Using synthetic biology to overcome barriers to stable expression of nitrogenase in eukaryotic organelles

Nan Xiang[a], Chenyue Guo[a], Jiwei Liu[a], Hao Xu[a], Ray Dixon[b,1], Jianguo Yang[a,1], and Yi-Ping Wang[a,1]

[a]State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences & School of Advanced Agricultural Sciences, Peking University, 100871 Beijing, China; and [b]Department of Molecular Microbiology, John Innes Centre, NR4 7UH Norwich, United Kingdom

Edited by Éva Kondorosi, Hungarian Academy of Sciences, Biological Research Centre, Szeged, Hungary, and approved May 29, 2020 (received for review February 7, 2020)

Engineering biological nitrogen fixation in eukaryotic cells by direct introduction of *nif* genes requires elegant synthetic biology approaches to ensure that components required for the biosynthesis of active nitrogenase are stable and expressed in the appropriate stoichiometry. Previously, the NifD subunits of nitrogenase MoFe protein from *Azotobacter vinelandii* and *Klebsiella oxytoca* were found to be unstable in yeast and plant mitochondria, respectively, presenting a bottleneck to the assembly of active MoFe protein in eukaryotic cells. In this study, we have delineated the region and subsequently a key residue, NifD-R98, from *K. oxytoca* that confers susceptibility to protease-mediated degradation in mitochondria. The effect observed is pervasive, as R98 is conserved among all NifD proteins analyzed. NifD proteins from four representative diazotrophs, but not their R98 variants, were observed to be unstable in yeast mitochondria. Furthermore, by reconstituting mitochondrial-processing peptidases (MPPs) from yeast, *Oryza sativa*, *Nicotiana tabacum*, and *Arabidopsis thaliana* in *Escherichia coli*, we demonstrated that MPPs are responsible for cleavage of NifD. These results indicate a pervasive effect on the stability of NifD proteins in mitochondria resulting from cleavage by MPPs. NifD-R98 variants that retained high levels of nitrogenase activity were obtained, with the potential to stably target active MoFe protein to mitochondria. This reconstitution approach could help preevaluate the stability of Nif proteins for plant expression and paves the way for engineering active nitrogenase in plant organelles.

nitrogenase stability | mitochondria | processing peptidase | synthetic biology | plant organelle

Nitrogenase is a complex enzyme that requires multiple components for biosynthesis and activity. The catalytic components of nitrogenase consist of two metalloproteins: the *nifH*-encoded Fe protein and the *nifDK*-encoded MoFe protein (1). However, in addition to the structural subunits of the core enzyme encoded by *nifHDK*, components for metallocluster assembly—specifically, the [Fe_4S_4] cluster, P-cluster, and FeMo-co [Mo–7 Fe–9S–C–homocitrate]—are also required for the biosynthesis of active nitrogenase (2). Assembly of the Fe protein is dependent on the maturase protein NifM and an [Fe_4S_4] cluster provided by NifU and NifS, which bridges two NifH subunits to confer functionality (3, 4). Maturation of MoFe protein requires incorporation of the P cluster, with a composition [Fe_8S_7] located at the NifD (α) and NifK (β) interfaces and insertion of FeMo-co into the α-subunit to form the NifDK heterotetramer (5). The biosynthesis pathway of FeMo-co, the most complex heterometal cluster in biology, requires at least nine Nif proteins (H, E, N, B, U, S, V, Q, Y) in vivo (6). In contrast, biosynthesis of P-cluster is simpler, facilitated by NifH and the assembly factors NafH, NifW, and NifZ (7). Additional gene products are required to provide nitrogenase with electrons; the flavodoxin NifF and the pyruvate-flavodoxin oxidoreductase NifJ fulfill this role in *Klebsiella oxytoca* (8).

Engineering of nonlegume crops that can "fix" their own nitrogen is an important approach to reducing the use of industrial nitrogen fertilizers in agriculture. This could be achieved by synthetic engineering of the nitrogenase system into plants, especially cell organelles (e.g., mitochondria, chloroplasts) (9, 10). These energy-conversion organelles can potentially provide the reducing power and ATP required for the nitrogen-fixation process (10). Recent attempts at engineering nitrogenase into eukaryotic organelles have suggested mitochondria or chloroplasts as suitable locations for expression of nitrogenase in eukaryotes (11–16). The assembly of functional Fe protein in mitochondria of aerobically grown yeast has demonstrated that mitochondria can provide a suitable location for engineering of oxygen-sensitive nitrogenase (12). In addition, some nitrogenase Fe protein activity has been detected in plants when *nifH* was coexpressed with *nifM* in tobacco plastids under low-oxygen conditions (11), and functional Fe protein has been obtained under aerobic conditions when NifU and NifS, in addition to NifH and NifM, were targeted to chloroplasts (16). Taken together, these studies show that under aerobic conditions, coexpression of NifH with the NifM maturase is sufficient to support Fe protein activity in yeast mitochondria and also in tobacco plastids when these proteins are expressed together with NifU and NifS.

A soluble form of the NifB protein from *Methanocaldococcus infernus* also has been purified from aerobically grown yeast and found to be active for in vitro FeMo-co biosynthesis when coexpressed with NifU, NifS, NifX, and FdxN proteins (14, 17). Several studies have demonstrated that more Nif proteins can accumulate in mitochondria. All 16 Nif proteins (B, D, E, F, H, J, K, M, N, Q, S, U, V, X, Y, and Z) from *K. oxytoca* were

Significance

Stable expression of each component of the nitrogenase system in an active form is a prerequisite for engineering nitrogen fixation in eukaryotic cells. Mitochondria provide an oxygen-depleted environment for the expression of active nitrogenase in plants, but signal peptides are required to target nuclear encoded Nif proteins to this organelle. We demonstrate that one of the structural subunits of nitrogenase, NifD, is itself susceptible to cleavage by mitochondrial processing peptidases from a variety of plant origins, presenting a major challenge to engineering nitrogen fixation in mitochondria. To overcome this issue, we have engineered NifD variants that are resistant to cleavage and retain high levels of nitrogenase activity, thus providing a potential solution for engineering active MoFe protein in plants.

Author contributions: R.D., J.Y., and Y.-P.W. designed research; N.X., C.G., J.L., H.X., and J.Y. performed research; N.X., R.D., J.Y., and Y.-P.W. analyzed data; and N.X., R.D., J.Y., and Y.-P.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: ray.dixon@jic.ac.uk, yangjg@pku.edu.cn, or wangyp@pku.edu.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2002307117/-/DCSupplemental.

First published June 29, 2020.

detectable using a transient expression system in *Nicotiana benthamiana* when these components were targeted to the mitochondrial matrix (13). However, in this case, NifD protein was not highly expressed in a stable form, and this could be partially solved by construction of an NifDK fusion. Expression of a *nif* cluster containing nine genes (*nifHDKUSMBEN*) from *Azotobacter vinelandii* was achieved in yeast when targeted to mitochondria (15). However, due to the instability of NifD expressed in yeast mitochondria, only a low level of NifDK tetramer purified from yeast mitochondria was detected. These studies all indicate that it is potentially feasible to achieve nitrogen fixation in eukaryotic cells, especially in organelles. However, the instability of NifD when targeted to yeast and tobacco mitochondria creates a bottleneck for obtaining functional MoFe protein in eukaryotic organelles.

Based on the pattern of the NifD digestion products and the process of mitochondrial protein targeting, we speculated that the aforementioned instability of NifD could be a consequence of the import and processing of the protein in mitochondria. A large number of proteases are required to maintain mitochondrial homeostasis (18, 19). These proteases can be broadly classified into two groups: specific processing peptidases that remove N-terminal targeting signals from nuclear-encoded mitochondrial preproteins (20) and proteases that exert regulatory functions (19). The processing peptidase of mitochondria contains three members: mitochondrial processing peptidase (MPP), mitochondrial intermediate peptidase (MIP; Oct1), and the intermediate cleaving peptidase of 55 kDa, Icp55 (21). MPP, the main protease in the mitochondrial matrix, is responsible for removing leader peptides of mitochondrial protein precursors. MPP usually consists of two subunits, alpha and beta, encoded by two separated genes (22). Oct1 and Icp55 are secondary processing peptidases functioning to cleave an octapeptide and one amino acid, respectively, from intermediates generated by MPP (23, 24). In addition to proteases responsible for peptide processing, other proteases with chaperone-like properties in mitochondria, such as Pim1/LON and the mtClpXP complex, function to remove damaged proteins and thus help maintain mitochondrial protein homeostasis (25).

