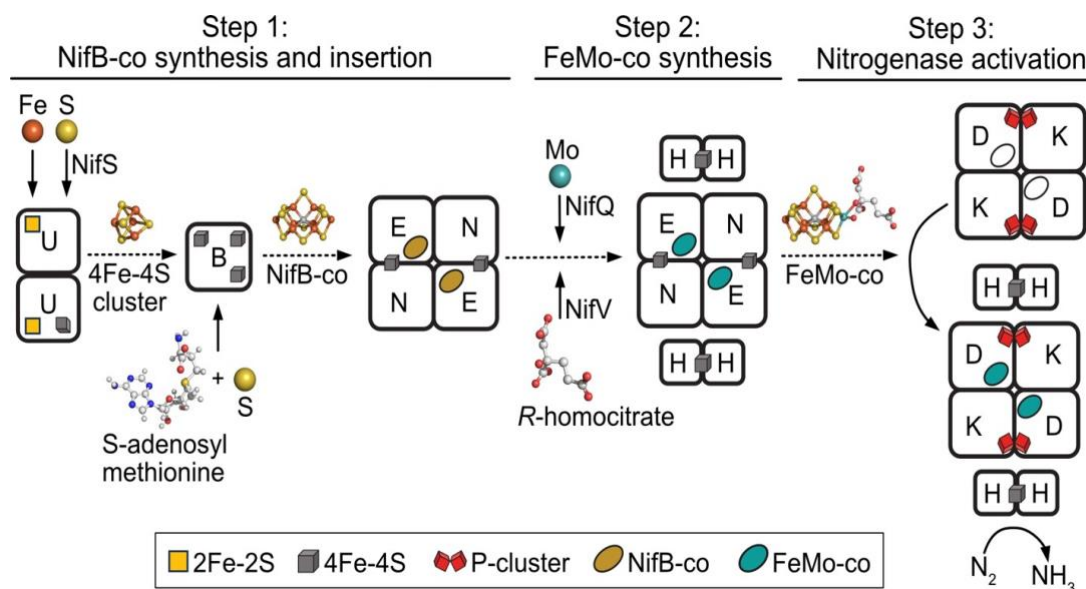


Figure 11. Simplified representation of FeMo-co synthesis and its insertion into the apo-form of NifDK.

In vitro FeMo-co synthesis assays can be initiated in different ways (Buren et al., 2020): 1) the NifB-dependent assay uses NifB together with a source of [4Fe-4S] clusters (either delivered by NifU or directly assembled on NifB using Fe and S) and with presence of SAM, 2) the NifB-co-dependent assay uses NifB-co directly (usually isolated from a diazotrophic bacterium in complex with the NifX carrier protein), or 3) the NifEN-

dependent assay that uses NifEN containing the FeMo-co precursor derived from NifB-co (holo-NifEN) (Figure 12).

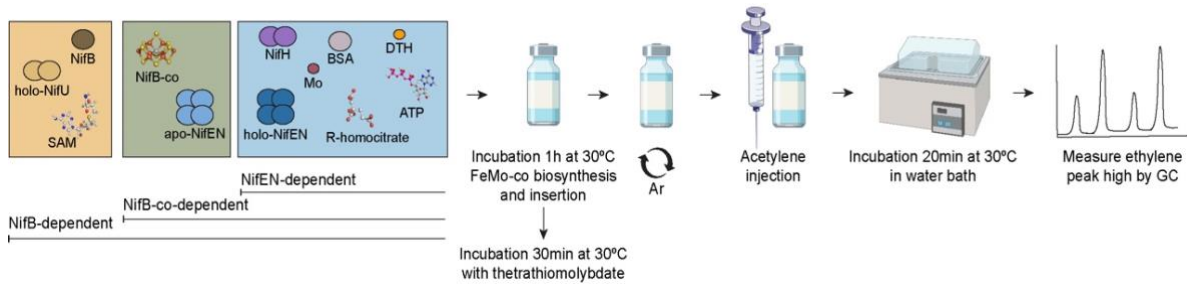


Figure 12. *In vitro* FeMo-co synthesis and insertion assay procedure. Representation of the steps followed to perform the NifB-, NifB-co- and NifEN- dependent FeMo-co synthesis labeled due to the indicated starting components. Synthesis and insertion of FeMo-co into apo-NifDK were performed under an anaerobic atmosphere. After that, the mix was transferred to vials, sealed, and the gas phase was exchanged for argon. Acetylene was injected into the argon gas phase; vials were incubated at 30°C, and the reaction stopped before measuring ethylene formation in a gas chromatograph.

All steps of the *in vitro* nitrogenase activity assays were performed inside an anaerobic chamber (Coy Laboratories). To maintain nitrogenase activity, the reaction required an ATP regenerating mixture consisting of 1.23 mM ATP, 18 mM phosphocreatine disodium salt, 2.2 mM MgCl₂, 3 mM DTH, and 40 µg/ml creatine phosphokinase in 100 mM MOPS buffer (pH 7.0) (Curatti et al., 2006). The performance of the ATP regenerating mixture was tested for each experiment using pure *A. vinelandii* NifH and NifDK proteins. *In vitro* FeMo-co synthesis was carried out in 100 µl reactions in Eppendorf tubes at 30°C with shaking at 600 rpm for 60 min. The reaction mixtures varied in composition depending on the specific type of assay, as detailed in the subsequent sections 3.10.2.1, 3.10.2.2, and 3.10.2.3.

Following the FeMo-co synthesis reaction, in all type of activity assays, 17.5 µM (NH₄)₂MoS₄ was added to the reaction mixture and incubated for 30 min at 30°C to inhibit further insertion of FeMo-co (which could be synthesized by NifEN^{xx} and NifH^{Av}) into apo-NifDK (Shah et al., 1985). For the determination of reconstituted NifDK, the FeMo-co synthesis and insertion mixture was transferred to a 9 ml glass vial containing 500 µl of ATP regenerating mixture to which 1.0 µM NifH^{Av} was added (corresponding to a final molar ratio of NifH: NifDK of 20:1). Then, the N₂ atmosphere in the headspace was exchanged for Ar and 0.75 ml acetylene was injected. The vials were incubated with shaking in a water bath for 20 min at 30°C. The reactions were then stopped and the concentrations of ethylene in the vial headspace were measured using the GC method using a flame ionization detector (GC-2014, Shimadzu).

3.10.2.1 NifEN-dependent *in vitro* FeMo-co synthesis and insertion assays

NifEN-dependent *in vitro* FeMo-co synthesis assays were performed with NifEN variants isolated from *A. vinelandii*. The 100 µl FeMo-co synthesis reaction mixture contained 17.5 µM Na₂MoO₄, 175 µM *R*-homocitrate, 0.6 µM NifEN^{xx}, 3.0 µM NifH^{Av}, 0.3 µM apo-NifDK^{Av}, 1 mg/ml BSA and ATP-regenerating mixture. Positive control reactions contained 0.6 µM NifB-co-loaded NifEN^{Av} isolated from *A. vinelandii* strain DJ1041 ($\Delta nifHDKTYorf1orf2 his-nifEN$) (Soboh et al., 2006), and for negative control reactions no NifEN was added.

3.10.2.2 NifB-co-dependent *in vitro* FeMo-co synthesis assays

For *in vitro* FeMo-co synthesis and apo-NifDK activation using NifEN proteins isolated from *E. coli* or *S. cerevisiae*, the 100 µl FeMo-co synthesis reaction mixture contained GST-tagged NifX-NifB-co (9.6 µM Fe), 17.5 µM Na₂MoO₄, 175 µM *R*-homocitrate, NifEN^{xx} isolated from *E. coli* (4.8 µM Fe) or 0.6 µM NifEN^{xx} isolated from *S. cerevisiae*, 3.0 µM NifH^{xx} isolated from *S. cerevisiae*, 0.3 µM apo-NifDK^{Av}, 1 mg/ml BSA and ATP-regenerating mixture. Positive control reactions contained 0.6 µM apo-NifEN^{Av} isolated from *A. vinelandii* strain DJ1118 ($\Delta nifB::kan^R PnifH::polyHis::nifEN$), which corresponds to the *A. vinelandii* strain DJ1041 (expresses His-tagged NifEN) (Goodwin et al., 1998) with a deleted *nifB* gene. For negative control reactions, no NifEN was added.

3.10.2.3 NifB-dependent *in vitro* FeMo-co synthesis and insertion assay

NifB-dependent FeMo-co synthesis assays using NifH, NifEN, and NifB proteins isolated from yeast were performed in the 100 µl FeMo-co synthesis reaction mixture contained 1.2 µM ScNifB^{xx} isolated from *S. cerevisiae* (Burén et al., 2019), 10.8 µM [4Fe-4S] cluster loaded NifU^{Av} isolated from *E. coli* and reconstituted using NifS, Fe, and cysteine as previously described (Jiang et al., 2022), 125 µM SAM, 1.2 µM NifX^{Av} isolated from *E. coli*, 17.5 µM Na₂MoO₄, 175 µM *R*-homocitrate, 0.6 µM ScNifEN^{xx}, 3.0 µM ScNifH^{xx}, 0.3 µM apo-NifDK^{Av}, 1 mg/ml BSA and ATP-regenerating mixture. Positive control reactions contained 0.6 µM apo-NifEN^{Av} isolated from *A. vinelandii* strain DJ1118 ($\Delta nifB::kan^R PnifH::polyHis::nifEN$), and for negative control reactions, no NifEN was added.

3.11 Bioinformatic methods

3.11.1 Protein sequence and structure analysis

For primary sequence alignments, multiple sequence alignments were performed using the Clustal Omega Multiple Sequence Alignment online tool (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers et al., 2011). The natural NifEN fusion proteins from *A. variabilis* and *G. metallireducens* were first aligned to the *A. vinelandii* NifE and NifN protein sequences to determine which parts within the natural fusions were relevant for alignment with the *A. vinelandii* NifE and NifN sequences. For structural alignments, the individual NifE and NifN structures were obtained using AlphaFold (Jumper et al., 2021, Varadi et al., 2022). The structures were fitted to the *A. vinelandii* NifEN structure (Kaiser et al., 2011) using the MatchMaker function of the Structure Comparison tool in UCSF Chimera (v. 1.16, <https://www.rbvi.ucsf.edu/chimera>) (Pettersen et al., 2004). Matching was iterated by pruning long atom pairs until no pairs exceeded 2.0 Å, and the root-mean-square deviation (RMSD) obtained for the pruned atom pairs was used to determine the structural similarity between the modelled NifEN variant and the NifEN^{Av}. Unpaired t-tests were used to test for structural similarity between complementing and non-complementing NifE or NifN variants, respectively, with the corresponding *A. vinelandii* subunit.

NifEN protein sequences obtained from UniProt (<https://www.uniprot.org/>) were analyzed using the MUSCLE tool (**M**Ultiple **S**equences **C**omparison by **L**og- **E**xpectation).

3.11.2 Statistical analysis

Statistical analyses such as linear regression and unpaired t-tests were performed using Prism 6 (GraphPad). Technical replicates signify individual vials containing protein from a single purification procedure, or individual vials containing cells originating from the same culture. Biological replicates signify individual vials containing protein from separate purification procedures, or individual vials containing cells originating from different cultures.

4. RESULTS

4.1 Identification of a suitable NifEN protein for expression in *S. cerevisiae* mitochondria

4.1.1 Selection of tags for NifEN protein complex detection and purification

A polyhistidine tag (His-tag) placed at the N-terminus of the NifE polypeptide is commonly used to isolate NifEN protein from *A. vinelandii* (Goodwin et al., 1998) or from the heterologous host *E. coli* (Solomon et al., 2020) by Immobilized Metal Affinity Chromatography (IMAC). However, to detect the expression of both subunits of the NifEN protein complex in yeast mitochondria, it was necessary to tag the NifEN protein with two different epitopes. Therefore, to facilitate the detection and protein purification of NifEN protein from yeast, the tolerance of placing a Strep-tag at the NifN polypeptide was first tested in *E. coli*.

E. coli cells were transformed with plasmids expressing NifEN^{Av} protein tagged with, respectively 1) a His-tag at the N-terminus of NifE as a control (His-NifE/NifN), 2) a His-tag at the N-terminus of NifE and a Strep-tag at N-terminus of NifN (His-NifE/Strep-NifN), and 3) a His-tag at N-terminus of NifE and a Strep-tag at C-terminus of NifN (His-NifE/NifN-Strep) (Figure 13).

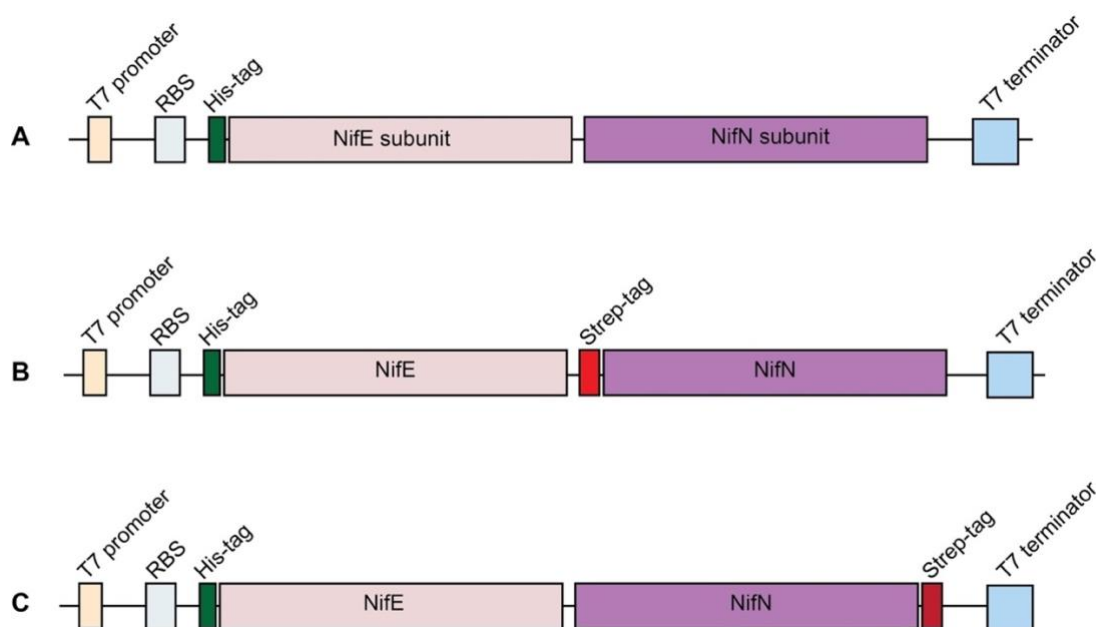


Figure 13. Schematic diagram of the strategy for tagging NifEN proteins expressed in *E. coli* cells. Three constructs of the NifEN protein expressed under the T7 promoter and tagged with (A) a His-tag placed at the N-terminus of NifE, (B) a His-tag placed at the N-terminus of NifE and a Strep-tag placed at the N-terminus of NifN, and (C) a His-tag placed at the N-terminus of NifE and Strep-tag placed at the C-terminus of NifN.

All NifEN proteins were produced in *E. coli* strains that additionally co-expressed NifU^{Av} and NifS^{Av} that provide [4Fe-4S] clusters to Nif proteins. The solubility of the two subunits was determined by analyzing the amount of each protein in the soluble protein fractions (Figure 14).

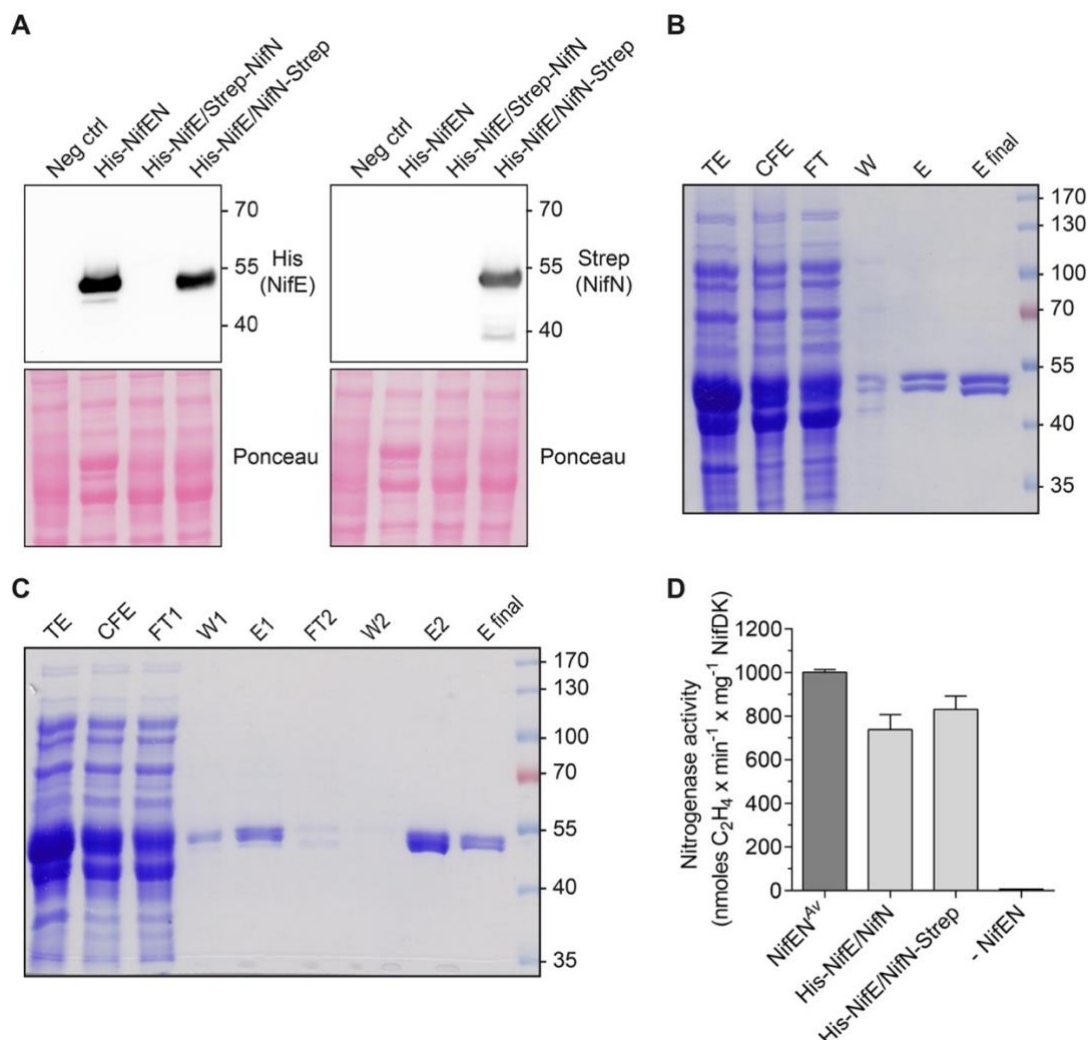


Figure 14. Solubility and functionality of epitope-tagged NifEN proteins isolated from *E. coli*. (A) Immunoblots analysis showing accumulation of epitope-tagged variants of NifEN^{Av} in soluble *E. coli* protein extracts. (B, C) Immobilized metal affinity chromatography (IMAC) purification of His-NifE/NifN (B) and IMAC purification (fractions labeled with “1”) followed by STAC purification (fractions labeled with “2”) of His-NifE/NifN-Strep (C) from *E. coli* cells co-expressing NifU^{Av} and NifS^{Av}. Fractions were analyzed by SDS-PAGE followed by Coomassie staining, TE = total extract, CFE = soluble cell-free extract, FT = flow-through fraction, W = wash fraction, E = elution fraction, E final = final protein fraction after desalting and concentration. Molecular weight markers (kDa) and antibodies used are indicated. (D) NifB-co-dependent FeMo-co synthesis and apo-NifDK activation using His-NifE/NifN and His-NifE/NifN-Strep proteins isolated from *E. coli*. NifEN^{Av} denotes His-NifE/NifN isolated from *A. vinelandii* strain DJ1118 (apo-NifEN). Activities are represented as nanomoles of ethylene produced per minute and milligram of NifDK. Error bars indicate mean $\pm$ SD; $n = 3$ technical replicates.

