PNAS RESEARCH ARTICLE | BIOCHEMISTRY

Heterologous synthesis of the complex homometallic cores of nitrogenase P- and M-clusters in *Escherichia coli*

Robert Quechol[a,1], Joseph B. Solomon[a,b,1], Yiling A. Liu[a,1], Chi Chung Lee[a], Andrew J. Jasniewski[a], Kamil Górecki[a], Paul Oyala[c], Britt Hedman[d,2], Keith O. Hodgson[d,e,2], Markus W. Ribbe[a,b,2], and Yilin Hu[a,2]

Contributed by Keith O. Hodgson; received August 26, 2023; accepted September 28, 2023; reviewed by Sean J. Elliott and Russ Hille

Nitrogenase is an active target of heterologous expression because of its importance for areas related to agronomy, energy, and environment. One major hurdle for expressing an active Mo-nitrogenase in *Escherichia coli* is to generate the complex metalloclusters (P- and M-clusters) within this enzyme, which involves some highly unique bioinorganic chemistry/metalloenzyme biochemistry that is not generally dealt with in the heterologous expression of proteins via synthetic biology; in particular, the heterologous synthesis of the homometallic P-cluster ([Fe_8S_7]) and M-cluster core (or L-cluster; [Fe_8S_9C]) on their respective protein scaffolds, which represents two crucial checkpoints along the biosynthetic pathway of a complete nitrogenase, has yet to be demonstrated by biochemical and spectroscopic analyses of purified metalloproteins. Here, we report the heterologous formation of a P-cluster-containing NifDK protein upon coexpression of *Azotobacter vinelandii nifD*, *nifK*, *nifH*, *nifM,* and *nifZ* genes, and that of an L-cluster-containing NifB protein upon coexpression of *Methanosarcina acetivorans nifB*, *nifS,* and *nifU* genes alongside the *A. vinelandii fdxN* gene, in *E. coli*. Our metal content, activity, EPR, and XAS/EXAFS data provide conclusive evidence for the successful synthesis of P- and L-clusters in a nondiazotrophic host, thereby highlighting the effectiveness of our metallocentric, divide-and-conquer approach that individually tackles the key events of nitrogenase biosynthesis prior to piecing them together into a complete pathway for the heterologous expression of nitrogenase. As such, this work paves the way for the transgenic expression of an active nitrogenase while providing an effective tool for further tackling the biosynthetic mechanism of this important metalloenzyme.

nitrogenase | NifB | NifDK | EPR | XAS

Significance

The heterologous synthesis of the two complex metalloclusters of nitrogenase has yet to be demonstrated through characterization of purified, cluster-containing proteins. Here, we specifically tackle the heterologous synthesis of the homometallic cores of the nitrogenase clusters in *E. coli* and provide conclusive biochemical and spectroscopic evidence for the successful synthesis of these structurally unique metallocenters in a foreign host. This work highlights the effectiveness of our metallocentric, divide-and-conquer approach that implements critical checkpoints along the nitrogenase biosynthetic pathway for the systematic development of a heterologous expression system of a complete nitrogenase enzyme. Moreover, it provides an alternative yet important platform for us to probe how nitrogenase metalloclusters are built naturally and how their chemistry could be harnessed in the future.

Nitrogenase catalyzes the ambient conversion of N_2 to NH_3 as a key step in the global nitrogen cycle (1). The classical Mo-nitrogenase from *Azotobacter vinelandii* (*SI Appendix*, Fig. S1) is a two-component system comprising a reductase (designated the Fe protein, NifH) and a catalytic component (designated the MoFe protein, NifDK) (2, 3). NifH is a homodimer containing a subunit-bridging [Fe_4S_4] cluster and a magnesium adenosine triphosphate (MgATP)-binding site per subunit, whereas NifDK is an $\alpha_2\beta_2$-tetramer containing a pair of complex metalloclusters per αβ-dimer: a P-cluster ([Fe_8S_7]) that is bridged at the α/β-subunit interface and an M-cluster (or cofactor; [(*R*-homocitrate)$MoFe_7S_9C$]) that is situated within the α-subunit (4–8). The two components form a functional complex during catalysis, allowing electrons to flow from the [Fe_4S_4] cluster of NifH, through the P-cluster, to the M-cluster of NifDK to enable substrate reduction (*SI Appendix*, Fig. S1) (2, 4, 5). Underlying the catalytic prowess of the Mo-nitrogenase are its P- and M-clusters, arguably two of the most complicated metallocenters found in nature. The complexity of these metallocenters, as well as their ability to undergo facile redox changes, renders the Mo-nitrogenase highly versatile in catalysis. Other than N_2, nitrogenase can reduce a wide range of small molecules, including CO, C_2H_2, CN^-, and N_3^- (9–11). Of particular note is the ability of nitrogenase to convert CO to hydrocarbons (12), such as C_2H_4, C_2H_6, C_3H_6, and C_3H_8, in a reaction that is analogous to the conversion of N_2 to NH_3 by the same enzyme (13). Importantly, the reactions of CO- and N_2-reduction by nitrogenase parallel the industrial Fischer–Tropsch (14, 15) and Haber–Bosch (16, 17) processes for the production of carbon fuels and ammonia, respectively; however, contrary to their industrial parallels, the enzymatic processes occur at ambient conditions (instead of high temperatures and/or pressures) and use H^+/e^- (instead of H_2) as the reducing equivalents (13), making nitrogenase an attractive candidate for heterologous expression in genetically amenable hosts, such as *Escherichia coli*, for the future development of bioreactors to harness the reducing power of this unique metalloenzyme.

One major hurdle for the expression of an active Mo-nitrogenase in *E. coli* is to generate fully assembled P- and M-clusters in the catalytic NifDK component of this enzyme. Previous efforts to express nitrogenase in *E. coli* (18–20) involve transfer of a whole set of

Author contributions: B.H., K.O.H., M.W.R., and Y.H. designed research; R.Q., J.B.S., Y.A.L., C.C.L., A.J.J., K.G., and P.O. performed research; R.Q., J.B.S., Y.A.L., C.C.L., A.J.J., K.G., P.O., B.H., K.O.H., M.W.R., and Y.H. analyzed data; and B.H., K.O.H., M.W.R., and Y.H. wrote the paper.

Reviewers: S.J.E., Boston University; and R.H., University of California Riverside.

The authors declare no competing interest.

[1]R.Q., J.B.S., and Y.A.L. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: bhedman@stanford.edu, hodgsonk@stanford.edu, mribbe@uci.edu, or yilinh@uci.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2314788120/-/DCSupplemental.

Published October 23, 2023.

nif (or equivalent) genes into this nondiazotrophic host, followed by analysis of whole-cell expression of nitrogenase by immunoblotting, C_2H_2 reduction assays, and $^{15}N/^{14}N$ ratio-derived calculations. Despite progress made using this approach (21), the heterologously expressed nitrogenase proteins are yet to be purified for combined biochemical and spectroscopic analyses, which is crucial for addressing this metalloprotein-specific problem. As such, the feasibility of synthesizing the high-nuclearity metalloclusters of nitrogenase in a non-nitrogen-fixing organism remains elusive. Moreover, transfer of a complete *nif* gene set into a non-native host like *E. coli* does not allow for the detection of any jam points of nitrogenase assembly that are caused by a) the inability of FeS-cluster synthesis to keep up with peptide synthesis and b) the fact that certain *nif* genes encoding key components of the assembly machinery (e.g., *nifB* from *A. vinelandii*) simply do not express as functional proteins in *E. coli*. An alternative strategy that could circumvent the problems associated with the existing approach is to individually target the heterologous expression of the P- and M-clusters prior to combining them for the expression of a complete nitrogenase enzyme. Such a divide-and-conquer approach implements crucial checkpoints along the biosynthetic pathway of nitrogenase and, coupled with a metallocentric theme that focuses on the spectroscopic and biochemical analyses of the purified metalloproteins, allows for a conclusive demonstration of the successful expression of nitrogenase proteins in a foreign host.

Of the two NifDK-associated clusters, biosynthesis of the P-cluster occurs in situ at its target location between the α- and β-subunits of NifDK; additionally, it occurs prior to the incorporation of the M-cluster, resulting in a P-cluster-containing NifDK species with unoccupied cofactor-binding sites (Fig. 1*A*) (10, 22). In the case of *A. vinelandii*, formation of the P-clusters on NifDK begins with synthesis of small $[Fe_4S_4]$ clusters by NifS and NifU, followed by delivery of a pair of $[Fe_4S_4]$ clusters to each α/β interface of the tetrameric NifDK (Fig. 1*A*, ①) (10, 22). Subsequently, the two $[Fe_4S_4]$-like cluster pairs (P*-clusters, or precursors) are coupled into two $[Fe_8S_7]$ clusters (P-clusters) at the two α/β interfaces, resulting in a P-cluster-replete, yet M-cluster-depleted apo-form of NifDK (designated $NifDK^{apo}$) (Fig. 1*A*, ②–④) (23–25). Interestingly, formation of the two homometallic P-clusters occurs one at a time in the two αβ-dimers of NifDK, with NifH (the reductase) required for the synthesis of both P-clusters (Fig. 1*A*, ①, ③), and NifZ (a nonmetalloprotein) specifically involved in the synthesis of the second P-cluster (Fig. 1*A*, ②) (26–29). Moreover, maturation of the P-clusters induces a conformational change that opens the cofactor-binding site in the α-subunit for the insertion of the externally synthesized M-cluster, thereby completing the assembly of a P- and M-cluster-replete, holo-form of NifDK (designated $NifDK^{holo}$) (Fig. 1*A*, ⑤) (10, 22). As such, generation of $NifDK^{apo}$ represents the first crucial step, or checkpoint, toward the successful generation of the catalytically competent $NifDK^{holo}$ conformation.