In the current study, we found that an internal sequence within NifD from *K. oxytoca* confers susceptibility to cleavage by a yeast mitochondrial protease, and that a key amino acid, NifD-R98, is necessary for this degradation. We further demonstrate that the observed instability of NifD is pervasive, with respect to both diazotrophic origin of NifD and mitochondria from diverse eukaryotic sources. Sequence alignments showed that NifD-R98 is highly conserved among NifD proteins from different origins and four representative NifDs were observed to be unstable in yeast mitochondria. Heterologous reconstitution of MPPs from yeast and three model plants in *Escherichia coli* reveals that MPP is responsible for the degradation of the NifD protein. Our results thus clearly demonstrate the root cause of the instability of NifD in mitochondria. By identifying NifD variants that retain nitrogenase activity and are resistant to degradation by the MPP, we provide efficient solutions for this instability. These studies provide a powerful approach to bypass obstacles to the engineering of diazotrophy in plant organelles.

Results

Identification of a Key Amino Acid Residue Required for NifD Degradation. As reported recently, N-terminal degradation of the NifD protein was observed in both yeast and tobacco when NifD was expressed in the nuclear genome and targeted to mitochondria with specific leader peptides (13, 15). According to the migration rate (between 10 and 15 kDa) of the degraded portion of the NifD protein (NifDn; Fig. 1*B*), we predicted that protease mediated digestion might occur within an N-terminal region of NifD (81 to 130 aa; Fig. 1*A*).

To identify the precise position responsible for the protease cleavage, internal deletions at 10-aa intervals within the NifD protein were constructed (*Materials and Methods* and *SI Appendix*, Fig. S1). When expressed in yeast from the galactose inducible promoter *GAL1* and targeted to mitochondria with the Su9 (residues 1 to 69 of ATP synthase subunit 9, *Neurospora crassa*) leader peptide (26), we found that deletion of residues 91 to 100 prevented NifD degradation (Fig. 1*B*). Subsequently, two 5-aa internal deletions within residues 91 to 100 were constructed. Examination of the protein sequence of *K. oxytoca* NifD revealed three arginine (R) residues located within this region, corresponding to R94, R97, and R98 (Fig. 1*A*). Because an arginine residue is responsible for the recognition and cleavage by proteases in many cases, we substituted these three arginine residues for lysine (K). Protein stability assays showed that both the 96- to 100-aa deletion variant and the R98K variant increased the tolerance of NifD to protease-mediated degradation when expressed in yeast mitochondria (Fig. 1*C*). Thus, R98 is specifically required for the degradation of NifD in yeast mitochondria.

We next analyzed more than 70 NifD sequences representing diversity in diazotrophic origin and NifD phylogeny (Dataset S1). We found that R98 is conserved among almost all NifD sequences analyzed, except three NifDs from the phylum Chloroflexi (*Oscillochloris trichoides* DG-6, *Roseiflexus* sp. RS-1, and *Roseiflexus castenholzii* DSM), which contain a proline (P) instead of arginine at position 98 (Fig. 1*D*). However, these strains do not encode NifE and NifN, which are required to synthesize FeMo-co, and their NifDK sequences have a discrete phylogeny (27, 28). Therefore, it is possible that these NifD-like sequences are not bona fide Mo nitrogenases.

Since R98 is conserved in conventional nitrogenases, we tested NifDs from other diazotrophs. To determine whether this residue is also the key amino acid for the stability of NifD in mitochondria, NifD proteins and their corresponding R98K variants from *A. vinelandii*, *Rhodobacter capsulatus*, and *Anabaena* sp. PCC7120 were cloned and tested for stability following import into yeast mitochondria. We found that native NifDs from different origins were also degraded, although at varying levels (Fig. 1*E*). In contrast, their R98K variants escaped from the degradation process (Fig. 1*E*). Clearly the degradation of NifD in yeast mitochondria is generic to NifDs from different origins, and R98 is essential for protease cleavage.

Screening for Stable NifD R98 Variants That Retain Nitrogenase Functionality. In an attempt to maintain NifD activity, we originally screened variants in which the arginine at position 98 was substituted by lysine, an amino residue with similar side chain properties. Unexpectedly, when this mutation was introduced into the operon-based *nif* gene cluster in *E. coli*, only 6.6% nitrogenase activity was obtained when assayed by the acetylene reduction assay, a proxy for nitrogen fixation (Fig. 2*A*). To solve this problem, a saturated library at this position was constructed (*Materials and Methods* and *SI Appendix*, Fig. S1) to screen for an appropriate amino acid substitution that both retained nitrogenase activity and conferred tolerance to degradation in yeast mitochondria. The R98P substitution retained ~70% activity, whereas two amino acid substitutions, R98H and R98N, resulted in ~55% nitrogenase activity (Fig. 2*A*). Interestingly, the amino acid substitution R98P with the highest activity is identical to the naturally existing amino acid residue (proline) at this position from *Chloroflexi* species (Figs. 1*D* and 2*A*). However, this may be just a coincidence, since the NifD-like proteins in *Chloroflexi* might not catalyze nitrogen fixation.