Only when the NifN polypeptide was tagged at its C-terminus a soluble NifEN protein was detected. Epitope-tagged NifEN^{Av} variants identified as soluble were then

purified under anaerobic conditions by IMAC purification (for His-NifEN) and IMAC purification followed by STAC purification (for His-NifE/NifN-Strep) as described in section 3.8.3 and 3.8.4 (Figure 14B and Figure 14C). The pure proteins were assayed for NifB-co-dependent FeMo-co synthesis and apo-NifDK activation to test if epitope-tagging of the NifN polypeptide affects NifEN functionality. The activity of this dual-tagged NifEN (His-NifE/NifN-SS) was similar to the His-tagged NifEN control protein (His-NifE/NifN) when isolated from *E. coli* (Figure 14D). These results indicate that the soluble complex formation of the NifEN variants in yeast could be evaluated *in vitro* by pull-down experiments targeting each of the NifEN subunits. However, in the final constructs of yeast strains affinity His-tag was replaced with epitope hemagglutinin (HA-tag), as the HA-tag has been shown to enhance the solubility of some proteins expressed in yeast cells and does not interfere with protein folding and functionality (Zhao et al., 2013). To our experience, the HA-tag is also more specific for binding and detection assays in yeast extracts.

4.1.2 Design of the *nifEN* library

Efforts to express NifEN originating from *A. vinelandii* in the yeast mitochondria generated a protein that, as confirmed by immunoblotting analysis, migrated as a single band. This observation led to the interpretation that perhaps only one subunit can be expressed and/or remain stable in yeast mitochondria underscoring the non-trivial nature of forming soluble NifEN complex in eukaryotic cells (Buren et al., 2017). This finding highlights the need to generate a codon-optimized library of 17 *nifEN* genes originating from diverse bacterial and archaeal origins for expression in *S. cerevisiae*. The selection was biased towards organisms with proven ability to fix nitrogen, thermophilic organisms that could harbor proteins with the ability to withstand the elevated temperatures of the mitochondria (Chrétien et al., 2018), aerobic and/or photosynthetic organisms that could produce more O₂-resistant NifEN variants and organisms that had served as successful sources of soluble and stable yeast and plant NifH and NifB proteins in previous studies. Two genes that produce natural “NifEN fusion proteins” in *A. variabilis* and *G. metallireducens* were also included, as previous work by Allen and colleagues showed that a synthetic gene fusion between *nifD* and *nifK* generated a NifDK protein that was soluble and more resistant to degradation in *N. benthamiana* mitochondria (Allen et al., 2020) (Annex I). All NifEN candidates contained the cysteine (Cys) ligands for binding of the permanent [4Fe-4S] cluster (corresponding to Cys³⁷, Cys⁶² and Cys¹²⁴ in *A. vinelandii* NifE, and Cys⁴⁴ in *A. vinelandii* NifN) and the NifB-co precursor (Cys²⁵ in *A. vinelandii* NifE) (Figure 15).

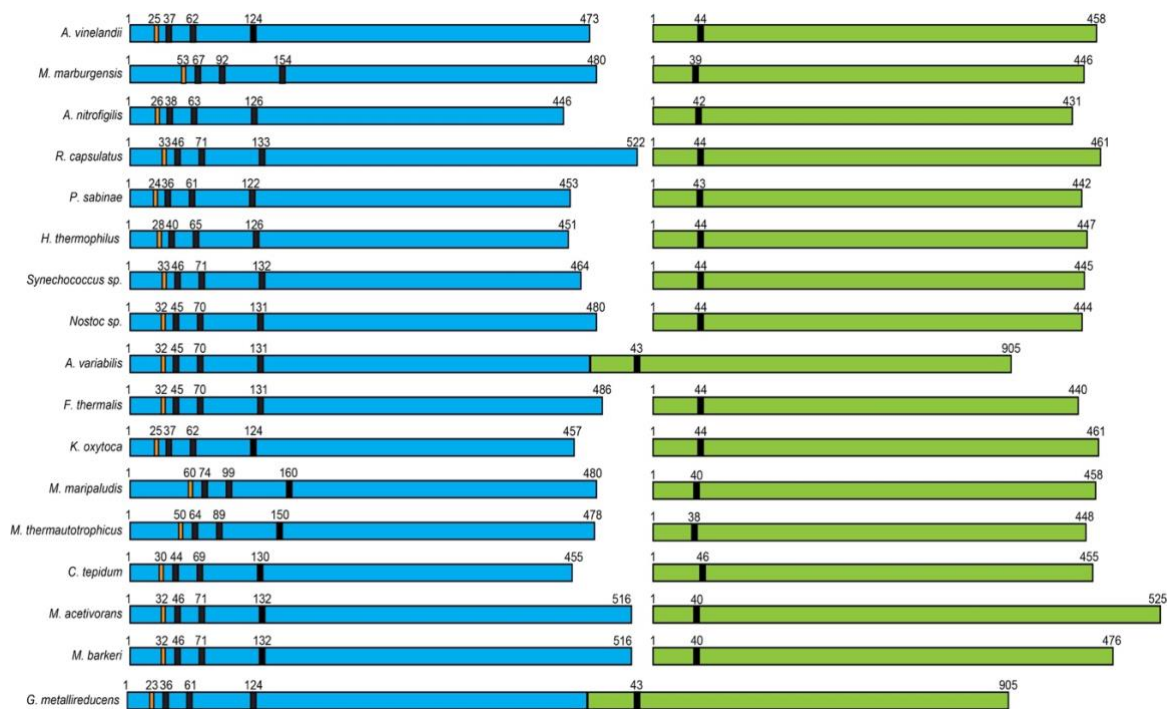


Figure 15. Alignment of NifE and NifN proteins depicting conserved cluster-binding Cys residues. NifE and NifN are represented as blue and green rectangular boxes, respectively. The ligands for binding of the permanent [4Fe-4S] clusters (corresponding to Cys³⁷, Cys⁶², and Cys¹²⁴ in *A. vinelandii* NifE, and Cys⁴⁴ in *A. vinelandii* NifN) are marked in black. The putative ligand to the FeMo-co precursor (Cys²⁵ in *A. vinelandii* NifE) is marked in orange.

4.1.3 *In vivo* functionality of NifEN variants in *A. vinelandii*

In parallel with the yeast expression experiments, screening of the NifEN candidates was also conducted in *A. vinelandii*. Since NifEN plays a critical role in the assembly of the FeMo-co cluster, which is required for nitrogenase activity, assessing the *in vivo* activity and cross-compatibility of NifEN variants with other *A. vinelandii* proteins allows for examining if selected NifEN protein variants can interact and cooperate with the other essential proteins. This analysis also provides valuable insights into their effectiveness for nitrogen fixation. To test this assessment, each of the 17 *nifE* and *nifN* genes were cloned into a mutant strain of *A. vinelandii* lacking the endogenous *nifEN* genes and the genes encoding the structural components of the two alternative nitrogenases. This ensures that the nitrogen-fixing capability of the *A. vinelandii* mutant strain can be restored specifically by introduced NifEN variants.

The *A. vinelandii* mutant strain DJ2898 ($\Delta vnfDGK$, $\Delta anfHDGK$: *Gent*^R, $\Delta nifEN$: *Kan*^R, *Str*^R in *scr* region) was used to test *in vivo* functionality and compatibility of NifEN candidates with other components of *A. vinelandii* FeMo-co biosynthesis pathway. To enable specific and inducible expression of the heterologous *nifEN* genes, the 17 *nifEN* variants were integrated into the sucrose region of the *A. vinelandii* genome, downstream of the *Para*. A Strep-tag affinity tag (SS-tag) was strategically engineered into the N-

terminal coding region of the *nifE* gene to facilitate the efficient purification of the NifEN protein, as this location does not interfere with the expression or activity of the NifEN protein. Eleven out of the 17 species-specific *nifEN* constructs successfully rescued the null diazotrophic growth phenotype of DJ2898 in both solid and nitrogen-free and arabinose-containing media (Figure 16).

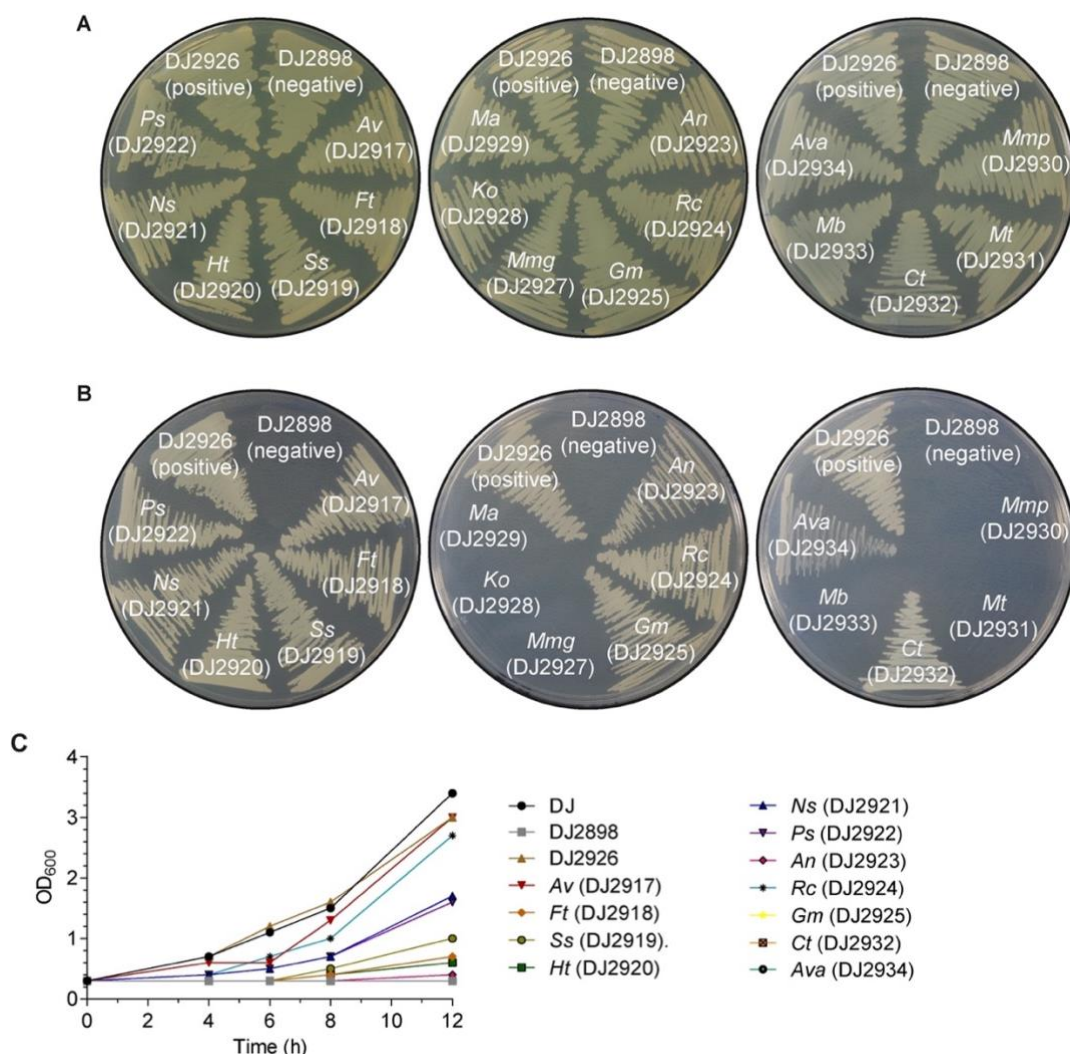


Figure 16. In vivo functionality of NifEN candidates. (A, B) Growth on solid media with nitrogen source (non-diazotrophic growth) (A) or on nitrogen-free and arabinose-containing solid media (non-diazotrophic growth) (B) of *A. vinelandii* DJ2898 transformed with *nifEN* genes. (C) Growth of all strains with diazotrophic capacity in nitrogen-free liquid media. *A. vinelandii* DJ2898 complemented with *nifEN* genes from *F. thermalis*, *H. thermophilus*, *A. nitrofigilis*, *G. metallireducens*, *C. tepidum*, and *A. variabilis* presented slower diazotrophic growth and significant turbidities were only observed at later time points (*F. thermalis*, OD₆₀₀ of 4.0 after 25 h; *H. thermophilus*, OD₆₀₀ of 5.8 after 25 h; *A. nitrofigilis*, OD₆₀₀ of 1.8 after 25 h; *G. metallireducens*, OD₆₀₀ of 1.5 after 25 h; *C. tepidum*, OD₆₀₀ of 1.9 after 45 h; *A. variabilis*, OD₆₀₀ of 1.1 after 54 h). No growth was observed in *A. vinelandii* DJ2898 complemented with *nifEN* genes from *M. marburgensis*, *K. oxytoca*, *M. acetivorans*, *M. maripaludis*, *M. thermoautotrophicus*, or *M. barkeri*. Av, *A. vinelandii*; Ft, *F. thermalis*; Ss, *Synechococcus* sp. JA-2-3B'a; Ht, *H. thermophilus*; Ns, *Nostoc* sp.; Ps, *P. sabinae*; An, *A. nitrofigilis*; Rc, *R. capsulatus*; Gm, *G. metallireducens*; Mmg, *M. marburgensis*; Ko, *K. oxytoca*; Ma, *M. acetivorans*; Mmp, *M. maripaludis*; Mt, *M. thermoautotrophicus*; Ct, *C. tepidum*; Mb, *M. barkeri*; Ava, *A. variabilis*.

In vivo complementation was further verified by the *in vivo* acetylene reduction assay described in section 3.10.1., as a proxy for N₂ reduction. This additional analysis confirmed the functional integrity of the heterologously produced NifEN protein, as it demonstrated the ability to support nitrogen reduction in the rescued strains (Figure 17).

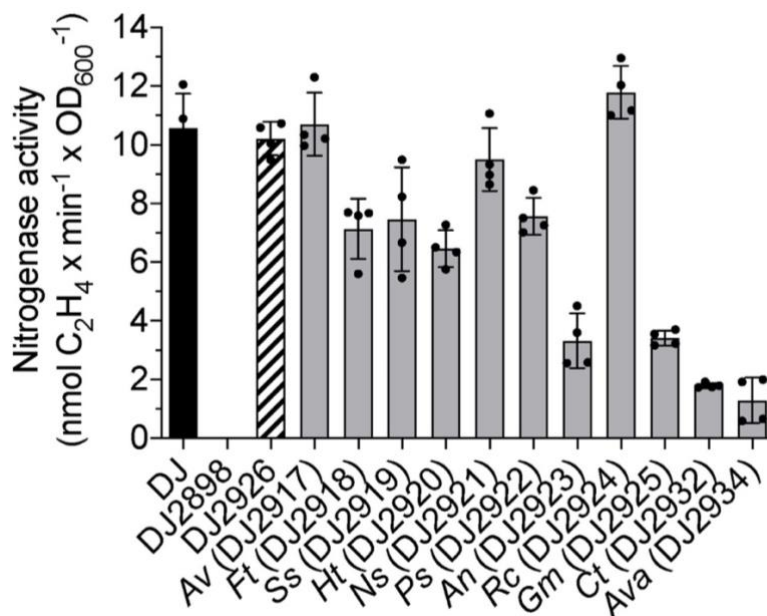


Figure 17. *In vivo* acetylene reduction activities of strains with restored Nif⁺ phenotype. DJ2926 results from the transformation of DJ2898 with the native *A. vinelandii* nifEN genes. Error bars indicate mean activity $\pm$ SD; $n = 2$ biological replicates shown by dots (2 technical replicates each). Note that the individual activity values contributing to the mean were identical for DJ2926 and DJ2932 (Ct). Av, *A. vinelandii*; Ft, *F. thermalis*; Ss, *Synechococcus* sp. JA-2-3B'a; Ht, *H. thermophilus*; Ns, *Nostoc* sp.; Ps, *P. sabinae*; An, *A. nitrofigilis*; Rc, *R. capsulatus*; Gm, *G. metallireducens*; Ct, *C. tepidum*; Ava, *A. variabilis*.

4.1.4 Identification of soluble NifEN protein complex in the yeast mitochondria

The yeast-codon-optimized gene sequences encoding the NifE and NifN variants were cloned into distinct yeast pESC vectors using GAL10 promoters to produce both proteins at similar levels. To target all expressed Nif proteins to yeast mitochondria, an SU9 presequence was used (Westermann and Neupert, 2000). For detection and purification purpose the NifE polypeptide was expressed containing an N-terminal HA-tag and the NifN polypeptide with a C-terminal SS-tag (i.e., HA-NifE/NifN-SS). Only one tag was required for the two natural NifEN fusion proteins originating from *A. variabilis* and *G. metallireducens*. These variants were instead expressed with a duplicated Strep-tag II sequence (Twin-Strep, or TS) located at the N-terminus of NifEN (i.e., TS-NifEN). All DNA and protein sequences used for heterologous NifEN expression in *S. cerevisiae* are listed in Annex I. All NifEN proteins were produced in an *S. cerevisiae* background

strain (SB320Y) that contains an auxiliary vector for co-expression of *A. vinelandii* NifU and NifS. Also these proteins are targeted to the mitochondria using SU9 (Figure 18A).