Biosynthesis of the M-cluster, on the other hand, occurs ex situ prior to the incorporation of this cofactor into its target location within the α-subunit of the P-cluster-containing, but cofactor-deficient $NifDK^{apo}$ (Fig. 1*B*) (10, 22, 29, 30). Sharing the early events with the P-cluster assembly pathway that involve the synthesis of small $[Fe_4S_4]$ clusters by NifS/U, the M-cluster assembly pathway continues with transfer of a pair of $[Fe_4S_4]$ clusters to NifB for the subsequent conversion to an $[Fe_8S_9C]$ cofactor core (8, 31–35). This step is followed by maturation of the homometallic, $[Fe_8S_9C]$ cofactor core (8, 36–38) into a fully assembled M-cluster on NifEN via NifH-mediated insertion of Mo and homocitrate (39–41), and delivery of the M-cluster from NifEN to $NifDK^{apo}$ to yield a catalytically competent $NifDK^{holo}$ protein (25, 39). Of all events that occur during cofactor biosynthesis, the reactions catalyzed by NifB are the most crucial, as they dictate the transformation of the standard $[Fe_4S_4]$ building blocks into a highly unusual cofactor core and therefore represent another crucial checkpoint for the successful synthesis of a catalytically competent $NifDK^{holo}$ species.

A radical *S*-adenosyl-*L*-methionine (SAM) enzyme, NifB contains a SAM-binding $[Fe_4S_4]$ cluster (designated SAM- or RS-cluster) that is coordinated by Cys ligands from the canonical CXXXCXX motif, as well as two $[Fe_4S_4]$ modules of a so-called K-cluster (designated K1- and K2-clusters, respectively) that are coordinated by additional, conserved Cys and His ligands (42). Biochemical, spectroscopic, and structural studies of the NifB species from *A. vinelandii*, *Methanosarcina acetivorans,* and *Methanobacterium thermoautotrophicum* have revealed a flexible conformation of this protein and an undercoordinated ligation pattern of its three $[Fe_4S_4]$ clusters that accommodate the dynamic cluster transformation. Moreover, these studies have led to the proposal of a pathway of NifB-catalyzed cofactor-core formation (Fig. 1*B*) (10, 42), which involves an S_N2-type methyl-transfer from one SAM molecule to the K2-cluster, followed by hydrogen atom abstract from the K2-bound methyl group by a 5′-dA• radical that is derived from the homolytic cleavage of a second SAM molecule. The resultant, K2-bound methylene radical then initiates a radical-based coupling/rearrangement of K1- and K2-clusters while undergoing deprotonation/dehydrogenation to yield a μ_6 interstitial carbide, and this event is accompanied by the insertion of a sulfite-derived ninth belt sulfide, leading to the formation of an $[Fe_8S_9C]$ cofactor core (designated L-cluster) that is structurally indistinguishable from a mature M-cluster but has an Fe atom in place of Mo/homocitrate at one end of the cluster.

The early appearances and crucial roles of P- and L-clusters in nitrogenase biosynthesis make them the logical first targets of heterologous expression in a foreign host like *E. coli*. However, while the essential set of genes required for the expression of an $NifDK^{apo}$ species has been well established in *A. vinelandii*, the successful transfer of such a feat to *E. coli* has yet to be demonstrated through purification and characterization of a P-cluster-replete form of $NifDK^{apo}$. With respect to NifB, despite the successful expression of methanogen NifB species in *E. coli* (31), none of these heterologously expressed proteins carried L-clusters in the as-purified forms and, consequently, required in vitro FeS reconstitution and cluster maturation procedures prior to acquiring the ability to serve as a competent donor of L-clusters that could be subsequently maturated and used for the reconstitution and activation of $NifDK^{apo}$ (32–35, 43).

In this work, we report the successful heterologous formation of P- and L-clusters, two high-nuclearity, homometallic metallocenters essential for the functionality of Mo-nitrogenase, in *E. coli*. Specifically, we demonstrate that a P-cluster-replete, yet M-cluster-depleted $NifDK^{apo}$ protein is generated upon coexpression of the *nifD* and *nifK* genes alongside the *nifH*, *nifM* and *nifZ* genes from *A. vinelandii* in *E. coli*, which can be activated upon cofactor incorporation. Moreover, we show that an L-cluster-containing NifB protein is generated upon coexpression of the *nifB*, *nifS,* and *nifU* genes from *M. acetivorans* alongside the *fdxN* gene from *A. vinelandii* in *E. coli*, which can directly serve as an L-cluster donor in the as-purified state. Through combined metal content, activity, electron paramagnetic resonance (EPR) and X-ray absorption spectroscopy (XAS)/extended X-ray absorption fine structure (EXAFS) analyses of these proteins, we provide conclusive evidence for the heterologous formation of a P-cluster ($[Fe_8S_7]$) and an L-cluster (a $[Fe_8S_9C]$ core of the M-cluster) in a nondiazotrophic host, thereby illustrating the effectiveness of our metallocentric, divide-and-conquer approach that individually tackles the key events of nitrogenase

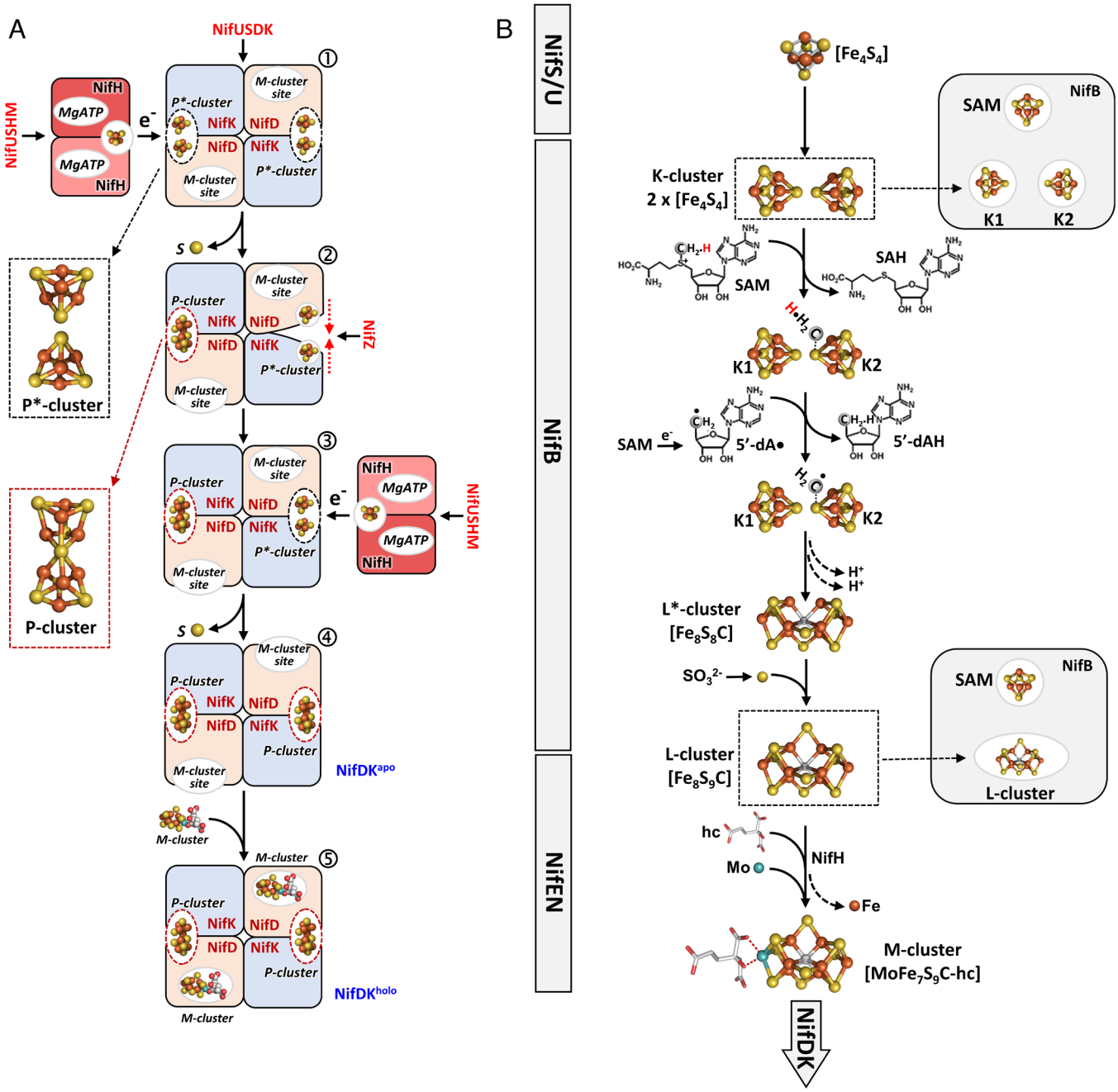

Fig. 1. Biosynthesis of the Mo-nitrogenase. Biosynthesis of the Mo-nitrogenase involves (*A*) the in situ assembly of the P-cluster at its target binding site in NifDK and (*B*) the ex situ assembly of the M-cluster outside its target binding site in NifDK prior to its incorporation. (*A*) The in situ assembly of P-clusters begins with the generation of a P*-cluster (or precursor) replete, yet M-cluster depleted form of NifDK (①), followed by stepwise conversion of the two P*-clusters (2×[Fe_4S_4]) into two mature P-clusters ([Fe_8S_7]) in the two αβ-dimers of NifDK, with NifH required for the maturation of both P-clusters (①, ③), and NifZ required specifically for the maturation of the second P-cluster (②). These events result in a P-cluster replete, yet M-cluster depleted form of NifDK (designated $NifDK^{apo}$; ④), which has its cofactor-binding sites opened up via a conformational change induced by P-cluster maturation. Subsequent insertion of the M-clusters into $NifDK^{apo}$ results in the formation of a P- and M-cluster replete form of NifDK (designated $NifDK^{holo}$; ⑤), thereby completing the biosynthesis of the catalytic component of Mo-nitrogenase. The Nif proteins involved in this process are indicated in red fonts. The atoms of the metalloclusters are colored as described in *SI Appendix*, Fig. S1. (*B*) The ex situ assembly of M-clusters centers on the radical SAM-dependent transformation of a pair of [Fe_4S_4] clusters (K-cluster, comprising K1- and K2-modules) into a [Fe_8S_9C] cofactor core (L-cluster) on NifB, which begins with an S_N2-type methyltransfer from one SAM molecule to a K2-associated sulfur atom. Subsequently, a second SAM molecule undergoes homolytic cleavage to yield a 5′-deoxyadenosyl radical (5′-dA•) for the hydrogen atom abstraction from the K2-bound methyl group, resulting in a K2-bound methylene radical. Continued deprotonation of the K2-bound methylene radical gives rise to an interstitial carbide concomitant with the coupling and rearrangement of K1 and K2 into a [Fe_8S_8C] cluster (L*-cluster), followed by insertion of a sulfite-derived ninth sulfur that leads to the formation of a complete [Fe_8S_9C] core (L-cluster). Further maturation of the L-cluster to an M-cluster occurs on NifEN, where an apical Fe atom of the L-cluster is replaced by Mo/homocitrate (hc) in a NifH-dependent process. Upon completion on NifEN, the M-cluster is transferred to its target binding site in $NifDK^{apo}$, thereby completing the biosynthesis of $NifDK^{holo}$ (Fig. 1*A*). The Nif genes involved in this process are indicated in gray tabs. The atoms of the metalloclusters are colored as described in *SI Appendix*, Fig. S1.

assembly prior to piecing them together into a complete pathway for the heterologous expression of nitrogenase. As such, our work paves the way for the transgenic expression of an active, purifiable nitrogenase enzyme while providing an effective tool for further elucidating details of the biosynthetic mechanism of nitrogenase.