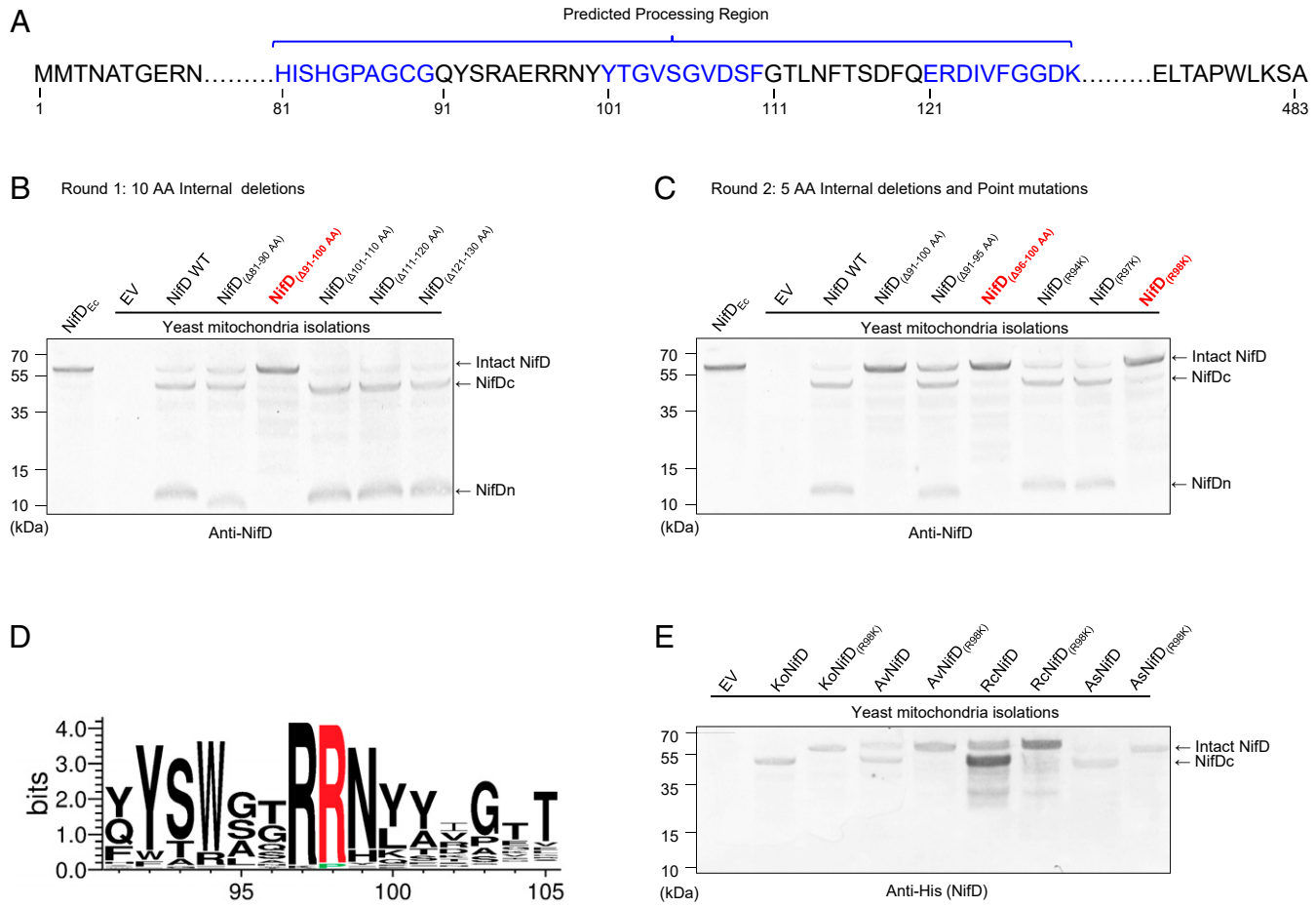

Fig. 1. Analysis of the protease processing site within NifD. (*A*) Predicted protease processing region within NifD from *K. oxytoca*, colored in either blue or black at 10-residue intervals. (*B* and *C*) Two rounds of internal deletions within NifD. Intact NifD expressed in *E. coli* ($NifD_{Ec}$) was used as a positive control, and empty vector (EV) was used as a negative control. (*B*) 10-aa internal deletions within the predicted region to narrow down the processing region. Intact (unprocessed) NifD expressed in yeast exhibited a single band; see the lane marked $NifD_{(\Delta 91–100\ AA)}$. Processed NifD exhibited three bands, including intact NifD, $NifD_c$ (C-terminal portion of processed NifD), and a much smaller fragment labeled $NifD_n$ (N-terminal portion of processed NifD). (*C*) 5-aa internal deletions within NifD region 91 to 100 aa and point mutations at specific amino acids. The R98K variant protects NifD from protease mediated degradation. (*D*) Conserved sequence around position R98 (indicated in red) among diverse NifD proteins (protein IDs provided in Dataset S1). WebLogo 3 was used to draw the sequence logo. (*E*) Testing the tolerance of NifDs and their corresponding R98K variants from *A. vinelandii* DJ (Av), *R. capsulatus* SB1003 (Rc), and *Anabaena* sp. PCC7120 (As) to protease-mediated degradation. Each NifD was labeled with a C-terminal 6×His tag. The Western blot assays shown in *A*, *B*, and *E* were conducted using extracts derived from purified yeast mitochondria as indicated above the lanes; mitochondrial targeting was achieved using the Su9 signal peptide.

To exclude the possibility that the R98 substitutions in *K. oxytoca* NifD influence dinitrogen reduction (since R98 is located close to FeMo-co in nitrogenase MoFe protein) (30), diazotrophic growth experiments were carried out to provide direct evidence for nitrogen fixation. Similar to the results observed in the acetylene reduction assays, variants with activities > 50% could support diazotrophic growth of *E. coli*, whereas variants with <30% activity could barely support growth (Fig. 2*B* and *SI Appendix*, Fig. S2). Subsequently, NifD proteins carrying the, R98P, R98H, and R98N substitutions were cloned and expressed in yeast. All three variants showed increased tolerance to protease-mediated degradation in yeast mitochondria compared with native NifD (Fig. 2*C*). To exclude the possibility of artifacts resulting from cross-reaction with the polyclonal NifD antibody, an anti-His monoclonal antibody was also used. As shown in *SI Appendix*, Fig. S3, all three variants showed complete resistance to degradation by the mitochondrial protease.

To investigate the general applicability of the R98P substitution, NifD proteins from *A. vinelandii*, *R. capsulatus*, and *Anabeana* with equivalent substitutions were constructed and expressed in yeast. On Western blot assays, all variants equivalent to *K. oxytoca* R98P showed resistance to the mitochondrial protease (Fig. 2*D*). Also, since Nif proteins from *A. vinelandii* have been widely used for expression in eukaryotic mitochondria (12, 14, 15), the functionality of the equivalent substitution (R97P) in *A. vinelandii* NifD was assayed in *E. coli*. This was achieved by replacing the original *K. oxytoca nifHDKTY* operon with the *A. vinelandii nifHDKTY* operon or its NifD-R97P variant. Compared with wild-type *A. vinelandii* NifD, the NifD-R97P variant retained 70% activity in the acetylene reduction assay (Fig. 2*E*). These results suggest that equivalent substitutions to R98P in *K. oxytoca* NifD have the potential for universal application in engineering active nitrogenase in mitochondria.

The crystal structure of *K. oxytoca* MoFe protein shows that R98 in NifD forms a hydrogen bond with D516 in NifK, which might be important for stabilizing the NifDK protein complex (31) (Fig. 2*F*). Amino acid substitutions at R98 may disrupt hydrogen-bonding interactions with NifD-S443 and NifD-S447, as well as the interaction with NifK-D516, leading to decreased