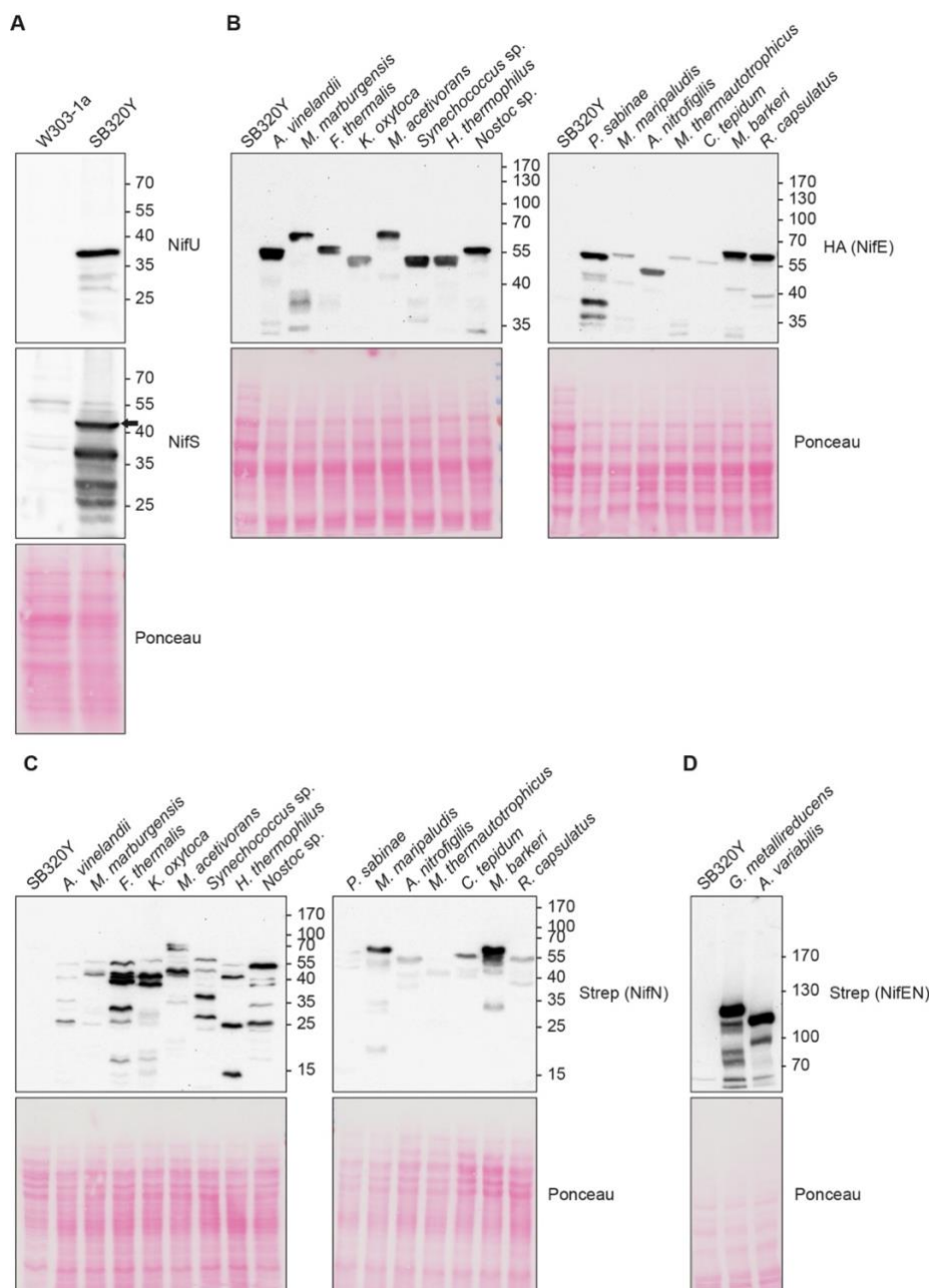


Figure 18. Expression of NifEN variants in *S. cerevisiae*. (A) Immunoblot analysis of total protein yeast extracts showing expression of NifU^{Av} and NifS^{Av} in yeast strain SB320Y (parental strain for the NifEN library). The black arrow indicates full-length NifS (51,3 kDa), where the faster migrating polypeptides likely represent NifS degradation products. (B, C, D) Accumulation of HA-NifE (B), NifN-Strep (C), and TS-NifEN (D) in total protein yeast extracts. Molecular weight markers (kDa) and antibodies used are indicated.

S. cerevisiae cells expressing NifU^{Av} and NifS^{Av} were transformed with either two plasmids containing *su9-ha-nifE* and *su9-nifN-ss*, or with a single plasmid for expression of *su9-ts-nifEN* (NifEN fusion protein). The majority of the NifE and NifN proteins could

be detected in total protein extracts (Figure 18B-D). However, many variants generated polypeptides with different migration patterns.

For assessing solubility and NifEN complex formation, purification and co-purification experiments were performed. For this, cell-free protein extracts were prepared under anaerobic conditions as described in section 3.8.1 and then incubated with streptavidin-coated magnetic beads for NifN or NifEN fusion purification, anti-HA magnetic beads for NifE purification (Figure 19).

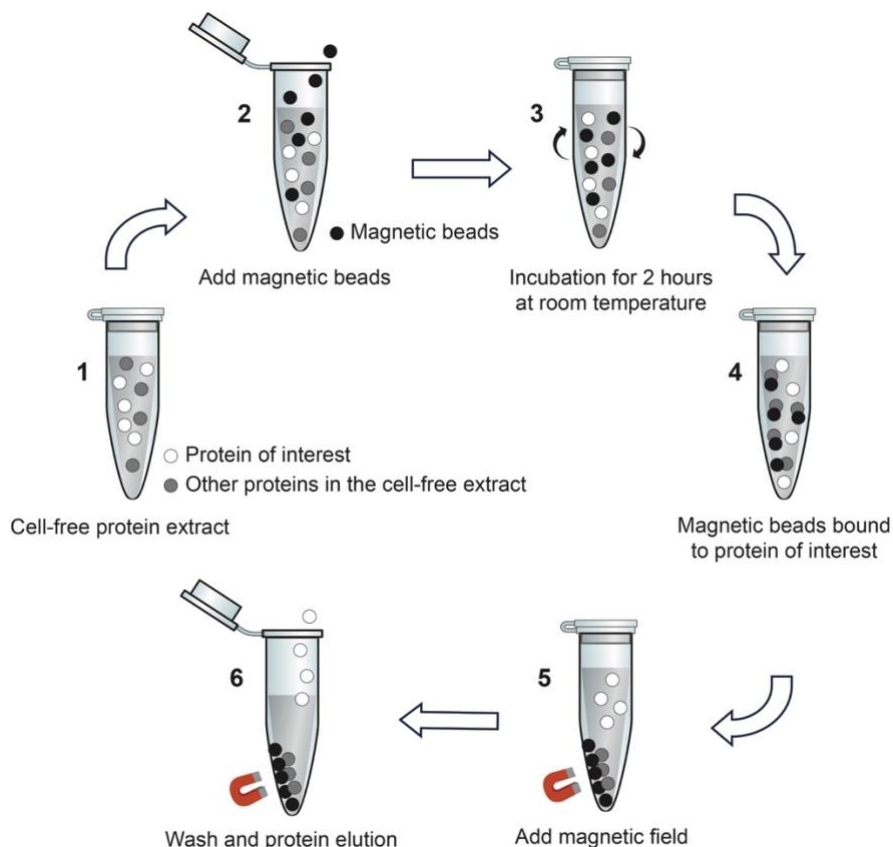


Figure 19. The workflow for small-scale NifEN purification from cell-free extracts using magnetic beads. 1) The cell-free protein extract (obtained after lysis) contains the proteins of interest along with other proteins produced by the cells. 2) A suitable amount of magnetic beads is added to the cell-free protein extract. 3) The mixture of cell-free protein extract and magnetic beads is incubated for 2h at room temperature with gentle shaking. 4) Proteins of interest have bound to the magnetic beads. 5) The application of a magnetic field enables the separation of the magnetic beads, carrying the bound proteins of interest, from the remaining cell-free protein extract. 6) The bound proteins are washed and eluted with appropriate buffers to obtain the purified protein of interest.

NifEN complex formation was investigated by immunoblot detection of NifN co-eluting with purified NifE, or NifE co-eluting with purified NifN. The results of the co-purification experiments showed that NifEN complexes from *Synechococcus* sp. JA-2-3B'a, *H. thermophilus*, and *G. metallireducens* could consistently be detected in soluble protein extracts (Figure 20).

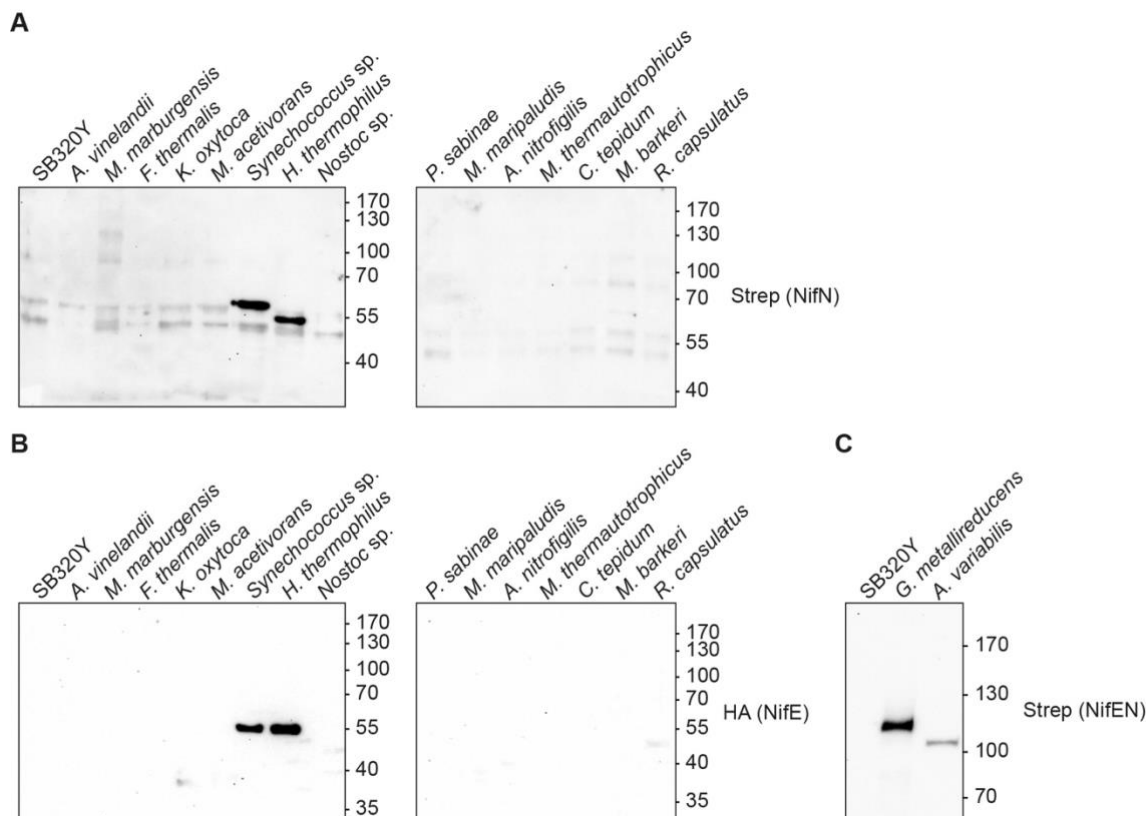


Figure 20. Co-purification and purification experiments to identify soluble NifEN variants accumulating in the mitochondria of aerobically cultured *S. cerevisiae*. (A, B, C) Immunoblot analysis to detect NifN-SS co-purifying with HA-NifE (A), HA-NifE co-purifying with NifN-SS (B), and Strep-tag purified TS-NifEN (C). The presented immunoblots are representative of three independent experiments. Molecular weight markers (kDa) and antibodies used are indicated.

For simplicity, throughout the following chapters, selected NifEN variants expressed in *S. cerevisiae*, *A. vinelandii*, and *E. coli* will be labeled preceding NifEN with either Sc, Av or Ec, respectively, and followed by initial letters of the species name NifEN^{Sc} (*Synechococcus* sp. JA-2-3B'a), NifEN^{Ht} (*H. thermophilus*) and NifEN^{Gm} (*G. metallireducens*).

4.1.5 Analysis of NifEN protein structures

Based on primary sequence analysis, it was observed that NifE and NifN variants with higher primary sequence similarity to *A. vinelandii* NifEN were capable of rescuing diazotrophic growth *in vivo* (Figure 21A). Specifically, candidates with an identity over 40% to NifE^{Av} and 35% to NifN^{Av} were functional. However, there was one exception to this. The *nifEN* genes from *K. oxytoca* did not restore the Nif⁺ phenotype in DJ2898, despite the high similarity of the *K. oxytoca* NifEN protein to *A. vinelandii* NifEN (about 65% and 45% for NifE and NifN, respectively).

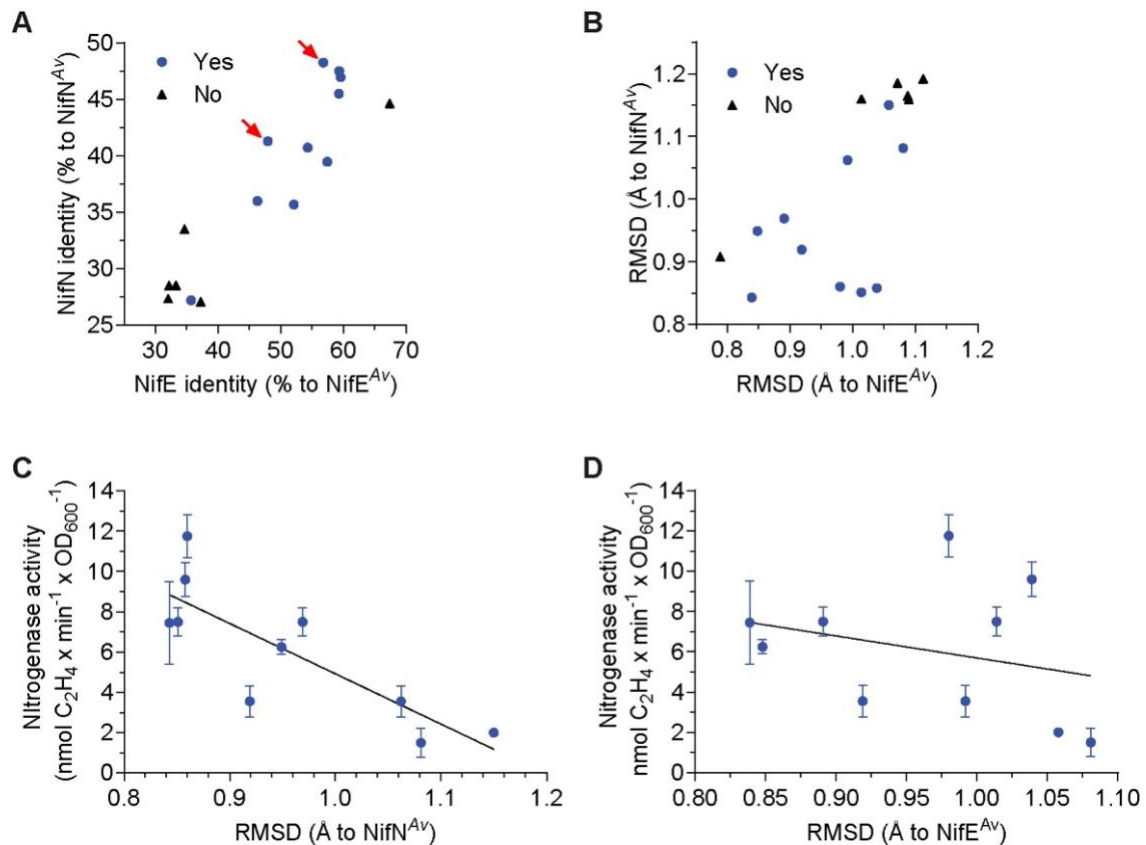


Figure 21. Sequence and structural divergence analysis of NifEN proteins in the library. (A) Identity of NifE and NifN primary sequences to the respective *A. vinelandii* sequence and their capacity to restore diazotrophic growth in the *A. vinelandii* mutant strain DJ2898. The NifE and NifN sequences were analyzed using Clustal Omega. Blue dots represent NifEN proteins that restored diazotrophic growth, black triangles symbolize those variants that did not restore diazotrophic growth. The two natural NifEN fusion proteins are highlighted with red arrows. (B) Structural alignments of the NifE and NifN subunits of the NifEN library against the *A. vinelandii* NifE and NifN. The root-mean-square deviation (RMSD) between the predicted structures obtained using AlphaFold, with NifE and NifN in the *A. vinelandii* NifEN are plotted. Blue dots indicate NifEN variants that restored diazotrophic growth in DJ2898 while black triangles indicate those that did not. (C) The RMSD between the predicted NifN structures and the *A. vinelandii* NifN in *A. vinelandii* NifEN (seen in panel B) plotted against the measured nitrogenase activity in the strains expressing the corresponding NifEN (Figure 17). The line shows linear regression analysis ($R^2 = 0.64$, $p < 0.0001$). (D) The RMSD between the predicted NifE structures and NifE in *A. vinelandii* NifEN (seen in panel B) plotted against the measured nitrogenase activity in the strains expressing the corresponding NifEN protein (Figure 17). The line indicates linear regression analysis ($R^2 = 0.08$, $p = 0.2350$).

The structures of the NifE and NifN subunits, as predicted by AlphaFold, were compared to those of the crystal structure of *A. vinelandii* NifEN. Among the NifEN variants, those that could effectively replace *A. vinelandii* NifEN *in vivo* displayed the highest similarity to *A. vinelandii* NifN ($p = 0.008$) (Figure 21B).

Additionally, a strong correlation was observed between the level of similarity to NifN^{Av} among the functional NifEN variants and *in vivo* nitrogenase activity ($p < 0.0001$,

Figure 21C). Interestingly, no such correlation was found when analyzing the NifE structures ($p = 0.235$, Figure 21D).

Based on the previous results that demonstrated the successful selection of three NifEN variants capable of forming NifEN protein complexes when expressed in yeast mitochondria (section 4.1.4), further biochemical and functionality analyses were conducted on NifEN^{SS}, NifEN^{Ht}, and NifEN^{Gm} expressed in *A. vinelandii*. To ensure high Fe-S cluster content in the NifEN protein, *ss-nifEN* genes were integrated into the *A. vinelandii* strain DJ under the control of the *nifH* promoter by replacing *nifHDKTYnafABnifEN* with the *nifEN* genes. Because NifH is required for the conversion of NifB-co into FeMo-co once it is bound to NifEN (Rangaraj et al., 1999, Hu et al., 2005), the various NifEN scaffold proteins isolated from this genetic background were expected to produce NifEN proteins with bound VK-cluster.

All three NifEN variants were isolated using STAC, as described in section 3.8.6. Notably, once purified, the NifEN proteins exhibited a dark color, which is often indicative of the presence of Fe-S clusters (Figure 22B). This observation suggested that the purified NifEN proteins retain their native Fe-S clusters and likely harbor VK-clusters.

To characterize the purified NifEN proteins and confirm their identities, mass spectrometry analysis was performed. This analysis aimed to identify the polypeptides that co-eluted with the NifEN proteins during the purification process. While conducting the mass spectrometry analysis it was observed that certain polypeptides, identified as AvNifEN^{SS}, exhibited lower abundance and faster migration patterns. These AvNifEN^{SS} polypeptides were likely derived from partial degradation of the expressed NifEN^{SS} in *A. vinelandii*. Two common contaminants in STAC purifications from *A. vinelandii* cell extracts (PycB and AccB) were also detected in the eluates of the three NifEN variants (Figure 22A).