Results and Discussion

Heterologous Synthesis of an [Fe_8S_7] P-Cluster on NifDK in *E. coli*. Our previous work on the biosynthesis of the Mo-nitrogenase of *A. vinelandii* has led to the identification of NifH and NifZ as two essential factors for the maturation of P-clusters on the catalytic NifDK component. Based on this knowledge, we set out to generate NifDK^{apo} from *A. vinelandii* heterologously in *E. coli* by coexpressing *Av*NifDK with *Av*IscS/U and i) *Av*NifZ alone, ii) *Av*NifH/M alone or iii) both *Av*NifZ and *Av*NifH/M in *E. coli*. Such a strategy takes advantage of the recent success in the heterologous expression of a fully active *Av*NifH upon coexpression with *Av*NifM in *E. coli* as well as the well-estalished ability of IscS/U (a homolog to NifS/U) to supply small FeS building blocks for the synthesis of metallocenters in both *Av*NifH and *Av*NifDK. Excitingly, all three heterologously expressed NifDK species were isolated as soluble, brown proteins at a yield of ~100 mg protein per 50 g wet cells. As expected, all three proteins were $\alpha_2\beta_2$-tetramers comprising α- and β-subunits of ~56 kDa and ~59 kDa, respectively (*SI Appendix*, Fig. S2*A*); additionally, all of them were free of M-clusters due to a lack of the cofactor assembly machinery in *E. coli*.

Biochemical and spectroscopic evidence for P-cluster formation on NifDK. Upon reconstitution with solvent-extracted M-clusters, the NifDK proteins coexpressed with *Av*NifZ (designated NifDK^{NifZ}), *Av*NifH/M (designated NifDK^{NifHM}), and both *Av*NifH/M and *Av*NifZ (designated NifDK^{NifHMZ}) showed i) 1%, 13%, and 81%, respectively, of C_2H_2-reducing activity; and ii) 1%, 9%, and 70%, respectively, of N_2-reducing activity, relative to those of the P-cluster-replete, yet M-cluster-depleted *Av*NifDK^{apo} protein isolated from the native *A. vinelandii* host (Fig. 2*A* and *SI Appendix*, Table S1). As incorporation of the M-clusters only occurs upon formation of the P-clusters (Fig. 1*A*, ⑤), the activities of the M-cluster-reconstituted NifDK proteins should reflect their P-cluster contents. Indeed, the oxidized NifDK^{NifZ}, NifDK^{NifHM}, and NifDK^{NifHMZ} proteins displayed the P^{2+} (or P^{OX})-specific, g = 11.8 signal (44) at 1%, 12% and 81%, respectively, of the signal intensity of *Av*NifDK^{apo} (Fig. 2*B*), which aligned well with the percentages of C_2H_2- and N_2-reducing activities of these proteins relative to those of *Av*NifDK^{apo} (Fig. 2*A*). Moreover, the reduced NifDK^{NifHM} and NifDK^{NifHMZ} proteins displayed a precursor (a [Fe_4S_4]-like cluster pair)-specific, S=1/2 signal (24, 27) at 78% and 11%, respectively, of the signal intensity of NifDK^{NifZ} (Fig. 2*C*), consistent with a lack of precursor conversion in NifDK^{NifZ} that contrasted a moderate- or high-level precursor conversion in NifDK^{NifHM} or NifDK^{NifHMZ}. Taken together, these results are in strong agreement with those derived from the previous studies of the native *A. vinelandii* system (Fig. 1*A*); specifically, they reveal a dual requirement of NifH and NifZ for P-cluster assembly as well as a prerequisite for NifH to act prior to NifZ in this process, which would account for a lack of precursor conversion in the sole presence of NifZ (as the action of NifH precedes that of NifZ) and a partial precursor conversion in the sole presence of NifH (as NifZ is required alongside NifH to complete this process).

Subsequent Fe K-edge EXAFS analyses provided further insights into the structures of the cluster species in these heterologously expressed NifDK proteins (Fig. 3 and *SI Appendix*, Tables S2–S5). Consistent with the activity- and EPR-based observations of a successful conversion of the precursor to a mature P-cluster in NifDK^{NifHMZ}, the EXAFS fit of NifDK^{NifHMZ} (Fig. 3 *A* and *B*, red traces) is similar to that of the P-cluster-replete, yet M-cluster-depleted *Av*NifDK^{apo} (Fig. 3 *A* and *B*, blue traces) in that the first main component (as displayed in the FT) contains one Fe---Fe scatterer at ~2.5 Å, and one Fe–S component at ~2.3 Å (*SI Appendix*, Tables S2 and S3), while an Fe---Fe scatter component at ~2.7 Å is not visible in the *Av*NifDK^{apo} FT (and shows high disorders in the fits). However, contrary to *Av*NifDK^{apo} (Fig. 3*A*, blue trace), the spectrum of NifDK^{NifHMZ} displayed an additional FT feature at R+Δ ~2.3 Å (Fig. 3*A*, red trace) that was also present in the spectra of NifDK^{NifZ} (Fig. 3*A*, black trace) and NifDK^{NifHM} (Fig. 3*A*, green trace), although the magnitude of this feature was substantially reduced in the spectrum of NifDK^{NifHMZ} as compared to those in the spectra of NifDK^{NifZ} and NifDK^{NifHM}. The EXAFS data of NifDK^{NifZ} and NifDK^{NifHM} were best fit with one Fe---Fe scatterer at ~2.7 Å, one Fe---Fe scatterer at ~2.5 Å, and one Fe–S component at ~2.3 Å (*SI Appendix*, Tables S4 and S5). Given the close resemblance of these fits to that of a [Fe_4S_4] cluster (such as

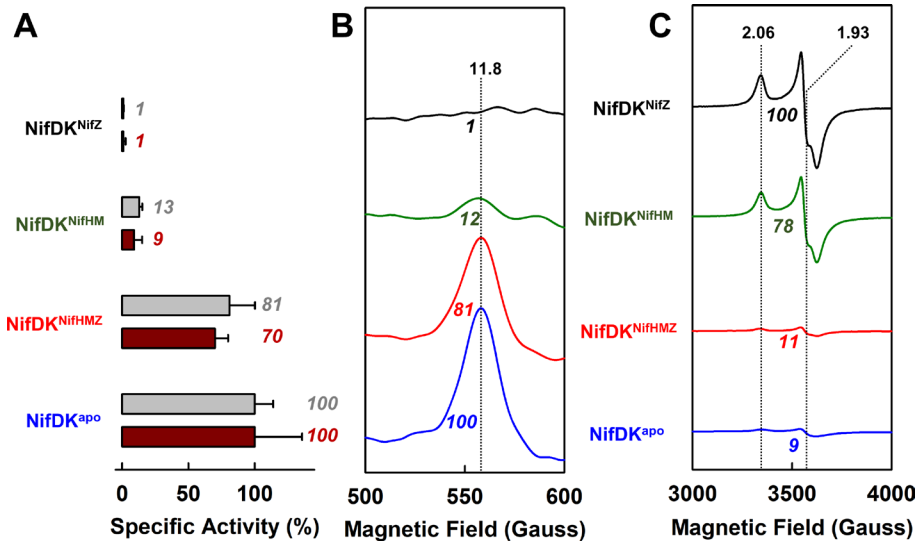

Fig. 2. Biochemical and spectroscopic analyses of NifDK proteins. (*A*) Specific activities of C_2H_2-reduction (to C_2H_4; light gray) and N_2-reduction (to NH_3; dark red) by NifDK^{NifZ}, NifDK^{NifHM}, and NifDK^{NifHMZ} as compared to those by NifDK^{apo} upon reconstitution with M-clusters. The activities (see *SI Appendix*, Table S1, for details) are expressed in percentages in this figure, with the activities of NifDK^{apo} set as 100% and the activities of NifDK^{NifZ}, NifDK^{NifHM}, and NifDK^{NifHMZ} calculated relative to those of NifDK^{apo}. (*B* and *C*) Parallel (*B*) and perpendicular (*C*) mode EPR spectra of NifDK^{NifZ} (black), NifDK^{NifHM} (green), and NifDK^{NifHMZ} (red) as compared to those of NifDK^{apo} (blue) in the IDS-oxidized (*B*) and dithionite-reduced (*C*) states. The specific activities (*A*) and signal intensities of the EPR spectra (*B* and *C*) are normalized based on the Fe contents of NifDK^{NifZ} (4.0 ± 0.4 mol Fe/mol protein), NifDK^{NifHM} (4.8 ± 0.7 mol Fe/mol protein), NifDK^{NifHMZ} (8.3 ± 1.2 mol Fe/mol protein), and NifDK^{apo} (12.7 ± 0.2 mol Fe/mol protein).