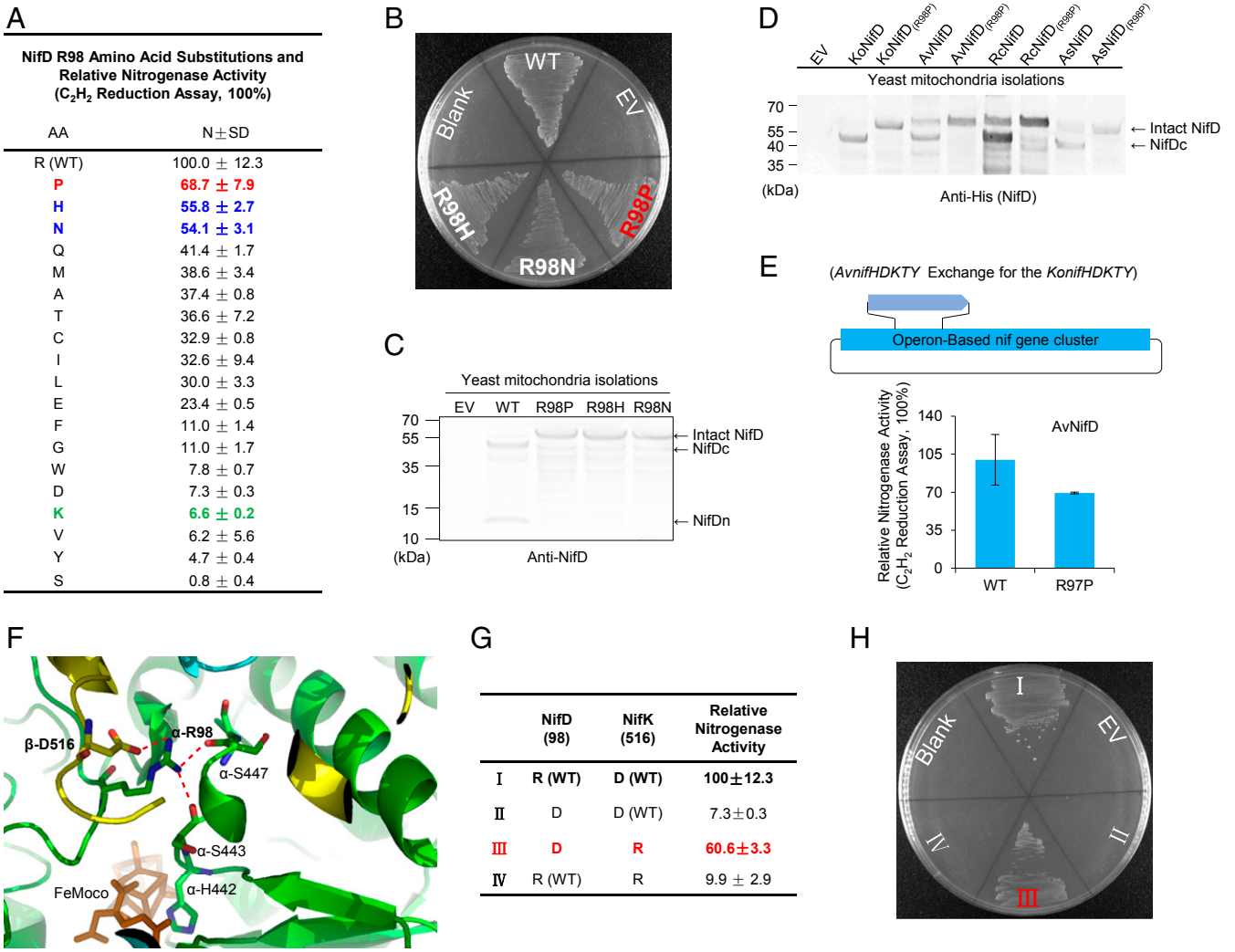

Fig. 2. Screening NifD from *K. oxytoca* for R98 substitutions that retain nitrogenase activity. (*A*) NifD R98 saturation mutagenesis and nitrogenase activities relative to wild-type NifD (normalized as 100%, 28.5 ± 3.4 nmol C_2H_4/min/mg total protein) using the *K. oxytoca* reconstituted operon-based Nif system in *E. coli* (29) (Fig. 3*C*). (*B*) Diazotrophic growth of *K. oxytoca* NifD R98 variants in *E. coli*. (*C*) Stability analysis in yeast mitochondria of the three *K. oxytoca* NifD variants with the highest nitrogenase activities. (*D*) Stability of NifD proteins from different diazotrophs and their R98P variants in yeast mitochondria. The Western blot assays in *C* and *D* were conducted using extracts derived from purified yeast mitochondria as indicated above the lanes. Ko, *Klebsiella oxytoca* M5al; Av, *A. vinelandii* DJ; Rc, *R. capsulatus* SB1003; As, *Anabaena* sp. PCC7120. Each NifD and its R98P variants were labeled with a C-terminal 6xHis tag, and an anti-His monoclonal antibody was used for protein detection. (*E*) Acetylene reduction assay of *A. vinelandii* NifD and its R97P variant when the *A. vinelandii nifHDKYT* operon is introduced into the *K. oxytoca* operon-based system. The activity obtained from native *A. vinelandii* NifD in this system was set at 100% (13.0 ± 2.9 nmol C_2H_4/min/mg total protein). (*F*) Structure of *K. oxytoca* nitrogenase MoFe protein (Protein Data Bank ID code 1QGU) showing interactions between NifD R98 and adjacent residues. Relevant features are shown as sticks; FeMo-co is in brown; side chains of α-R98, α-H442, α-S443, α-S447 from NifD are in green; and β-D516 from NifK is in yellow. (*G* and *H*) Acetylene reduction (*G*) and diazotrophic growth (*H*) of the residue exchange variant NifD R98D/NifK D516R in *E. coli*. The Roman numerals in *H* correspond to the rows indicated in *G*. The SDs for *A*, *E*, and *G* were calculated based on at least two biological replicates.

nitrogenase activity (Fig. 2*A*). To test whether the interaction between NifD-R98 and NifK-D516 is important, we constructed reciprocal charge-change substitutions to create the double variant NifD-R98D/NifK-D516R. When assayed for acetylene reduction, we found that this charge swapped MoFe protein variant recovered 60% activity (Fig. 2*G*), which is sufficient to support diazotrophic growth of *E. coli* (Fig. 2*H*). In addition, both the NifD-R98D and the NifK-D516R variants are resistant to MPP cleavage (*SI Appendix*, Fig. S4). The incomplete (<100%) recovery of activity is not surprising, since NifD-R98 is also involved in intramolecular interactions with nearby NifD residues S443 and S447 in addition to the intermolecular interaction with NifK-D516. It is also possible that charge changes in this region of the protein may influence FeMo-co occupancy (32). Nevertheless, these results suggest that rational design and/or artificial evolution of nitrogenase may be feasible through the modification of intersubunit interactions.

Reconstitution of MPP in *E. coli*. Although several candidate enzymes might be responsible for proteolytic cleavage of NifD in mitochondria, we focused our attention on MPP, which usually recognizes an arginine residue within its target cleavage site (21). MPP consists of two subunits in *Saccharomyces cerevisiae* encoded by the MAS1 and MAS2 genes. However, these two genes are essential in yeast and their deletion mutants are lethal (33, 34). Therefore, it was not possible for us to confirm their

involvement in NifD degradation in the original yeast host. To overcome this problem, we decided to reconstitute the yeast MPP in *E. coli*. The expression module was constructed by arranging the two MAS genes as an operon under the control of the Ptac promoter (Fig. 3*A* and *Materials and Methods*). The proteolytic function of the reconstituted MPP in *E. coli* was confirmed by the efficient cleavage of the specific mitochondria-leading sequence, Su9, fused to green fluorescent protein (GFP) (Fig. 3*B*). Subsequently, the yeast MPP expression module was coexpressed with the operon-based nif gene cluster carrying either the wild-type nifD gene or its R98P derivative under diazotrophic conditions (Fig. 3*C*). When expressed together with all of the Nif components required for nitrogenase biosynthesis and activity, native NifD was degraded in the presence of the yeast MPP but not when empty vector was present (Fig. 3*D*). This suggests that MPP can access the unfolded form of NifD in yeast mitochondria, rather than a misfolded version. In contrast, the NifD-R98P variant was not degraded (Fig. 3*D*).

As anticipated from these results, coexpression of yeast MPP with the wild-type nitrogenase system abolished diazotrophic growth. In contrast, diazotrophic growth of the reconstituted nitrogenase system expressing the NifD-R98P variant protein was not affected by coexpression with the yeast MPP (*SI Appendix*, Fig. S6). Therefore, our data clearly demonstrate that yeast MPP is responsible for the degradation of NifD protein when targeted to mitochondria with a leader peptide.

To engineer nitrogenase in crop plants, it is important to establish whether the key arginine residue at position 98 of NifD also results in susceptibility to MPP-mediated protease in higher plants. A similar strategy was used to reconstitute MPPs from three model plants—*Oryza sativa*, *Nicotiana tabacum* (Nt), and *Arabidopsis thaliana* (At)—in *E. coli* (*Materials and Methods*). The three reconstituted plant-type MPPs were functional in *E. coli*, since they could efficiently cleave the Su9 leader peptide cotranslated with GFP (Fig. 3*B*). Likewise, NifD and its R98P variant expressed from the operon-based *nif* gene cluster were tested for their tolerance to the plant-type MPPs. Similar results to the yeast MPPs were observed, since native NifDs were digested by plant-type MPPs, while their R98P variants were resistant to cleavage (Fig. 3*D*), although the plant-type MPPs were apparently less efficient in cleaving native NifD proteins compared with their yeast MPP counterpart (Fig. 3*D*). MPPs from yeast and *O. sativa* were selected as representatives to assess other essential Nif components (H, K, M, E, N, J, U, S, V, B,