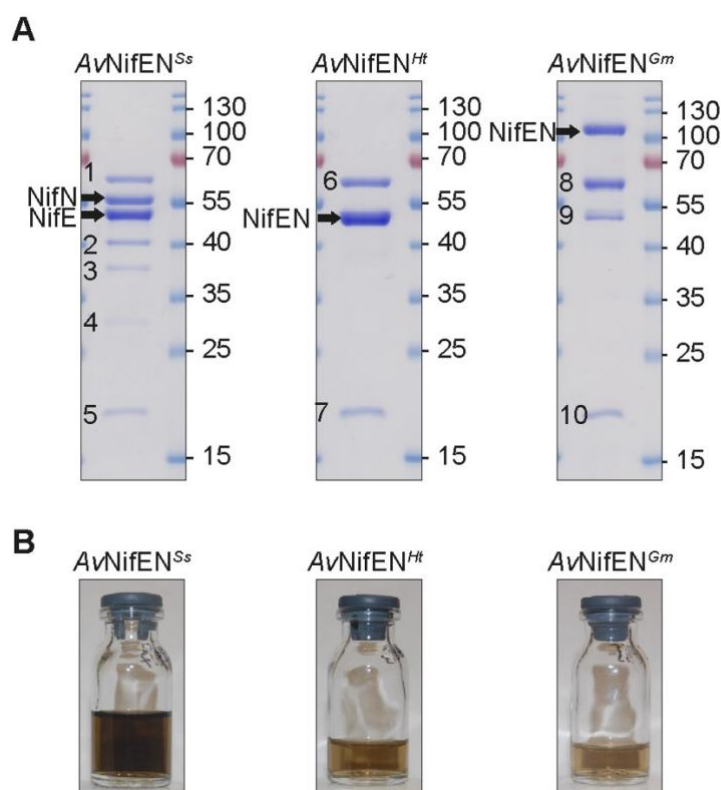


Figure 22. Isolation of NifEN variants from *A. vinelandii*. (A) Coomassie staining of AvNifEN^{Ss}, AvNifEN^{Ht}, and AvNifEN^{Gm} isolated by STAC. Polypeptides numbered 1-10 were identified by peptide mass fingerprinting (1, 6, 8 and 9, pyruvate carboxylase subunit B (PycB, UniProt C1DH60); 2, 3 and 4, degradation products of AvNifN^{Ss}; 5, 7 and 10, biotin carboxyl carrier protein of acetyl-CoA carboxylase (AccB, UniProt C1DLJ8)). Intact AvNifE^{Ht} and AvNifH^{Ht} migrate as a single band due to their similar size. Molecular weight markers (kDa) are indicated for each panel. (B) Appearance of AvNifEN^{Ss}, AvNifEN^{Ht}, and AvNifEN^{Gm} isolated from *A. vinelandii*. Note that the differences in color between the samples may reflect differences in protein concentration rather than cluster occupancy/composition.

Considering that NifEN heterotetramer is equipped with two permanent [4Fe-4S] clusters and two [8Fe-9S-C] NifB-co, the ideal Fe content of a fully loaded protein is 24 iron atoms per protein. The iron content of the purified NifEN proteins ranged from 21 to 28 atoms per protein, as shown in Table 17. The Fe composition of the isolated NifEN proteins highlights their potential functionality in the *in vitro* FeMo-co synthesis pathway.

Table 17. Fe atom per NifEN protein and yield of NifEN from *A. vinelandii* cells.

Protein variant	Yield (mg NifEN per 100 g cells)	Fe atoms per NifEN protein
AvNifEN ^{Ss}	20	26
AvNifEN ^{Ht}	6,7	28
AvNifEN ^{Gm}	5,8	21

* Data obtained from one protein purification (n=1).

4.1.6 *In vitro* FeMo-co synthesis using NifEN protein variants isolated from *A. vinelandii*

The three NifEN variants isolated from *A. vinelandii* were subsequently tested for their ability to support *in vitro* FeMo-co formation in a reaction containing Nif components isolated from *A. vinelandii*. This ability was tested by performing NifEN-dependent FeMo-co synthesis assays with as-isolated AvNifEN^{Xx}, NifH, homocitrate, and Mo.

Consistent with their high Fe content, all three NifEN variants could activate apo-NifDK without further addition of free Fe or Fe-S clusters (Figure 23), confirming that the NifEN proteins were isolated from *A. vinelandii* with bound NifB-co (or VK-cluster).

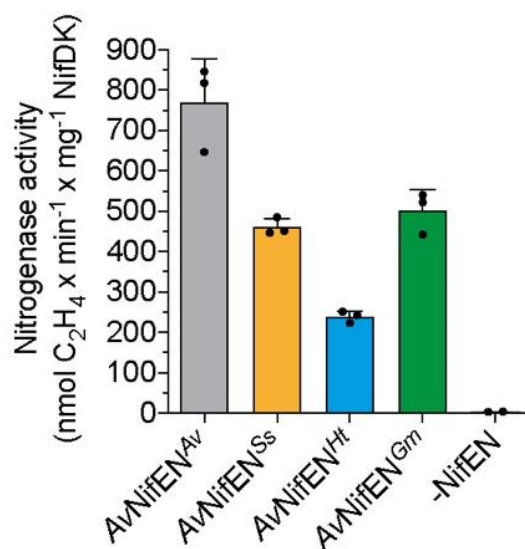


Figure 23. NifEN-dependent FeMo-co synthesis and apo-NifDK activation using AvNifEN^{Ss}, AvNifEN^{Ht}, and AvNifEN^{Gm}. Activity units are nanomoles of ethylene produced per minute and milligram of reconstituted NifDK. Error bars indicate mean $\pm$ SD; $n = 3$ technical replicates ($n = 2$ for the -NifEN control reaction) shown by dots.

4.1.7 Cross-compatibility of selected NifEN variants with NifH proteins isolated from yeast mitochondria

4.1.7.1 Purification of EcNifEN^{Ss}, EcNifEN^{Ht}, and EcNifEN^{Gm} variants from *E. coli* cells

To obtain enough protein for functional analysis, the three NifEN variants identified to form soluble protein complexes in yeast mitochondria were expressed in *E. coli*. For this reason, yeast codon-optimized *ha-nifE* and *nifN-ss* or *ts-nifEN* constructs were therefore transferred to bacterial pETDuet-1 vector designed to co-express the two *nif* genes, each under the control of separate T7 promoter (Table 12). *E. coli* cells were additionally co-expressing with NifU and NifS proteins from *A. vinelandii*, to obtain NifEN proteins harboring the permanent [4Fe-4S] clusters.

STAC was used to purify the three NifEN proteins, and the purification process was monitored by taking samples of the total extract (TE), cell-free extract (CFE), flow-through (FT), wash (W), elution (E), and final protein fraction after desalting and concentration (E final) for analysis by SDS-PAGE and immunoblotting (Figure 24). SDS-PAGE and Coomassie staining of the isolated NifEN^{Ht} migrated as a single band due to the similar size of both subunits, while NifEN^{Gm} showed an approximately 103 kDa band that corresponded to the fused NifE-N subunit.

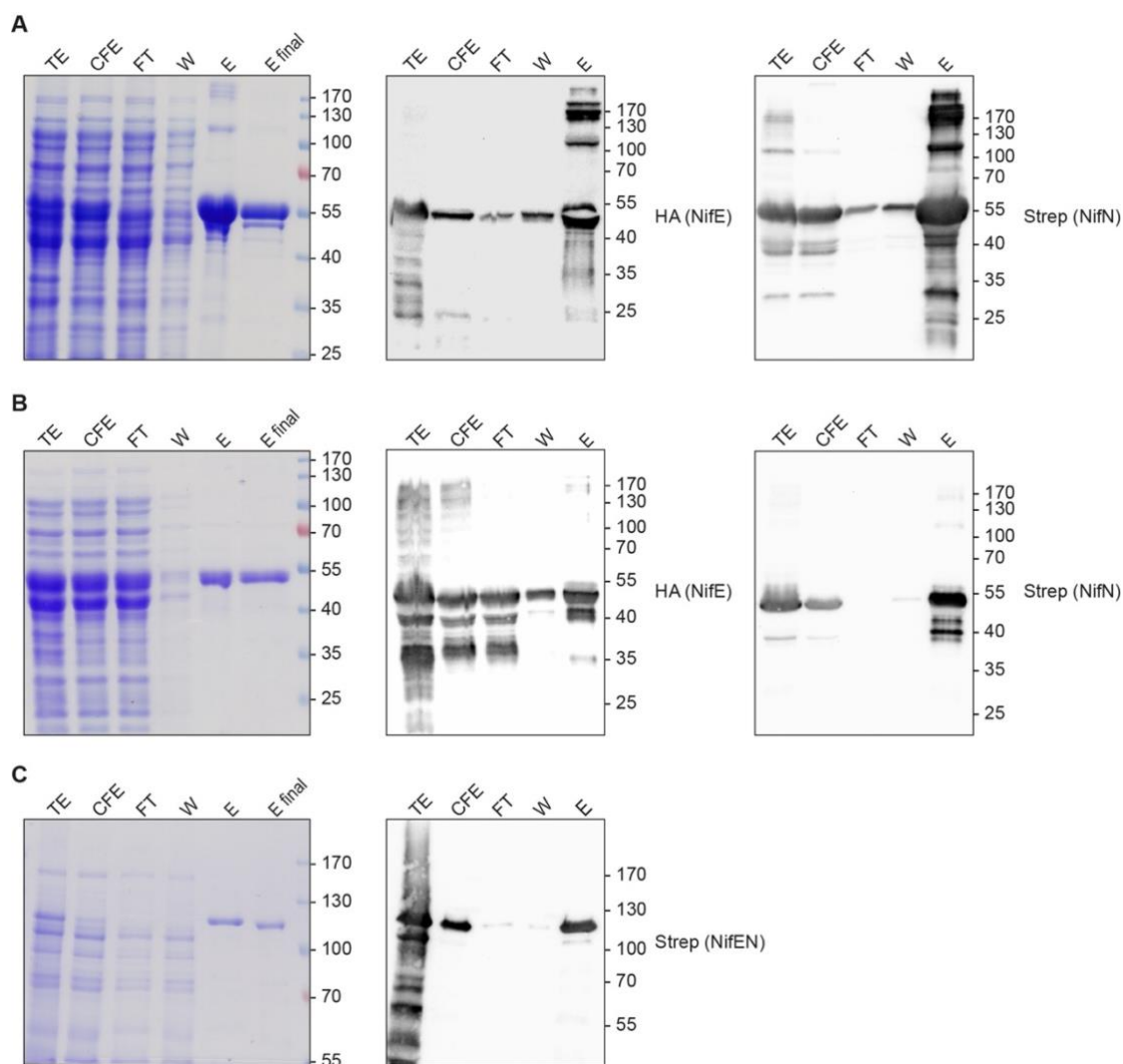


Figure 24. STAC purification of NifEN variants expressed in *E. coli*. *EcNifEN^{Ss}* (A), *EcNifEN^{Ht}* (B), and *EcNifEN^{Gm}* (C) were isolated from *E. coli* cells co-expressing NifU^{Av} and NifS^{Av}. Fractions were analyzed by SDS-PAGE followed by Coomassie staining or immunoblot analysis. TE = total extract, CFE = soluble cell-free extract, FT = flow-through fraction, W = wash fraction, E = elution fraction, E final = final protein fraction after desalting and concentration. Molecular weight markers (kDa) and antibodies used are indicated.

The yields of the three NifEN proteins were relatively high as expected, which facilitated an in-depth analysis. Iron (Fe) quantification of purified samples indicated that NifEN variants were isolated as apo-proteins containing the two permanent [4Fe-4S] clusters (Table 18).

Table 18. Fe atoms per protein and yield of NifEN from *E. coli* cells.

Protein variant	Yield (mg NifEN per 100 g cells)	Fe atoms per NifEN protein	Fe-cluster occupancy of NifEN^{Av} from native host
<i>EcNifEN^{Ss}</i>	60	5,6	8
<i>EcNifEN^{Ht}</i>	152	3,0	8
<i>EcNifEN^{Gm}</i>	30	4,7	8

* Data obtained from one protein purification (n=1).

4.1.7.2 *In vitro* FeMo-co biosynthesis combined distinct NifH variants with three NifEN variants isolated from *E. coli*.

NifEN proteins isolated from *E. coli* were used to test their compatibility with eight distinct NifH variants (denoted here ScNifH^{Xx}) previously identified as soluble proteins accumulating in yeast mitochondria (Paya-Tormo et al., 2022). In this study, the screening of 35 mitochondria-targeted NifH proteins generated eight NifH variants that could be successfully isolated from yeast extracts. Interestingly, the purification yield for selected variants was significantly higher than that of NifH^{Av}, which previously was also isolated from yeast. These findings suggest that these diverse NifH proteins are better suited for FeMo-co biosynthesis engineering.

The functionality of the purified *EcNifEN^{Xx}* proteins was assessed by NifB-co-dependent *in vitro* FeMo-co synthesis assays and the insertion of the synthesized FeMo-co into apo-NifDK containing fully matured P-cluster, as described in section 3.10.2.2. The activation of NifDK was measured by the standard acetylene reduction assay. In this reaction, FeMo-co precursor (isolated NifB-co bound to the cluster-transfer protein NifX) was delivered to NifEN protein that converts NifB-co into FeMo-co by incorporating Mo and homocitrate. Importantly, this conversion depends on transient NifH-NifEN association, as the interaction of those two proteins was suggested to be important for electron transfer and insertion of Mo in FeMo-co synthesis (Hernandez et al., 2008). Moreover, to be able to better compare the different NifEN activities, *EcNifEN^{Xx}* variant proteins were added to the reactions based on Fe concentration instead of protein concentration (as the isolated NifEN proteins had incomplete FeS cluster occupancy), see Table 18.

Importantly, all three as-isolated *EcNifEN^{Xx}* proteins proved to be active when combined with its corresponding best NifH variant (Figure 25). The results demonstrated that the ability of different NifH variants to promote FeMo-co biosynthesis exhibited significant variation across different NifEN sources. Maximum activities were obtained when *Roseiflexus* sp. ScNifH^{Rs} was combined with *EcNifEN^{Ss}*, when *Dehalococcoides*

ethenogenes ScNifH^{De} was combined with EcNifEN^{Ht}, and when *G. sulfurreducens* ScNifH^{Gs} was combined with EcNifEN^{Gm}. Importantly, the maximum FeMo-co assembly activities obtained using EcNifEN^{Ht} and EcNifEN^{Gm} approached the activity seen using native NifEN and NifH proteins isolated from *A. vinelandii*.

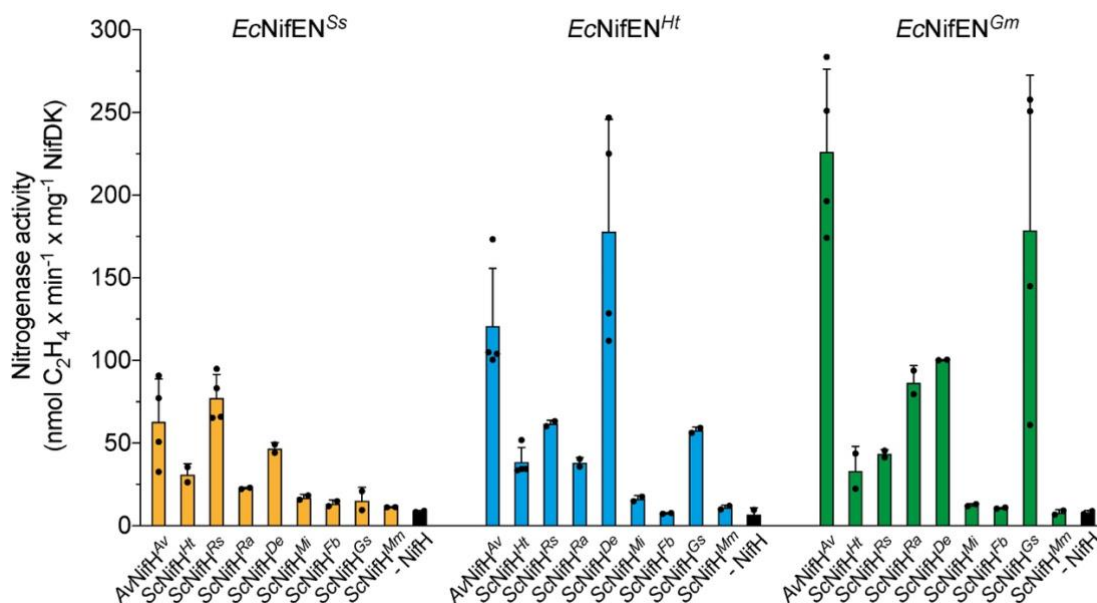


Figure 25. Compatibility of NifEN candidates with distinct NifH variants. NifB-co-dependent FeMo-co synthesis and apo-NifDK activation using EcNifEN^{Ss}, EcNifEN^{Ht}, or EcNifEN^{Gm} variants combined with distinct ScNifH^{Xx} isolated from yeast. NifH^{Av} isolated from *A. vinelandii* was used during the acetylene reduction step to eliminate the effect of different compatibilities between NifH variants and NifDK^{Av}. The activity using apo-NifEN^{Av} and NifH^{Av} isolated from *A. vinelandii* strains DJ1118 and DJ, respectively, was 480 ± 110 units. Error bars indicate mean $\pm$ SD; $n \geq 2$ technical replicates (shown by dots). Av, *A. vinelandii*; Ht, *H. thermophilus*; Rs, *Roseiflexus* sp.; Ra, *R. albus*; De, *D. ethenogenes*; Mi, *M. infernus*; Fb, *Firmicutes bacterium*; Gs, *G. sulfurreducens*; Mm, *M. marburgensis*.

4.1.8 FeMo-co biosynthesis using NifEN expressed in *S. cerevisiae*

4.1.8.1 Purification of ScNifEN^{Ss}, ScNifEN^{Ht}, and ScNifEN^{Gm}

Based on the promising results of the cross-compatibility experiments showing that NifEN variant proteins could produce FeMo-co and activate apo-NifDK *in vitro* when combined with distinct NifH variants, these three NifEN candidates were then isolated from *S. cerevisiae*. To achieve this, the *S. cerevisiae* strains were grown in fermenters under aerobic conditions, and cultured cells were lysed and used in the STAC purification process. Notably, the purification experiment was performed twice using a new batch of grown cells each time and conducted in anaerobic chambers. To obtain the protein extract, the cells were lysed using a high-pressure homogenizer operating at 20,000 psi, effectively disrupting the yeast cell wall. Furthermore, the cell-free extract obtained via

centrifugation was loaded onto Strep-affinity columns, and the proteins were eluted with a biotin-supplemented buffer. The concentrated and desalted proteins were sampled for immunoblot analysis (Figure 26), and stored in LN₂ to subsequently evaluate their functionality in *in vitro* FeMo-co biosynthesis assays.