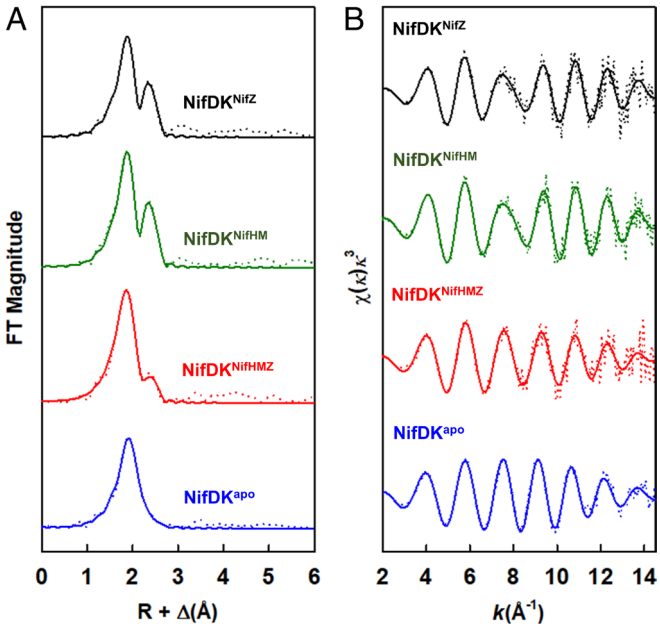

Fig. 3. Fe K-edge XAS analysis of NifDK proteins. Shown are the Fourier-transformed (FT) (*A*) and k^3-weighted (*B*) EXAFS data (dotted) and best fits (solid) of NifDK^{NifZ} (black), NifDK^{NifHM} (green), and NifDK^{NifHMZ} (red) as compared to those of NifDK^{apo} (blue). See *SI Appendix*, Tables S2–S5, for details of fits.

that in *Av*NifH) (45), the notable spectral deviation of NifDK^{NifHMZ} from NifDK^{NifZ} and NifDK^{NifHM}—as reflected by the substantially decreased intensity of its R+Δ ~2.3 Å feature—aligned well with a high-percentage (up to ~80%) conversion of the available precursors to mature P-clusters in NifDK^{NifHMZ} as suggested by the EPR and activity data (Fig. 2).

Dual requirement of NifH and NifZ for P-cluster synthesis. Interestingly, the Fe contents of NifDK^{NifHM} and NifDK^{NifZ} were only 58% and 48%, respectively, of that of NifDK^{NifHMZ} (*SI Appendix*, Table S1). Such a discrepancy highlights the instability of the partially assembled, intermediary conformations of NifDK (i.e., NifDK^{NifHM} and NifDK^{NifZ}) wherein the P-cluster sites at the α/β-subunit interfaces contain, at least in part, modular [Fe_4S_4]-like cluster pairs that are prone to separation and easily lost from the protein. This problem is particularly pronounced in a non-nitrogen-fixing host like *E. coli*, which does not have the sophisticated oxygen-protection mechanism that is usually employed by a native nitrogen-fixing host like *A. vinelandii*. The increased oxygen lability in *E. coli*, coupled with a potential shortage in the FeS-cluster supply that is in part caused by unoptimized ratios of the expressed *nif* (and related) gene products in this heterologous expression host, could very well contribute to a lower P-cluster content of NifDK^{NifHM} (13%) than that of its half-assembled *A. vinelandii* counterpart (50%). Regardless, the fact that the P-cluster is absent from NifDK^{NifZ} but present in NifDK^{NifHM} or NifDK^{NifHMZ} clearly illustrates a sequential participation of NifH prior to NifZ in P-cluster assembly (Fig. 1*A*). More importantly, the observation of a substantially increased P-cluster yield of NifDK^{NifHMZ} to as high as 80% of that of its native *A. vinelandii* counterpart demonstrates a concerted effort of NifH and NifZ in preventing loss of precursors and promoting formation of P-clusters (Fig. 1*A*).

The crucial importance of NifH and NifZ for P-cluster maturation raises the question of their respective roles in this process. As for NifH, it is likely that this protein interacts with NifDK^{apo} during P-cluster assembly in a similar manner to that with NifDK^{holo} during substrate reduction, both of which enable adenosine triphosphate (ATP)-dependent electron transfer from NifH to its NifDK partner; only in the case of the former, the electrons donated by NifH are used for the reductive coupling of the [Fe_4S_4] pair to an [Fe_8S_7] P-cluster at the α/β subunit interface. With respect to NifZ, a high-confidence structural model generated with AlphaFold reveals a pseudodimeric architecture of this protein, with a small B-barrel present in each monomeric half (*SI Appendix*, Fig. S3). Of note, the two B-barrels of NifZ show a nearly perfect structural overlap with each other, which is typical for proteins with binding partner(s). Such a pseudodimeric conformation of NifZ would be consistent with our previous proposal that this small protein serves as a chaperone that zips together the second αβ-dimer of NifDK that is pulled apart upon fusion of the first [Fe_4S_4] pair into a P-cluster at the first αβ-dimer interface, thereby assisting NifH in the coupling of the second [Fe_4S_4] cluster pair into a mature P-cluster at the second α/β-dimer interface (Fig. 1*A*). The exact mode-of-action of NifZ in this process, however, is yet to be fully elucidated.

Identification of NifW as a potential negative effector of P-cluster assembly. The successful in vivo formation of P-clusters in *E. coli* not only lends strong support to the P-cluster assembly mechanism derived from studies of the native *A. vinelandii* system but also provides a clean platform for evaluating the impact of other *nif* gene products on this process without the interference of nonessential, *nif*-encoded or related proteins and/or the complication originating from the arrangement of *nif* gene clusters in the native nitrogen-fixing organisms. One such *nif* gene product that has been implicated in the functionality of nitrogenase is NifW, the encoding gene of which is located immediately adjacent to, and upstream of those encoding NifZ and NifM in an arrangement of *nifWZM* in the genomes of diazotrophic organisms like *A. vinelandii* and *Klebsiella pneumoniae*. Previous studies demonstrated that disruptions of the *nifW* gene in *A. vinelandii* and *K. pneumoniae* resulted in accumulation of NifDK species with decreased substrate-reducing activities (46), leading to the proposed role of NifW in augmenting the functionality of NifDK. Introduction of NifW into our NifDK^{apo}-expressing *E. coli* strains, however, revealed an opposite, negative impact of NifW on P-cluster assembly.

Upon reconstitution with M-clusters, the NifDK protein coexpressed with *Av*NifH/M and *Av*NifW (designated NifDK^{NifHMW}) showed an 8% decrease in substrate-reducing activity relative to its counterpart coexpressed without *Av*NifW (i.e., NifDK^{NifHM}); likewise, the NifDK protein coexpressed with *Av*NifH/M/Z and *Av*NifW (designated NifDK^{NifHMZW}) showed a 30% decrease in substrate-reducing activity relative to its counterpart coexpressed without *Av*NifW (i.e., NifDK^{NifHMZ}) (*SI Appendix*, Fig. S4*A*). The decrease in the activities of these NifDK species correlated specifically with the reduction of their P-cluster contents, as NifDK^{NifHMW} and NifDK^{NifHMZW} showed a decrease by 10% and 26%, respectively, in the magnitude of the P-cluster-specific signal relative to NifDK^{NifHM} (*SI Appendix*, Fig. S4*B*, blue vs. green) and NifDK^{NifHMZ} (*SI Appendix*, Fig. S4*B*, brown vs. red). Moreover, accompanying the decrease in the magnitude of the P-cluster-specific signal, NifDK^{NifHMW} and NifDK^{NifHMZW} displayed an increase by 20% and 45%, respectively, in the magnitude of the precursor-specific signal relative to NifDK^{NifHM} (*SI Appendix*, Fig. S4*C*, blue vs. green) and NifDK^{NifHMZ} (*SI Appendix*, Fig. S4*C*, brown vs. red).

Our observation of a seemingly contradictory effect of NifW to that reported previously (46) could be rationalized by the well-known polar effect, or an impact on the expression of the downstream genes, upon mutation of the upstream gene. In this case, a disruption of the upstream *nifW* gene in the *A. vinelandii* genome as described in the earlier study (46) may very well down-regulate the expression of the downstream *nifZ*

and/or *nifM* gene(s), particularly given the apparent, coupled transcription of *nifW* and *nifZ* in this organism, as well as the recent observation of association of NifW to the NifDK species expressed in a *nifZ*-deletion strain of *A. vinelandii* (47). As such, the possible false-positive effect of NifW—as indicated by the decreased activity of NifDK upon disruption of NifW—could be an indirect effect of a decreased expression of NifZ and/or NifM (and consequently, NifH, in the latter case), the key protein factors for P-cluster assembly. In comparison, introduction of NifW into a nondiazotrophic expression host like *E. coli* allows for a direct assessment of the role of NifW as a potential negative effector for P-cluster assembly, although caution should be taken when comparing results derived from the non-native and native hosts given the lack of regulatory mechanisms for gene expression in the case of the former. The specific target of NifW, be it NifZ, NifM/H or NifDK, awaits further investigation.

Heterologous Synthesis of an [Fe_8S_9C] L-Cluster on NifB in *E. coli*. Our previous efforts to coexpress the *M. acetivorans* NifB protein with *A. vinelandii* IscS/U in *E. coli* under aerobic conditions resulted in a NifB species (designated $NifB^{IscSU}$) with a low FeS content and no detectable L-cluster donor activity in the as-purified state. This observation suggests an insufficient amount of RS- and/or K-clusters in this NifB protein to sustain the K- to L-cluster conversion in vivo despite the presence of SAM in the *E. coli* host cell. To circumvent the problem of an insufficient FeS content caused by oxygen damage, we expressed $NifB^{IscS/U}$ in *E. coli* under anaerobic conditions (*SI Appendix*, Fig. S2*B*) and observed a minor activity of the as-isolated $NifB^{IscSU}$ to directly serve as an L-cluster donor (i.e., without the *in vitro* incubation of $NifB^{IscSU}$ with SAM) in the maturation assay (Fig. 4*A*). A closer examination revealed an increased metal content (2.5 ± 0.4 mol Fe/mol protein) when $NifB^{IscSU}$ was expressed anaerobically (Fig. 4*B*), which could account for the in vivo formation of a minor amount of L-clusters in this protein; however, the activity of this $NifB^{IscSU}$ species as an L-cluster donor in the maturation assay was still very low, ~5% relative to that of the same protein reconstituted with synthetic [Fe_4S_4] clusters (designated $NifB^{recon}$) (34) and treated with SAM to allow for the K- to L-cluster conversion under in vitro conditions (Fig. 4*A*).