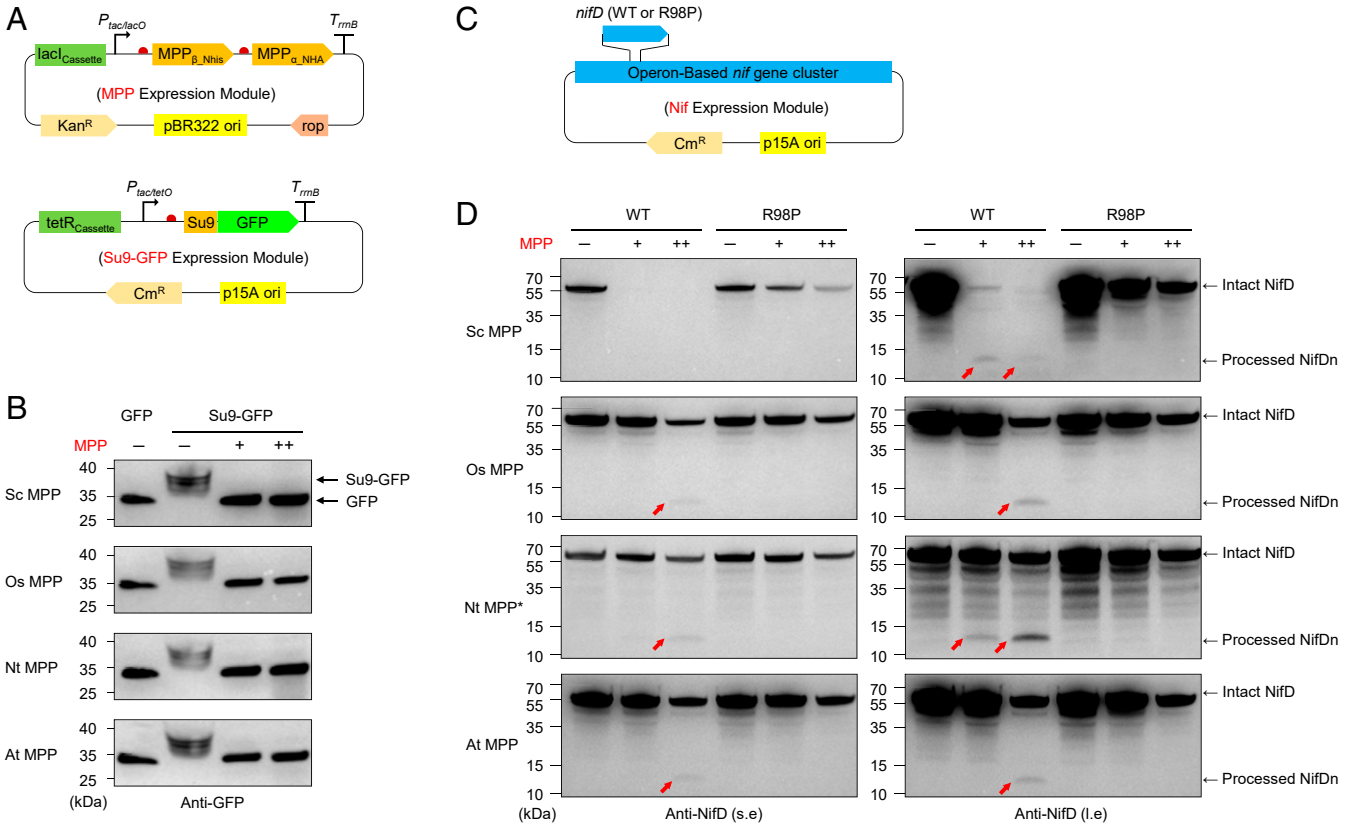

Fig. 3. Reconstitution of eukaryotic MPPs and determination of the stability of Nif proteins toward MPPs in *E. coli*. (*A*) Schematic representation of plasmid construction. The red domes indicate ribosome binding sites. MPP_β and MPP_α were arranged to form an operon under the control of $P_{tac/lacO}$ and induced with IPTG for expression; Su9-GFP was cloned downstream of the $P_{LtetO-1}$ promoter and induced with aTc (anhydrotetracycline) for expression (*B*) Functionality test of reconstituted MPPs from yeast and plants in *E. coli*. GFP expressed in *E. coli* was used as control (lane 1). For lanes 2–4, Su9-GFP was induced for expression with 400 ng/mL of aTc. "-" indicates no MPP expression modules were present (lanes 1 and 2); "+" indicates MPPs were expressed at basal level; "++" represents the expression of MPPs when induced with 200 μM of IPTG. Protein levels of the MPPs are shown in *SI Appendix*, Fig. S5. (*C*) Overview of NifD or NifD R98P variant expression plasmids. These were expressed in the operon-based *nif* gene cluster (29) together with the other components of the nitrogenase system from *K. oxytoca* under anaerobic conditions. (*D*) Comparison of the protein stability of NifD and its R98P variants in the intact nitrogenase system when expressed with reconstituted MPPs. The Nif expression module (expressing native NifD or its NifD R98P variant as in *C*) was cotransformed with either empty vector ("-"; lanes 1 and 4) or the MPP expression modules depicted in *A* (lanes 2, 3, 5, and 6; "+" indicates MPP expression at the basal level; "++" represents the expression of MPPs when induced with 200 μM [for Sc, *O. sativa* Os, and At MPPs] or 50 μM [for Nt MPP] IPTG). Red arrows refer to the $NifD_n$ fragment (defined in the Fig. 1 legend), which indicates in this instance that NifD was processed by the corresponding MPP. The $NifD_c$ fragment was not observed in many cases, perhaps because it is less stable in *E. coli* than in yeast. s.e., short-term exposure; l.e., long-term exposure.

F, Y, W) for their tolerance to digestion. No cleavage of any of these 13 Nif proteins was observed when coexpressed with either yeast MPP (*SI Appendix*, Fig. S7*A*) or *O. sativa* MPP (*SI Appendix*, Fig. S7*B*). In conclusion, MPP-mediated digestion of NifD proteins is likely to be prevalent among eukaryotic organisms, and the R98P mutation could enable engineering of stable NifD in plant mitochondria.

Discussion

The possibility of engineering nitrogen fixation in crops by introducing *nif* genes into plant cells has been proposed ever since the first step of transferring *nif* genes from a diazotrophic bacterium to a nondiazotrophic bacterium was realized more than 45 y ago (35). However, this is not a facile goal, and there are many well-accepted challenges to face, including the multiplicity of genes involved, the requirement for maintaining balanced expression of Nif components, the demand for energy and reducing power, and the oxygen sensitivity of nitrogenase. The development of synthetic biology tools has enabled *nif* cluster redesign and high-throughput combinatorial libraries of gene parts to express complex combinations of Nif proteins in prokaryotes or eukaryotes (15, 29). To achieve stoichiometric expression of Nif components, a polyprotein-based strategy has been used to reduce gene numbers, resulting in a balanced expression of protein components required for nitrogenase biosynthesis and activity (36). To investigate whether plant electron transport chains can replace their prokaryotic counterparts as electron donors to support nitrogenase catalysis, plant organelle-derived electron transport modules were reconstituted in *E. coli* and found to serve as electron donors to nitrogenase (37). Since plants can potentially provide the energy and reducing power required to support nitrogenase activity, this may further reduce the gene numbers required for engineering nitrogen fixation in crops.

Plant organelles have long been considered suitable locations for nitrogenase in view of their roles in energy conversion and thus could potentially provide the high concentrations of ATP and reducing power required for nitrogenase activity (10). The respiratory activity of mitochondria and the presence of an iron-sulfur cluster assembly pathway in this organelle suggest that the mitochondrial matrix is likely to be oxygen-depleted and thus could provide a suitable environment for nitrogenase. This was confirmed by the seminal demonstration that nitrogenase Fe protein, the most oxygen-sensitive component of nitrogenase, can be purified in an fully active form from aerobically grown yeast when expressed in mitochondria (12). Thus, the oxygen problem has been solved, at least for this organelle. In contrast, as perhaps expected, lower Fe protein activity could be recovered from chloroplasts, and in this case the inclusion of the Nif-specific iron-sulfur cluster assembly components NifU and NifS was required to achieve activity under aerobic conditions (16).

Although the assembly of active Fe protein requires relatively few Nif components, the requirements for biosynthesis of MoFe protein are far more complicated, involving at the very least a minimal gene set of nine Nif proteins, with the additional requirement of supplying molybdenum, homocitrate, and S-adenosyl methionine to support the biosynthesis of FeMo-co. Although many Nif proteins have been expressed in yeast and plant mitochondria, aiming for suitable stoichiometric expression of components, solubility and stability issues have been encountered (13, 15). In some cases, it has been possible to take advantage of the biodiversity of Nif proteins originating from different diazotrophs. Using this strategy, a soluble and in vitro functional NifB from *M. infernus* has been obtained when expressed in yeast mitochondria (14, 17). However, expression and assembly of the NifDK tetramer has been problematic, possibly due to aberrant proteolytic processing of NifD. This creates a considerable barrier to achieving expression of active MoFe protein in mitochondria.