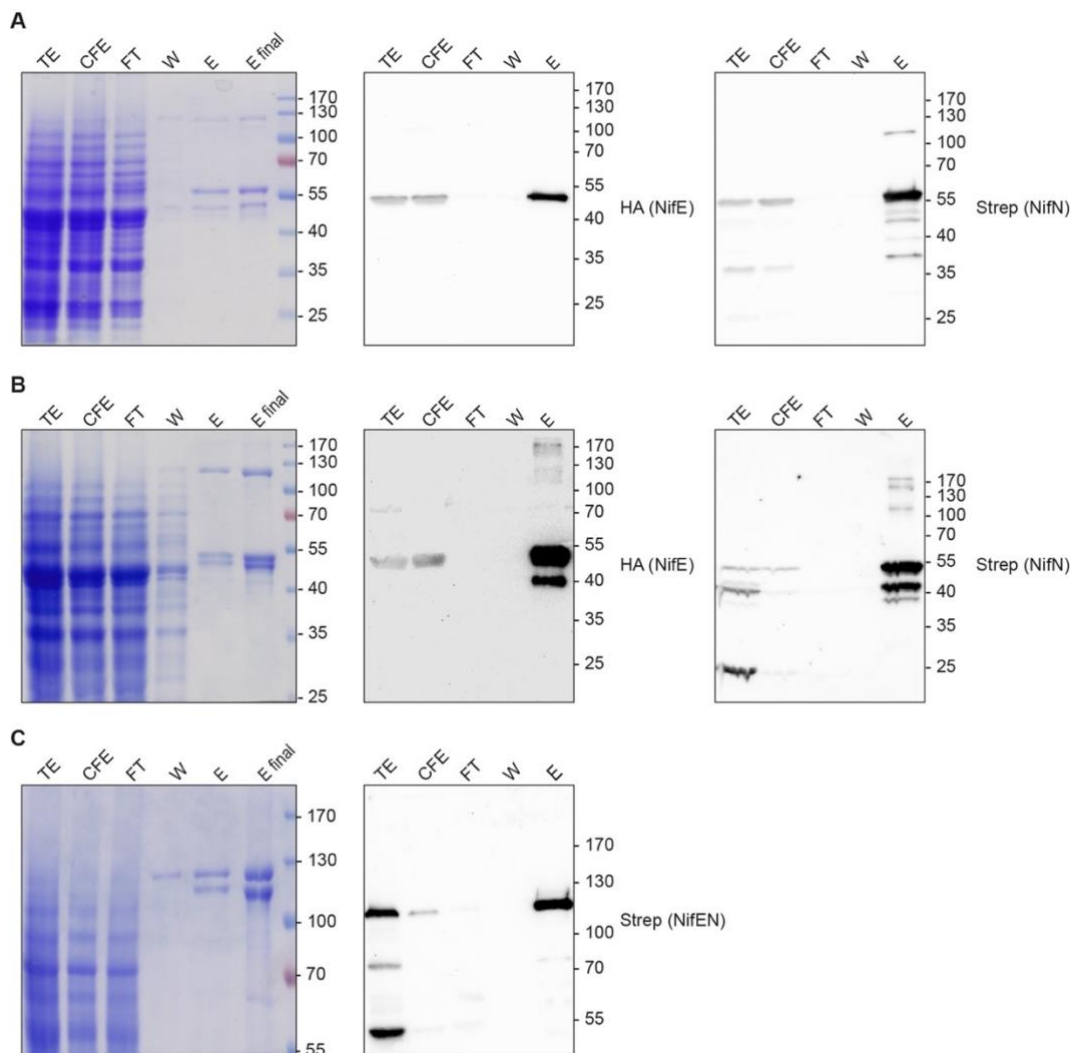


Figure 26. STAC purification of NifEN variants expressed in aerobic cultures of *S. cerevisiae*. ScNifEN^{Ss} (A), ScNifEN^{Ht} (B), and ScNifEN^{Gm} (C) were isolated from yeast cells co-expressing NifU^{Av} and NifS^{Av}. Fractions were analyzed by SDS-PAGE followed by Coomassie staining and immunoblot analysis. TE = total extract, CFE = soluble cell-free extract, FT = flow-through fraction, W = wash fraction, E = elution fraction, E final = final protein fraction after desalting and concentration. Molecular weight markers (kDa) and antibodies used are indicated.

SDS-PAGE analysis of as-isolated NifEN variants showed the additional migrating co-eluting proteins. Except for low-abundant and faster migrating NifEN polypeptides, likely resulting from partial degradation of this protein when expressed in mitochondria of yeast, the other protein migrating as a band of 120 kDa corresponded to pyruvate carboxylase protein. This protein contains biotin as a prosthetic group, which

explains its affinity to the Strep-Tactin resin. As the elution fractions contained additional proteins, the concentrations of isolated NifEN proteins were estimated from SDS-PAGE and Coomassie staining gels, where final purification samples were loaded and NifEN proteins were quantified by densitometric analysis using ImageJ.

As expected, the yields of the three yeast-isolated NifEN proteins were lower than when they were purified from *E. coli* (Table 19). As a comparison, the amount of Strep-tagged NifEN^{Av} purified from *A. vinelandii* cells was about 20 mg per 100 g of cells.

Table 19. The yield of NifEN from *S. cerevisiae* cells.

Protein variant	Yield (mg NifEN per 100 g cells)
ScNifEN ^{Ss}	0.4
ScNifEN ^{Ht}	0.9
ScNifEN ^{Gm}	0.7

* Data is the average of two independent protein purifications (n=2).

4.1.8.2 Acetylene reduction assays using NifEN variants isolated from *S. cerevisiae*

Once ScNifEN^{Ss}, ScNifEN^{Ht}, and ScNifEN^{Gm} were isolated from the mitochondria of aerobically cultured *S. cerevisiae*, the purified proteins were first used to evaluate their ability to support *in vitro* NifB-co-dependent FeMo-co biosynthesis. The Fe quantification measurements have shown that the isolated ScNifEN^{xx} proteins exhibited a low Fe content. Additionally, the yield of NifEN-isolated variants was also low (Table 19). Consequently, the [4Fe-4S] cluster occupancy could not be accurately quantified, making it impossible to use Fe concentration to assess NifEN activity. Therefore, the activity assays were based on the concentration of isolated protein rather than Fe concentration.

These assays employed NifX with bound NifB-co, isolated from *K. oxytoca*. Each NifEN protein was combined with its corresponding optimum NifH variant. To evaluate the activation of apo-NifDK, ScNifH^{Rs} was combined with ScNifEN^{Ss}, ScNifH^{De} was combined with ScNifEN^{Ht}, and ScNifH^{Gs} was combined with ScNifEN^{Gm}. As demonstrated in Figure 27, substantial FeMo-co synthesis and apo-NifDK activation was observed when ScNifH^{Gs} was combined with ScNifEN^{Gm}, while lower levels were seen when combining ScNifH^{De} and ScNifEN^{Ht}. No activation above background was observed when ScNifH^{Rs} was combined with ScNifEN^{Ss}, suggesting either poor [4Fe-4S] cluster insertion and/or instability of this NifEN variant in the mitochondrial environment. Therefore, ScNifEN^{Ss} was excluded for further analysis, and later FeMo-co biosynthesis assays were performed only with the active proteins ScNifEN^{Ht} and ScNifEN^{Gm}.

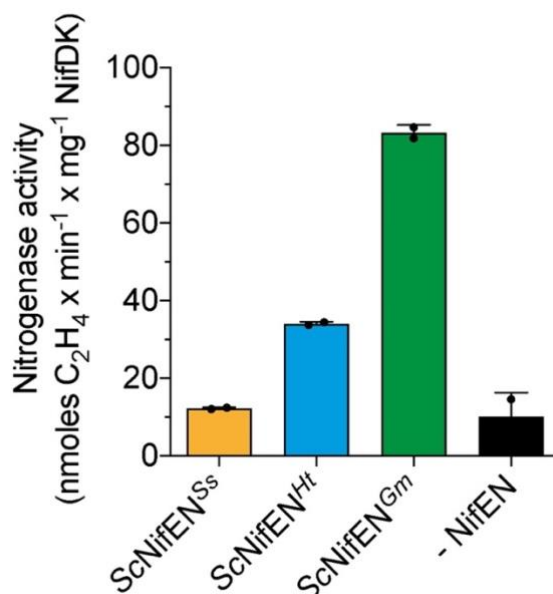


Figure 27. FeMo-co biosynthesis using NifEN isolated from *S. cerevisiae*. NifB-co-dependent FeMo-co synthesis and apo-NifDK activation using NifEN and NifH proteins isolated from yeast (ScNifEN^{Ss} combined with ScNifH^{Rs}, ScNifEN^{Ht} combined with ScNifH^{De}, and ScNifEN^{Gm} combined with ScNifH^{Gs}). The activity using apo-NifEN^{Av} and NifH^{Av} isolated from *A. vinelandii* strains DJ1118 and DJ, respectively, was 380 ± 53 units. Error bars indicate mean $\pm$ SD; $n = 2$ biological replicates (≥ 4 technical replicates each).

4.1.8.3 FeMo-co biosynthesis pathway using proteins expressed in *S. cerevisiae*

As two active forms of NifEN proteins could be isolated from the mitochondria of yeast, the functionality of these NifEN variants was then assessed by the NifB-dependent *in vitro* FeMo-co synthesis and activation of apo-NifDK protein assay. In addition to NifEN and NifH, this assay requires the NifB-co synthesis protein NifB combined with S-adenosylmethionine and holo-NifU (which delivers two [4Fe-4S] units that are the substrates clusters for NifB-co formation). The main purpose of this experiment was to evaluate the *in vitro* FeMo-co biosynthesis pathway by using phylogenetically diverse NifB, NifEN, and NifH proteins, all produced in the *S. cerevisiae* mitochondria.

The assay was performed using apo-NifDK purified from *A. vinelandii* and *A. vinelandii* NifU expressed in the heterologous host *E. coli*. NifU was previously reconstituted for [4Fe-4S] clusters using cysteine, Fe, DTT, and NifS to obtain holo-NifU. Two variants of NifB, from *M. infernus* (NifB^{Mi}) and *M. thermautotrophicus* (NifB^{Mt}), were selected. It was previously described that these NifB variants can be produced in active form when targeted to the mitochondria of yeast (Burén et al., 2019). Using the same approach described in the previous FeMo-co synthesis assay, ScNifH^{De} was combined with ScNifEN^{Ht}, and ScNifH^{Gs} was combined with ScNifEN^{Gm}. Importantly, both ScNifEN^{Ht} and ScNifEN^{Gm} could activate apo-NifDK when combined with their corresponding optimum NifH partner and either of the two NifB variants (Figure 28).

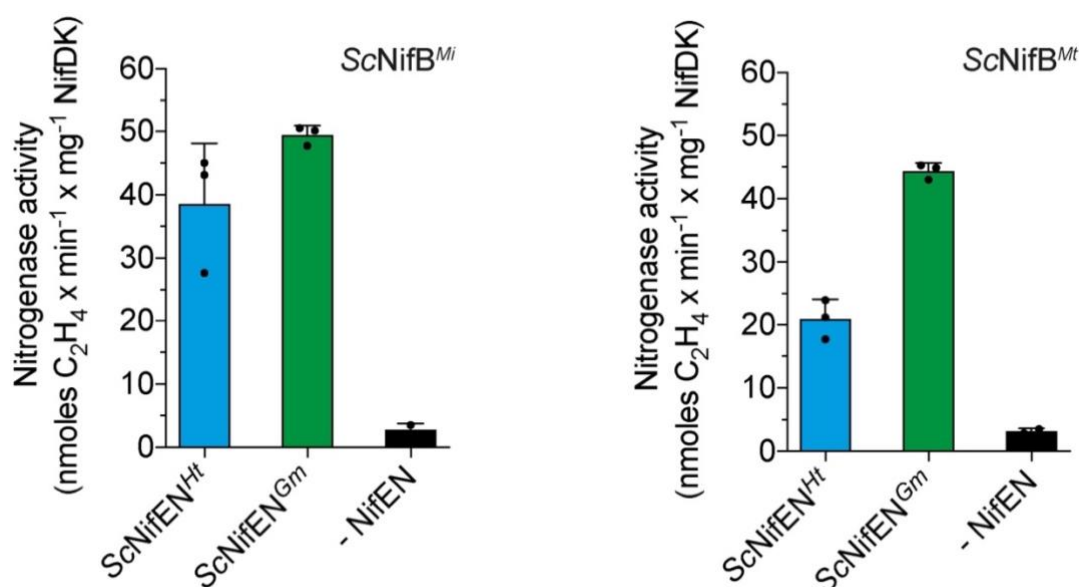


Figure 28. NifB-dependent FeMo-co synthesis and apo-NifDK activation using NifB, NifEN, and NifH proteins isolated from yeast. ScNifEN^{Ht} combined with ScNifH^{De} and ScNifEN^{Gm} with ScNifH^{Gs} were tested with two different NifB variants, *M. infernus* (ScNifB^{Mi}) and *M. thermautotrophicus* (ScNifB^{Mt}). The activity using apo-NifEN^{Av} and NifH^{Av} isolated from *A. vinelandii* strains DJ1118 and DJ, respectively, was 873 ± 124 units (using ScNifB^{Mi}) and 653 ± 91 units (using ScNifB^{Mt}). Error bars indicate mean $\pm$ SD; $n = 3$ technical replicates ($n = 2$ for - NifEN, shown by dots). All activities are represented as nanomoles of ethylene produced per minute and milligram of NifDK.

5. DISCUSSION

Challenges of engineering FeMo-co biosynthesis into eukaryotic cells

In diazotrophic organisms, the presence of *nifEN* genes, along with *nifB* and the structural genes *nifHDK*, constitutes what is considered to be the minimal gene set required for having a functional Mo-nitrogenase (Dos Santos et al., 2012). It is therefore essential to identify variants of these genes that produce functional proteins in eukaryotic organisms, and specifically in plants, when trying to generate active nitrogenase in crops. Among these six genes, *nifE*, *nifN*, and *nifB* encode proteins that are required for the biosynthesis of FeMo-co, the metallocluster located at the active site of the Mo-nitrogenase.

FeMo-co is arguably the most complex metallocluster known in Nature. The biosynthesis of FeMo-co takes place at the scaffold protein NifEN. In short, NifEN accepts a precursor cluster known as NifB-co, which is produced by NifB, and incorporates the organic moiety homocitrate and a Mo atom. The reactions happening at NifEN convert the [8Fe-9S-C] FeMo-co precursor from NifB into the mature [7Fe-9S-C-Mo-R-homocitrate] FeMo-co. In addition to NifEN, the reaction(s) at NifEN additionally requires a functional interaction with NifH.

Numerous obstacles confront researchers who aim to develop crop plants with N₂-fixing abilities, with one of the most prominent being the susceptibility of nitrogenase metalloclusters to O₂. Aside from managing the O₂-generating reactions within the plant cell, nitrogenase necessitates certain heterometals (Mo and V) to construct the Mo and V-nitrogenase active-site cofactors. Lastly, N₂-fixation places a significant demand on energy (ATP) and reducing power (Curatti and Rubio, 2014). Previous studies have utilized the strategy of safeguarding O₂-sensitive constituents through organellar targeting. This approach involves the targeting of Nif proteins to the mitochondria of actively respiring cells, which presumably would provide adequate levels of ATP and reducing power, and at the same time overcome the challenge of O₂. Employing this strategy, an active form of NifH, the most O₂-sensitive nitrogenase catalytic component, has been produced in the mitochondria of *S. cerevisiae* and *N. benthamiana* (López-Torrejón et al., 2016, Paya-Tormo et al., 2022, Jiang et al., 2021). Additionally, it has been shown that soluble NifB can be produced in the mitochondria of *S. cerevisiae*, *N. benthamiana*, and *O. sativa* (Burén et al., 2019, Jiang et al., 2022, He et al., 2022). By co-expressing NifU, NifS, and FdxN with NifB, accumulation of NifB-co was demonstrated in the mitochondria of yeast (Burén et al., 2019). While these results suggest that the initial steps of the FeMo-co biosynthetic pathway could be completed in yeast mitochondria, achieving the complete pathway necessitates the expression of an active NifEN

protein. Unfortunately, the first studies showed that NifEN originating from the most common diazotrophic model organisms *A. vinelandii* and *K. oxytoca* did not express well in the mitochondria of *S. cerevisiae* and *N. benthamiana* (Okada et al., 2020a, Buren et al., 2017). For this reason, the work reported in this PhD thesis aimed at identifying a better NifEN candidate for expression in the mitochondria of yeast.

Exploring NifEN variants for expression in yeast mitochondria

The first step of this study involved identifying NifEN variants that accumulated as soluble and stable proteins within the mitochondria of aerobically respiring yeast. Only soluble and stable proteins will form functional complexes, and only functional complexes will perform FeMo-co synthesis. Previous studies have shown that NifEN from *Paenibacillus polymyxa* WLY78 could be expressed in the cytosol of the yeast cell. Immunoblot analysis of yeast protein extracts confirmed that both NifE and NifN subunits could be detected (Liu et al., 2019), but the work did not investigate whether NifEN tetramer was formed, or whether this NifEN was functional. Indeed, the exclusive localization of this NifEN protein in the cytosol of *S. cerevisiae* makes it unlikely that this protein would contain intact Fe-S clusters and be active. Moreover, attempts to express NifEN from *K. oxytoca* in *N. benthamiana* demonstrated that the NifE polypeptide only accumulated in the insoluble fraction, and no soluble NifEN complex was reported (Okada et al., 2020a). In addition, attempts in our laboratory to express NifEN from *A. vinelandii* in the yeast mitochondria generated a protein that migrated as a single band when analyzed by SDS-PAGE (Buren et al., 2017). Using polyclonal antibodies against NifEN raised doubts about whether the detected band represented a fully assembled complex or a single subunit, casting uncertainty on the expression and stability of both subunits.

Taking into consideration the results from these previous studies, we decided to generate a library consisting of 17 *nifEN* genes originating not only from bacteria but also from archaeal organisms. It has been hypothesized that Nif proteins from archaea might be better suited for expression in an eukaryotic environment because of the closer evolutionary relationship between Archaea and Eukaryota (Koonin, 2015). Moreover, the *nifEN* library was biased towards genes from thermophilic organisms as these proteins could have the ability to withstand the elevated temperatures inside the mitochondria (Chrétien et al., 2018). We also included genes producing naturally occurring NifEN fusion proteins as we hypothesized that these genes could facilitate the formation of NifEN “heterotetramer” complexes. This hypothesis originated from work by Allen and Gregg and colleagues, that reported that fusing NifD to NifK using a polypeptide linker improved the association of NifD with NifK. They showed that the NifDK fusion not only enabled the formation of soluble NifDK complex, but also that

the complex was more resistant to degradation in the *N. benthamiana* mitochondrial matrix (Allen et al., 2020).