Identification of a matching set of NifS/U as an FeS donor for NifB. To further improve the in vivo formation of L-clusters on NifB, we focused on enhancing the efficiency of FeS-cluster delivery from NifU to NifB during the assembly process (Fig. 1*B*). Considering that *Ma*NifB was previously paired with IscS/U (a NifS/U homolog) from a different organism (*A. vinelandii*) for the expression of $NifB^{IscSU}$, we turned our attention to identifying a matching pair of NifS/U and NifB from the same origin in hopes of improving the interaction between NifU and NifB and thereby improving the efficiency of FeS-cluster delivery from NifU to NifB. Based on the three sets of *Ma*NifS/U (designated NifS1/U1, NifS2/U2, and NifS3/U3, respectively) identified through a search of the genomic database (*SI Appendix*, Fig. S5), we coexpressed each set of *Ma*NifS/U with *Ma*NifB (designated $NifB^{NifS1U1}$, $NifB^{NifS2U2}$, and $NifB^{NifS3U3}$, respectively) in *E. coli* under anaerobic conditions (*SI Appendix*, Fig. S2*B*). Strikingly, the FeS contents of the as-isolated $NifB^{NifS1U1}$, $NifB^{NifS2U2}$, and $NifB^{NifS3U3}$ (Fig. 4*B*) were approximately the same as that of the anaerobically expressed $NifB^{IscSU}$ (Fig. 4*B*); however, the activity of the as-isolated $NifB^{NifS3U3}$ (Fig. 4*A*) as an L-cluster doner was considerably higher (by ~threefold to sevenfold) than those of its as-isolated $NifB^{IscSU}$, $NifB^{NifS1U1}$, and $NifB^{NifS2U2}$ counterparts (Fig. 4*A*). This result implies that the interaction between NifU and NifB not only facilitates the FeS-cluster delivery from the former to the latter, but also impacts the efficiency of the in vivo formation of L-clusters on NifB, likely through a favored distribution of the delivered [Fe_4S_4] clusters toward utilization as K-clusters in the case of $NifB^{NifS3U3}$.

Biochemical and spectroscopic evidence for L-cluster formation on NifB. Having identified *Ma*NifS3/U3 as the most effective partner for *Ma*NifB in the in vivo formation of L-clusters in *E. coli*, we cloned the genes encoding *Ma*NifB and *Ma*NifS3/U3 into to an expression vector with a higher copy number (see *SI Appendix*, Table S6, for constructs used in this work), and coexpressed *Ma*NifB and *Ma*NifS3/U3 with FdxN from *A. vinelandii*, an electron donor proposed to be specifically involved in the cofactor assembly process (48). Given the high degree of sequence identity/homology between *Ma*NifB and *Av*NifB, it is likely that *Av*FdxN would cross-react well with *Ma*NifB and that overexpression of *Av*FdxN could increase the size of the electron pool that is specifically dedicated to the assembly of the

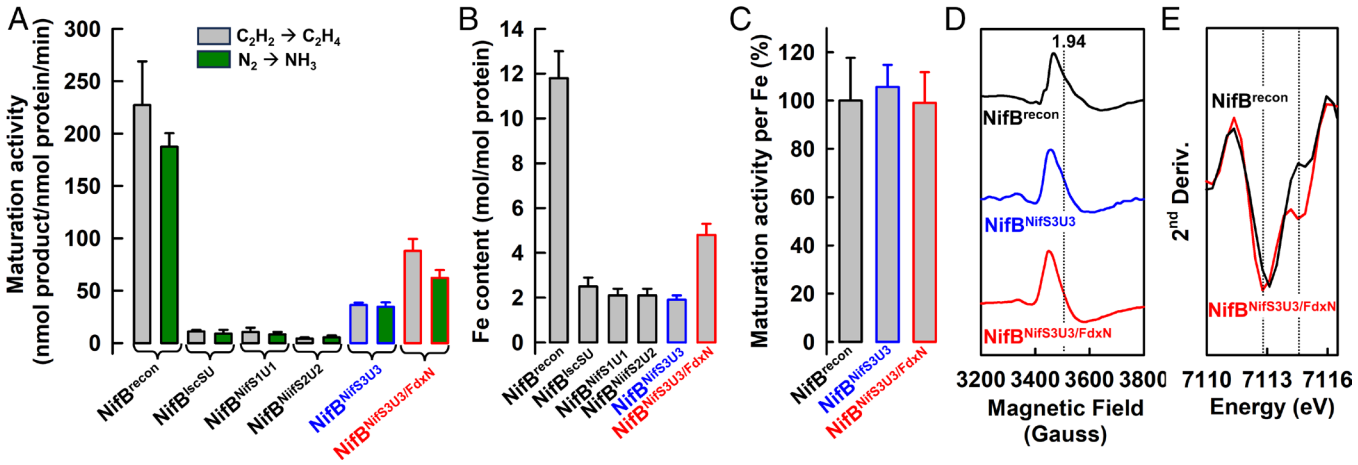

Fig. 4. Biochemical and spectroscopic analyses of NifB proteins. (*A*) Maturation activities (C_2H_2-reduction to C_2H_4, light gray; N_2-reduction to NH_3, green) and (*B*) Fe contents of $NifB^{recon}$, $NifB^{IscSU}$, $NifB^{NifS1U1}$, $NifB^{NifS2U2}$, $NifB^{NifS3U3}$, and $NifB^{NifS3U3/FdxN}$. (*C*) Maturation activities normalized based on the Fe contents (shown in *B*) of $NifB^{recon}$, $NifB^{NifS3U3}$, and $NifB^{NifS3U3/FdxN}$. The normalized activities are expressed in percentages, with the normalized activity of $NifB^{recon}$ set as 100% and the normalized activities of $NifB^{NifS3U3}$ and $NifB^{NifS3U3/FdxN}$ calculated relative to that of $NifB^{recon}$. (*D*) Perpendicular mode EPR spectra of the IDS-oxidized $NifB^{recon}$, $NifB^{NifS3U3}$, and $NifB^{NifS3U3/FdxN}$. The signal intensities of the EPR spectra are normalized based on the Fe contents (shown in *B*). The *g* values are indicated. (*E*) Smoothed second derivatives of the pre-edge regions of the Fe K-edge XAS spectra of $NifB^{recon}$ and $NifB^{NifS3U3/FdxN}$. The peaks at ~7,112.6 eV and ~7,114.5 eV are indicated by gray lines. Note the absence of the feature at ~7,114.5 eV from the spectrum of the FeS-reconstituted, yet SAM-untreated $NifB^{recon}$.

cofactor core on *Ma*NifB. Indeed, compared to its counterpart expressed without FdxN (i.e., NifB^{NifS3U3}; Fig. 4 *A* and *B*), the NifB species coexpressed with FdxN (designated NifB$^{NifS3U3/FdxN}$; Fig. 4 *A* and *B*) demonstrated a concurrent increase of its Fe content (by 2.4-fold) and maturation activity (by 1.7-fold) in the as-isolated state, reflecting a further increase in the in vivo formation of L-clusters that was accomplished through the use of a specific electron donor (i.e., FdxN) for this process. Yet, when normalized based on the Fe content, the as-isolated NifB$^{NifS3U3/FdxN}$ protein was nearly indistinguishable from its NifB^{NifS3U3} counterpart (with a lower Fe content) as well as its SAM-treated, in vitro FeS-reconstituted NifB^{recon} counterpart (with a higher Fe content) as an L-cluster donor (Fig. 4*C*). This observation points to an almost identical conversion yield of the available K-clusters to L-clusters in NifB$^{NifS3U3/FdxN}$, NifB^{NifS3U3} and SAM-treated NifB^{recon} regardless of their disparate cluster occupancies, an assignment supported further by the nearly identical intensities of the L-cluster-specific, g = 1.94 EPR signals displayed by the three IDS-oxidized NifB species upon normalization of their respective Fe contents (Fig. 4*D*) (31).

The presence of L-clusters on NifB$^{NifS3U3/FdxN}$ was further verified by comparing the smoothed second derivative of the pre-edge XAS data of this protein with that of the in vitro reconstituted, but SAM-untreated NifB^{recon} (Fig. 4*E*). Consistent with the presence of unconverted K-cluster (i.e., [Fe_4S_4] clusters) in NifB^{recon} prior to treatment with SAM, the pre-edge spectrum of NifB^{recon} displayed a major peak at ~7,112.6 eV that was characteristic of protein-associated FeS clusters with tetrahedral Fe site geometries (33, 49). In contrast, the pre-edge spectrum of NifB$^{NifS3U3/FdxN}$ displayed, in addition to the peak at ~7,112.6, an additional peak at ~7,114.5 eV that was characteristic of the distinct, intermediary geometry of the L-cluster ([Fe_8S_9C]) between tetrahedral and trigonal pyramidal (33). Clearly, there is a major structural rearrangement of the cluster species on NifB$^{NifS3U3/FdxN}$ upon in vivo K- to L-cluster transformation at the physiological concentration of SAM within the *E. coli* host; yet, the distinct feature of NifB$^{NifS3U3/FdxN}$ at ~7,114.5 eV was somewhat weaker than that reported previously for the SAM-treated NifB^{recon} (33), suggesting a possible presence of unconverted K-clusters (comprising K1- and K2-clusters) in NifB$^{NifS3U3/FdxN}$ in the as-purified state.

Assessment of unconverted K-clusters on NifB. Consistent with the XAS-derived observation, NifB$^{NifS3U3/FdxN}$ demonstrated mixed EPR signals originating from the L-clusters and individual [Fe_4S_4] cluster modules, with features associated with both K1- and K2-clusters being notable in the spectrum of this protein in the dithionite-reduced state (Fig. 5*A*). To assess the presence of K1-clusters in NifB$^{NifS3U3/FdxN}$, we performed pulsed EPR analyses of this protein species in comparison with its in vitro reconstituted, but SAM-untreated NifB^{recon} counterpart (Fig. 5 *B* and *C* and *SI Appendix*, Fig. S6). Previous three-pulse electron spin echo envelope modulation (3P-ESEEM) and two-dimensional hyperfine sublevel correlation (HYSCORE) experiments have led to the assignment of a histidine-derived nitrogen ligand to the K1-cluster, which is lost upon coupling and rearrangement of the K1- and K2-clusters into an L-cluster (32). As expected, the SAM-untreated NifB^{recon} displayed deep modulations in the time domain of the ESEEM spectrum (Fig. 5*B*) and the corresponding intensity between 1 and 8Â MHz in the FFT (Fig. 5*C*), consistent with the hyperfine and quadrupole couplings of a K1-ligated ^{14}N nucleus (32, 33). In contrast, NifB$^{NifS3U3/FdxN}$ demonstrated shallower modulations and intensities in its ESEEM and FFT spectra (Fig. 5 *B* and *C*), reflecting a partial loss of the nitrogen coupling to K1 upon the K- to L-cluster transformation. However, the modulations and intensities were still clearly visible in the ESEEM and FFT spectra of NifB$^{NifS3U3/FdxN}$ (Fig. 5 *B* and *C*), indicating the presence of residual K1-clusters following the in vivo K- to L-cluster conversion in this protein species.