Since many proteins within organelles are encoded in the nuclear genome, engineering strategies frequently use a targeting approach, in which leader peptides are used to transport proteins expressed from nuclear transgenes into organelles. Protein targeting is particularly important when engineering mitochondria, since a stable and convenient transformation method for these organelles has not yet been developed. This is also pertinent to engineering chloroplasts of crops plants, since although plastid transformation is routine in model dicotyledonous plants, it is not well established in monocots. The organelle-targeting approach requires efficient translocation of the protein across both mitochondrial membranes and appropriate processing and cleavage of the leader peptide once the preprotein has entered the mitochondrial matrix. Usually this process results in processing of the protein to the correct size, although it is often important to trial different leader peptides to achieve accurate processing. Although we considered that a number of candidate endopeptidases could be responsible for the degradation of NifD in the mitochondrial matrix, once we had identified the region required for cleavage, we focused on the possibility that the mitochondrial processing peptidase was carrying out secondary processing of NifD close to residue R98. Although there is significant degeneracy in the amino acid sequences recognized by MPPs, the importance of R98 and the positioning of an aromatic residue two residues downstream (Y100) was suggestive of cleavage by the MPP (38). The complete conservation of the residue equivalent to R98, in combination with the highly conserved Y100 residue in NifD proteins from diverse sources, implies that the degradation problem cannot be resolved by exploiting the biodiversity of NifD proteins (38). Indeed, we found this to be so when examining the stability of NifD from four different diazotrophs in yeast mitochondria and discovered that in each case, cleavage could be prevented by constructing the equivalent R98K substitution (Fig. 1*E*).

To confirm unequivocally that the MPP was responsible for the cleavage, we reconstituted yeast MPP in *E. coli* and demonstrated that the protein carries out endoproteolytic cleavage of NifD to generate the same processed fragments observed in yeast. This activity is not specific to yeast MPP, since it was also observed with three plant-derived MPPs, with processing prevented by the NifD R98 variant in each case (Fig. 2). In yeast, both MPP subunits are localized to the matrix (39), whereas in plants, both the MPP_{α} and MPP_{β} subunits are completely integrated into the cytochrome bc_1 complex of the respiratory chain as Core1 and Core2 proteins, respectively (40). Since *E. coli* does not possess genes encoding a *cyt* bc_1 complex (41), the reconstitution of plant-type MPPs in this host potentially could affect processing activity. However, we observed that the plant-type MPPs efficiently processed the Su9 leader peptide in *E. coli* (Fig. 3*B*), supporting previous evidence that electron transfer within the *cyt* bc_1 complex is not required for processing (42).

Although we have identified an internal cleavage site for MPP within NifD, the accessibility of this cleavage site after targeting to yeast mitochondria remains unclear. Translocation across the mitochondrial membranes requires that NifD be in an unfolded state, and during this period, the protein is exposed to MPP to remove the signal peptide. We suspect that MPP also cleaves the internal site within NifD during this period, before the protein is folded into the correct form. This secondary cleavage event could occur before the interaction with NifK, thus explaining the difficulty in assembling stable NifDK heterotetramers in yeast mitochondria (15). Alternatively, it is possible that MPP recognizes NifD that is misfolded after translocation into the mitochondrial matrix. However, if this were the case, it would be difficult to rationalize how the R98P variant induces correct folding. Moreover, the protease is active against wild-type NifD

in the intact bacterial system, providing strong evidence that misfolding is not the primary cause of the degradation. Nevertheless, it is important to note that cleavage of NifD in the bacterial system does not necessarily imply that the MPP recognizes the mature form of MoFe protein replete with metalloclusters. MPP cleavage could occur at a stage before the maturation of nitrogenase MoFe protein, depending on the kinetics of assembly. For example, apo-NifDK might be cleaved by the protease, but the holo NifDK tetramer, replete with metalloclusters, may have a conformation in which the cleavage site is inaccessible. In summary, it is difficult to distinguish whether MPP cleavage at the internal site in NifD occurs in yeast mitochondria before the folding of this protein or whether the cleavage site remains accessible during maturation of the holo NifDK heterotetramer.

Plant organelles are important targets for introducing heterologous metabolic pathways or complex biological systems for plant genetic engineering. Our results highlight issues associated with the correct processing of foreign proteins targeted to organelles and pinpoint the risk of aberrant processing as a consequence of subliminal signal peptide cleavage sites present in nonhost protein sequences. Although the occurrence of such sequences is probably serendipitous, signal peptide cleavage proteases can be considered "gate keepers" for foreign proteins. Since there is considerable degeneracy in the amino acid sequences recognized by the processing peptidases, it is difficult to accurately predict whether foreign proteins will be degraded after entry into organelles. Our approach to reconstitute MPPs in *E. coli* provides a simple biological assay to preevaluate stability before engineering in eukaryotes and, if necessary, carry out directed evolution for stable expression in mitochondria.

Materials and Methods

Strains and Media. The *E. coli* Top10 strain was used for routine cloning and plasmid propagation. *E. coli* JM109 was the strain background used to test the functionality of MPPs and measure nitrogenase activity via an acetylene reduction assay. *E. coli* NCM3722 strain was used as a host for the diazotrophic growth experiments as described previously (36). Luria-Bertani broth for *E. coli* growth contains 10 g/L of tryptone, 5 g/L of yeast extract and 10 g/L of NaCl (with or without 15 g/L agar). All nitrogen fixation assays were performed in KPM minimal medium (10.4 g/L Na_2HPO_4, 3.4 g/L KH_2PO_4, 26 mg/L $CaCl_2{\cdot}2H_2O$, 30 mg/L $MgSO_4$, 0.3 mg/L $MnSO_4$, 36 mg/L ferric citrate, 10 mg/L para-aminobenzoic acid, 5 mg/L biotin, 1 mg/L vitamin B1, 0.05% casamino acids, and 0.8% [wt/vol] glucose), supplied with 10 mM ammonium sulfate (KPM-HN) for pregrowth or 0.1% glutamate (KPM-LN) for nitrogenase activity assays. Diazotrophic growth experiments were carried out with solid KPM-NN minimal medium (KPM-LN medium without 0.1% glutamate, para-aminobenzoic acid and casamino acids, containing highly purified agarose (TsingKe BioTech; TSJ001) instead of agar). Both kanamycin and chloramphenicol were used at a concentration of 25 μg/mL.

S. cerevisiae W303-1a strain (*MATα ade2-1 leu2-3,112 trp1-1 his3-11,15 ura3-1*) was used as the host to express NifDs and their variants from different diazotrophs. YDS medium for *S. cerevisiae* growth containing 20 g/L peptone, 10 g/L yeast extract, 20 g/L glucose, and 100 mg/L adenine, with or without 20 g/L agar. Solid synthetic dropout medium (13.4 g/L yeast nitrogen base [BD Biosciences; 291920], 0.69 g/L dropout mixture [-Leu Do supplement; BD Biosciences, 630414], 20 g/L glucose, and 20 g/L agar) was used to select the correct colony after transformation. For protein expression, *S. cerevisiae* W303-1a was grown at 30 °C in YPDG medium containing 20 g/L peptone, 10 g/L yeast extract, 10 g/L glucose, 10 g/L galactose, and 100 mg/L adenine.

Plasmid Construction. The plasmids used in this study are listed in Dataset S1. All plasmids were verified by sequencing before use in further experiments. NifD encoding sequences used for expression in yeast were chemically synthesized by GenScript according to the codon bias of S. *cerevisiae* nuclear genes; sequences are provided in Dataset S2. Plasmids used for the expression of NifDs and their variants were constructed by assembling the Sc GAL1 promoter, Su9-*nifD* sequences, and Sc ADH2 terminator using Golden Gate assembly with pBDS1549 as a vector. The NifD deletion mutants or single amino acid replacement mutants were constructed by dividing the NifD encoding sequences into two parts and assembling them with Sc GAL1 promoter and ADH2 terminator using Golden Gate assembly (*SI Appendix*, Fig. S1). For *K. oxytoca* NifD-R98 saturation amino acid replacement, we used pBDS1091 (carrying the *nifHDKTY* operon from *K. oxytoca*) as a template to perform PCR to build a pool of P_{nifH}-*nifHD*$_n$ fragments and obtain the *nifD*$_c$*KTY* fragment (details in *SI Appendix*, Fig. S1). To introduce random mutations in the R98 site of *nifD*, we used a forward primer with NNK at the site of R98 in the NifD coding sequence. The P_{nifH}-*nifHD*$_n$ fragments and *nifD*$_c$*KTY* fragement were random fused and inserted into the pBDS402K vector by Golden Gate assembly.