The compatibility of NifEN with other Nif components available in purified form is essential to assess their functionality using enzymatic assays *in vitro*. In the context of this study, it means that our NifEN candidates should be compatible with NifH and NifB originating from organisms that in previous studies produced soluble and stable proteins in yeast and plants. Another consideration is that the NifEN variants should preferably be compatible with Nif components from *A. vinelandii*, as this is the model diazotroph that we usually use in our laboratory to isolate proteins needed for our *in vitro* assays, for example 1) NifU and NifS, that synthesize [4Fe-4S] clusters, 2) the cluster-carrier protein NifX, that facilitate the transfer of NifB-co from NifB to NifEN, and 3) NafY, that transfer mature FeMo-co from NifEN to apo-NifDK.

Identification of NifEN proteins that formed a soluble complex in yeast mitochondria

For most of the seventeen NifEN candidates, polypeptides corresponding to NifE and NifN could be detected in total protein extracts. However, as could be seen from immunoblot analysis, many of the NifN variants generated multiple polypeptides with different migration patterns. This was probably due to degradation of subunits that were not properly assembled into a functional heterotetramer, or from polypeptides that were not correctly targeted into mitochondria. It has previously been demonstrated in *K. oxytoca* that NifE protein is unstable in the absence of NifN protein, and vice versa, suggesting that neither subunit is stable outside the assembled protein complex (Roberts et al., 1978). It should be noted that the NifEN variants were always co-expressed with NifU^{Av} and NifS^{Av}, which provide [4Fe-4S] clusters to NifEN. This is important as it is known that Nif proteins are generally more unstable as [Fe-S] cluster-free apo-proteins.

Three NifEN variants, from *Synechococcus* sp. (ScNifEN^{Ss}), *H. thermophilus* (ScNifEN^{Ht}), and *G. metallireducens* (ScNifEN^{Gm}) were consistently detected in soluble protein extracts when targeted to the yeast mitochondria. Those from *Synechococcus* sp. and *H. thermophilus* were also shown to form a protein complex, as NifE was co-purified with NifN and vice versa (the NifEN from *G. metallireducens* is a natural NifEN fusion and this analysis was therefore not possible to perform for that variant). The NifEN protein from *A. variabilis* NifEN (NifEN^{Ava}) was also seen in the soluble protein extract but at lower levels. As for NifEN^{Gm}, NifEN^{Ava} is also a naturally occurring NifEN fusion protein. Early studies in *A. vinelandii* demonstrated that translationally fused *nifE-nifN* and *nifD-nifK* genes produced functional NifEN and NifDK proteins capable of supporting nitrogen fixation (Suh et al., 2003, Suh et al., 2002). These studies have shown the potential of fused proteins and indicated that

translational fusions of Nif proteins can both simplify the genetic componentry and retain nitrogenase function. Building upon these findings, Yang and colleagues employed a polyprotein-based approach by reorganizing operon-based *nif* genes from *K. oxytoca* into giant genes, which were subsequently cleaved to release individual components in non-diazotrophic *E. coli* cells (Yang et al., 2018). This strategy, known as posttranslational splicing, is naturally observed in RNA viruses, where a single giant gene is translated as a polyprotein and then cleaved by viral proteases to generate functional viral components with consistent stoichiometry (Palmenberg, 1990). The polyprotein combination effectively reproduces the protein stoichiometry observed with native *nif* operons, which is especially significant for the formation of tetrameric complexes of NifEN and NifDK proteins. Importantly, this study demonstrated the feasibility of expressing large Nif polyproteins in heterologous non-diazotrophic organisms and highlighted the potential of polyprotein systems to simplify the expression of prokaryotic *nif* genes in eukaryotes by reducing the number of translational units required. Taking into account these findings, Allen and colleagues then aimed to express *K. oxytoca* NifDK in the mitochondrial matrix of eukaryotic cells (Allen et al., 2020). Their findings revealed that a translational fusion of NifD to NifK enabled the complete processing and import of the polyprotein into mitochondria, even though the synthetic polyprotein approached the upper size limit of endogenous proteins transported to plant mitochondria. This suggests that mitochondria possess the capacity to import and process larger synthetic proteins, opening up new possibilities for expressing functional nitrogenase components in eukaryotic organisms. Additionally, since the translational fusion of NifD and NifK resulted in the solubility of the polyprotein within mitochondria, it can be suggested that, for structurally similar NifEN proteins, NifE and NifN polypeptides also prefer to be in close proximity upon entry into the mitochondrial matrix. This proximity could facilitate protein folding, improve solubility, and enhance resistance to degradation. This might explain why both of the two NifEN fusion proteins in the library were soluble when targeted to the yeast mitochondria.

The NifEN^{Ss} and NifEN^{Ht} proteins that accumulated at high levels in the mitochondria of aerobically respired yeast originate from thermophilic organisms (Annex I). It was previously reported that the temperature inside the respiring mitochondria of cultured human cells is around 50°C (Chrétien et al., 2018). While the internal temperature of yeast mitochondria remains unknown, also in previous library studies from our laboratory mainly Nif proteins from thermophiles generated soluble and functional proteins while Nif components from mesophilic diazotrophs were difficult to produce. It is therefore likely that also in this work, the thermophilic origin of these two NifEN variants was the reason for their stability and solubility in the yeast mitochondria.

At this point, it is worth considering the potential of establishing the effect of co-expressing *nifEN* genes of selected variants and a set of *nif* genes derived from some thermophilic archaea that lack *nifEN* and consist only of *nifHBDK*. It is believed that six genes is the minimum requirement for Mo-nitrogenase biosynthesis and functionality. However, recent discoveries have uncovered diazotrophs lacking either *nifE* or *nifN* genes, suggesting alternative mechanisms for FeMo-co assembly compared to the established Mo-nitrogenase system (Dekas et al., 2009). Notably, *Roseiflexus sp.* possesses a simplified and more ancient set of nitrogen fixation genes, where the resulting gene products are sufficient to accomplish the *in vitro* synthesis of an active nitrogenase, along with the substrates homocitrate, Mo, and SAM (Payá Tormo, 2020). This bacterium is proposed to directly insert NifB-co into apo-NifDK, where FeMo-co maturation is completed. This hypothesis is supported by laboratory findings demonstrating that NifDK can form a complex with NifB-co and catalyze acetylene reduction to ethylene *in vitro*, even in the absence of NifEN (Soboh et al., 2010). Phylogenetic analysis indicates that *nifEN* genes arose from a duplication of *nifDK*, leading to the compartmentalization of cofactor maturation and nitrogenase activity into distinct proteins (Mus et al., 2019). This duplication event is thought to have improved the efficiency of the nitrogenase machinery. It is hypothesized that nitrogenases with the minimal gene set may exhibit lower efficiency and slower reaction rates compared to their counterparts with complete *nif* genes, possibly representing a primitive form of the enzyme that predated the *nifDK* duplication. This primitive form would have combined cofactor maturation and reductase functions into a single protein. Considering these findings, it would be intriguing to investigate the impact of co-expressing *nifEN* genes with the *Roseiflexus sp. nifHBDK* genes on the nitrogenase biosynthetic pathway to test whether the resulting NifDK protein has changed the cofactor composition and so on exhibits enhanced stability or activity.

Expression of the NifEN library in $\Delta nifEN$ mutant of *A. vinelandii*

Screening of the NifEN candidates was performed in parallel using the *A. vinelandii* strain DJ2898. This strain is Nif⁻ because it is incapable of Mo-dependent diazotrophic growth due to a deletion of the *nifEN* genes. It also lacks the genes for the structural components of the alternative (V and Fe-only) nitrogenases. Eleven of the 17 NifEN variants could restore the diazotrophic growth on a nitrogen-free solid medium, demonstrating *in vivo* compatibility with the *A. vinelandii* FeMo-co biosynthetic machinery. In addition, *in vivo*, acetylene reduction assays showed that neither the use of the arabinose promoter to control heterologous *nifEN* gene expression, the yeast codon optimization, nor the N-terminal Strep-tag fused to these NifEN proteins impaired the functionality of these eleven NifEN variants. In this regard, strain DJ2917 (DJ2898 complemented with yeast codon-optimized and Strep-tagged *A. vinelandii*

NifEN) presented similar growth and nitrogenase activity as DJ2926 (DJ2898 complemented with the wild-type *A. vinelandii* *nifEN* genes).

As expected, primary sequence analysis of the NifE and NifN variants suggested that those NifEN proteins that restored diazotrophic growth showed higher similarity to NifEN from *A. vinelandii*. The only exception was NifEN from *K. oxytoca* which did not rescue the DJ2898 Nif phenotype although the protein is very similar to *A. vinelandii* NifEN. Two factors could explain this incompatibility. First, the replacement of the sequence between the *nifE* and *nifN* genes with the native intergenic sequence found in *A. vinelandii* resulted in a few amino acids being altered in the NifEN protein sequence, and these could be essential for the formation and/or stability of the *K. oxytoca* NifEN protein. Second, changes to the *nifE* and *nifN* gene sequences, due to the yeast-codon optimization, may influence protein folding and stability, and potentially impact the function of *K. oxytoca* NifEN protein. It was also observed that the structures of NifN in the NifEN protein variants capable of replacing *A. vinelandii* NifEN *in vivo* had a high similarity to *A. vinelandii* NifN. On the other hand, no similar correlation was found when comparing the structures of NifE. The lack of correlation regarding NifE structures is likely because NifE has three conserved cysteines coordinating the permanent [4Fe-4S] cluster and one conserved cysteine ligating the FeMo-co precursor. Perhaps this reduces the possibilities for structural variations of NifE. Furthermore, the structural variations in NifN subunits could negatively influence compatibility with the *A. vinelandii* nitrogenase machinery *in vivo*.

The three NifEN variants identified in the yeast solubility/complex formation study were isolated from an *A. vinelandii* strain lacking *nifHDK* and used for initial biochemical characterization and functionality assays. All three variants could activate apo-NifDK *in vitro* without the need for further protein reconstitution, suggesting that these NifEN proteins were isolated with bound FeMo-co precursor. This further demonstrated that these selected NifEN variants were compatible with the *A. vinelandii* Nif components in FeMo-co synthesis.

FeMo-co synthesis using proteins produced in yeast

In addition to identifying NifEN variants that formed soluble heterotetramers in yeast mitochondria, this work also aimed at identifying Nif partners compatible with the identified NifEN proteins. These Nif partners should be proteins involved in FeMo-co biosynthesis (i.e., NifH and NifB) that have already been obtained as soluble and functional proteins in the yeast mitochondria. Such demonstration of functional interactions between NifB, NifEN, and NifH isolated from mitochondria would undeniably demonstrate the possibility of engineering FeMo-co biosynthesis in eukaryotic organisms growing under aerobic conditions.

FeMo-co biosynthesis at NifEN requires constructive interactions with NifH to process NifB-co, produced by NifB, into FeMo-co. Therefore, identifying NifH proteins compatible with the three identified NifEN variants was first tested. It is worth noting that *in vivo* FeMo-co synthesis using the three NifEN variants combined with NifH from *A. vinelandii* was demonstrated both *in vitro* and *in vivo*. However, it has been shown that soluble NifH^{Av} is only produced at low levels in yeast mitochondria (Jiang et al., 2021). To address this issue, Paya-Tormo and colleagues identified alternative NifH variants that showed higher solubility and maintained activity in mitochondria and are therefore likely better suited for the FeMo-co biosynthesis engineering (Paya-Tormo et al., 2022).

Significant variability in the capacity of distinct NifH variants to support FeMo-co biosynthesis was observed depending on the NifEN used. This observation could be in line with the model for FeMo-co synthesis proposed by Hernandez and collaborators (followed in this thesis work) which suggests that NifEN forms a complex with NifH during FeMo-co synthesis, allowing for insertion of required substrates (Mo and homocitrate) and induction of a conformational change in NifEN that promotes FeMo-co formation (Hernandez et al., 2008). Notably, the compatibility results presented in this thesis suggest that constructive interactions between NifH and NifEN are essential for FeMo-co synthesis. The protein structure compatibility between NifH and NifEN holds significant implications. As shown, the highest activity was observed when NifEN^{Gm} was combined with NifH^{Gs} (90% identity to NifH^{Gm}). This finding is understandable given that these two Nif components originate from the same bacterial genus, potentially facilitating the interaction between Nif proteins involved in FeMo-co assembly. Moreover, the NifH from *G. sulfurreducens* has been shown to have an increased O₂ tolerance (Paya-Tormo et al., 2022), which is an additional advantage of this NifH protein as O₂-resistance is an inhibitory factor when engineering FeMo-co biosynthesis in the mitochondria of actively respiring cells.

The combination of NifEN^{Ht} with NifH originating also from *H. thermophilus* demonstrated significantly lower activity compared to NifH^{De}. This discrepancy could be attributed to the optimal growth temperature of *H. thermophilus* TK-6 (range between 70°C and 75°C) (Zeytun et al., 2011). At 30°C, the temperature used in the FeMo-co biosynthesis assays, NifH^{Ht} may be less active or not interact effectively with NifEN^{Ht}, impacting its optimal functioning. However, studies examining NifH^{Ht} ability to activate apo-NifDK have demonstrated that it can perform this function at lower temperatures. Nonetheless, it is noteworthy that NifH from *D. ethenogenes* offers another advantage in that it does not require NifM for its maturation, ultimately facilitating the transfer of genes to eukaryotic cells (Paya-Tormo et al., 2022).

NifEN^{SS} was also tested for compatibility with distinct NifH proteins, however the activity of this specific variant when combined with NifH from *Roseiflexus sp.* was significantly lower compared to the other two NifEN candidates. Furthermore, the *in vitro* FeMo-co activity assays using NifEN proteins isolated from yeast showed that this candidate was incapable of supporting the activation of apo-NifDK. Because of these results, this NifEN variant was excluded from further analysis.

It is worth mentioning the differences in activities using NifEN proteins isolated from bacteria and yeast. The significant lower activities obtained using yeast-isolated NifEN proteins could be attributed to either the protein stability and/or, more likely, to inefficient insertion of [4Fe-4S] clusters into the protein within yeast mitochondria. This hypothesis is supported by the fact that it was not possible to detect Fe in yeast-isolated NifEN proteins. This suggests that only a small population of the isolated NifEN proteins from yeast mitochondria harbor [4Fe-4S] clusters, making them inactive or partially active. There are several possible explanations for this observation, including the availability of Fe-S cluster assembly proteins within the mitochondria and the stability of Fe-S clusters in the mitochondria environment.

In this point, it is noteworthy to highlight that although some NifEN proteins exhibit low Fe content and likely lack permanent [4Fe-4S] clusters, they nonetheless maintain stability in the mitochondrial environment and demonstrated activity in *in vitro* FeMo-co biosynthesis assays without requiring an additional source of Fe or S (other than NifB-co) in the reaction mixture. This suggests that the formation of permanent Fe-S clusters may not be a stringent requirement for the stability and activity of these specific NifEN proteins. This is an important finding to consider when engineering Nif protein expression in eukaryotic cells. Furthermore, these results hold promise for the expression of NifEN proteins in plant cells, where Fe-S clusters are particularly vulnerable to damage from O₂ exposure. If the stability and activity of NifEN proteins can be maintained even with incomplete Fe-S cluster occupation, their expression within the O₂-rich environment of plant cells could become a viable reality.

In this study, the initial *in vitro* FeMo-co biosynthesis activity assays were performed using isolated NifX with bound NifB-co. This was to focus solely on the functionality of selected NifEN variants and to avoid eventual incompatibility with NifB. This also ensured that the as-isolated NifEN proteins used in the assay contained the permanent [4Fe-4S] clusters, as the addition of Fe or Fe-S clusters used in NifB-dependent assays (e.g., [4Fe-4S] cluster-loaded NifU) could result in re-activation of otherwise non-functional NifEN. Once the functionality of the as-isolated ScNifEN candidates was verified, FeMo-co synthesis assays were performed in which NifB-co was instead produced within the reaction mixture by ScNifB (NifB-dependent assays). Buren and colleagues have identified two NifB proteins from the archaea *M. infernus* and *M. thermautotrophicus* that are expressed in mitochondria in a soluble form together with

A. vinelandii accessory components. They were also shown to accumulate as functional proteins as they supported FeMo-co synthesis *in vitro* (Burén et al., 2019). The NifB-dependent FeMo-co synthesis assays enabled us to demonstrate the compatibility of the functional yeast-isolated NifEN proteins with these two NifB variants and to recapitulate the FeMo-co biosynthetic pathway using only components produced in *S. cerevisiae*.

The results presented in this thesis show that both ScNifEN^{Ht} and ScNifEN^{Gm} can support *in vitro* FeMo-co synthesis and apo-NifDK activation when combined with selected NifH partners and either of the two NifB variants. This demonstrates that combining diverse interspecies nitrogenase FeMo-co assembly components could be an effective approach to achieving and optimizing nitrogen fixation in a eukaryotic organism.

However, it is important to mention that in this thesis work, the *in vitro* FeMo-co biosynthesis assays depend on the activation of apo-NifDK originating from *A. vinelandii*. Buren and colleagues showed that although apo-NifDK^{Av} targeted yeast mitochondria produced NifDK tetramer, they observed that NifD was susceptible to degradation, forming a non-functional NifDK protein (Buren et al., 2017). Similarly, targeting *K. oxytoca* NifD into plant mitochondria resulted in protein degradation, suggesting that apart from exploiting a biodiversity approach, additional strategies may be needed to stabilize NifDK protein for successful expression in eukaryotes. Allen and colleagues have shown that a translational fusion strategy could be a promising solution for NifDK. Another potential solution proposed by Yang and colleagues is the use of a posttranslational protein splicing strategy (Yang et al., 2018). These two strategies could not only decrease the number of *nif* genes required to be transferred to eukaryotes but also ensure correct stoichiometry between the NifD and NifK subunits, essential for the functionality of the NifDK protein. Upon identification of a suitable NifDK protein for expression in a model eukaryote, the FeMo-co biosynthetic modules presented in this study can be employed to generate functional nitrogenase *in vivo*.

To conclude, NifEN^{Ss}, NifEN^{Ht}, and NifEN^{Gm} could accumulate in soluble forms in the mitochondria of aerobically cultured *S. cerevisiae*. These NifEN proteins are functional *in vivo* in *A. vinelandii*, proving their compatibility with the *A. vinelandii* Nif machinery. All three candidates were active when isolated from *E. coli* cells and combined with a compatible NifH partner, however, only NifEN^{Ht} and NifEN^{Gm} were active in *in vitro* FeMo-co biosynthesis when purified from yeast mitochondria. The yeast-isolated NifEN proteins produced the nitrogenase active-site cluster FeMo-co *in vitro* when combined with NifH and NifB proteins also isolated from yeast mitochondria. This is an important finding as expression of the NifEN^{Av} protein in the mitochondria has previously proven very difficult. Therefore, future efforts should be devoted to finding a competent form of apo-NifDK that can be produced in yeast mitochondria.