To assess the presence of excess K2-clusters in NifB$^{NifS3U3/FdxN}$, we then examined the products generated upon treatment of this protein with additional SAM supplied to the in vitro assay. High-performance liquid chromatography (HPLC) and gas chromatograph-mass spectrometry (GC-MS) analyses revealed formation of *S*-adenosyl-L-homocysteine (SAH) by NifB$^{NifS3U3/FdxN}$ upon incubation with SAM (Fig. 6*A*), as well as formation of methanethiol upon acid quenching of the SAM-treated NifB$^{NifS3U3/FdxN}$ (Fig. 6*B*), indicating transfer of the methyl group from SAM (leaving behind SAH) to an acid-labile sulfide

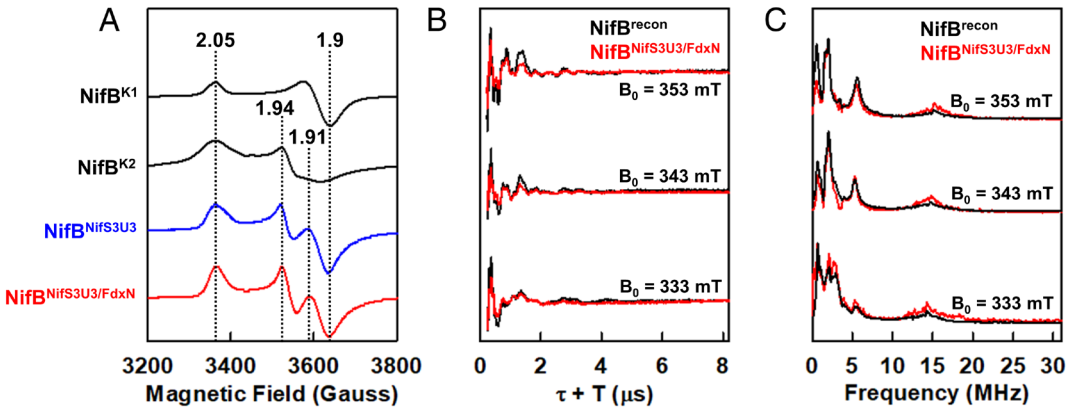

Fig. 5. CW and pulse EPR analyses of NifB proteins. (*A*) CW EPR spectra of NifB^{K1} (black), NifB^{K2} (black), NifB^{NifS3U3} (blue), and NifB$^{NifS3U3/FdxN}$ (red). The spectra were collected as described in *Materials and Methods*, and the g values are indicated. (*B*) Field-dependent X-band 3-pulse ESEEM time domain data of the FeS-reconstituted, yet SAM-untreated NifB^{recon} (black) and the as-isolated NifB$^{NifS3U3/FdxN}$ (red). The time domain spectra have modulations from ^{14}N that appear as peaks in the fast Fourier-transformed (FFT) spectra between 1 and 8 MHz (*C*). The sharp modulations between 0.25 and 0.5 μs in the time domain and the resulting broad peak near 14 MHz in the FFT are from nearby weakly coupled protons. All time domain data are scaled to maximum intensity. ESEEM modulations present in both samples are identical to those previously reported (32), but the modulation depth of NifB$^{NifS3U3/FdxN}$ is approximately 50% of that observed for NifB^{recon}, consistent with a smaller percentage of the precursor K1-cluster with a histidine ligation. Acquisition parameters: temperature = 10 K; MW frequency = 9.338 GHz; $\pi/2$, pulse length = 8 ns; τ = 134 ns (353 mT), 136 ns (343 mT), 142 ns (333 mT); ΔT = 16 ns; shot repetition time = 5 ms. (*C*) Field-dependent X-band Fourier-transformed 3-pulse ESEEM data of NifB^{recon} (black) and the as-isolated NifB$^{NifS3U3/FdxN}$ (red). All FFT data are scaled to maximum intensity. Acquisition parameters: temperature = 10 K; MW frequency = 9.338 GHz; $\pi/2$, pulse length = 8 ns; τ = 134 ns (353 mT), 136 ns (343 mT), 142 ns (333 mT); ΔT = 16 ns; shot repetition time = 5 ms. NifB^{K1}, NifB^{K2}, and NifB^{recon} were reconstituted with synthetic [Fe_4S_4] clusters as described in *Materials and Methods*. Also see *SI Appendix*, Fig. S6, for HYSCORE spectra of NifB^{recon} and NifB$^{NifS3U3/FdxN}$.

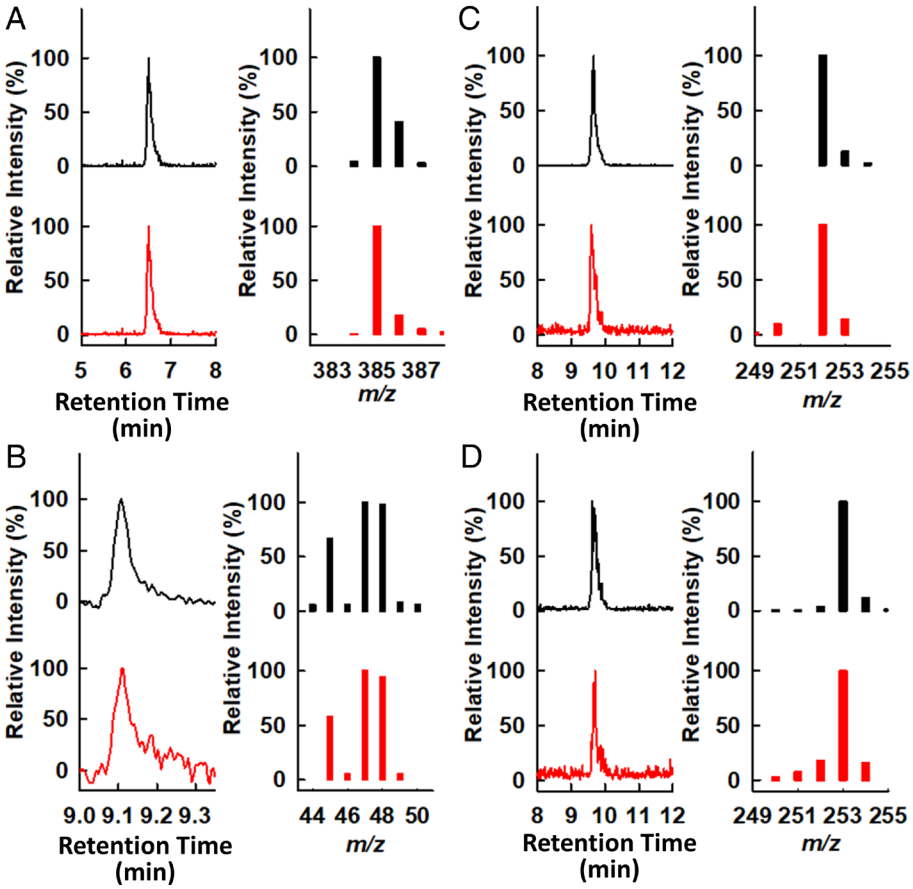

Fig. 6. HPLC and GC–MS analyses of products generated by SAM-treated NifB proteins. HPLC elution profiles (*Left*) and GC-MS fragmentation patterns (*Right*) of (*A*) SAH generated upon in vitro incubation of $NifB^{recon}$ (black) and $NifB^{NifS3U3/FdxN}$ (red) with unlabeled SAM; (*B*) methanethiol (CH_3-SH) generated upon in vitro incubation of $NifB^{recon}$ (black) and $NifB^{NifS3U3/FdxN}$ (red) with unlabeled SAM, followed by acid quenching; and (*C* and *D*) 5′-dAH and 5'-dAD generated upon in vitro incubation of $NifB^{recon}$ (black) and $NifB^{NifS3U3/FdxN}$ (red) with unlabeled SAM (*C*) and [methyl-*d3*] SAM (*D*), respectively.

of the K2-cluster (resulting in methanethiol following acid treatment) in this protein. Additionally, HPLC and liquid chromatography-mass spectrometry (LC-MS) experiments demonstrated formation of 5'-deoxyadenosine (5'-dAH) and deuterated 5'-dAH (5'-dAD) (Fig. 6 *C* and *D*), respectively, upon incubation of $NifB^{NifS3U3/FdxN}$ with unlabeled SAM and [d_3-methyl] SAM, consistent with hydrogen abstraction from the K2-bound methyl group in this protein. However, the intensities of the peaks corresponding to the mass and HPLC retention time of methanethiol in $NifB^{NifS3U3/FdxN}$ were only ~30% of those of the in vitro reconstituted $NifB^{recon}$, suggesting that a portion of the K2-clusters was left over alongside the K1-clusters after the in vivo K- to L-cluster conversion in $NifB^{NifS3U3/FdxN}$.