The coding sequences of MPPs (α subunit labeled with HA-tag and β subunit labeled with His-tag) were also chemically synthesized by GenScript; sequences are provided in Dataset S2. Each of the MPP subunits was amplified by PCR, and a ribosome-binding site (RBS) was added. The promoter element ($P_{tac/lacO}$), MPP_β and MPP_α subunits, and terminator element (T_{rrnB}) were assembled by Golden Gate assembly to form an operon in the pBDS402K vector. To test the effectiveness of the reconstituted MPPs in the *E. coli* host, we used Su9-GFP as a reporter system, expressed under the control of the $P_{LtetO-1}$ promoter and integrated into the pBDS1549 vector.

To construct the Flag-tag or His-tag labeled Nif components to allow coexpression with MPPs, Flag- or His-tag coding sequences were added to the 5′ or 3′ end of each gene by PCR. Subsequently, the PCR products were assembled with the *K. oxytoca nifH* promoter and artificial terminator L3S2P21 (43) by the Golden Gate assembly method with pBDS1549 as a vector to generate the expression cassette. The promoter element (P_{tet}), *K. oxytoca nifA* coding sequence, and terminator elements (T_{rrnB}) were assembled to form the NifA expression cassette in the vector pBDS402K by Golden Gate assembly (Dataset S2 and *SI Appendix*, Fig. S8). Subsequently, plasmids carrying the MPP expression cassette and the NifA expression cassette were further assembled in the pBDS1609 vector by Golden Gate assembly (Dataset S2 and *SI Appendix*, Fig. S8) to generate MPP/NifA dual-expression plasmids.

Yeast Transformation and Mitochondria Extraction. For yeast transformation, *S. cerevisiae* W303-1a was first grown in YPD medium at 30 °C. Yeast transformations were carried out according to the lithium acetate method (44), and transformants were selected on solid synthetic dropout medium. For mitochondria extraction, *S. cerevisiae* W303-1a and its derivative strains were grown in flasks in YPDG medium at 30 °C and 200 rpm for 36 h. Yeast cells were collected, and mitochondrial enriched protein extracts were basically prepared as described previously (45). In brief, yeast cells were collected, washed with ddH_2O, and resuspended in specific zymolyase buffer with an appropriate amount of zymolyase to convert the cells to spheroplasts. Subsequently, the spheroplasts were broken using a glass homogenizer, and total extracts were centrifuged at 1,500 × *g* for 5 min to remove the cellular debris and unbroken cells. Then the supernatants were centrifuged twice at 4,000 × *g* for 5 min to remove other needless cellular contents. The supernatants from the last step were centrifuged at 12,000 × *g* for 15 min to collect the crude mitochondria. Precipitates (mitochondria pellets) from 10 mL of yeast cells (~0.5 g wet weight) were resuspended in 200 μL of PBS buffer. After boiling for 20 min, samples were cooled to room temperature and then centrifuged at 12,000 rpm for 2 min. Western blot assays were run using 20 μL of each sample.

Acetylene Reduction and Diazotrophic Growth. The C_2H_2 reduction assay was modified from a previously described method (46). To measure the activity of *E. coli* JM109 derivatives, cells were initially grown overnight in KPM-HN medium. The cells were then diluted into 2 mL of KPM-LN medium in 20-mL sealed tubes to a final OD_{600} of ~0.3, with or without 200 μM isopropyl-β-D-thiogalactoside (IPTG) to induce the expression of MPPs. For the acetylene reduction assay, air in the tubes was repeatedly evacuated and replaced with argon. After incubation at 30 °C for 4 h, 2 mL of C_2H_2 was injected, and the gas phase was analyzed ~16 h later with a Shimadzu GC-2014 gas chromatograph. Data presented are mean values based on at least two biological replicates. Immediately after the acetylene reduction assay, 4 mL (two parallel samples, OD_{600} = ~1.0) of *E. coli* cells carrying the appropriate plasmids were collected and suspended in 300 μL of PBS buffer supplied with 75 μL of 5× SDS loading buffer (Sangon; C508320). After boiling for 20 min, samples were cooled to room temperature and centrifuged at 12,000 rpm for 2 min. Western blot assays were performed using 10 to 25 μL of each sample (according to the sensitivity of the antibody for each Nif protein).

For diazotrophic growth, reconstructed plasmids were transformed into the *E. coli* NCM3722 strain, and the transformants were spread on LB plates with appropriate concentration of antibiotics. After incubation at 37 °C for

MICROBIOLOGY

16 h, single colonies were picked and streaked onto KPM-NN plates. Then the plates were moved to a 2.5-L anaerobic jar (Oxoid AG0025A; Thermo Fisher Scientific) equipped with anaerobic gas-generating sachets (Oxoid AN0025A; Thermo Fisher Scientific) and an oxygen indicator (Oxoid BR0055B; Thermo Fisher Scientific). Anaerobic jars were immediately locked and incubated at 30 °C for 3 to 4 d.

Stability Assay of the Tag-Labeled Nif Components. The tag-labeled Nif protein expression plasmids were cotransformed with the MPP/NifA dual-expression plasmids into *E. coli* JM109. The transformants were inoculated in LB medium with an appropriate concentration of IPTG (0, 50, or 200 μM) for inducing the expression of MPP at 30 °C and 200 rpm for 24 h. Then 1 mL of *E. coli* cells were collected and suspended in 300 μL of PBS buffer supplied with 75 μL of 5× SDS loading buffer (Sangon; C508320). After boiling for 20 min, samples were cooled to room temperature and then centrifuged at 12,000 rpm for 2 min. Western blot assays were run with 20 μL of each sample.

Western Blot Assays. Samples were loaded on 10% SDS-polyacrylamide gels (Thermo Fisher Scientific; NP0301, NP0302) with 6 μL of PageRuler Prestained Protein Ladder (Thermo Fisher Scientific; 26616) as a marker. Proteins on the gels were subsequently transferred to PVDF membranes (Thermo Fisher Scientific; IB24002) using the iBolt 2 (Thermo Fisher Scientific). The membranes were blocked with 5% skim milk (BD Difco 232100; BD Biosciences) in PBS buffer. The antibodies for each Nif protein were used at a dilution of 1:500 to 1:3,000 (according to the sensitivity of the antibody for each Nif protein). The secondary antibody goat anti-rabbit IgG-HRP (ZSGB Biotech; ZB-2301) or goat anti-mouse IgG-HRP (ZSGB Biotech; ZB-2305) was used at 1:3,000 dilution. Development was done by using two different methods: using an enhanced chemiluminescent substrate for HRP (Thermo Fisher Scientific; 34077) and captured by an automatic chemiluminescence imaging analysis system (Tanon-4200) or directly developed on the PVDF membranes using the DAB Chromogen/HRP Substrate Kit (ZSGB Biotech; ZLI-9018). The antibodies used for immunoblotting and quantification in this study were as described previously (36).

Data and Materials Availability. All data are available in the main text, *SI Appendix*, and Datasets S1 and S2. Plasmids and strains constructed in this study, listed in the main text and the *SI Appendix*, are available on request.

ACKNOWLEDGMENTS. We thank Jilun Li (China Agriculture University) for providing the antisera against nitrogenase proteins and Weifeng Liu (Shandong University) for providing the *S. cerevisiae* W303-1a strain. This research was supported by National Key R&D Program of China (research grant 2019YFA09004700), the National Science Foundation of China (research grant 31530081), startup funding from the Peking University School of Advanced Agricultural Science, the SLS-Qidong Innovation Fund, and the State Key Laboratory of Protein and Plant Gene Research (research grant B02). Y.-P.W. is a recipient of a grant from the National Science Fund for Distinguished Young Scholars (grant 39925017). R.D. was funded by the UK Research and Innovation Biotechnology and Biological Sciences Research Council (research grant BBS/E/J/000PR9797) and the Royal Society (International Collaboration Award ICA\R1\180088).