6. CONCLUSIONS

1. Seventeen NifEN proteins from diverse evolutionary lineages were expressed in *S. cerevisiae* mitochondria and tested for function in the diazotrophic bacterium *A. vinelandii*.
2. Eleven of the tested NifEN proteins restored diazotrophic growth and nitrogenase activity in an *A. vinelandii* strain lacking *nifEN* genes, indicating their compatibility with other Nif components.
3. NifEN^{Ss}, NifEN^{Ht}, and NifEN^{Gm} obtained from recombinant *A. vinelandii* strains carried FeMo-co precursors and were functional in FeMo-co biosynthesis.
4. Three NifEN variants (NifEN^{Ss}, NifEN^{Ht} and NifEN^{Gm}) formed soluble protein complexes in *S. cerevisiae* mitochondria.
5. Optimal combinations of NifH and NifEN for FeMo-co production were identified as *Roseiflexus* sp. NifH with *Synechococcus* sp. JA-2-3B'a NifEN, *D. ethenogenes* NifH with *H. thermophilus* NifEN, and *G. sulfurreducens* NifH with *G. metallireducens* NifEN.
6. A complete *in vitro* FeMo-co synthesis pathway was established using NifEN, NifH, and NifB produced in *S. cerevisiae* mitochondria.

REFERENCES

- ALLEN, R. M., CHATTERJEE, R., LUDDEN, P. W. & SHAH, V. K. 1995. Incorporation of Iron and Sulfur from NifB Cofactor into the Iron-Molybdenum Cofactor of Dinitrogenase (*). *Journal of Biological Chemistry*, 270, 26890-26896.
- ALLEN, R. M., HOMER, M. J., CHATTERJEE, R., LUDDEN, P. W., ROBERTS, G. P. & SHAH, V. K. 1993. Dinitrogenase reductase- and MgATP-dependent maturation of apodinitrogenase from *Azotobacter vinelandii*. *Journal of Biological Chemistry*, 268, 23670-23674.
- ALLEN, R. S., GREGG, C. M., OKADA, S., MENON, A., HUSSAIN, D., GILLESPIE, V., JOHNSTON, E., DEVILLA, R., WARDEN, A. C., TAYLOR, M., BYRNE, K., COLGRAVE, M. & WOOD, C. C. 2020. Plant expression of NifD protein variants resistant to mitochondrial degradation. *Proc Natl Acad Sci U S A*, 117, 23165-23173.
- AYALA-CASTRO, C., SAINI, A. & OUTTEN, F. W. 2008. Fe-S Cluster Assembly Pathways in Bacteria. *Microbiology and Molecular Biology Reviews*, 72, 110-125.
- BARAHONA, E., JIANG, X., JIMÉNEZ-VICENTE, E., RUBIO, L. M. & GONZÁLEZ-GUERRERO, M. 2021. *Azotobacter vinelandii* scaffold protein NifU transfers iron to NifQ as part of the iron-molybdenum cofactor biosynthesis pathway for nitrogenase. *bioRxiv*, 2021.12.06.471366.
- BERTANI, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of bacteriology*, 62, 293-300.
- BOYD, E. S., COSTAS, A. M., HAMILTON, T. L., MUS, F. & PETERS, J. W. 2015. Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J Bacteriol*, 197, 1690-9.
- BRIGLE, K., WEISS, M. C., NEWTON, W. E. & DEAN, D. R. 1987. Products of the Iron-Molybdenum Cofactor-Specific Biosynthetic Genes, *nifE* and *nifN*, Are Structurally Homologous to the Products of the Nitrogenase Molybdenum-Iron Protein Genes, *nifD* and *nifK*. *Journal of Bacteriology*, 169, 1547-1553.
- BRODERICK, J. B., DUFFUS, B. R., DUSCHENE, K. S. & SHEPARD, E. M. 2014. Radical S-adenosylmethionine enzymes. American Chemical Society.
- BUREN, S., JIMENEZ-VICENTE, E., ECHAVARRI-ERASUN, C. & RUBIO, L. M. 2020. Biosynthesis of Nitrogenase Cofactors. *Chem Rev*, 120, 4921-4968.
- BURÉN, S., PRATT, K., JIANG, X., GUO, Y., JIMÉNEZ-VICENTE, E., ECHAVARRI-ERASUN, C., DEAN, D. R., SAAEM, I., GORDON, D. B., VOIGT, C. A. & RUBIO, L. M. 2019. Biosynthesis of the nitrogenase active-site cofactor precursor NifB-co in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*, 116, 25078-25086.
- BUREN, S., YOUNG, E. M., SWEENEY, E. A., LOPEZ-TORREJON, G., VELDHUIZEN, M., VOIGT, C. A. & RUBIO, L. M. 2017. Formation of Nitrogenase NifDK Tetramers in the Mitochondria of *Saccharomyces cerevisiae*. *ACS Synth Biol*, 6, 1043-1055.
- BURGESS, B. K. & LOWE, D. J. 1996. Mechanism of molybdenum nitrogenase. *Chemical Reviews*, 96, 2983-3011.
- BURRIS, R. H. & ROBERTS, G. P. 1993. Biological Nitrogen Fixation. *Annual Review of Nutrition*, 13, 317-335.
- CHEN, J. S., TOTH, J. & KASAP, M. 2001. Nitrogen-fixation genes and nitrogenase activity in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *J Ind Microbiol Biotechnol*, 27, 281-6.
- CHRÉTIEN, D., BÉNIT, P., HA, H.-H., KEIPERT, S., EL-KHOURY, R., CHANG, Y.-T., JASTROCH, M., JACOBS, H. T., RUSTIN, P. & RAK, M. 2018. Mitochondria are physiologically maintained at close to 50 °C. *PLOS Biology*, 16, e2003992-e2003992.

- CHRISTIANSEN, J., GOODWIN, P. J., LANZILOTTA, W. N., SEEFELDT, L. C. & DEAN, D. R. 1998. Catalytic and biophysical properties of a nitrogenase Apo-MoFe protein produced by a *nifB*-deletion mutant of *Azotobacter vinelandii*. *Biochemistry*, 37, 12611-23.
- COTTON, M. S., RUPNIK, K., BROACH, R. B., HU, Y., FAY, A. W., RIBBE, M. W. & HALES, B. J. 2009. VTVH-MCD study of the Delta *nifB* Delta *nifZ* MoFe protein from *Azotobacter vinelandii*. *J Am Chem Soc*, 131, 4558-9.
- CRACK, J. C., GREEN, J., THOMSON, A. J. & LE BRUN, N. E. 2014. Techniques for the Production, Isolation, and Analysis of Iron-Sulfur Proteins. Humana Press, Totowa, NJ.
- CURATTI, L., HERNANDEZ, J. A., IGARASHI, R. Y., SOBOH, B., ZHAO, D. & RUBIO, L. M. 2007. In vitro synthesis of the iron-molybdenum cofactor of nitrogenase from iron, sulfur, molybdenum, and homocitrate using purified proteins. *Proc Natl Acad Sci U S A*, 104, 17626-31.
- CURATTI, L., LUDDEN, P. W. & RUBIO, L. M. 2006. NifB-dependent in vitro synthesis of the iron-molybdenum cofactor of nitrogenase. *Proceedings of the National Academy of Sciences*, 103, 5297-5301.
- CURATTI, L. & RUBIO, L. 2014. Challenges to develop nitrogen-fixing cereals by direct *nif*-gene transfer. *Plant Science*, 225, 130-137.
- DAVID, K. A. V., APTE, S. K., BANERJI, A. & THOMAS, J. 1980. Acetylene Reduction Assay for Nitrogenase Activity: Gas Chromatographic Determination of Ethylene Per Sample in Less Than One Minute. *Applied and Environmental Microbiology*, 39, 1078-1080.
- DEKAS, A. E., PORETSKY, R. S. & ORPHAN, V. J. 2009. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science*, 326, 422-6.
- DIXON, R. A. & POSTGATE, J. R. 1972. Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature*, 237, 102-3.
- DOS SANTOS, P. C. & DEAN, D. R. 2011. Co-ordination and fine-tuning of nitrogen fixation in *Azotobacter vinelandii*. *Mol Microbiol*, 79, 1132-5.
- DOS SANTOS, P. C., DEAN, D. R., HU, Y. & RIBBE, M. W. 2004. Formation and Insertion of the Nitrogenase Iron-Molybdenum Cofactor. *Chemical Reviews*, 104, 1159-1173.
- DOS SANTOS, P. C., FANG, Z., MASON, S. W., SETUBAL, J. C. & DIXON, R. 2012. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics*, 13, 162.
- DOS SANTOS, P. C., JOHNSON, D. C., RAGLE, B. E., UNCIULEAC, M. C. & DEAN, D. R. 2007. Controlled expression of *nif* and *isc* iron-sulfur protein maturation components reveals target specificity and limited functional replacement between the two systems. *J Bacteriol*, 189, 2854-62.
- DUVAL, S., DANYAL, K., SHAW, S., LYTLE, A. K., DEAN, D. R., HOFFMAN, B. M., ANTONY, E. & SEEFELDT, L. C. 2013. Electron transfer precedes ATP hydrolysis during nitrogenase catalysis. *Proceedings of the National Academy of Sciences*, 110, 16414-16419.
- DYER, D. H., RUBIO, L. M., THODEN, J. B., HOLDEN, H. M., LUDDEN, P. W. & RAYMENT, I. 2003. The three-dimensional structure of the core domain of Naf Y from *Azotobacter vinelandii* determined at 1.8-Å resolution. *J Biol Chem*, 278, 32150-6.
- ECHAVARRI-ERASUN, C., ARRAGAIN, S., SCANDURRA, A. A. & RUBIO, L. M. 2014. Expression and Purification of NifB Proteins from Aerobic and Anaerobic Sources. In: FONTECILLA-CAMPS, J. C. & NICOLET, Y. (eds.) *Metalloproteins: Methods and Protocols, Methods in Molecular Biology*. Humana Press: Totowa, NJ.
- FANI, R., GALLO, R. & LIO, P. 2000. Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD*, *nifK*, *nifE*, and *nifN* genes. *J Mol Evol*, 51, 1-11.

- FAY, A. W., WIIG, J. A., LEE, C. C. & HU, Y. 2015. Identification and characterization of functional homologs of nitrogenase cofactor biosynthesis protein NifB from methanogens. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 14829-33.
- FREY, P. A., HEGEMAN, A. D. & RUZICKA, F. J. 2008. The Radical SAM Superfamily. *Crit Rev Biochem Mol Biol*, 43, 63-88.
- FU, W., JACK, R. F., MORGAN, T. V., DEAN, D. R. & JOHNSON, M. K. 1994. *nifU* gene product from *Azotobacter vinelandii* is a homodimer that contains two identical [2Fe-2S] clusters. *Biochemistry*, 33, 13455-63.
- GALLON, J. R. 1992. Reconciling the incompatible: N₂ fixation And O₂. *New Phytologist*, 122, 571-609.
- GIETZ, R. D. & SCHIESTL, R. H. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc*, 2, 31-4.
- GOODWIN, P. J., AGAR, J. N., ROLL, J. T., ROBERTS, G. P., JOHNSON, M. K. & DEAN, D. R. 1998. The *Azotobacter vinelandii* NifEN complex contains two identical [4Fe-4S] clusters. *Biochemistry*, 37, 10420-10428.
- HAGEMAN, R. V. & BURRIS, R. H. 1978. Nitrogenase and nitrogenase reductase associate and dissociate with each catalytic cycle. *Proc Natl Acad Sci U S A*, 75, 2699-702.
- HE, W., BUREN, S., BAYSAL, C., JIANG, X., CAPELL, T., CHRISTOU, P. & RUBIO, L. M. 2022. Nitrogenase Cofactor Maturase NifB Isolated from Transgenic Rice is Active in FeMo-co Synthesis. *ACS Synth Biol*, 11, 3028-3036.
- HERNANDEZ, J. A., CURATTI, L., AZNAR, C. P., PEROVA, Z., BRITT, R. D. & RUBIO, L. M. 2008. Metal trafficking for nitrogen fixation: NifQ donates molybdenum to NifEN/NifH for the biosynthesis of the nitrogenase FeMo-cofactor. *Proc Natl Acad Sci U S A*, 105, 11679-84.
- HERNANDEZ, J. A., IGARASHI, R. Y., SOBOH, B., CURATTI, L., DEAN, D. R., LUDDEN, P. W. & RUBIO, L. M. 2007. NifX and NifEN exchange NifB cofactor and the VK-cluster, a newly isolated intermediate of the iron-molybdenum cofactor biosynthetic pathway. *Mol Microbiol*, 63, 177-92.
- HILL, H. D. & STRAKA, J. G. 1988. Protein determination using bicinchoninic acid in the presence of sulfhydryl reagents. *Analytical biochemistry*, 170, 203-208.
- HOFFMAN, B. J., BROADWATER, J. A., JOHNSON, P., HARPER, J., FOX, B. G. & KENEALY WILLIAM, R. 1995. Lactose Fed-Batch Overexpression of Recombinant Metalloproteins in Escherichia coli BL21 (DE3): Process, Control Yielding High Levles of Metal-incorporated, Soluble Protein. *Protien Expression and Purification*, 6, 646-654.
- HOFFMAN, B. M., LUKOYANOV, D., YANG, Z. Y., DEAN, D. R. & SEEFELDT, L. C. 2014. Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem Rev*, 114, 4041-62.
- HOOVER, T. R., IMPERIAL, J., LUDDEN, P. W. & SHAH, V. K. 1989. Homocitrate Is a Component of the Iron-Molybdenum Cofactor of Nitrogenase. *Biochemistry*, 28, 2768-2771.
- HORTON, R. M., CAI, Z., HO, S. M. & PEASE, L. R. 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques*, 54, 129-33.
- HU, Y., CORBETT, M. C., FAY, A. W., WEBBER, J. A., HODGSON, K. O., HEDMAN, B. & RIBBE, M. W. 2006. FeMo cofactor maturation on NifEN. *Proc Natl Acad Sci U S A*, 103, 17119-24.
- HU, Y., FAY, A. W., DOS SANTOS, P. C., NADERI, F. & RIBBE, M. W. 2004. Characterization of *Azotobacter vinelandii* nifZ deletion strains: Indication of stepwise MoFe protein assembly. *Journal of Biological Chemistry*, 279, 54963-54971.

- HU, Y., FAY, A. W., LEE, C. C. & RIBBE, M. W. 2007. P-cluster maturation on nitrogenase MoFe protein. *Proceedings of the National Academy of Sciences*, 104, 10424-10429.
- HU, Y., FAY, A. W. & RIBBE, M. W. 2005. Identification of a nitrogenase FeMo cofactor precursor on NifEN complex. *Proc Natl Acad Sci U S A*, 102, 3236-41.
- HU, Y., YOSHIZAWA, J. M., FAY, A. W., LEE, C. C., WIIG, J. A. & RIBBE, M. W. 2009. Catalytic activities of NifEN: implications for nitrogenase evolution and mechanism. *Proc Natl Acad Sci U S A*, 106, 16962-6.
- JACOBSON, M. R., BRIGLE, K. E., BENNETT, L. T., SETTERQUIST, R. A., WILSON, M. S., CASH, V. L., BEYNON, J., NEWTON, W. E. & DEAN, D. R. 1989a. Physical and genetic map of the major nif gene cluster from *Azotobacter vinelandii*. *J Bacteriol*, 171, 1017-1027.
- JACOBSON, M. R., CASH, V. L., WEISS, M. C., LAIRD, N. F., NEWTON, W. E. & DEAN, D. R. 1989b. Biochemical and genetic analysis of the nifUSVWZM cluster from *Azotobacter vinelandii*. *MGG Molecular & General Genetics*, 219, 49-57.
- JENNER, L. P., CHERRIER, M. V., AMARA, P., RUBIO, L. M. & NICOLET, Y. 2021. An unexpected P-cluster like intermediate en route to the nitrogenase FeMo-co. *Chemical Science*, 12, 5269-5274.
- JIANG, X., COROIAN, D., BARAHONA, E., ECHAVARRI-ERASUN, C., CASTELLANOS-RUEDA, R., ESEVERRI, A., AZNAR-MORENO, J. A., BUREN, S. & RUBIO, L. M. 2022. Functional nitrogenase cofactor maturase NifB in mitochondria and chloroplasts of *Nicotiana benthamiana*. *mBio*, 13, e0026822.
- JIANG, X., PAYA-TORMO, L., COROIAN, D., GARCIA-RUBIO, I., CASTELLANOS-RUEDA, R., ESEVERRI, A., LOPEZ-TORREJON, G., BUREN, S. & RUBIO, L. M. 2021. Exploiting genetic diversity and gene synthesis to identify superior nitrogenase NifH protein variants to engineer N(2)-fixation in plants. *Commun Biol*, 4, 4.
- JIMENEZ-VICENTE, E., MARTIN DEL CAMPO, J. S., YANG, Z. Y., CASH, V. L., DEAN, D. R. & SEEFELDT, L. C. 2018. Application of affinity purification methods for analysis of the nitrogenase system from *Azotobacter vinelandii*. *Methods Enzymol*, 613, 231-255.
- JIMÉNEZ-VICENTE, E., NAVARRO-RODRÍGUEZ, M., POZA-CARRIÓN, C. & RUBIO, L. M. 2014. Role of *Azotobacter vinelandii* FdxN in FeMo-co biosynthesis. 588, 512-516.
- JIMENEZ-VICENTE, E., YANG, Z. Y., MARTIN DEL CAMPO, J. S., CASH, V. L., SEEFELDT, L. C. & DEAN, D. R. 2019. The NifZ accessory protein has an equivalent function in maturation of both nitrogenase MoFe protein P-clusters. *Journal of Biological Chemistry*, 294, 6204-6213.
- JIMÉNEZ-VICENTE, E., YANG, Z. Y., RAY, W. K., ECHAVARRI-ERASUN, C., CASH, V. L., RUBIO, L. M., SEEFELDT, L. C. & DEAN, D. R. 2018. Sequential and differential interaction of assembly factors during nitrogenase MoFe protein maturation. *J Biol Chem*, 293, 9812-9823.
- JOERGER, R. D. & BISHOP, P. E. 1988. Nucleotide Sequence and Genetic Analysis of the *nifB-nifQ* Region from *Azotobacter vinelandii*. *Journal of Bacteriology*, 170, 1475-1487.
- JOHNSON, B. & GOLDBLATT, C. 2015. The nitrogen budget of Earth. *Earth Science Reviews*, 148, 150-173.
- JOHNSON, D. C., DOS SANTOS, P. C. & DEAN, D. R. 2005. NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in *Azotobacter vinelandii*. *Biochemical Society Transactions*, 33, 90-93.
- JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ZIDEK, A., POTAPENKO, A., BRIDGLAND, A., MEYER, C., KOHL, S. A. A., BALLARD, A. J., COWIE, A., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., ADLER, J., BACK, T., PETERSEN, S., REIMAN, D., CLANCY, E., ZIELINSKI, M., STEINEGGER, M.,