Surprisingly, incubation of $NifB^{NifS3U3/FdxN}$ with SAM did not result in an increase in the cofactor maturation activity (*SI Appendix*, Fig. S7*A*). Such an effect could be explained by portions of $NifB^{NifS3U3/FdxN}$ carrying either excess K1- or excess K2-clusters and neither of which contributing to the formation of additional L-clusters upon incubation with additional SAM (*SI Appendix*, Fig. S7*B*, Scenario 1). Alternatively, it could reflect an intrinsic limit of the maximum percentage of L-cluster formation that results in a fixed percentage of unconverted K-clusters (*SI Appendix*, Fig. S7*B*, Scenario 2), which aligns well with our observation of the same K- to L-cluster conversion yields for the *in vitro* and *in vitro* matured NifB proteins (Fig. 4 *C* and *D*). In either case, the presence of excess K1- and K2-modules alongside L-clusters on $NifB^{NifS3U3/FdxN}$, coupled with the previously observed capability of the single-module $NifB^{K1}$ or $NifB^{K2}$ variant to carry either the K1- or K2-module (Fig. 5*A*), strongly refutes the previously proposed, sequential insertion of the K1-cluster prior to the K2-cluster and the subsequent, highly unlikely event wherein a fused, P-cluster-type intermediate between the K1- and K2-clusters is generated spontaneously in the absence of SAM (50). Moreover, the fact that the K- to L-cluster transformation on $NifB^{NifS3U3/FdxN}$ cannot proceed beyond methyl transfer and hydrogen atom abstraction at the leftover K2-cluster sites makes this species a potential template for capturing intermediates of cofactor assembly, further highlighting the utility of this heterologous expression system in the mechanistic investigation of nitrogenase biosynthesis. But most importantly, the in vivo formation of an L-cluster in *E. coli* that already has the unique core structure of a mature M-cluster in place represents a crucial step toward the successful expression of an active nitrogenase in a nondiazotrophic organism.

Conclusions

The P- and M-clusters of Mo-nitrogenase are arguably two of the most complex metalloclusters that have thus-far evaded successful chemical synthesis and effective biogenesis in a non-native host. Taking a unique divide-and-conquer approach, we specifically tackled the heterologous synthesis of the P-cluster (a [Fe_8S_7] cluster) and L-cluster (the [Fe_8S_9C] core of the M-cluster) in *E. coli* and verified the formation of these structurally unique and functionally crucial homometallic metallocenters through combined metal content, activity, EPR, and XAS/EXAFS analyses. As such, our approach is fundamentally different from the traditional, whole *nif* gene-set transfer approach (18–20, 51, 52)

in that it systematically implements critical checkpoints along the complex biosynthetic pathway of nitrogenase to improve the chance of success in achieving the heterologous expression of a complete nitrogenase. Moreover, our approach has a distinct, metallocentric focus that aims to conclusively demonstrate the successful expression of biosynthetic components through spectroscopic and biochemical characterization of purified metalloproteins, which is a drastic departure from the traditional method that centers on whole-cell or crude-extract analyses featuring immunoblotting assays and alternative activity measurements. Building on our success in heterologously generating the P- and L-clusters in *E. coli*, our ongoing efforts are focused on completing the M-cluster assembly pathway in hopes of combining it with the P-cluster assembly pathway for the heterologous expression of a functional nitrogenase enzyme in the near future. Additionally, we will gauge efforts toward expanding our knowledge of nitrogenase assembly and catalysis, taking advantage of the utility of the *E. coli* expression sytems in quickly generating large libraries of variants for the mechanistic investigations of this important metalloenzyme. Together, these efforts could facilitate the mechanistic understanding and biotechnological adaptations of nitrogenase, both of which could prove beneficial for areas related to agronomy, energy, and environment in the long run.

Materials and Methods

All chemicals were purchased from Sigma-Aldrich and Thermo Fisher Scientific unless specified otherwise. All experiments were conducted in a glove box or on a Schlenk line under an Ar atmosphere, with an O_2 concentration of <3 ppm. Strain construction, cell growth, and protein purification procedures are described below, and biochemical and spectroscopic analyses are detailed in *SI Appendix*.

Strain Construction. For the heterologous synthesis of P-clusters on NifDK, the genes encoding the *A. vinelandii* NifH, NifM, NifZ, NifW, NifD, and NifK proteins were codon-optimized for *E. coli* expression, synthesized, and cloned into pCDF-Duet-1 or pRSFDuet-1 as summarized in *SI Appendix*, Table S6 (GenScript). These constructs were cotransformed with a plasmid harboring *iscSUA* and *hscABfdx* genes from *A. vinelandii*, an ensemble of genes encoding FeS cluster assembly proteins (53–57), into the *E. coli* strain BL21(DE3). This procedure resulted in strains expressing His-tagged NifDK^{NifZ} (strain YM387EE), NifDK^{NifHM} (strain YM388EE), NifDK^{NifHMZ} (strain YM332EE), NifDK^{NifHMW} (strain YM422EE), and NifDK^{NifHMZW} (strain YM423EE) upon induction with isopropyl β-D-1-thiogalactopyranoside (IPTG). The plasmid carrying *iscSUA* and *hscABfdx* genes was a generous gift from Silke Leimkhüler, University of Potsdam, Germany.

For the heterologous synthesis of L-clusters on NifB, the genes encoding the *A. vinelandii* FdxN protein and the *M. acetivorans* NifB, NifS1, NifU1, NifS2, NifU2, NifS3, and NifU3 proteins were codon-optimized for *E. coli* expression, synthesized, and cloned into pET-26b(+), pCDFDuet-1 or pRSFDuet-1 as summarized in *SI Appendix*, Table S6 (GenScript). These constructs were transformed, with or without the plasmid harboring *iscSUA* and *hscABfdx* genes from *A. vinelandii*, into the *E. coli* strain BL21(DE3). This procedure resulted in strains expressing His-tagged NifB^{IscSU} (YM395EE), NifB^{NifS1U1} (YM291EE), NifB^{NifS2U2} (YM292EE), NifB^{NifS3U3} (YM293EE), and NifB$^{NifS3U3/FdxN}$ (YM434EE) upon induction with IPTG.

Cell Growth and Protein Purification. *E. coli* strains were grown in 10-L batches in LB medium (Difco) supplemented with 50 mM MOPS/NaOH (pH 7.4), 25 mM glucose, 2 mM ferric ammonium citrate, 19 mg/L kanamycin (for YM387EE, YM388EE, YM332EE, YM422EE, YM423EE, YM395EE, YM291EE, YM292EE, YM293EE, and YM434EE), 28 mg/L chloramphenicol (for YM387EE, YM388EE, YM332EE, YM422EE, YM423EE, and YM395EE), and 26 mg/L streptomycin (for YM387EE, YM388EE, YM332EE, YM422EE, YM423EE, YM291EE, YM292EE, and YM293EE) in a BIOFLO 415 fermenter (New Brunswick Scientific) at 37 °C with 200 rpm agitation and 10 L/min airflow. When OD_{600} reached 0.5, the airflow was terminated and the fermenter was purged with N_2 (ultrahigh purity) at a rate of 1.5 L/min; additionally, the temperature was lowered to 24 °C. Once the culture reached 24 °C, 25 mM sodium fumarate and 2 mM cysteine were added, and the expression of His-tagged NifDK^{NifZ}, NifDK^{NifHM}, NifDK^{NifHMZ}, NifDK^{NifHMW}, NifDK^{NifHMZW}, NifB^{IscSU}, NifB^{NifS1U1}, NifB^{NifS2U2}, NifB^{NifS3U3}, or NifB$^{NifS3U3/FdxN}$ was induced by the addition of 250 μM IPTG. Each protein was expressed for 16 h prior to harvesting of cells by centrifugation using a Thermo Fisher Scientific Legend XTR centrifuge. The heterologously expressed, His-tagged NifDK or NifB proteins were purified by immobilized metal affinity chromatography (IMAC) using a method adapted from the purification of the His-tagged nitrogenase proteins from *A. vinelandii* (58).

A. vinelandii strains DJ1162, DJ1141, DJ1143, DJ1041, and YM9A expressing His-tagged *Av*NifH, *Av*NifDK, *Av*NifDK^{apo}, *Av*NifEN, and *Av*NifEN^{apo} (28), respectively, were grown in 180-L batches in Burke's minimal medium (supplemented with 2 mM ammonium acetate) in a 200-L fermenter (New Brunswick Scientific) at 30 °C with 100 rpm agitation and 30 L/min airflow. Cell growth was monitored at OD_{436} using a Spectronic 20 Genesys spectrometer (Spectronic Instruments), and, upon depletion of ammonia, cells were derepressed for 3 h prior to harvesting by a flow-through centrifugal harvester (Cepa). Published methods were used to purify His-tagged *Av*NifH, *Av*NifDK, *Av*NifDK^{apo}, *Av*NifEN, and *Av*NifEN^{apo} (28).

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by the National Institute of Health (NIH), National Institute of General Medical Sciences (NIGMS) grant GM67626 (to M.W.R. and Y.H.), which funded research related to nitrogenase assembly, and the Department of Energy (DOE) Basic Energy Sciences (BES) grant DE-SC0016510 (to Y.H. and M.W.R.), which funded work related to the mechanistic investigation of ammonia formation through engineering nitrogenase proteins. Stanford Synchrotron Radiation Lightsource (SSRL), SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the NIH, National Institute of General Medical Sciences (P30GM133894) (to K.O.H. and B.H.).