1. M. M. Georgiadis *et al*., Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science* **257**, 1653–1659 (1992).
2. L. M. Rubio, P. W. Ludden, Maturation of nitrogenase: A biochemical puzzle. *J. Bacteriol.* **187**, 405–414 (2005).
3. P. C. Dos Santos *et al*., Iron-sulfur cluster assembly: NifU-directed activation of the nitrogenase Fe protein. *J. Biol. Chem.* **279**, 19705–19711 (2004).
4. K. S. Howard *et al*., *Klebsiella pneumoniae nifM* gene product is required for stabilization and activation of nitrogenase iron protein in *Escherichia coli*. *J. Biol. Chem.* **261**, 772–778 (1986).
5. L. M. Rubio, P. W. Ludden, Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu. Rev. Microbiol.* **62**, 93–111 (2008).
6. Y. Hu, M. W. Ribbe, Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *J. Biol. Chem.* **288**, 13173–13177 (2013).
7. E. Jimenez-Vicente *et al*., The NifZ accessory protein has an equivalent function in maturation of both nitrogenase MoFe protein P-clusters. *J. Biol. Chem.* **294**, 6204–6213 (2019).
8. S. Hill, E. P. Kavanagh, Roles of *nifF* and *nifJ* gene products in electron transport to nitrogenase in *Klebsiella pneumoniae*. *J. Bacteriol.* **141**, 470–475 (1980).
9. L. Curatti, L. M. Rubio, Challenges to develop nitrogen-fixing cereals by direct nif-gene transfer. *Plant Sci.* **225**, 130–137 (2014).
10. P. H. Beatty, A. G. Good, Plant science. Future prospects for cereals that fix nitrogen. *Science* **333**, 416–417 (2011).
11. N. B. Ivleva, J. Groat, J. M. Staub, M. Stephens, Expression of active subunit of nitrogenase via integration into plant organelle genome. *PLoS One* **11**, e0160951 (2016).
12. G. López-Torrejón *et al*., Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* **7**, 11426 (2016).
13. R. S. Allen *et al*., Expression of 16 nitrogenase proteins within the plant mitochondrial matrix. *Front Plant Sci* **8**, 287 (2017).
14. S. Burén, X. Jiang, G. López-Torrejón, C. Echavarri-Erasun, L. M. Rubio, Purification and in vitro activity of mitochondria targeted nitrogenase cofactor maturase NifB. *Front Plant Sci* **8**, 1567 (2017).
15. S. Burén *et al*., Formation of nitrogenase NifDK tetramers in the mitochondria of *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **6**, 1043–1055 (2017).
16. Á. Eseverri *et al*., Use of synthetic biology tools to optimize the production of active nitrogenase Fe protein in chloroplasts of tobacco leaf cells. *Plant Biotechnol. J.*, 10.1111/pbi.13347 (2020).
17. S. Burén *et al*., Biosynthesis of the nitrogenase active-site cofactor precursor NifB-co in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25078–25086 (2019).
18. L. Van Dyck, T. Langer, ATP-dependent proteases controlling mitochondrial function in the yeast *Saccharomyces cerevisiae*. *Cell. Mol. Life Sci.* **56**, 825–842 (1999).
19. P. M. Quirós, T. Langer, C. López-Otín, New roles for mitochondrial proteases in health, ageing and disease. *Nat. Rev. Mol. Cell Biol.* **16**, 345–359 (2015).
20. O. Gakh, P. Cavadini, G. Isaya, Mitochondrial processing peptidases. *Biochim. Biophys. Acta (BBA) Mol. Cell Res.* **1592**, 63–77 (2002).
21. P. F. Teixeira, E. Glaser, Processing peptidases in mitochondria and chloroplasts. *Biochim. Biophy. Acta (BBA) Mol. Cell Res.* **1833**, 360–370 (2013).
22. M. Yang, R. E. Jensen, M. P. Yaffe, W. Oppliger, G. Schatz, Import of proteins into yeast mitochondria: The purified matrix processing protease contains two subunits which are encoded by the nuclear MAS1 and MAS2 genes. *EMBO J.* **7**, 3857–3862 (1988).
23. G. Isaya, F. Kalousek, L. E. Rosenberg, Sequence analysis of rat mitochondrial intermediate peptidase: Similarity to zinc metallopeptidases and to a putative yeast homologue. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8317–8321 (1992).
24. A. Naamati, N. Regev-Rudzki, S. Galperin, R. Lill, O. Pines, Dual targeting of Nfs1 and discovery of its novel processing enzyme, Icp55. *J. Biol. Chem.* **284**, 30200–30208 (2009).
25. W. Voos, Chaperone-protease networks in mitochondrial protein homeostasis. *Biochim. Biophys. Acta* **1833**, 388–399 (2013).
26. B. Westermann, W. Neupert, Mitochondria-targeted green fluorescent proteins: Convenient tools for the study of organelle biogenesis in *Saccharomyces cerevisiae*. *Yeast* **16**, 1421–1427 (2000).
27. P. C. Dos Santos, Z. Fang, S. W. Mason, J. C. Setubal, R. Dixon, Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* **13**, 162 (2012).
28. B. Soboh, E. S. Boyd, D. Zhao, J. W. Peters, L. M. Rubio, Substrate specificity and evolutionary implications of a NifDK enzyme carrying NifB-co at its active site. *FEBS Lett.* **584**, 1487–1492 (2010).
29. X. Wang *et al*., Using synthetic biology to distinguish and overcome regulatory and functional barriers related to nitrogen fixation. *PLoS One* **8**, e68677 (2013).
30. J. Christiansen, V. L. Cash, L. C. Seefeldt, D. R. Dean, Isolation and characterization of an acetylene-resistant nitrogenase. *J. Biol. Chem.* **275**, 11459–11464 (2000).
31. S. M. Mayer, D. M. Lawson, C. A. Gormal, S. M. Roe, B. E. Smith, New insights into structure-function relationships in nitrogenase: A 1.6-Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe protein. *J. Mol. Biol.* **292**, 871–891 (1999).
32. B. Schmid *et al*., Structure of a cofactor-deficient nitrogenase MoFe protein. *Science* **296**, 352–356 (2002).
33. C. Witte, R. E. Jensen, M. P. Yaffe, G. Schatz, MAS1, a gene essential for yeast mitochondrial assembly, encodes a subunit of the mitochondrial processing protease. *EMBO J.* **7**, 1439–1447 (1988).
34. R. E. Jensen, M. P. Yaffe, Import of proteins into yeast mitochondria: The nuclear MAS2 gene encodes a component of the processing protease that is homologous to the MAS1-encoded subunit. *EMBO J.* **7**, 3863–3871 (1988).
35. R. A. Dixon, J. R. Postgate, Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature* **237**, 102–103 (1972).
36. J. Yang *et al*., Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8509–E8517 (2018).

37. J. Yang, X. Xie, M. Yang, R. Dixon, Y.-P. Wang, Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2460–E2465 (2017).
38. R. Allen *et al*., Engineering a functional NifDK polyprotein resistant to mitochondrial degradation. *bioRxiv*. 10.1101/755116 (3 September 2019).
39. G. Hawlitschek *et al*., Mitochondrial protein import: Identification of processing peptidase and of PEP, a processing enhancing protein. *Cell* **53**, 795–806 (1988).
40. H.-P. Braun, U. K. Schmitz, Are the "core" proteins of the mitochondrial bc1 complex evolutionary relics of a processing protease? *Trends Biochem. Sci.* **20**, 171–175 (1995).
41. B. Schoepp-Cothenet *et al*., Menaquinone as pool quinone in a purple bacterium. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8549–8554 (2009).
42. A. C. Eriksson, S. Sjöling, E. Glaser, Characterization of the bifunctional mitochondrial processing peptidase (MPP)/bc1 complex in *Spinacia oleracea*. *J. Bioenerg. Biomembr.* **28**, 285–292 (1996).
43. Y.-J. Chen *et al*., Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **10**, 659–664 (2013).
44. R. D. Gietz, R. H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).
45. K. Diekert, A. I. P. M. de Kroon, G. Kispal, R. Lill, "Isolation and subfractionation of mitochondria from the yeast *Saccharomyces cerevisiae*" in *Methods in Cell Biology*, L. A. Pon, E. A. Schon, Eds. (Academic Press, 2001), Vol. 65, pp. 37–51.
46. F. C. Cannon, R. A. Dixon, J. R. Postgate, Derivation and properties of F-prime factors in *Escherichia coli* carrying nitrogen fixation genes from *Klebsiella pneumoniae*. *J. Gen. Microbiol.* **93**, 111–125 (1976).