- PACHOLSKA, M., BERGHAMMER, T., BODENSTEIN, S., SILVER, D., VINIYALS, O., SENIOR, A. W., KAVUKCUOGLU, K., KOHLI, P. & HASSABIS, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
- KAISER, J. T., HU, Y., WIIG, J. A., REES, D. C. & RIBBE, M. W. 2011. Structure of Precursor-Bound NifEN: A Nitrogenase FeMo Cofactor Maturase/Insertase. *Science*, 331, 91-94.
- KENNEDY, C. & DEAN, D. 1992. The *nifU*, *nifS* and *nifV* gene products are required for activity of all three nitrogenases of *Azotobacter vinelandii*. *Mol Gen Genet*, 231, 494-8.
- KENNEDY, C., GAMAL, R., HUMPHREY, R., RAMOS, J., BRIGLE, K. & DEAN, D. 1986. The *nifH*, *nifM* and *nifN* genes of *Azotobacter vinelandii*: Characterisation by Tn 5 mutagenesis and isolation from pLAFR1 gene banks. *MGG Molecular & General Genetics*, 205, 318-325.
- KESSLER, P. S., BLANK, C. & LEIGH, J. A. 1998. The *nif* gene operon of the methanogenic archaeon *Methanococcus maripaludis*. *J Bacteriol*, 180, 1504-11.
- KOONIN, E. V. 2015. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Philos Trans R Soc Lond B Biol Sci*, 370, 20140333.
- KOSKELA, E. V. & FREY, A. D. 2015. Homologous Recombinatorial Cloning Without the Creation of Single-Stranded Ends: Exonuclease and Ligation-Independent Cloning (ELIC). *Molecular Biotechnology*, 57, 233-240.
- LI, X. X., LIU, Q., LIU, X. M., SHI, H. W. & CHEN, S. F. 2016. Using synthetic biology to increase nitrogenase activity. *Microb Cell Fact*, 15, 43.
- LIU, C. W., SUNG, Y., CHEN, B. C. & LAI, H. Y. 2014. Effects of nitrogen fertilizers on the growth and nitrate content of lettuce (*Lactuca sativa* L.). *Int J Environ Res Public Health*, 11, 4427-40.
- LIU, X., WANG, M., SONG, Y., LI, Y., LIU, P., SHI, H., LI, Y., HAO, T., ZHANG, H., JIANG, W., CHEN, S. & LI, J. 2019. Combined Assembly and Targeted Integration of Multigene for Nitrogenase Biosynthetic Pathway in *Saccharomyces cerevisiae*. *ACS Synth Biol*, 8, 1766-1775.
- LÓPEZ-TORREJÓN, G., JIMÉNEZ-VICENTE, E., BUESA, J. M., HERNÁNDEZ, J. A., VERMA, H. K. & RUBIO, L. M. 2016. Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat Commun*, 7, 11426-11426.
- MAIER, R. J. & MOSHIRI, F. 2000. Role of the *Azotobacter vinelandii* nitrogenase-protective shethna protein in preventing oxygen-mediated cell death. *J Bacteriol*, 182, 3854-7.
- MALINAS, A., VIDICAN, R., ROTAR, I., MALINAS, C., MOLDOVAN, C. M. & PROOROCU, M. 2022. Current Status and Future Prospective for Nitrogen Use Efficiency in Wheat (*Triticum aestivum* L.). *Plants (Basel)*, 11.
- MARTIN DEL CAMPO, J. S., RIGSBEE, J., BUENO BATISTA, M., MUS, F., RUBIO, L. M., EINSLE, O., PETERS, J. W., DIXON, R., DEAN, D. R. & DOS SANTOS, P. C. 2022. Overview of physiological, biochemical, and regulatory aspects of nitrogen fixation in *Azotobacter vinelandii*. *Crit Rev Biochem Mol Biol*, 57, 492-538.
- MCLEAN, P. A. & DIXON, R. 1981. Requirement of *nifV* gene for production of wild-type nitrogenase enzyme in *Klebsiella pneumoniae*. *Nature*, 292, 655-656.
- MOSS, M. L. & MELLON, M. G. 1942. Colorimetric Determination of Iron with 2,2'-Bipyridyl and with 2,2',2'-Terpyridyl. *Industrial & Engineering Chemistry Analytical Edition*, 14, 862-865.
- MUS, F., COLMAN, D. R., PETERS, J. W. & BOYD, E. S. 2019. Geobiological feedbacks, oxygen, and the evolution of nitrogenase. *Free Radic Biol Med*, 140, 250-259.

- NICOLET, Y. 2020. Structure–function relationships of radical SAM enzymes. *Nature Catalysis*, 3, 337-350.
- OKADA, S., GREGG, C. M., ALLEN, R. S., MENON, A., HUSSAIN, D., GILLESPIE, V., JOHNSTON, E., BYRNE, K., COLGRAVE, M. L. & WOOD, C. C. 2020a. A Synthetic Biology Workflow Reveals Variation in Processing and Solubility of Nitrogenase Proteins Targeted to Plant Mitochondria, and Differing Tolerance of Targeting Sequences in a Bacterial Nitrogenase Assay. *Frontiers in Plant Science*, 11, 1388-1388.
- OKADA, S., GREGG, C. M., ALLEN, R. S., MENON, A., HUSSAIN, D., GILLESPIE, V., JOHNSTON, E., BYRNE, K., COLGRAVE, M. L. & WOOD, C. C. 2020b. A Synthetic Biology Workflow Reveals Variation in Processing and Solubility of Nitrogenase Proteins Targeted to Plant Mitochondria, and Differing Tolerance of Targeting Sequences in a Bacterial Nitrogenase Assay. *Front Plant Sci*, 11, 552160.
- PAGE, W. J. & VON TIGERSTROM, M. 1979. Optimal conditions for transformation of *Azotobacter vinelandii*. *J Bacteriol*, 139, 1058-61.
- PALMENBERG, A. C. 1990. Proteolytic processing of picornaviral polyprotein. *Annu Rev Microbiol*, 44, 603-23.
- PAUSTIAN, T. D., SHAH, V. K. & ROBERTS, G. P. 1989. Purification and characterization of the nifN and nifE gene products from *Azotobacter vinelandii* mutant UW45. *Proc Natl Acad Sci U S A*, 86, 6082-6.
- PAUSTIAN, T. D., SHAH, V. K. & ROBERTS, G. P. 1990. Apodinitrogenase: Purification, Association with a 20-Kilodalton Protein, and Activation by the Iron-Molybdenum Cofactor in the Absence of Dinitrogenase Reductase. *Biochemistry*, 29, 3515-3522.
- PAYÁ TORMO, L. 2020. *Insights into nitrogenase biosynthesis obtained from thermophilic prokaryotes*. Biology Doctoral, E.T.S. de Ingeniería Agronómica, Alimentaria y de Biosistemas (UPM).
- PAYA-TORMO, L., COROIAN, D., MARTIN-MUNOZ, S., BADALYAN, A., GREEN, R. T., VELDHUIZEN, M., JIANG, X., LOPEZ-TORREJON, G., BALK, J., SEEFELDT, L. C., BUREN, S. & RUBIO, L. M. 2022. A colorimetric method to measure in vitro nitrogenase functionality for engineering nitrogen fixation. *Sci Rep*, 12, 10367.
- PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 25, 1605-12.
- RANGARAJ, P., RYLE, M. J., LANZILOTTA, W. N., LUDDEN, P. W. & SHAH, V. K. 1999. In vitro biosynthesis of iron-molybdenum cofactor and maturation of the nif-encoded apodinitrogenase. Effect of substitution for NifH with site-specifically altered forms of NifH. *J Biol Chem*, 274, 19778-84.
- RETTBERG, L. A., WILCOXEN, J., LEE, C. C., STIEBRITZ, M. T., TANIFUJI, K., BRITT, R. D. & HU, Y. 2018. Probing the coordination and function of Fe₄S₄ modules in nitrogenase assembly protein NifB. *Nature Communications*, 9.
- ROBERTS, G. P., MACNEIL, T., MACNEIL, D. & BRILL, W. J. 1978. Regulation and characterization of protein products coded by the nif (nitrogen fixation) genes of *Klebsiella pneumoniae*. *Journal of Bacteriology*, 136, 267-279.
- RODRÍGUEZ-QUIÑONES, F., BOSCH, R. & IMPERIAL, J. 1993. Expression of the nifBfdxNnifOQ region of *Azotobacter vinelandii* and its role in nitrogenase activity. *J Bacteriol*, 175, 2926-35.
- RUBIO, L. M. & LUDDEN, P. W. 2008. Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase. *Annual Review of Microbiology*, 62, 93-111.
- RUBIO, L. M., RANGARAJ, P., HOMER, M. J., ROBERTS, G. P. & LUDDEN, P. W. 2002. Cloning and mutational analysis of the γ gene from *Azotobacter vinelandii* defines a new family of proteins capable of metallocluster binding and protein stabilization. *Journal of Biological Chemistry*, 277, 14299-14305.

- SABRA, W., ZENG, A. P., LÜNSDORF, H. & DECKWER, W. D. 2000. Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. *Applied and environmental microbiology*, 66, 4037-4044.
- SAMBROOK, J. & RUSSELL, D. W. 2001. *Molecular cloning : a laboratory manual*, Cold Spring Harbor Laboratory Press.
- SAMMAURIA, R., KUMAWAT, S., KUMAWAT, P., SINGH, J. & JATWA, T. K. 2020. Microbial inoculants: potential tool for sustainability of agricultural production systems. *Archives of Microbiology*, 202, 677-693.
- SCHLESIER, J., ROHDE, M., GERHARDT, S. & EINSLE, O. 2016. A Conformational Switch Triggers Nitrogenase Protection from Oxygen Damage by Shethna Protein II (FeSII). *J Am Chem Soc*, 138, 239-47.
- SENDRA, M., OLLAGNIER DE CHOUDENS, S., LASCOUX, D., SANAKIS, Y. & FONTECAVE, M. 2007. The SUF iron-sulfur cluster biosynthetic machinery: sulfur transfer from the SUFS-SUFE complex to SUFA. *FEBS Lett*, 581, 1362-8.
- SHAH, V. K., DAVIS, C. L. & BRILL, W. J. 1972. Nitrogenase I. Repression and derepression of the iron-molybdenum and iron proteins of nitrogenase in *Azotobacter vinelandii*. *Biochim Biophys Acta*, 256, 498-511.
- SHAH, V. K., RANGARAJ, P., CHATTERJEE, R., ALLEN, R. M., ROLL, J. T., ROBERTS, G. P. & LUDDEN, P. W. 1999. Requirement of NifX and Other nif Proteins for In Vitro Biosynthesis of the Iron-Molybdenum Cofactor of Nitrogenase. *Journal of Bacteriology*, 181, 2797-2801.
- SHAH, V. K., UGALDE, R. A., IMPERIAL, J. & BRILL, W. J. 1985. Inhibition of iron-molybdenum cofactor binding to component I of nitrogenase. *J Biol Chem*, 260, 3891-3894.
- SIBOLD, L., QUIVIGER, B., CHARPIN, N., PAQUELIN, A. & ELMERICH, C. 1983. Cloning and expression of a DNA fragment carrying a his nifA fusion and the nifBQ operon from a nif constitutive mutant of *Klebsiella pneumoniae*. *Biochimie*, 65, 53-63.
- SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*, 7, 539.
- SMIL, V. 2004. The roots of nitrogen fixation. *Nature*, 431, 909-911.
- SMITH, A. D., JAMESON, G. N. L., DOS SANTOS, P. C., AGAR, J. N., NAIK, S., KREBS, C., FRAZZON, J., DEAN, D. R., HUYNH, B. H. & JOHNSON, M. K. 2005. NifS-mediated assembly of [4Fe-4S] clusters in the N- and C-terminal domains of the NifU Scaffold protein. *Biochemistry*, 44, 12955-12969.
- SMITH, P. K., KROHN, R. I., HERMANSON, G. T., MALLIA, A. K., GARTNER, F. H., PROVENZANO, M. D., FUJIMOTO, E. K., GOEKE, N. M., OLSON, B. J. & KLENK, D. C. 1985. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, 150, 76-85.
- SOBOH, B., BOYD, E. S., ZHAO, D., PETERS, J. W. & RUBIO, L. M. 2010. Substrate specificity and evolutionary implications of a NifDK enzyme carrying NifB-co at its active site. *FEBS Lett*, 584, 1487-92.
- SOBOH, B., IGARASHI, R. Y., HERNANDEZ, J. A. & RUBIO, L. M. 2006. Purification of a NifEN protein complex that contains bound molybdenum and a FeMo-Co precursor from an *Azotobacter vinelandii* DeltanifHDK strain. *J Biol Chem*, 281, 36701-9.
- SOLOMON, J. B., LEE, C. C., JASNIEWSKI, A. J., RASEKH, M. F., RIBBE, M. W. & HU, Y. 2020. Heterologous Expression and Engineering of the Nitrogenase Cofactor Biosynthesis Scaffold NifEN. *Angew Chem Int Ed Engl*, 59, 6887-6893.
- SUH, M.-H., PULAKAT, L. & GAVINI, N. 2002. Functional expression of the FeMo-cofactor-specific biosynthetic genes nifEN as a NifE-N fusion protein synthesizing unit in *Azotobacter vinelandii*. *Biochemical and Biophysical Research Communications*, 299, 233-240.

- SUH, M. H., PULAKAT, L. & GAVINI, N. 2003. Functional expression of a fusion-dimeric MoFe protein of nitrogenase in *Azotobacter vinelandii*. *J Biol Chem*, 278, 5353-60.
- TAKAHASHI, Y. & TOKUMOTO, U. 2002. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *J Biol Chem*, 277, 28380-3.
- THIEL, T. & PRATTE, B. S. 2014. Regulation of Three Nitrogenase Gene Clusters in the Cyanobacterium *Anabaena variabilis* ATCC 29413. *Life (Basel)*, 4, 944-67.
- UEKI, T. & LOVLEY, D. R. 2010. Novel regulatory cascades controlling expression of nitrogen-fixation genes in *Geobacter sulfurreducens*. *Nucleic Acids Res*, 38, 7485-99.
- UNIPROT, C. 2023. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res*, 51, D523-D531.
- VAN STAPPEN, C., JIMENEZ-VICENTE, E., PEREZ-GONZALEZ, A., YANG, Z. Y., SEEFELDT, L. C., DEBEER, S., DEAN, D. R. & DECAMPS, L. 2022. A conformational role for NifW in the maturation of molybdenum nitrogenase P-cluster. *Chem Sci*, 13, 3489-3500.
- VARADI, M., ANYANGO, S., DESHPANDE, M., NAIR, S., NATASSIA, C., YORDANOVA, G., YUAN, D., STROE, O., WOOD, G., LAYDON, A., ZIDEK, A., GREEN, T., TUNYASUVUNAKOOL, K., PETERSEN, S., JUMPER, J., CLANCY, E., GREEN, R., VORA, A., LUTFI, M., FIGURNOV, M., COWIE, A., HOBBS, N., KOHLI, P., KLEYWEGT, G., BIRNEY, E., HASSABIS, D. & VELANKAR, S. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*, 50, D439-D444.
- WANG, L., ZHANG, L., LIU, Z., ZHAO, D., LIU, X., ZHANG, B., XIE, J., HONG, Y., LI, P., CHEN, S., DIXON, R. & LI, J. 2013. A Minimal Nitrogen Fixation Gene Cluster from *Paenibacillus* sp. WLY78 Enables Expression of Active Nitrogenase in *Escherichia coli*. *PLoS Genetics*, 9, 1-11.
- WESTERMANN, B. & NEUPERT, W. 2000. Mitochondria-targeted green fluorescent proteins: convenient tools for the study of organelle biogenesis in *Saccharomyces cerevisiae*. *Yeast*, 16, 1421-7.
- WIIG, J. A., HU, Y. & RIBBE, M. W. 2011. NifEN-B complex of *Azotobacter vinelandii* is fully functional in nitrogenase FeMo cofactor assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 8623-8627.
- WILCOXEN, J., ARRAGAIN, S., SCANDURRA, A. A., JIMENEZ-VICENTE, E., ECHAVARRI-ERASUN, C., POLLMANN, S., BRITT, R. D. & RUBIO, L. M. 2016. Electron Paramagnetic Resonance Characterization of Three Iron-Sulfur Clusters Present in the Nitrogenase Cofactor Maturase NifB from *Methanocaldococcus infernus*. *Journal of the American Chemical Society*, 138, 7468-7471.
- WOLK, C. P. 1996. Heterocyst formation. *Annu Rev Genet*, 30, 59-78.
- YANG, J., XIE, X., XIANG, N., TIAN, Z. X., DIXON, R. & WANG, Y. P. 2018. Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. *Proc Natl Acad Sci U S A*, 115, E8509-E8517.
- YUVANIYAMA, P., AGAR, J. N., CASH, V. L., JOHNSON, M. K. & DEAN, D. R. 2000. NifS-directed assembly of a transient [2Fe-2S] cluster within the NifU protein. *Proceedings of the National Academy of Sciences*, 97, 599-604.
- ZEYTUN, A., SIKORSKI, J., NOLAN, M., LAPIDUS, A., LUCAS, S., HAN, J., TICE, H., CHENG, J. F., TAPIA, R., GOODWIN, L., PITLUCK, S., LIOLIOS, K., IVANOVA, N., MAVROMATIS, K., MIKHAILOVA, N., OVCHINNIKOVA, G., PATI, A., CHEN, A., PALANIAPPAN, K., NGATCHOU-DJAO, O. D., LAND, M., HAUSER, L., JEFFRIES, C. D., HAN, C., DETTER, J. C., UBLER, S., ROHDE, M., TINDALL, B. J., GOKER, M., WIRTH, R., WOYKE, T., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ, P., KLENK, H. P. & KYRPIDES, N. C. 2011.

- Complete genome sequence of *Hydrogenobacter thermophilus* type strain (TK-6). *Stand Genomic Sci*, 4, 131-43.
- ZHAO, X., LI, G. & LIANG, S. 2013. Several affinity tags commonly used in chromatographic purification. *J Anal Methods Chem*, 2013, 581093.
- ZHENG, L., CASH, V. L., FLINT, D. H. & DEAN, D., R. 1998. Assembly of Iron-Sulfur Clusters. Identification of an iscSUA-hscBA-fdx gene cluster from *Azotobacter vinelandii*. *The Journal of Biological Chemistry*, 273, 13264-13272.
- ZHENG, L. & DEAN, D. R. 1994. Catalytic formation of a nitrogenase iron-sulfur cluster. *Journal of Biological Chemistry*, 269, 18723-18726.
- ZHENG, L., WHITE, R. H., CASH, V. L., JACK, R. F. & DEAN, D. 1993. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proceedings of the National Academy of Sciences*, 90, 2754-2758.
- ZHENG, L., WHITE, R. H. & DEAN, D. R. 1997. Purification of the *Azotobacter vinelandii* nifV-encoded homocitrate synthase. *J Bacteriol*, 179, 5963-6.

ANNEX

ANNEX I: NIFEN LIBRARY (EXCEL FILE NIFEN LIBRARY)

ANNEX II: PUBLICATION