Author affiliations: [a]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3900; [b]Department of Chemistry, University of California, Irvine, CA 92697-2025; [c]Department of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125; [d]Stanford Synchrotron Radiation Lightsource, Stanford Linear Accelerator Center National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025; and [e]Department of Chemistry, Stanford University, Stanford, CA 94305

1. B. K. Burgess, D. J. Lowe, Mechanism of molybdenum nitrogenase. *Chem. Rev.* **96**, 2983–3012 (1996).
2. H. L. Rutledge, F. A. Tezcan, Electron transfer in nitrogenase. *Chem. Rev.* **120**, 5158–5193 (2020).
3. S. T. Stripp *et al.*, Second and outer coordination sphere effects in nitrogenase, hydrogenase, formate dehydrogenase, and CO dehydrogenase. *Chem. Rev.* **122**, 11900–11973 (2022).
4. H. L. Rutledge, B. D. Cook, H. P. M. Nguyen, M. A. Herzik Jr., F. A. Tezcan, Structures of the nitrogenase complex prepared under catalytic turnover conditions. *Science* **377**, 865–869 (2022).
5. H. Schindelin, C. Kisker, J. L. Schlessman, J. B. Howard, D. C. Rees, Structure of ADP x AlF4–stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387**, 370–376 (1997).
6. T. Spatzal *et al.*, Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* **334**, 940 (2011).
7. K. M. Lancaster *et al.*, X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science* **334**, 947–974 (2011).
8. J. A. Wiig, Y. Hu, C. C. Lee, M. W. Ribbe, Radical SAM-dependent carbon insertion into the nitrogenase M-cluster. *Science* **337**, 1672–1675 (2012).
9. Y. Hu *et al.*, Enzymatic Fischer-Tropsch-type reactions. *Chem. Rev.* **123**, 5755–5797 (2023).
10. A. J. Jasniewski, C. C. Lee, M. W. Ribbe, Y. Hu, Reactivity, mechanism, and assembly of the alternative nitrogenases. *Chem. Rev.* **120**, 5107–5157 (2020).
11. C. C. Lee, Y. Hu, M. W. Ribbe, Catalytic reduction of CN^-, CO, and CO_2 by nitrogenase cofactors in lanthanide-driven reactions. *Angew. Chem. Int. Ed. Engl.* **54**, 1219–1222 (2015).
12. C. C. Lee, Y. Hu, M. W. Ribbe, Vanadium nitrogenase reduces CO. *Science* **329**, 642 (2010).
13. C. C. Lee, Y. Hu, M. W. Ribbe, Tracing the hydrogen source of hydrocarbons formed by vanadium nitrogenase. *Angew. Chem. Int. Ed. Engl.* **50**, 5545–5547 (2011).
14. C. K. Rofer-DePoorter, A comprehensive mechanism for the Fischer-Tropsch synthesis. *Chem. Rev.* **81**, 447–474 (1981).
15. K. T. Rommens, M. Saeys, Molecular views on Fischer-Tropsch synthesis. *Chem. Rev.* **123**, 5798–5858 (2023).
16. C. Smith, A. K. Hill, L. Torrente-Murciano, Current and future role of Haber-Bosch ammonia in a carbon-free energy landscape. *Energy Environ. Sci.* **13**, 331–344 (2020).

17. J. W. Erisman, M. A. Sutton, J. Galloway, Z. Klimont, W. Winiwarter, How a century of ammonia synthesis changed the world. *Nat. Geosci.* **1**, 636–639 (2008).
18. R. A. Dixon, J. R. Postgate, Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature* **237**, 102–103 (1972).
19. L. Wang *et al.*, A minimal nitrogen fixation gene cluster from *Paenibacillus sp*. WLY78 enables expression of active nitrogenase *in Escherichia coli*. *PLoS Genet.* **9**, e1003865 (2013).
20. J. Yang, X. Xie, X. Wang, R. Dixon, Y. P. Wang, Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3718–E3725 (2014).
21. Q. Li, S. Chen, Transfer of nitrogen fixation (*nif*) genes to non-diazotrophic hosts. *Chembiochem* **21**, 1717–1722 (2020).
22. Y. Hu, M. W. Ribbe, Biosynthesis of the metalloclusters of nitrogenases. *Annu. Rev. Biochem.* **85**, 455–483 (2016).
23. M. C. Corbett *et al.*, Comparison of iron-molybdenum cofactor-deficient nitrogenase MoFe proteins by X-ray absorption spectroscopy: Implications for P-cluster biosynthesis. *J. Biol. Chem.* **279**, 28276–28282 (2004).
24. C. C. Lee *et al.*, Stepwise formation of P-cluster in nitrogenase MoFe protein. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18474–18478 (2009).
25. B. Schmid *et al.*, Structure of a cofactor-deficient nitrogenase MoFe protein. *Science* **296**, 352–356 (2002).
26. K. Rupnik, C. C. Lee, Y. Hu, M. W. Ribbe, B. J. Hales, A VTVH MCD and EPR spectroscopic study of the maturation of the "second" nitrogenase P-cluster. *Inorg. Chem.* **57**, 4719–4725 (2018).
27. Y. Hu, A. W. Fay, C. C. Lee, M. W. Ribbe, P-cluster maturation on nitrogenase MoFe protein. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10424–10429 (2007).
28. Y. Hu, A. W. Fay, P. C. Dos Santos, F. Naderi, M. W. Ribbe, Characterization of *Azotobacter vinelandii nifZ* deletion strains. Indication of stepwise MoFe protein assembly. *J. Biol. Chem.* **279**, 54963–54971 (2004).
29. Y. Hu, M. W. Ribbe, Biosynthesis of the metalloclusters of molybdenum nitrogenase. *Microbiol. Mol. Biol. Rev.* **75**, 664–77 (2011).
30. Y. Hu, M. W. Ribbe, Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *J. Biol. Chem.* **288**, 13173–13177 (2013).
31. A. W. Fay, J. A. Wiig, C. C. Lee, Y. Hu, Identification and characterization of functional homologs of nitrogenase cofactor biosynthesis protein NifB from methanogens. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14829–14833 (2015).
32. L. A. Rettberg *et al.*, Probing the coordination and function of Fe4S4 modules in nitrogenase assembly protein NifB. *Nat. Commun.* **9**, 2824 (2018).
33. A. J. Jasniewski *et al.*, Spectroscopic characterization of an eight-iron nitrogenase cofactor precursor that lacks the "9th sulfur". *Angew. Chem. Int. Ed. Engl.* **58**, 14703–14707 (2019).
34. K. Tanifuji *et al.*, Tracing the "ninth sulfur" of the nitrogenase cofactor via a semi-synthetic approach. *Nat. Chem.* **10**, 568–572 (2018).
35. W. Kang *et al.*, X-Ray crystallographic analysis of NifB with a full complement of clusters: Structural insights into the radical SAM-dependent carbide insertion during nitrogenase cofactor assembly. *Angew. Chem. Int. Ed. Engl.* **60**, 2364–2370 (2021).
36. Y. Hu, A. W. Fay, M. W. Ribbe, Identification of a nitrogenase FeMo cofactor precursor on NifEN complex. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3236–3241 (2005).
37. M. C. Corbett *et al.*, Structural insights into a protein-bound iron-molybdenum cofactor precursor. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1238–1243 (2006).
38. J. T. Kaiser, Y. Hu, J. A. Wiig, D. C. Rees, M. W. Ribbe, Structure of precursor-bound NifEN: A nitrogenase FeMo cofactor maturase/insertase. *Science* **331**, 91–94 (2011).
39. Y. Hu *et al.*, FeMo cofactor maturation on NifEN. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17119–17124 (2006).
40. Y. Hu *et al.*, Nitrogenase Fe protein: A molybdate/homocitrate insertase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17125–17130 (2006).
41. J. M. Yoshizawa *et al.*, Optimization of FeMoco maturation on NifEN. *J. Am. Chem. Soc.* **131**, 9321–9325 (2009).
42. Y. A. Liu *et al.*, Radical SAM-dependent formation of a nitrogenase cofactor core on NifB. *J. Inorg. Biochem.* **233**, 111837 (2022).
43. K. Tanifuji *et al.*, Tracing the incorporation of the "ninth sulfur" into the nitrogenase cofactor precursor with selenite and tellurite. *Nat. Chem.* **13**, 1228–1234 (2021).
44. A. J. Pierik, H. Wassink, H. Haaker, W. R. Hagen, Redox properties and EPR spectroscopy of the P clusters of *Azotobacter vinelandii* MoFe protein. *Eur. J. Biochem.* **212**, 51–61 (1993).
45. M. A. Blank *et al.*, Structural models of the $[Fe_4S_4]$ clusters of homologous nitrogenase Fe proteins. *Inorg. Chem.* **50**, 7123–7128 (2011).
46. D. R. Dean, M. R. Jacobson, "Biochemical genetics of nitrogenase" in *Biological Nitrogen Fixation*, G. Stacey, R. H. Burris, H. J. Evans, Eds. (Springer, 1992), pp. 763–834.
47. C. Van Stappen *et al.*, A conformational role for NifW in the maturation of molybdenum nitrogenase P-cluster. *Chem. Sci.* **13**, 3489–3500 (2022).
48. E. Jiménez-Vicente, M. Navarro-Rodríguez, C. Poza-Carrión, L. M. Rubio, Role of *Azotobacter vinelandii* FdxN in FeMo-co biosynthesis. *FEBS Lett.* **588**, 512–516 (2014).
49. K. B. Musgrave, H. C. Angove, B. K. Burgess, B. Hedman, K. O. Hodgson, All-ferrous titanium(III) citrate reduced Fe protein of nitrogenase: an XAS study of electronic and metrical structure. *J. Am. Chem. Soc.* **120**, 5325–5326 (1998).
50. L. P. Jenner, M. V. Cherrier, P. Amara, L. M. Rubio, Y. Nicolet, An unexpected P-cluster like intermediate en route to the nitrogenase FeMo-co. *Chem. Sci.* **12**, 5269–5274 (2021).
51. R. Tsujimoto *et al.*, Functional expression of an oxygen-labile nitrogenase in an oxygenic photosynthetic organism. *Sci. Rep.* **8**, 7380 (2018).
52. D. Liu, M. Liberton, J. Yu, H. B. Pakrasi, M. Bhattacharyya-Pakrasi, Engineering nitrogen fixation activity in an oxygenic phototroph. *mBio* **9**, e01029–18 (2018).
53. J. Frazzon, D. R. Dean, Formation of iron-sulfur clusters in bacteria: An emerging field in bioinorganic chemistry. *Curr. Opin. Chem. Biol.* **7**, 166–173 (2003).
54. J. Frazzon, J. R. Fick, D. R. Dean, Biosynthesis of iron-sulphur clusters is a complex and highly conserved process. *Biochem. Soc. Trans.* **30**, 680–685 (2002).
55. K. Kriek, L. Peters, Y. Takahashi, P. L. Roach, Effect of iron-sulfur cluster assembly proteins on the expression of *Escherichia coli* lipoic acid synthase. *Protein. Expr. Purif.* **28**, 241–245 (2003).
56. R. M. Cicchillo *et al.*, *Escherichia coli* lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide. *Biochemistry* **43**, 11770–17781 (2004).
57. N. D. Lanz *et al.*, RlmN and AtsB as models for the overproduction and characterization of radical SAM proteins. *Methods Enzymol.* **516**, 125–152 (2012).
58. M. W. Ribbe, Y. Hu, M. Guo, B. Schmid, B. K. Burgess, The FeMoco-deficient MoFe protein produced by a nifH deletion strain of *Azotobacter vinelandii* shows unusual P-cluster features. *J. Biol. Chem.* **277**, 23469–23476 (2002).